UNIVERSITÉ DE NEUCHÂTEL
INSTITUT DE MICROTECHNIQUE

DATA PROCESSING OF A NAVIGATION MICROSYSTEM

Catherine Marselli

THÈSE

PRÉSENTÉE À LA FACULTÉ DES SCIENCES
POUR OBTENIR LE GRADE DE DOCTEUR ÈS SCIENCES

Décembre 1998

EN COTUTELLE AVEC L'UNIVERSITÉ DE FRANCHE-COMTÉ (FRANCE)
DANS LE CADRE DU PROJET PICS (Projet International de Coopération Scientifique):
"RECHERCHE DES LIMITES EN MICROTECHNOLOGIES ET MICROSYSTÈMES"

IMPRIMATUR POUR LA THÈSE

Data processing of a navigation microsystem

de Mme Catherine Marselli

UNIVERSITÉ DE NEUCHÂTEL

FACULTÉ DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel sur le rapport des membres du jury,

MM. F. Pellandini (directeur de thèse),N. de Rooij, H.-P. Amann et D. Hauden (Besançon)

autorise l'impression de la présente thèse.

Neuchâtel, le 1er avril 1999

Le doyen:

F. Stoeckli

F. Stoeckli

Abstract

This research is part of a Swiss French academic project whose goal was the determination of some limits in the design and use of microtechnologies and microsystems, using as a common thread example a navigation system based on microaccelerometers and angular rate microsensors (gyros). The entire project was divided into four parts, including design at the component level as well as at the system level. This PhD report describes the data processing of the navigation microsystem realised at the Electronics and Signal Processing Laboratory of the Institute of Microtechnology, University of Neuchâtel.

Current low-cost microsensors are less expensive but less accurate that mechanical or optical sensors. In a navigation system, the accelerometer and gyro outputs are integrated, leading to the accumulation of the errors. Thus, the measured trajectory becomes quickly wrong and a corrective system has to be designed. Hence, the goals of the data processing system is to compute the navigation parameters (position, velocity, orientation) while preventing the trajectory from diverging, following two approaches:

- reducing the sensor errors,
- updating regularly the trajectory using an aiding navigation system.

These two objectives are realised through Kalman filtering. The Kalman filter is an optimal estimator of the state of a system. It is formed of a set of recursive equations and requires

a state description of the system model and of the measurement model.

The first Kalman filter, designed in this project, is a novel algorithm suited to eliminate the random drift of the sensor scale factor and bias, provided a stochastic model is known. It is applied to reduce the random drift of an in-house developed gyro as well as of an off-the-shelf gyro.

The second Kalman filter is used to compute the trajectory while coupling the inertial data with those of an aiding system. Two navigation experiments have been carried out. First, a scale train trajectory is measured using one microaccelerometer and one gyro and updated using optical references, whose positions are known. This simple experiment allows studying how the sensor errors influence the trajectory shape. The second experiment is automotive navigation based on inertial sensors and DGPS (Differential Global Positioning System). In this case, the DGPS prevents the trajectory from diverging while the inertial system replaces the GPS, when the satellite emitted signals are not available (e.g., tunnels, trees). The study has also considered the influence of the sensor number on the trajectory accuracy. The usual configuration of two accelerometers (along the lateral and longitudinal axis) suitable for 2D navigation has been modified by removing the lateral accelerometer and changing accordingly the navigation algorithm. By removing a sensor, this configuration reduces cost and complexity and allows better accuracy of the trajectory measured.

As a conclusion, the sensor has nowadays to be considered as one component of a system rather than as the key item of the application. The question of the limits has to be answered considering not only the limits of the sensor itself but also the limits due to the configuration of the system as well as those of the signal processing.

Résumé

Ce travail de recherche s'inscrit dans un projet académique franco-suisse visant à déterminer les limites des microtechnologies et des microsystèmes à travers l'exemple d'un système de navigation inertiel basé sur des microaccéléromètres et des microcapteurs de vitesse angulaire (gyros). Il comprend quatre volets allant de la conception des composants au développement du système de traitement. La présente thèse réalisée à l'Institut de Microtechnique (Université de Neuchâtel, Suisse) concerne le traitement de l'information du microsystème de navigation.

Les microcapteurs actuels sont moins chers mais moins précis que les capteurs mécaniques ou optiques classiques. Dans un système de navigation, le signal de sortie des accéléromètres et des gyroscopes est intégré conduisant à l'accumulation des erreurs dans le temps. Ainsi, en l'absence de correction, la trajectoire mesurée devient rapidement fausse. Le rôle du système de traitement est donc de calculer les paramètres de navigation (position, vitesse et orientation) et de limiter l'erreur de trajectoire suivant deux moyens : réduire les imperfections des capteurs et recaler régulièrement la trajectoire en utilisant un autre moyen de navigation. Ces différents objectifs sont réalisés par des filtres de Kalman. Le filtre de Kalman est un estimateur optimal de l'état d'un système. Il se présente sous la forme d'un ensemble d'équations récurrentes et nécessite une description d'état du système et des mesures.

Le premier filtre de Kalman, développé dans ce travail, est un nouvel algorithme conçu pour éliminer les erreurs aléatoires du facteur d'échelle et du biais des capteurs, à condition d'en avoir un modèle statistique. Il est appliqué à la correction de la dérive aléatoire du capteur gyroscopique développé dans le cadre du projet ainsi qu'à celle d'un capteur du commerce.

Le deuxième filtre de Kalman est conçu pour calculer la trajectoire en fusionnant les informations inertielles avec celles d'un autre moyen de navigation. Son fonctionnement est illustré dans le cas de deux expériences. La trajectoire d'un train électrique est calculée d'après un microaccéléromètre et un microgyroscope et mise à jour par des repères optiques aux positions connues. Cette expérience simple permet notamment d'étudier comment les imperfections des capteurs influencent la forme de la trajectoire. La deuxième expérience concerne la navigation automobile basée sur les capteurs inertiels et un récepteur GPS (Global Positioning System). Dans ce cas, le GPS évite la divergence de la trajectoire alors que le système inertiel prend le relais lorsque le signal émis par les satellites n'est pas accessible (obstruction par des bâtiments, arbres,...).

L'étude a aussi considéré l'influence du nombre de capteurs sur la qualité de la trajectoire. La configuration usuelle de deux accéléromètres selon les axes longitudinal et latéral, en plus des gyros, a été modifiée en supprimant l'accéléromètre latéral et en adaptant en conséquence l'algorithme de navigation. Cette solution économise un capteur et donne de meilleurs résultats d'après les expériences menées.

En conclusion, l'amélioration en précision d'un système basé sur des capteurs ne s'obtient que par une réflexion au niveau du système et non seulement des composants. La précision intrinsèque du capteur rentre bien sûr en compte, mais est à considérer par rapport aux performances potentielles du traitement du signal. La configuration du système est aussi importante

Notations and abbreviations

Notations and abbreviations used throughout this thesis are listed in the following. A same letter may unfortunately have different meanings. However, the context and additional explanations in the text should avoid confusion. Concepts, which are defined in the text, are notified with the chapter concerned.

Notations

Vectors

a	acceleration
B	magnetic field
d	shift vector
e	unit vector
f	specific force
g	gravity field
i, j, k	unit vectors of a cartesian frame
m	mean vector
v	velocity
q	quaternion (Chapter 4)
v, w	white noise
x	position
x	state vector (Chapter 3)
z	measurement/observation vector (Chapter 3)

θ	angular position
ω	angular rate
Θ	skew-matrix of the angular vector

Scalars

a	semi major axis of an ellipse
b	semi minor axis of an ellipse
b	sensor bias (Chapter 2)
c	speed of light
h	co-ordinate height
k	sensor scale factor (Chapter 2)
K	Kalman gain (Chapter 3)
m	mass
p	pitch angle (Chapter 2)
r	roll angle (Chapter 2)
s	signal
x,y,z	position co-ordinate
y	yaw angle (Chapter 2)
y	output signal
Δt	time increment
α	Euler angle (Chapter 4)
α, ς, β	correlation time of a first-order Gauss-Markov (Chapter 3)
α, β, γ	rotation angles in the Helmert transformation (Chapter 6)
λ	geodetic longitude (Chapter 6)
μ	scale factor in the Helmert transformation (Chapter 6)
ϕ	geodetic latitude (Chapter 6)
ρ	pseudo-range (Chapter 6)
σ^2	variance

Matrices/functions

C	input-coupling matrix (Chapter 3)
D	input/output-coupling matrix (Chapter 3)
F/f	dynamic matrix/function (Chapter 3)

H/h	observation matrix/function (Chapter 3)
I	unit matrix
P	error covariance matrix (Chapter 3)
Q	state noise covariance matrix (Chapter 3)
R	measurement noise covariance matrix (Chapter 3)
R	rotation matrix
Y	transfer function
Φ	state transition matrix (Chapter 3)

Subscripts/Superscripts

a	accelerometer
b	body-fixed frame
c	Coriolis phenomenon (Chapter 2)
e	earth-fixed frame
g	angular rate sensor
i	input value
k	discrete time index
m	measured value
o	nominal value
w	WGS-84
x	x axis
y	y axis
z	z axis
^	estimated value
~	error value
(+)	a posteriori estimate (Chapter 3)
(-)	a priori estimate (Chapter 3)

Abbreviations

ARMA	Autoregressive Moving Average (Chapter 3)
c(x)	cos(x)
s(x)	sin(x)
DGPS	Differential GPS
DoD	Department of Defense, USA
DoT	Department of Transportation, USA
DPRAM	Dual-Port RAM (Appendix C)

DSP	Digital Signal Processor (Appendix C)
FOG	Fibre Optical Gyroscope (Chapter 2)
GPS	Global Positioning System (Chapter 6)
Gyro	angular rate sensor
INS	Inertial Navigation System
LPMO	Laboratoire de Physique et de Métrologie des Oscillateurs, Université de Franche-Comté, Besançon (F) (Chapter 1)
NAVSTAR	NAVigation System by Timing And Ranging (Chapter 6)
NMEA	National Marine Electronics Association (Chapter 6)
PICS	Projet International de Coopération Scientifique (Chapter 1)
PRN	Pseudorandom noise (Chapter 6)
RDS	Radio Data System (Chapter 6)
RLG	Ring Laser Gyroscope (Chapter 2)
RTCM	Radio Technical Commission for Maritime Services (Chapter 6)
SA	Selective Availability (Chapter 6)
SRAM	Static RAM
UTC	Universal Time Co-ordinated
WGS-84	World Geodetic System 1984 (Chapter 6)
2D	2 dimensions
3D	3 dimensions
N(0,x)	normal density

Miscellaneous

Web addresses are of September 1998

cc	hundredth of a grade
nmi	nautical mile

Table of Contents

Chapter 1
Introduction

This Ph.D. research is part of a project whose goal is the study of a navigation system based on inertial microsensors (microaccelerometers and angular rate microsensors) in order to determine some limits in the design and use of microsensors.

1.1. The PICS project: goals and partners

This work was part of a *PICS* project. PICS, standing for the French shortening *Projet International de Coopération Scientifique* (International Project of Scientific Cooperation), refers to projects, initiated by the French government, including a collaboration with a foreign country.

This PICS project involved Swiss and French academic institutions active in the field of microtechnology. These are:

- the *Institute of Microtechnology* (*IMT*), *University of Neuchâtel*, with both the *Sensors, Actuators and Microsystems Laboratory* (*SAMLAB*) and the *Electronics and Signal Processing Laboratory* (*ESPLAB*), located in Neuchâtel (Switzerland),
- the *Centre National de la Recherche Scientifique* (*CNRS*) with the *Laboratoire de Physique et Métrologie des Oscillateurs* (*LPMO*), which is attached to the *University of Franche-Comté*, located in Besançon (France),

- the *Swiss Federal Institute of Technology – Lausanne (EPFL)* with the *Institute of Microsystems (IMS)*, located in Lausanne (Switzerland).

The project was divided into four parts, giving rise to four different thesis works. It included design from component level to system level. Firstly, a silicon angular rate microsensor at IMT-SAMLAB and a silicon multi-axial microaccelerometer at CNRS-LPMO have been designed, fabricated and characterised. Techniques for the simulation of the sensor behaviour were developed at EPFL-IMS. The data processing of the navigation microsystem was designed at the IMT-ESPLAB and is described in this Ph.D. report.

The term *microsystem* in the title refers to a system based on microsensors rather than to a complete miniaturised system.

To restrain the problem, the navigation type of interest was rapidly chosen as automotive navigation. The motivation in studying microsensors, especially in the case of automotive navigation, is explained in the following.

1.2. Use of inertial microsensors in automotive applications

Recent progress in semiconductor design has led to an increasing symbiosis between mechanical and electronic devices, leading to the advent of microsensors. Thus, acceleration and later angular rate microsensors have become commercially available. On the contrary of classical inertial sensors, they are relatively inexpensive, robust and small. Hence, while older technology were met only in particular costly high-end applications (airborne, military), modern devices allow potentially the development of new applications as well as the improvement of existing applications, especially by reducing cost and size. For example, microaccelerometers and angular rate microsensors have a great interest in:

- consumer applications, such as free-space pointers, remote control devices, anti-jitter for video cameras,
- industrial applications, such as machine control, stabilised platforms,

- medical applications, such as wheel-chair navigation, monitoring of body movement.

However, there is no doubt that the main field, which drives inertial microsensor development, is automotive applications, where electronics takes a more and more important part. On one hand, electric wires, sensors and actuators replace mechanical linkages. On the other hand, electronic systems based on sensors enable new functions for safe driving.

As high-g sensitive accelerometers came before, one of the first microsensor applications was shock detection in the airbag systems. With the advent of low-g and angular rate microsensors, many other applications were developed [John95], [Macd90], [Reid95]. Inertial microsensors are used in *ABS* (*Advanced/Anti-lock Braking Systems*): when the car slips, wheel speed sensors (*odometers*) give no longer a valid information on the contrary of inertial sensors, which enable to correct the motion of the car and avoid wheel lock. Inertial microsensors are also used in four wheel steering systems to compare and correct the motion of each wheel. They also improve the passenger's comfort in active suspension systems and levelling systems. Finally, the most recent navigation systems include also angular rate microsensors.

Navigation systems

Navigation systems have been commercially available for a few years. They estimate the car position and give advice to the driver regarding the desired destination. What seems to be a gadget for the private user represents time – and sometimes money! – saving for delivery, home and emergency services. Besides, navigation systems improve the safety in the case of bad weather conditions and facilitate driving in unknown roads. Thus, hired cars are often equipped with such systems.

Bosch was a pioneer by developing in 1978 a first prototype (*EVA*) and in 1989 the *Travel Pilot*. However, commercial devices became available only in 1994, first from high-end car manufacturers (*BMW*, *Mercedes*), and can be obtained nowadays from almost all manufacturers. Generally, car manufacturers do not build their own systems but subcontract electronic designers. Two of the most common systems are the

Travel Pilot from *Bosch-Blaupunkt*, which is proposed by *Mercedes*, *Ford*, *Volkswagen* and other, and the *Carin* from *Philips*, which is proposed by *Opel*, *BMW*, *Renault*, *Nissan* and other.

What made possible these developments is, of course, the development of the GPS (Global Positioning System) (see Chapter 6) but clearly also the advent of cheap, but powerful, processors as well as data storage facilities.

Current navigation systems always include a GPS receiver, which gives the starting point of the drive and provides absolute positions of the car with a basic accuracy of approximately 100 m. This information is no longer available downtown due to the distortions of the satellite signals by trees, buildings and tunnels (urban canyon). Thus, the translation motion information relies mainly on the odometers.

Orientation information is also necessary to know the heading of the car. Former systems used an electronic compass. Due to interference problems, compasses have been replaced by differential odometers. Recent systems include an angular rate microsensor, which simplifies the system mounting and calibration.

The information given by the sensors is fused to compute the car position. However, to reach the required accuracy (less than 20 m), systems use *map-matching* techniques, which consist of comparing the position measured with possible locations on a map database in order to find out the actual car position.

Finally, these systems include *route planning*, which consists of defining the best way from the current position to the final destination regarding the driver needs (e.g., rapid drive, panoramic drive, gas stations). [Zhao97] gives an overview of the techniques used in automotive navigation systems.

All the commercially available systems give satisfactory results and are more or less equivalent in terms of accuracy. Just the human-machine interface is worth comparing. They are relatively expensive, ranging from 1500 US$ to 4500 US$, depending on the product and on whether they are sold as accessories or factory mounted. Additional costs such as system

mounting, calibration as well as map updates are also significant.

As previously mentioned, angular rate microsensors are now used in current navigation systems. Accelerometers could potentially replace odometers, reducing the cost of mounting and calibration. Besides, provided inertial microsensors were accurate enough, they could enable a navigation solution without using sophisticated map matching algorithms, whose computational burden is still too heavy for high-speed navigation applications.

Orientation of the present research work

The goal of this work was not improving current automotive navigation systems. Otherwise, techniques such as map matching would have been considered. This project aimed at determining some limits of microsensors and this application was just used as a common thread example, which allowed delimiting the research field.

1.3. State of the art

Navigation is an old human problem and early inertial sensors are old technology. As old as it is, it has been the province of small groups of specialists, probably because of the particularities of the applications concerned (airborne and unfortunately military). Thus, the research has been done in very confidential circles. However, due to the advent of GPS and new inertial technologies, the subject becomes democratised. Next to industrial product brochures, which teach nothing, and academic literature, which is impenetrable after sixty years of development behind closed doors, it is now cheerful to notice the very recent augmentation of comprehensive books on the subject (see references throughout the text).

Inertial sensors enable measuring the acceleration and the angular rate of a body with respect to an inertial frame (see Chapter 2). They have always been a key part of navigation systems due to their inherent characteristics. They are self-

contained and require no information on the environment. Thus, they are indifferent to weather conditions and can be used even in unsurveyed areas. They cannot be jammed or interfered with and radiate no energy. Finally, they can be used anywhere, including the entire surface of the globe, even underground, undersea as well as in the air, provided gravitation field is known.

Although an inertial navigation system is conceptually a self-contained device, it cannot be used in a stand-alone mode. As a matter of fact, since the sensors provide rate information, the navigation system exhibits a drift with time due to the integration of sensor non-idealities and must be used in conjunction with other navigation systems.

For fifteen years, the additional system of interest has been the GPS. The INS-GPS association seems to be optimal because the GPS is used to calibrate the INS whereas the INS replaces GPS during signal jams. One of the most popular sensor fusion techniques is Kalman filtering. First academic papers on the subject concerned surveying applications (see the publications of the *Department of Geomatics and Engineering - University of Calgary*, Ca). With the advent of DGPS (Differential GPS) (see Chapter 6), the trend goes toward giving up INS in geodesy. Nevertheless, more recently, the development of microsensors made the subject again a current research topic.

The main difference with respect to past time developments is the computer revolution. On one hand, it had a profound effect on sensor technology. Analysis and simulation tools facilitate better initial design and subsequent refinement. On the other hand, the sensor may be less accurate since an intelligent signal processing has the potentiality to compensate errors. This is a main feature of modern inertial technology: the sensor is not the unique central piece. Electronics and signal processing are also of importance and have to be designed as much as possible in connection with the sensor. The research has now to be done at the system level rather than at the component level. The PICS project original goal was based on this consideration. On one hand, it aimed at mingling people active in sensor design with people involved in signal processing. On the other hand, it supported the development of

new approaches in both the sensor and system design regarding the new possibilities of computing.

1.4. Organisation of the report

Chapter 2 is devoted to inertial sensors. It details principles of measuring acceleration and angular rate and gives also the basis of navigation based on inertial sensors. While mentioning classical devices, it focuses on microsensor design. Besides, it presents the main features of the sensors used in the project. Two are those realised in the PICS project and two are off-the-shelf sensors.

Chapter 3 is a theoretical chapter dedicated to Kalman filtering. Navigation is definitely a multi-sensor applications and fusion methods are of concern, whatever the sensors used. At the beginning of the project, it was decided to use a Kalman filter, as it was usually done with older technologies, even if microsensors are quite different from classical sensors. Chapter 3 does not contain a complete derivation of the filter but tries to present the main aspects of Kalman filtering, which are useful to understand the design of the signal processing involved in this research. It is disconnected from the rest of the report and can be skipped by the reader who is already familiar with Kalman filtering or who is not interested in the theoretical details and in the filter implementation.

Chapter 4 and 5 present the design of two Kalman filters. The first one – referred to as *navigation filter* – blends inertial information with an aiding source, in order to output the navigation results. For this purpose, Chapter 4 develops firstly inertial navigation equations. Starting from the general case of 3D motion, it proposes also a configuration for 2D motion with attitude changes, involving a reduced set of sensors and allowing better accuracy of the trajectory measured.

The second filter is a new algorithm designed to correct sensor errors, such as the offset drift, independently of an aiding source. Thus, it is called *local filter*. As mentioned above, classical methods have to be modified to suit technology changes (less accurate sensors, more powerful electronics). Thus, a method to correct a sensor error with intelligent

processing is proposed. Kalman filtering has become naturally the ground of this algorithm since it has been studied because of the navigation problem. This method is applied to reduce PICS gyro drifts.

Chapter 6 can be seen as a base for Chapter 7. It gives the principles of GPS and basic geographical concepts. Chapter 7 addresses practical issues of Chapter 4 with positioning examples based on the microsensors. Besides scale train experiments, measurements of a car trajectory by means of inertial sensors and GPS are presented. Details on the prototype implementation are also given.

Final considerations are given in Chapter 8.

1.5. References

[John95] B. Johnson, "Vibrating Rotation Sensors", Sensors and Actuators '95, SAE International Congress & Exposition, Detroit, Michigan, pp. 41-47, February 27-March 2, 1995.

[Macd90] G. Macdonald, "A Review of Low Cost Accelerometers for Vehicle Dynamics", Sensors and Actuators, A21-A23, pp. 303-307, 1990.

[Reid95] E. Reidemeister, "Capacitive Acceleration Sensor for Vehicle Applications", Sensors and Actuators '95, SAE International Congress & Exposition, Detroit, Michigan, pp. 29-34, February 27-March 2, 1995.

[Zhao97] Y. Zhao, "Vehicle Location and Navigation Systems", Artech House, 1997.

Chapter 2
Inertial sensors

Accelerometers and gyros form the heart of any inertial system. This chapter describes the basic principles for measuring acceleration and angular rate and is enlarged on the mechanisation of the inertial system.

The focus is put on microsensors but other technologies are also briefly reviewed due to their large importance in the navigation field. As a complement, an overview of the market is given, including a short classification of both technologies and manufacturers.

Next, methods for characterising the sensor behaviour are outlined besides giving the terminology of a sensor description.

Finally, both in-house developed and off-the-shelf sensors, which have been used in this project, are presented, including a principle description and the main results of the measurement campaign.

2.1. Basic principles of inertial measurements

2.1.1. Acceleration measurement

All forms of accelerometers operate by measuring the inertia force when a mass accelerates, according to Newton's second law of motion:

$$\mathbf{f} = \mathbf{m} \cdot \mathbf{a} \tag{2.1}$$

All acceleration sensors contain a seismic mass, referred to as *proof mass*, which experiences the force; a suspension, which locates the mass and a pick-off, which produces a signal related to the acceleration [Kell94], [Norl90].

Typically, the mass is allowed a single degree of freedom, which may be linear or rotary. Rotary deflecting devices are called *pendulous* devices. Most devices operate in a closed loop, which prevents the mass from moving. They are referred to as *force-balanced* sensors.

Practically, the position detection is accomplished by many different means, such as capacitive, inductive, optical, and resistive.

Actually, accelerometers do not exclusively measure acceleration but a combination of the force producing acceleration and the force due to the gravitational field. The total force f divided by the mass is referred to as *specific force*. In inertial algorithms, the sustained gravitational acceleration must be accounted for to recover the body acceleration $\mathbf{a}_b$ as:

$$\mathbf{a}_b = \frac{\mathbf{f}}{m} - \mathbf{g} \tag{2.2}$$

2.1.2. Angular measurement

Unlike accelerometers, gyroscope approaches use different physical principles [Kell94], [Smit90]. Early sensor designs used the properties of a spinning mass according to the principles of rotational mechanics. More recent techniques range from optical principles to the use of the Coriolis force.

Mechanical gyroscopes include a *spinning rotor*, which is a body exhibiting axial symmetry and spinning about that axis. From Newton's second law applied to purely rotating mass (Euler's equation), it can be shown that a rotor under no net torque will remain rotating about its line of symmetry (*rigidity* property). Hence, a disk spinning about its axis and isolated from external torques by a set of gimbals provides a stable angular reference.

It can also be demonstrated that, when a torque is applied normal to the spin axis, the rotor will rotate in a direction normal to both the spin and the applied torque. This phenomenon is called *precession* and is the basic principle used in mechanical gyros.

Figure 2.1 illustrates the example of a single axis mechanical gyro. If the platform on which the rotor is mounted rotates around the input axis (IA), the device will rock about the output axis (OA). This rotation is proportional to the applied torque and can be registered by a rotational transducer placed on the output axis. This measurement is fed back to a motor in order to drive the table in such a way as to nullify the gyro output. This important principle is the basis of maintaining a stable angular reference in inertial systems.

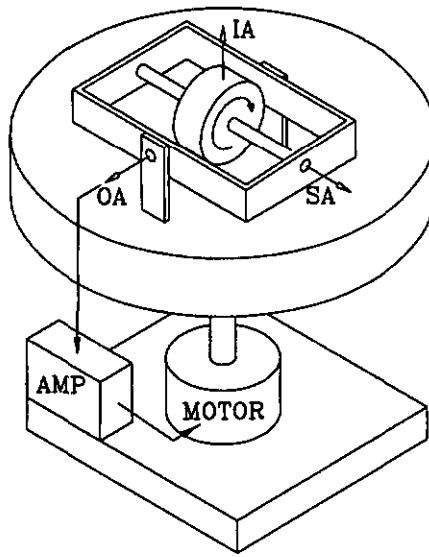

Figure 2.1: Single axis gyro precession [Kell94].

Numerous designs are derived from this basic configuration, e.g. to increase the number of degrees of freedom. On the other hand, devices can be realised such that the output angle is proportional either to the input rate or to the integral of the input rate. When reading on the subject, one

has to be aware that the terms are somewhat loose. The term *gyroscope* can characterise a device, which measure either orientation or angular rate. Some authors give the name *integrating gyro* to the former and *rate gyro* or *gyrometer* to the latter but it seems that it is not a rule of thumb.

Optical sensors are based on the *Sagnac effect*: if two rays of light are sent in opposite directions around a turntable, the light circulating in one direction will arrive at a detector sooner than the light going in the opposite direction. They all include a *ring interferometer*, which allows splitting a beam of coherent light into two counter-rotating components and recombining them after a complete transit of the loop to observe interference fringes.

The *ring laser gyroscope* (*RLG*) includes an optical path composed of mirrors to counter-propagate laser beams. The angular rate is proportional to the optical frequency difference of both beams.

The *fibre optical gyroscope* (*FOG*) has a ring interferometer constructed of optical fibre coiled into a compact loop and uses a laser. A beam splitter allows sending light in both directions simultaneously around the loop. At the end, beams are recombined and sent to a photo-detector.

The last phenomenon used to measure rotation is the *Coriolis force*, which is the basic principle of emerging microgyroscopes. When a reference frame is rotating with respect to another frame at some instantaneous angular velocity ω, the acceleration of a body having a velocity vector $\mathbf{v}$ will include, besides the acceleration measured in the second frame, the centripetal (or centrifugal) acceleration and the Coriolis acceleration $\mathbf{a}_c$:

$$\mathbf{a}_c = 2 \cdot (\omega \times \mathbf{v}) \tag{2.3}$$

These accelerations are derived from what is called *apparent forces* since no real forces are exerted on the object. One defines the Coriolis force $\mathbf{f}_c$ as:

$$\mathbf{f}_c = 2m \cdot (\omega \times \mathbf{v}) \tag{2.4}$$

Accounting for these apparent forces is a central issue in inertial navigation. Firstly, the Coriolis force and centrifugal force due to the earth's rotation must be considered in the

generic mechanisation of highly accurate inertial system, as will be explained in the following paragraph. Secondly, new vibratory gyroscopes directly use the principle of detecting the Coriolis force generated by a rotation (see § 2.2.1).

2.1.3. System mechanisation

Mechanisation refers to how sensors are arranged in the inertial system and how their readouts are processed to give navigation information.

Dead-reckoning

Positioning technique used in inertial systems is referred to as *dead-reckoning*, derived from *deduced reckoning*. It consists of adding up a number of differential displacement vectors. A current vehicle position can be calculated provided the starting location and all previous displacements are known. The differential displacements are obtained by integrating twice the accelerometer readouts. Since the displacement vectors are inherently expressed in a co-ordinate system rotating with the vehicle, they must first be expressed in a common co-ordinate system before being added. This is the reason why gyros are needed. Their readouts are used to provide the body orientation. The body rotations about the longitudinal, lateral and vertical axis are referred respectively to as *roll*, *pitch* and *yaw*.

Due to the integrations involved in the dead-reckoning process, an inertial system is inherently characterised by position errors that grow with time. Therefore, an inertial system is always used in conjunction with absolute navigation system. The whole system is usually referred to as an *integrated system*. Early integrated system simply switched from an absolute aiding source to the inertial system when the absolute source was no longer available. The last known position was simply used to update the dead-reckoning sensors. This methodology does not take full advantage of both systems because the absolute source is not used to improve the inertial

system when both systems are available. Hence, fusion methods, such as Kalman filtering, are used to optimally blend both kinds of information (see Chapters 3 and 4).

Strapdown and platform systems

The implementation of a particular navigation co-ordinate system involves keeping track of the accelerometer orientation as the vehicle moves. To this aim, the gyros are used either to maintain a desired orientation or to measure the body orientation. In the first case, the attitude is actively controlled and mechanical complexity characterises such systems. In the latter case, the system orientation is passively controlled and computational complexity is a main feature of the system.

The *gimballed systems*, also called *platform systems*, employ a stabilised platform servoed to a desired orientation. They were developed historically since navigational information was available directly from the gimbal angles (see § 2.1.2). Due to their mechanical complexity, they are increasingly replaced with *strapdown systems* (derived from strapped down). Here, the sensors are directly strapped on the vehicle and the system is computationally stabilised. The accelerometer readouts are converted to a common co-ordinate system using a matrix transformation. The strapdown approach has many advantages including elimination of the mass, volume and mechanical complexity of a gimballed platform amongst others. In [Lawr93], a comprehensive list of the respective advantages and disadvantages of both approaches is given. It is worth noting that the gyros in a strapdown system have to measure high rates (actually the body rates) whereas gyros in gimballed platforms must just be able to sense small orientation changes. The strapdown mechanisation is of concern for microsensor applications and will be presented in Chapter 4.

Dealing with the earth's rotation

An inertial frame of reference could be defined as a frame in which Newton's laws apply. Accelerometers and gyroscopes measure inherently inertial vectors. Nevertheless, the earth itself is not an inertial frame since it does rotate on its axis. The effect of this rotation is that the Coriolis force and the centrifugal acceleration are additional components of the acceleration (refer to any basic textbook on mechanics). In accurate applications, they must be accounted for in the inertial algorithm using the position and the velocity from the last cycle [Lapu90]. In this work, these effects have not been considered and the Coriolis force due to the earth's rotation has been neglected.

2.2. Micromachined sensors

2.2.1. Principle

Inertial microsensors, like classical sensors, are essentially based on a proof mass, which experiences the physical phenomena, and a transducer, which transforms the response of the mass into a signal that can be measured [Yazd98].

Usual structures for proof masses are membranes, beams or masses suspended by beams. There are two general principles of operating the proof mass. On one hand, it can be initially at rest and slowly deflect under the influence of the external phenomena. On the other hand, it is initially vibrating and changes either the vibration amplitude or frequency when experiencing the physical signal to be measured.

The first design is generally used for accelerometers, where the force resulting from the external acceleration causes the proof mass to displace. The second mode of operation concerns almost all microgyroscopes [John95]. Angular rate microsensors are based on the Coriolis force. A vibration provides the structure with a linear velocity component. Therefore, when the structure is rotated, the Coriolis force

causes the vibration motion to be transformed into another vibration mode or plane of the structure.

Vibrating structure design requires an excitation mechanism. Former vibrating structures were of quartz and the vibration was generated with electrodes, thanks to the piezoelectricity property. However, the trend now is replacing quartz with silicon, whose batch processing is based on IC fabrication technology. Hence, the actuation mechanisms to drive the vibrating structure into resonance are primarily electrostatic or electromagnetic excitation.

A variety of transduction mechanisms are used to convert the proof mass displacement into a measurable signal. They depend on the material of the structure. If the material is piezoelectric, the detection is operated by means of metallic electrodes. Piezoresistors are also commonly used in silicon devices. They are generally placed on the suspension frame of the proof mass and they elongate or shorten relative to its movement changing their resistivity. They are arranged in a half or full bridge, which allows picking up the signal. The detection can also be capacitive. When the proof mass moves from its rest position, the capacitance with respect to a fixed electrode is modified. In addition to the aforementioned methods, microsensors use other principles, including optical and electromagnetic.

Microsensors can be operated open-loop or closed-loop. In the latter case, an electromagnetic or electrostatic force is generated, which opposes the acceleration force. Thus, the mass maintains an equilibrium position and the energy required is a measurement of the applied signal. Closed-loop systems improve the overall sensor linearity whereas open-loop sensors require a simpler interface circuitry.

2.3. State of the art and manufacturers

Mechanical systems have been developed to a high state of refinement and have proved to be extremely reliable and accurate. However, advances in development have reached a plateau and the trend now is strapdown system. The computer revolution had a profound effect on sensor technology, allowing

the compensation of errors to be done numerically rather than mechanically or electronically.

Optical gyros (RLG, FOG) take over the place of spinning mass gyros for medium and even high accuracy applications [Lehm93].

Microsensors have been commercially available for a few years. Low cost microaccelerometers are now in large volume production for automotive navigation needs. Microgyroscopes have been the subject of extensive research over the past years and their performance has improved dramatically. Most are also targeted to automotive applications but effort is still under way for large volume production.

Though a continuing and amazing improvement over a rather short period, the development of inertial microsensors, which are both cheap and accurate, remains a challenge. It is foreseen that microsensors will be capable of providing *navigation-grade* performance, which is a severe requirement for navigation system (growth rate of the position error less than 0.1 nmi/hr).

Several companies are producing multiple sensor systems for angular rate and acceleration measurement. Products are clearly application dependent and their specifications cover a broad spectrum [Mars95]. Manufacturers can be grouped as follows:

- firms historically involved in navigation and guidance systems. They provide high-accuracy devices suitable for costly high-end guidance applications (airborne, military). They use mainly mechanical and optical technologies but also dedicate research to new approaches. Examples are: *British Aerospace Equipment* (www.base-usa.com), *Honeywell* (www.honeywell.com), *Kearfott Guidance and Navigation Corporation* (www.kearfott.com), *Litef* (www.litef.com), *Litton Industries Inc.* (www.litton.com), Sagem (www.sagem.com), *Systron Donner Inertial Division* (www.systron.com),
- firms active in the field of control and metrology. They provide a wide range of sensor (pressure, force) for particular applications. Recently, they also started to manufacture accelerometers, due to the development of less expensive technologies. Examples are: *Endevco* (www.endevco.com),

Entran (www.entran.com), *Kistler Instrument Corporation* (www.kistler.com), *Lucas Nova Sensor* (www.novasensor.com),

- firms specialised in semi-conductor design. They have been proposing microsensors mainly targeted for automotive applications for a few years. Examples are: *Analog Devices* (www.analogdevice.com), *Bosch* (www.bosch.de), *General Motors* (www.gm.com), *Motorola* (www.mot.com), *Mitsubishi* (www.mitsubishi.com), *Samsung* (www.samsung.com),
- firms for which inertial sensor development is far from their basic activity but that still provide a few inertial sensors suitable for particular application or because they have the know-how that can be used for some particular design. Examples are: *Andrew* (www.andrew.com), *Brüel & Kjaer* (www.bk.dk), *Murata* (www.murata.com).

The above classification reflects only the author's opinion and has by no means been endorsed by the firms concerned.

2.4. Sensor characterisation

The terminology to describe sensor performances as well as test procedures is rigorously defined in [ANSI84]. The main features are summarised in [Lawr83]. According to standards, models are established for every kind of sensor. [Lawr93] and [View96] give, among others, model examples for a pendulous accelerometer and a mechanical gyro. Models are usually defined as follows:

$$y = b + k \cdot s_m + f(s_m) + g(s_i) + h(\bar{s}) \qquad (2.5)$$

where y is the sensor output,
b is the *bias*, referred to as *zero offset*. This is the signal measured when the input is zero,
k is the *scale factor*. This is the ratio between a change in the output signal and the change in input,
s_m is the input signal applied along the sensitive axis of the sensor. Quadratic or cubic terms may be present in the output and are expressed through the function f,
s_i is the input signal, which is not applied along the sensitive axis but which nevertheless influence the sensor output. This is

typically the off-axis sensitivity specified for accelerometers. The effect is expressed through the function g,
$\bar{s}$ stands for the acceleration for gyros, respectively to as the angular rate for accelerometers. Life is such that accelerometers unfortunately respond to angular rate and gyros respond also to acceleration. This influence is described through the function h.

As shown in (2.5), the sensor behaviour is dynamic-dependent but there are anyway dynamic-independent factors that will modify the sensor output such as temperature, pressure, electromagnetic perturbation and sensor structure.

Depending on the application, the sensor behaviour has to be modelled with more or less accuracy. The unknown parameters of (2.5), which are accordingly selected, are characterised with either deterministic or stochastic components. Besides, it is worth measuring the sensor *dead band*, which is the lower limit below which the sensor cannot detect input changes, and the sensor *resolution*, which is the minimum input that produces the expected output value according to the scale factor.

The sensor behaviour, as it has been defined above, is determined through lab tests. The manufacturers first do this stage and provide the user with nominal values, characterising typically sensitivity, maximum operation range, frequency response, resolution, linearity, offset, off-axis sensitivity and shock survival. Since bias and scale factor may vary from day to day, these specifications must be redefined before use through the stage of calibration. Additionally, users may do their own test procedure to complete the model according to their needs.

Test procedures require basically that sensors be well aligned with respect to a known axis (e.g., pure vertical, earth rotation axis) and isolated from ground vibrations. They are done a certain time after the sensor is turned on to avoid warm-up trends. Precise rate tables are used both to rotate gyros and to apply centrifugal accelerations to accelerometers. Another common test for accelerometers is gravity sensing. The sensor is rotated about a horizontal axis so that it is exposed to known components of gravity. The test procedure also includes *tombstone* tests, where the sensor is allowed to run on a stable

base during a long period. The output will exhibit a random drift. The methods used in this work were limited to the above-described procedures. Though, test procedure may involve other steps and measurement devices [Lawr93].

In this work, the scale factor, resolution and dead band of the sensors considered have been calibrated in laboratory with use of rotating table and gravity tests and are given below. Besides, tombstone tests have been realised to characterise the in-run drift, considered in this thesis as the random part of the offset (see Chapter 5).

For in-field applications, the use of complex equipment is impractical. However, parameters like scale factor and bias have to be calibrated due to the day-to-day uncertainty. In [Ferr95], the proposed method is based on a set of rotations around a case. The accelerometers are calibrated using the usual tests based on local gravity sensing whereas the gyro output is integrated during a full revolution angle of the sensors.

2.5. PICS sensors

2.5.1. IMT angular rate microsensor

The principle of the IMT gyro is shown in Figure 2.2. The vibratory gyro consists of two proof masses suspended by four beams. The excitation is electromagnetic. When an AC current I flows along a u-shaped metallic conductor placed on top of the proof masses, the interaction with a perpendicular magnetic field **B** results in a Lorentz force that provides anti-phase vibration for the two masses. According to the Coriolis force principle, a rotation about an axis parallel with the mass plane will generate a force in the perpendicular direction and hence, an out of plane vibration. This vibration is sensed by means of four piezoresistors centred on top of the external suspensions and connected in a Wheatstone bridge configuration.

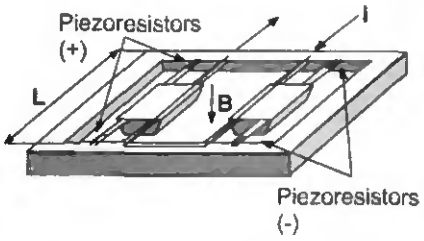

Figure 2.2: IMT gyroscope principle.

Prototypes with different masses (600 x 600 x 360 μm^3, 1000 x 1000 x 360 μm^3, 2000 x 1000 x 360 μm^3) have been fabricated (Figure 2.3).

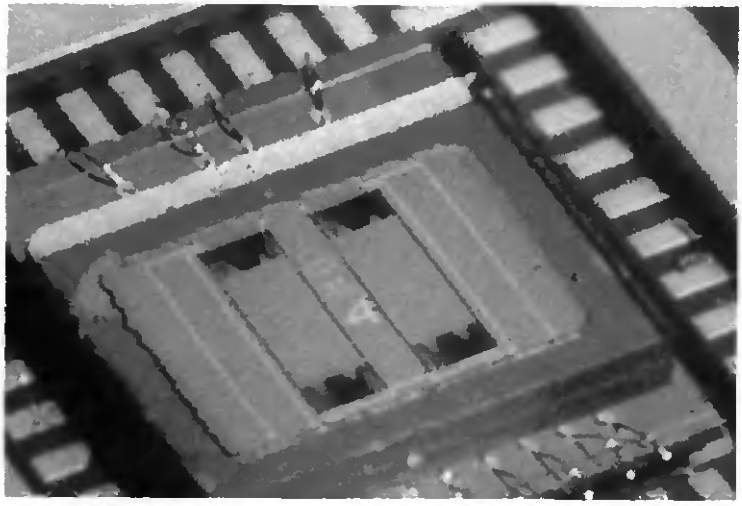

Figure 2.3: Picture of the packaged gyroscope. The two proof masses are 2000 μm long, 1000 μm wide and 360 μm thick.

Provided the sensor output is averaged during about 30 s and the mean offset level is removed, linearity has still been proved. The maximum sensitivity of 4 nV/deg/s has been obtained with the largest mass.

The gyroscope output is amplified 500 times and is thereafter filtered by a lock-in amplifier. The additional gain of

the lock-in is typically 200. The lock-in demodulates the signal around the carrier signal, whose frequency is chosen as the resonance frequency minus 100 Hz.

In the first generation design, the resolution fixed by the noise level is 20 deg/s. The sensor output exhibits a large in-run drift, due to parasite vibration modes. In Chapter 5, a corrective method of this random drift is presented.

Thanks to a second design of the gyro, the drift has been strongly reduced but the noise level is still 20 deg/s. The main perturbation source is the thermal noise in the piezoresistors.

Due to its poor accuracy, the IMT gyro does unfortunately not fulfil automotive navigation requirements. Thus, it is not considered in the navigation tests realised in this work.

More complete information about the gyro, including details on fabrication, can be found in [Paol96a], [Paol96b], [Gret98].

2.5.2. LPMO multi-axial microaccelerometer

The LPMO accelerometer consists of three independent seismic masses. Two masses sense the lateral accelerations (x- and y-direction), which are parallel to the substrate and one mass senses the out-of-plane acceleration, which is applied vertically (Figure 2.4).

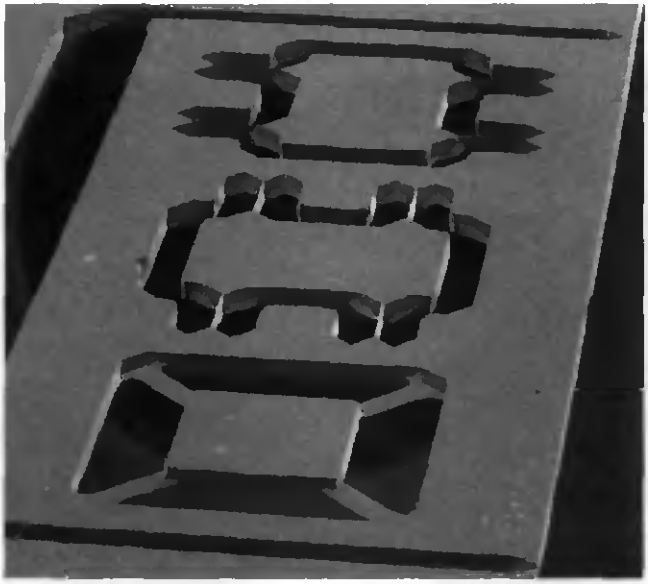

Figure 2.4: View of the monolithic seismic mass system.

Optical and capacitive transducers have been investigated. By forming a Fabry-Perot cavity between the seismic mass and the output of an optical fibre, the accelerations are sensed by measuring the optical path change. A 2D multiplexed optical accelerometer has been realised in this way.

A 3D capacitive accelerometer has been designed, fabricated and tested by implementing electrodes and building three differential capacitors.

Since the sensor was not ready at the time of this research, it has not been integrated in this study. Hence, no measurement is presented in this report. However, additional information is available in [Schr98a], [Schr98b].

2.5.3. Analog Devices microaccelerometer

The accelerometer ADXL05 manufactured by *Analog Devices* has been considered. Accelerometers from *Analog Devices* belong to the most successful sensors in the market. The firm began with the ADXL50 (range: 50 g) developed for air-bag

deployment and commercialised thereafter sensors measuring low accelerations. The ADXL05 is a force-balanced silicon uni-axial accelerometer with a capacitive detection. The principle is shown in Figure 2.5. When acceleration is applied, the central plate moves closer to one of the fixed plates. This creates a mismatch in the two capacitances, resulting in an output signal at the central plate, which varies directly with the amount of acceleration.

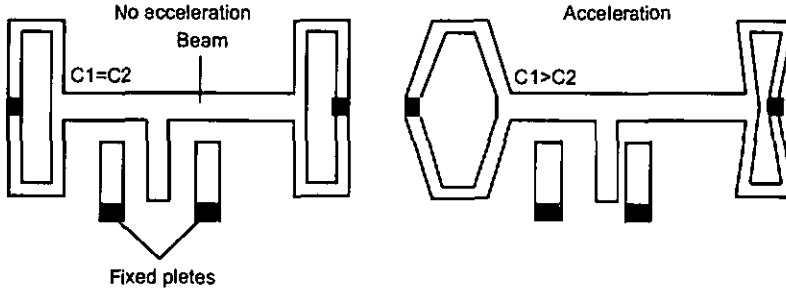

Figure 2.5: Principle of the ADXL05.

The full scale can be fixed externally from ±1 g to ±5 g and the sensitivity can range accordingly from 200 mV/g to 1 V/g. The chip provides connections to easily change gain, offset, level and bandwidth by adding suitable resistors and capacitors. The front-end electronic consists basically of low-pass filtering and amplification. Appendix A gives the designed circuit. The bandwidth has been set to 10 Hz and the basic sensitivity of 200 mV/g has been amplified to 3.88 V/g.

The measured noise level, limiting the resolution, is 0.03 g. The dead band is fixed by the same limitation.

The unit price of 25 $ in 1995 has dropped today to 15 $. The ADXL05 is going to be replaced by the ADXL202, a 2 g accelerometer.

2.5.4. Murata angular rate sensor

The Gyrostar ENC-05S manufactured by *Murata* has been chosen. It is a vibratory gyroscope consisting of a triangular prism made of elastic metal (Elinvar). A piezoelectric ceramic transducer is placed on each face (Figure 2.6).One is used as feedback to the driving circuit. The two others are used for detection. The left and the right detection values are subtracted giving the angular rate. This allows noise cancelling as well as large detection output.

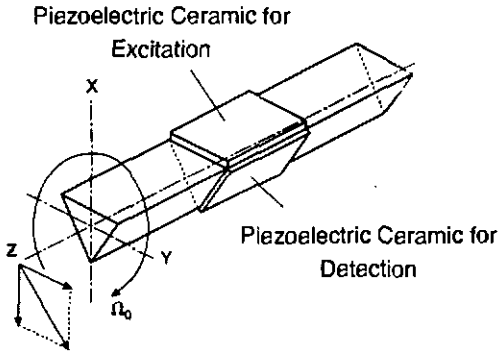

Figure 2.6: Structure of the Murata Gyrostar [Mura95].

The measurement range is 90 deg/s and the sensitivity is 0.8 mV/deg/s. The front-end electronic has been designed to amplify the sensor output 40 times and to get a bandwidth of 4 Hz (see Appendix A). Experiments have shown that the dead band is 5 deg/s and that the resolution is 0.4 deg/s. Chapter 5 presents corrective methods for the in-run drift.

The unit price was 75 $ in 1995. The ENC-05S is now no longer produced but similar products are still available from *Murata*.

2.6. Summary

Inertial sensors are relatively old technology with a wide variety of forms being developed over the years. Nowadays, the

development of inertial microsensors has progressed rapidly, mainly driven by the automotive market. Nevertheless, cheap as well as highly accurate sensors still do not exist.

The change of technology goes along with a change of the system mechanisation. Current inertial systems put emphasis on the computational processing instead of mechanical refinements.

A sensor description can be more or less sophisticated. It includes basically bias and scale factor determined mainly through the use of rotating table and gravity tests.

In the PICS project, two silicon microsensors have been fabricated: an mono-axial gyro and a multi-axial accelerometer. Due to timing constraints, only the first one is concerned in this work. A corrective method to reduce the drift of the gyro is presented in Chapter 5. However, the sensor performance does not fulfil automotive requirements. Thus, off-the-shelf sensors have also been investigated to allow the development of a complete automotive navigation system. The main sensor features obtained through measurement campaigns are given in this chapter. The use of off-the-shelf sensors in real navigation tests is presented in Chapter 7.

2.7. References

[ANSI84] ANSI/IEEE STD 528-1984, "Standard Inertial Sensor Terminology", 1984.

[Ferr95] F. Ferraris, "Procedure for Effortless In-field Calibration Three Axis Rate Gyros and Accelerometers", Sensors and Materials, Vol. 7:5, pp. 311-330, 1995.

[Gret98] F. Grétillat, "Silicon Micromachined Vibrating Gyroscopes with Piezoresistive Detection and Electromagnetic Excitation", PhD thesis, University of Neuchâtel, Institute of Microtechnology, Sensors Actuators and Microsystems Laboratory, Neuchâtel (CH), 1998.

[John95] B. Johnson, "Vibrating Rotation Sensors", Sensors and Actuators '95, SAE International Congress & Exposition, Detroit, Michigan, pp. 41-47, February 27-March 2, 1995.

[Kell94] A. Kelly, "Modern Inertial and Satellite Navigation Systems", Carnegie Mellon University, Robotics Institute, Internal report CMU-RI-TR-94-15, Pittsburgh, Peunsylvania, 1994.

[Lapu90] D. Lapucha, "Precise GPS/INS Positioning for a Highway Inventory System", PhD thesis, University of Calgary, Department of Surveying Engineering, Calgary, Alberta (Ca), 1990.

[Lawr93] A. Lawrence, "Modern Inertial Technology – Navigation, Guidance and Control", Springer-Verlag, New York, 1993.

[Lehm93] A. Lehmann, "Choosing a Gyro: a Comparison of DTG, RLG and FOG", Symposium Gyro Technology, Stuttgart (D), pp. 1.0-1.12, 1993.

[Mars95] C. Marselli, "Prospection du Marché des Accéléromètres et Capteurs Angulaires", University of Neuchâtel, Institute of Microtechnology, Electronics and Signal Processing Laboratory, Internal report n°389 PE06/95, Neuchâtel (CH), June 1995.

[Mura95] Murata Technical Manual, "Vibratory Gyroscope and Various Applications", Catalog n° T-14-A, 1995.

[Norl90] B. Norling, "Accelerometers: Current and Emerging Technologies", International Association of Geodesy Symposia, n°107, Banff, Alberta (Ca), pp. 70-84, September 10-13 1990.

[Paol96a] F. Paoletti, "A Silicon Vibrating Gyroscope with Piezoresistive Detection and Electromagnetic Excitation", IEEE MEMS Workshop '96, San Diego, California, pp. 162-167, 1996.

[Paol96b] F. Paoletti, "A Silicon Micromachined Tuning Fork Gyroscope", Symposium Gyro Technology, Stuttgart (D), pp. 5.0-5.8, 1996.

[Schr98a] G. Schröpfer, "Comparison between an Optical and a Capactive Transducer for a Novel Multi-Axial Bulk-Micromachined Accelerometer", SPIE Conference on Micromachined Devices and Components IV, Santa Clara, California, pp. 199-209, September 1998.

[Schr98b] G. Schröpfer, "Multi-Axial Bulk-Micromachined Accelerometers with Capacitive or Optical Read-out", PhD thesis, Université de Franche-Comté, Laboratoire de Physique et de Métrologie des Oscillateurs, Besançon (F), 1998.

[Smit90] R. Smith, "Gyroscopes: Current and Emerging Technologies", International Association of Geodesy Symposia, n°107, Banff, Alberta (Ca), pp. 59-69, September 10-13 1990.

[View96] S. Vieweg, "Modellierung Zeitvarianter Fehler von Inertialsensoren mit Hilfe Hochgenauer Ortungsreferenz", PhD thesis, Technische Universität Braunschweig, Institut für Flugführung, Braunschweig (D), 1996.

[Yazd98] N. Yazdi, "Micromachined Inertial Sensors", Proceedings of the IEEE, Vol. 86:8, pp. 1640-1655, August 1998.

Chapter 3
Kalman filtering

The Kalman filter estimates the state of a dynamic system having certain types of random behaviour. Its understanding and use requires the agreement with hypothesis (linearity and gaussian white noises of the system) and knowledge of a system formulation called state space approach.

This chapter gives a summary of these notions and a rough derivation of the discrete Kalman filter. First, the context of the discovery of the Kalman filter is illustrated. Then, the state space description is explained and the basic notions used when dealing with random phenomena are recalled. Thereafter, the derivation of the Kalman filter is given. Finally, its advantages and its range of applications are outlined.

Further information can be found in the references.

3.1. Short history

The Kalman filter is one of the most important discoveries in the history of estimation, touching on probability theory and estimation methods.

The probability theory, which aims at formalising the question of uncertainty, has been considered for 400 years (Pascal, Huygens and Bernouilli). However, the application of probabilistic models for the physical world expanded only in the 19th century (Maxwell) and the notions that we use today (e.g.,

random processes) were established in the 20th century (Kolmogorov and Wiener).

Estimation methods deal with the problem of obtaining an optimal estimate from noisy data. The need for estimation dates back to 16th century with awareness of measurement errors in astronomy. The first formal method is the well known Least-Mean Squares (LMS) attributed to Gauss (1795) who developed it for orbit determination needs.

The coupling between both fields arose in the early years of the Second World War when Wiener derived the LMS problem in terms of the autocorrelation of signal and noise. The Wiener filter was therefore the first method of estimation using statistical information. In 1958, Kalman applied the notion of state variables to the Wiener filtering problem. In 1960, assisted by Schmidt, he derived the extended Kalman filter, which was used for the Apollo spacecraft on-board guidance system and which is still employed today for real-time non-linear applications.

3.2. State space description of a linear dynamic system

A dynamic system, sometimes referred to as *plant*, is an assembly of interrelated entities, which are changing with time. Since Newton's work, mathematical models for dynamic systems have been based on differential equations of several orders. Using intermediate variables, the description can be expressed with a set of first-order differential equations, leading to the *state space* formulation (Figure 3.1):

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathrm{F}(t)\cdot\mathbf{x}(t)+\mathrm{C}(t)\cdot\mathbf{u}(t) \\ \mathbf{z}(t) &= \mathrm{H}(t)\cdot\mathbf{x}(t)+\mathrm{D}(t)\cdot\mathbf{u}(t) \end{aligned} \qquad (3.1)$$

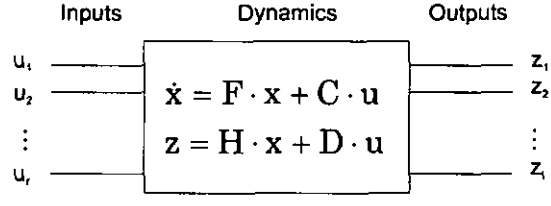

Figure 3.1: State space representation.

- $\mathbf{x}$ (n x 1) is called the *state vector*. It is composed of any set of variables sufficient to completely describe the unforced motion of a dynamic system. The *state variables* may be hidden and thus cannot be measured directly,
- $\mathbf{z}$ (l x 1) is called either the *output* of the system or the *measurement*. It concerns those variables that can be known through measurements,
- $\mathbf{u}$ (r x 1) is called the *input vector*. It is defined by being under the user's control,
- F (n x n) is the *dynamic matrix*,
- H (l x n) is the *measurement sensitivity matrix* or *observation matrix*,
- C (n x r) is the *input-coupling matrix*,
- D (l x r) is the *input/output-coupling matrix*.

The first set of equations of (3.1) is called *state equations*, sometimes referred to as *evolution equations* or *dynamic equations*. They express the evolution over time of the state variables with respect to the other state variables and the inputs. They define the *process model*. The second set is the *measurement equations*, also called *observation equations*. They define the *measurement process*.

There are many different possibilities when defining the state space model of a system. However, a choice is suitable only if the system is observable and controllable. *Observability* is the issue whether the state of the system is uniquely determinable from its input and output. *Controllability* is assessed if there is always an input that can drive a desired state. Mathematical tests of the resulting matrices exist to check these properties (see any book on the subject).

The discrete-time version of the equations is, similarly to (3.1):

$$\begin{aligned} \mathbf{x}_{k+1} &= \Phi_k \cdot \mathbf{x}_k + \Gamma_k \cdot \mathbf{u}_k \\ \mathbf{z}_k &= H_k \cdot \mathbf{x}_k + \Lambda_k \cdot \mathbf{u}_k \end{aligned} \tag{3.2}$$

The matrix Φ is called the *state transition matrix*. In fact, this is the fundamental solution of the homogeneous equation $\dot{\mathbf{x}}(t) = F(t) \cdot \mathbf{x}(t)$. If F is a constant function of time, it can be shown that:

$$\Phi(\tau, t) = e^{F \cdot (\tau - t)} \tag{3.3}$$

The matrix H in (3.2) is called the *design matrix* but is sometimes referred to as *measurement matrix* or *observation matrix*.

3.3. Random processes

Since the knowledge of the physical system is not certain, the system model has to include random processes. The Kalman filter will then realise the estimation using this statistical information. Perturbations of a system whose state is to be estimated by a Kalman filter must be expressed through gaussian white noise. Subject to these conditions, the first and second centred moment (variance) of the variables suffice to describe the random signals.

When dealing with the state space formulation, vector-valued random variables are used rather than scalar random variable. The variance of a vector-valued random process is a matrix called *covariance matrix*:

$$P = \begin{pmatrix} E\langle (x_1 - m_1)^2 \rangle & \cdots & E\langle (x_1 - m_1)\cdot(x_n - m_n) \rangle \\ \vdots & \ddots & \vdots \\ E\langle (x_n - m_n)\cdot(x_1 - m_1) \rangle & \cdots & E\langle (x_n - m_n)^2 \rangle \end{pmatrix} \quad (3.4)$$

where the x_i are the components of the random vector and the m_i are the components of its mean vector.

The covariance matrix is symmetric. The diagonal elements are the variances of the variables whereas the off-diagonal terms are indicators of the cross-correlation between the elements of **x**. A process is uncorrelated if its covariance is a Dirac function.

3.4. Shaping filter

Not all noise fulfil the properties of white gaussian noise processes. However, they can be modelled by a set of differential equations applied to white noise. Such systems are called *shaping filters*. They are driven by white noise with a flat

spectrum, which they shape to represent the spectrum of the actual random process.

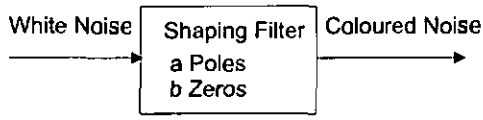

Figure 3.2: Shaping filter.

Table 3.1 summarises the most common shaping filters.

Random processes	State space models		Variance
	Continuous	Discrete	
Random walk	$\dot{\mathbf{x}}(t) = \mathbf{w}(t)$	$\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{w}_k$	$\sigma_x^2(0) = 0$
Random constant	$\dot{\mathbf{x}}(t) = 0$	$\mathbf{x}_{k+1} = \mathbf{x}_k$	$\sigma_x^2(0) = \sigma^2$
Exponentially correlated	$\dot{\mathbf{x}}(t) = -\alpha\mathbf{x}(t) + \alpha\sqrt{2\alpha}\mathbf{w}(t)$	$\mathbf{x}_{k+1} = e^{-\alpha}\mathbf{x}_k + \alpha\sqrt{1-e^{-2\alpha}}\mathbf{w}_k$	$\sigma_x^2(0) = \sigma^2$

Table 3.1: Shaping filters of random processes.

The *random walk* is also referred to as *Wiener* or *Brownian-motion process*. It is non-stationary and its variance increases linearly with time. The *exponentially correlated process*, called also *Gauss-Markov process*, is a stationary process whose autocorrelation is exponential.

Besides the above processes, common models for stochastic signals are *autoregressive moving average* (*ARMA*) models. They are expressed as follows:

$$y_k = -\sum_{n=1}^{M} a_n y_{k-n} + \sum_{n=0}^{M} b_n w_{k-n} \tag{3.5}$$

where a_n are the coefficients of the AR part (poles) and b_n are the coefficients of the MA part (zeros). An equivalent state space representation is among others:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{M-1} \\ x_M \end{bmatrix}_{k+1} = \begin{bmatrix} -a_1 & 1 & 0 & \cdots & \cdots & \cdots & 0 \\ -a_2 & 0 & 1 & 0 & \cdots & \cdots & 0 \\ -a_3 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 1 \\ -a_M & 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_{M-1} \\ x_M \end{bmatrix}_k + \begin{bmatrix} b_1 - b_0 a_1 \\ b_2 - b_0 a_2 \\ \vdots \\ b_m - b_0 a_m \\ -b_0 a_{m+1} \\ \vdots \\ -b_0 a_M \end{bmatrix} \cdot w_k$$

(3.6)

$$y_k = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{M-1} \\ x_M \end{bmatrix}_k + b_0 \cdot w_k \tag{3.7}$$

This description is not often presented in the textbooks but is used by *Matlab System Identification Toolbox*. The corresponding state noise covariance matrix is:

$$Q_k = \sigma^2 \cdot \mathbf{G}_k \cdot \mathbf{G}_k^T \tag{3.8}$$

where $\mathbf{G}_k$ is the vector multiplying w_k in (3.6), and σ^2 is the variance of w_k.

3.5. System identification

A model can be estimated from empirical data using *system identification* methods. The process comprises the following steps that have to be done repeatedly until a suitable model is selected:

- design an experiment to collect the data, including choice of the duration and of the sampling frequency for the acquisition,
- arrange the data, e.g. remove outliers and filter to select frequency ranges,
- define the model structure (AR, ARMA, ...) within which a model is to be found,
- given a criterion of fit and an existing identification method, set the parameters of the models,
- examine the properties of the obtained model to decide whether it is suitable or not.

In Chapter 5, this method is used to find a model of the random part of a gyro offset.

3.6. Discrete Kalman filter

The Kalman filter is also called *linear quadratic gaussian estimator*. It estimates the state of a linear system disturbed by gaussian white noise by using measurements, which are linearly related to the state and also corrupted by gaussian white noise. In the case of gaussian noises, the first and second moment suffice to describe the process. Therefore, the function of the Kalman filter consists of propagating the conditional probability density. The performance criterion is *the minimum squared error*.

In the following derivation, for the sake of simplicity, the system is assumed to have no deterministic inputs but only driving white noises $\mathbf{w}_k$ and $\mathbf{v}_k$:

$$\begin{aligned} \mathbf{x}_{k+1} &= \Phi_k \cdot \mathbf{x}_k + \mathbf{w}_k \\ \mathbf{z}_k &= \mathbf{H}_k \cdot \mathbf{x}_k + \mathbf{v}_k \end{aligned} \tag{3.9}$$

In (3.9), $\mathbf{w}_k$ and $\mathbf{v}_k$ include the Γ_k and Λ_k functions of (3.2). $\mathbf{w}_k$ and $\mathbf{v}_k$ are defined as white noise that are mutually uncorrelated. Therefore, their covariance matrices are:

$$\begin{aligned} E\left\langle \mathbf{w}_k \cdot \mathbf{w}_i^T \right\rangle &= \begin{cases} Q_k & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases} \\ E\left\langle \mathbf{v}_k \cdot \mathbf{v}_i^T \right\rangle &= \begin{cases} R_k & \text{if } k = i \\ 0 & \text{if } k \neq i \end{cases} \end{aligned} \tag{3.10}$$

$$E\left\langle \mathbf{w}_k \cdot \mathbf{v}_i^T \right\rangle = 0 \quad \forall\, k,i$$

Having an initial estimate of the process $\hat{\mathbf{x}}_k(-)$, the new measurement $\mathbf{z}_k$ is used to obtain a new estimate $\hat{\mathbf{x}}_k(+)$, according to the linear blending:

$$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + K_k \cdot (\mathbf{z}_k - H_k \cdot \hat{\mathbf{x}}_k(-)) \tag{3.11}$$

The optimal factor K_k, called the *Kalman gain*, minimises the covariance of the estimation error $\tilde{\mathbf{x}}_k$, defined as the difference between the true value of a random variable $\mathbf{x}_k$ and its estimate $\hat{\mathbf{x}}_k$:

$$\tilde{\mathbf{x}}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k \tag{3.12}$$

The error covariance matrix after update is:

$$P_k(+) = E\left\langle \tilde{\mathbf{x}}_k(+) \cdot \tilde{\mathbf{x}}_k^T(+) \right\rangle \tag{3.13}$$

Putting (3.11) and (3.12) into (3.13), it can be shown that:

$$P_k(+) = (I - K_k \cdot H_k) \cdot P_k(-) \cdot (I - K_k \cdot H_k)^T + K_k \cdot R_k \cdot K_k^T \tag{3.14}$$

By setting the derivative of (3.14) to zero, the optimal value for K_k can be found:

$$K_k = P_k(-) \cdot H_k^T \cdot \left(H_k \cdot P_k(-) \cdot H_k^T + R_k\right)^{-1} = P_k \cdot H_k^T \cdot R_k^{-1} \tag{3.15}$$

As shown in (3.15), the gain matrix is proportional to the uncertainty in the estimate, and inversely proportional to the measurement noise. Thus, if measurement noise is large and state estimate errors are small, only small changes in the estimates are made using (3.11). On the contrary, if the measurement noise is small and state estimate errors are large, the difference between the actual and the predicted measurement will be used as a basis for strong correction in the estimates during the update stage (3.11).

The error covariance matrix associated with the optimal estimate can be computed putting (3.15) into (3.14):

$$P_k(+) = (I - K_k \cdot H_k) \cdot P_k(-) \tag{3.16}$$

An equivalent form for (3.16) is:

$$P_k^{-1} = P_k^{-1}(-) + H_k^T \cdot R_k^{-1} \cdot H_k \tag{3.17}$$

Looking at (3.17), one sees that small measurement errors cause the error covariance matrix to decrease considerably whenever a measurement is used.

It remains to detail how $\hat{\mathbf{x}}_k(-)$ and $P_k(-)$ are computed during the *prediction* stage when no measurements are available. The prediction of the state vector is obtained by using the evolution equation and ignoring the white noise because it has zero-mean:

$$\hat{\mathbf{x}}_{k+1}(-) = \Phi_k \cdot \hat{\mathbf{x}}_k(+) \tag{3.18}$$

The prediction of the error covariance matrix comes from the expression for the a priori error:

$$\tilde{\mathbf{x}}_{k+1} = \Phi_k \cdot \tilde{\mathbf{x}}_k + \mathbf{w}_k \tag{3.19}$$

The computation of the error covariance matrix leads to:

$$P_{k+1}(-) = \Phi_k \cdot P_k(+) \cdot \Phi_k^T + Q_k \tag{3.20}$$

Looking at (3.20), one sees that the more the statistical parameters of the disturbance increase the more rapidly the error covariance increases.

(3.18) and (3.20) form the stage of prediction whereas (3.11), (3.15) and (3.16) form the stage of *update*. The basic steps of the computational procedure for the discrete-time Kalman estimator are summarised in Table 3.2:

step 1 System model Measurement model	 $\mathbf{x}_{k+1} = \Phi_k \cdot \mathbf{x}_k + \mathbf{w}_k \quad \mathbf{w}_k \approx N(0, Q_k)$ $\mathbf{z}_k = H_k \cdot \mathbf{x}_k + \mathbf{v}_k \quad \mathbf{v}_k \approx N(0, R_k)$
step 2 Initial conditions Other assumptions	 $E\left\langle (\mathbf{x}_0 - \hat{\mathbf{x}}_0) \cdot (\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T \right\rangle$ $P_0 = E\left\langle (\mathbf{x}_0 - \hat{\mathbf{x}}_0) \cdot (\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T \right\rangle$ $E\left\langle \mathbf{w}_k \cdot \mathbf{v}_j^T \right\rangle = 0 \quad \forall j,k$

step 3 State estimate prediction	$\hat{\mathbf{x}}_k(-) = \Phi_{k-1} \cdot \hat{\mathbf{x}}_{k-1}(+)$
Error cov. prediction	$P_k(-) = \Phi_{k-1} \cdot P_{k-1}(+) \cdot \Phi_{k-1}^T + Q_{k-1}$
step 4 State estimate update	$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + K_k \cdot (\mathbf{z}_k - H_k \cdot \hat{\mathbf{x}}_k(-))$
Error covariance update	$P_k(+) = (I - K_k \cdot H_k) \cdot P_k(-)$
Kalman gain matrix	$K_k = P_k(-) \cdot H_k^T \cdot (H_k \cdot P_k(-) \cdot H_k^T + R_k)^{-1}$

Table 3.2: Discrete Kalman filter equations.

Referring to Table 3.2, the following remarks might be useful:

- if the process noise in step 1 cannot be considered as white noise, it has to be modelled by a shaping filter and the state vector is augmented accordingly,
- when measurement data are not available or unreliable, the gain becomes zero and the update of the state vector in step **4** is simply replaced by:

$$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) \tag{3.21}$$

Figure 3.3 shows a block diagram of the Kalman filter algorithm.

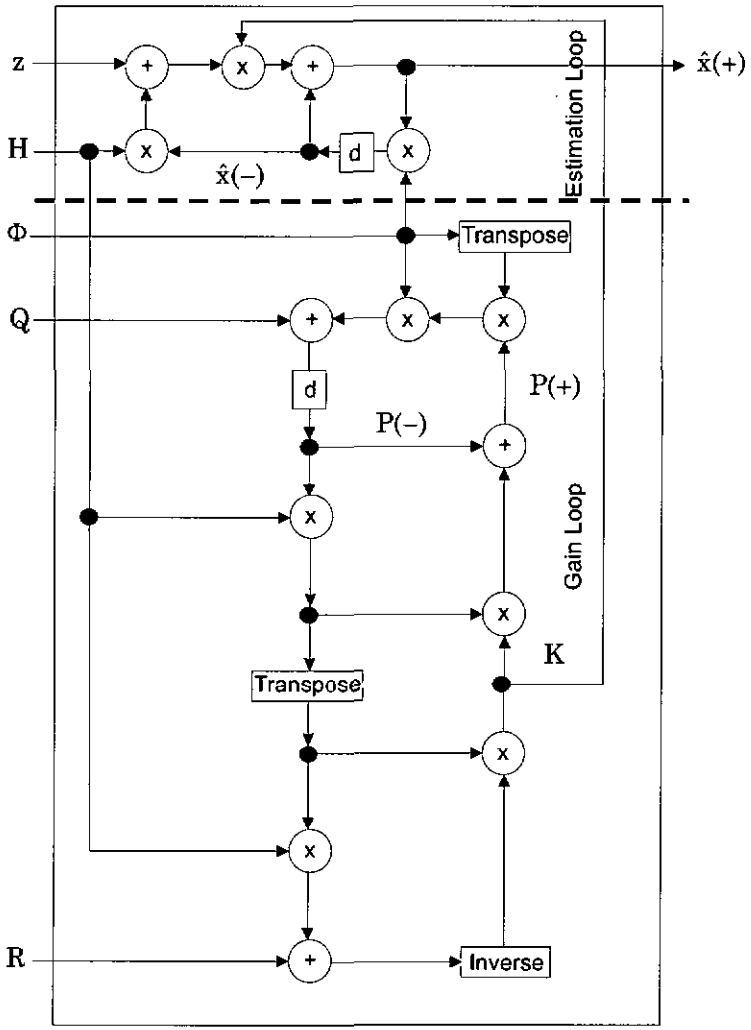

Figure 3.3: Block diagram of the Kalman filter algorithm. d is a one-order delay.

3.7. Application to non-linear problems

Many dynamic systems are such that the functional dependencies of the measurement or state dynamics on the system state are non-linear:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, k) + \mathbf{w}_k$$
$$\mathbf{z}_k = h(\mathbf{x}_k, k) + \mathbf{v}_k \qquad (3.22)$$

Two basic ways of linearising the problem can be considered. One is to linearise about some trajectory in state space, which does not depend on the measurement data. The resulting filter is referred to as *linearised Kalman filter*. The other method is to linearise about a trajectory that is continually updated with the state estimate resulting from the measurements. In the latter case, the filter is called *extended Kalman filter* or Kalman-Schmidt filter. The main advantage of the first approach is that, provided the system is stationary, K_k, Φ_k and H_k can be pre-computed off-line, thus reducing the real-time computational burden. However, its drawback is that the deviation of the actual trajectory from the nominal trajectory tends to increase with time and the linearisation assumption may no longer be valid. Therefore, the extended filter is preferred to the linearised filter in applications where the time mission is long and/or where no pre-programmed trajectory is known. In the following, for the sake of brevity, only the derivation of the extended filter is presented (Table 3.3). It can be shown that the state variables involved in an extended Kalman filter are total quantities rather than incremental quantities, contrary to the linearised filter.

step 1 System model Measurement model	$\mathbf{x}_{k+1} = f(\mathbf{x}_k, k) + \mathbf{w}_k \quad \mathbf{w}_k \approx N(0, Q_k)$ $\mathbf{z}_k = h(\mathbf{x}_k, k) + \mathbf{v}_k \quad \mathbf{v}_k \approx N(0, R_k)$
step 2 Initial conditions	$E\left\langle (\mathbf{x}_0 - \hat{\mathbf{x}}_0) \cdot (\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T \right\rangle$ $P_0 = E\left\langle (\mathbf{x}_0 - \hat{\mathbf{x}}_0) \cdot (\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T \right\rangle$

Other assumptions	$E\left\langle \mathbf{w}_k \cdot \mathbf{v}_j^T \right\rangle = 0 \quad \forall j,k$
step 3 State estimate prediction Error cov. prediction	$\hat{\mathbf{x}}_k(-) = f(\hat{\mathbf{x}}_{k-1}(+), k)$ $\Phi_{k-1} = \left.\frac{\partial f(\mathbf{x}, k)}{\partial \mathbf{x}}\right\|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1}(+)}$ $P_k(-) = \Phi_{k-1} \cdot P_{k-1}(+) \cdot \Phi_{k-1}^T + Q_{k-1}$
step 4 State estimate update Error covariance update Kalman gain matrix	$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + K_k \cdot (\mathbf{z}_k - h(\hat{\mathbf{x}}_k(-)))$ $H_k = \left.\frac{\partial h(\mathbf{x}, k)}{\partial \mathbf{x}}\right\|_{\mathbf{x}=\hat{\mathbf{x}}_k(-)}$ $P_k(+) = (I - K_k \cdot H_k) \cdot P_k(-)$ $K_k = P_k(-) \cdot H_k^T \cdot (H_k \cdot P_k(-) \cdot H_k^T + R_k)^{-1}$

Table 3.3: Extended discrete Kalman filter equations.

A careful match between the theoretical equations and their implementation is crucial to avoid the divergence of the estimator. Linearising about the right state vector is especially important for the correct functioning of the filter.

3.8. Applications and advantages

Since its discovery, the Kalman filter has been extensively used for navigation and tracking purposes, especially in military applications. It is also used in various other applications, among them: process control, flood prediction, traffic modelling and economics. It has many advantages. First, its mathematical background is relatively simple due to the state space formulation. Moreover, it is well suited to computer implementation since it requires a finite representation of the

problem and is recursive, avoiding keeping previous samples in memory. However, it assumes infinite precision and may have problems of numerical stability. Finally, it is derived for both continuous and discrete time and can also be used with non-stationary processes. In this latter case, state/measurement matrices as well as noise covariance matrices change with time.

3.9. Summary

The Kalman filter processes measurements to deduce a minimum error estimate of the state of a linear stochastic system. It is a temporal recursive algorithm.

The models of the system and measurement dynamics must be expressed using the state space formulation. Assumed statistics of systems and of measurement noise must be provided by means of their covariance matrices.

The Kalman filter algorithm involves two principal steps. In the prediction stage, the values of both the state vector and the error covariance matrix are propagated following the model of the system. In the update stage, the occurring measurement is used to modify the predicted estimates, regarding the Kalman gain. The Kalman gain takes into account the quality of the last estimate and of the new measurement.

The extended form of the Kalman filter allows the application of the algorithm to non-linear systems.

3.10. References

[Ande79] B. Anderson, "Optimal filtering", Prentice Hall, Englewood Cliffs, New Jersey, 1979.

[Brow92] R. G. Brown, "Introduction to Random Signals and Applied Kalman Filtering", John Wiley & Sons Inc., New York, 1992.

[Gelb74] A. Gelb, "Applied Optimal Estimation", MIT Press, Cambridge, Massachusetts, 1974.

[Grew93] M. Grewal, "Kalman Filtering: Theory and Practice", Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[Jazw70] A. H. Jazwinski, "Stochastic Process and Filtering Theory", Academic Press Inc., New York, 1970.

[Laba88] M. Labarrere, "Le Filtrage et ses Applications", Éditions Cépadues, Toulouse (F), 1988.

[Mars97] C. Marselli, "Kalman Filtering Overview", University of Neuchâtel, Institute of Microtechnology, Electronics and Signal Processing Laboratory, Internal report n°418 PE06/97, Neuchâtel (CH), June 1997.

[Mayb79] P. Maybeck, "Stochastic Models, Estimation and Control", Vol. 1, Academic Press Inc., Vol. 141, Mathematics in Science and Engineering, New York, 1979.

Chapter 4
Navigation filter

The navigation filter computes the navigation data (position, velocity and heading) from the inertial measurements while integrating aiding sources. This chapter reports first on the dynamical equations to deduce a 3D trajectory from a strapdown system. Emphasis is placed on the computation of the attitude matrix. Secondly, the use of an unusual set of sensors (only 1 accelerometer instead of 2 or 3) is proposed for the case of land navigation. Finally, the bases of integrating aiding sources in the inertial system are explained and the main features of the filter are outlined. Practical results related to these developments are given in Chapter 7.

4.1. Strapdown inertial process

In the most general configuration (3D motion), the measurements of an inertial strapdown unit come from two sensor triads instrumenting three near-orthogonal axes:

- an accelerometer block, measuring the three components of the specific force $\mathbf{f}_b$,
- a gyro assembly, measuring the three components of the body rotation $\boldsymbol{\omega}_b$. It is either aligned to the accelerometer axes or in a known relationship to them.

The sensors are attached to the vehicle and hence, measure the navigation components in a movable axis set. Thus, it is

required to distinguish two frames: the measurement frame, which is a body-fixed frame, and the navigation frame, which is an earth-fixed frame. A key part of the algorithm is then the computation of the transformation matrix R_b^e between both frames (see also Chapter 2).

Figure 4.1 shows the algorithmic flowchart of the strapdown inertial process.

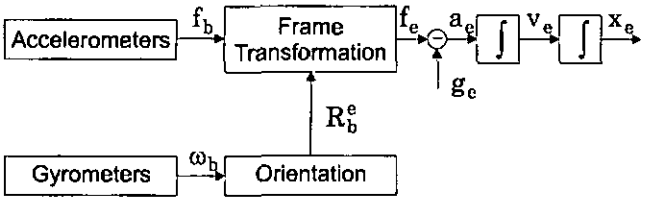

Figure 4.1: Strapdown inertial process.

First, body rates are integrated to obtain the transformation matrix R_b^e from the body-fixed frame to the earth-fixed frame. Thereafter, the specific forces measured by the accelerometers are transformed into the earth-fixed frame:

$$\mathbf{f}_e = R_b^e \cdot \mathbf{f}_b \tag{4.1}$$

Accelerations are obtained by subtracting the gravity vector from the specific force:

$$\mathbf{a}_e = \mathbf{f}_e - \mathbf{g} \tag{4.2}$$

They are integrated twice to compute first the velocity and then the position:

$$\begin{aligned} \mathbf{v}_e &= \int \mathbf{a}_e \cdot dt \\ \mathbf{x}_e &= \int \mathbf{v}_e \cdot dt \end{aligned} \tag{4.3}$$

4.2. Computation of the orientation matrix

The transformation matrix from the body-fixed frame to the earth-fixed frame R_b^e is actually a 3D rotation matrix whose angles describe the current orientation of the body with respect

to the earth. It can be seen as the product of three elementary rotations around the axis of the body. Freely attributing the x-axis to the roll axis, the y-axis to the pitch axis and the z-axis to the yaw axis (Figure 4.2), R_b^e is, provided usual sign rules are respected [Radi91]:

$$
\begin{aligned}
R_b^e &= R_z(y) \cdot R_y(p) \cdot R_x(r) \\
&= \begin{pmatrix} cy & -sy & 0 \\ sy & cy & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} cp & 0 & sp \\ 0 & 1 & 0 \\ -sp & 0 & cp \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & cr & -sr \\ 0 & sr & cr \end{pmatrix} \\
&= \begin{pmatrix} cp \cdot cy & sp \cdot sr \cdot cy - cr \cdot sy & sp \cdot cr \cdot cy + sr \cdot sy \\ sy \cdot cp & sp \cdot sr \cdot sy + cr \cdot cy & sp \cdot cr \cdot sy - sr \cdot cy \\ -sp & cp \cdot sr & cp \cdot cr \end{pmatrix}
\end{aligned}
\tag{4.4}
$$

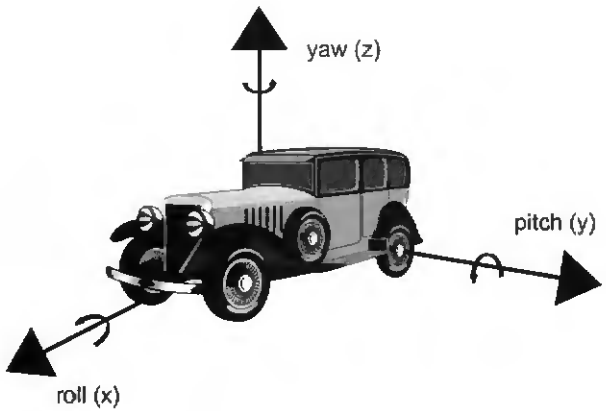

Figure 4.2: Body frame.

Preferably to this approach, *quaternions* are often chosen to avoid the use of trigonometric functions and thus, to reduce the load of computation [Frie78] [Vaga93]. The link between quaternions and pitch-roll-yaw angles is established using the Euler description of a rotation. A theorem due to Euler states that any sequence of rotations of a rigid body, which has one point fixed, is equivalent to a single rotation of the angle α about some axis passing through the fixed point and defined by the unit vector $\mathbf{e}_\alpha$:

$$\mathbf{e}_\alpha = l \cdot \mathbf{x}_e + m \cdot \mathbf{y}_e + n \cdot \mathbf{z}_e \tag{4.5}$$

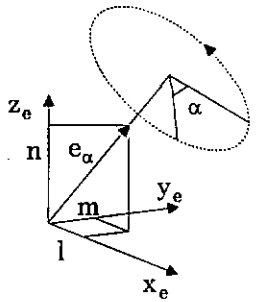

Figure 4.3: Euler representation of a rotation.

R_b^e can be expressed in terms of the *direct cosines* of the Euler axis l, m, n and the Euler angle α [Grub70]:

$$R_b^e = \begin{pmatrix} c\alpha + l^2 \cdot (1-c\alpha) & lm \cdot (1-c\alpha) - n \cdot s\alpha & ln \cdot (1-c\alpha) + m \cdot s\alpha \\ lm \cdot (1-c\alpha) + n \cdot s\alpha & c\alpha + m^2 \cdot (1-c\alpha) & mn \cdot (1-c\alpha) - l \cdot s\alpha \\ ln \cdot (1-c\alpha) - m \cdot s\alpha & mn \cdot (1-c\alpha) + l \cdot s\alpha & c\alpha + n^2 \cdot (1-c\alpha) \end{pmatrix} \tag{4.6}$$

A quaternion is a fourth dimension vector defined as [Kant89]:

$$\mathbf{q} = q_0 + q_1 \cdot \mathbf{i} + q_2 \cdot \mathbf{j} + q_3 \cdot \mathbf{k} \tag{4.7}$$

where q_0 is the real part (also referred to as scalar part) of the quaternion and q_1, q_2, q_3 are its imaginary part (also referred to as vector part). A quaternion represents a rotation of angle α and of the axis given by (4.5) by defining:

$$q_0 = \cos\left(\frac{\alpha}{2}\right) \qquad q_1 = l \cdot \sin\left(\frac{\alpha}{2}\right)$$

$$q_2 = m \cdot \sin\left(\frac{\alpha}{2}\right) \qquad q_3 = n \cdot \sin\left(\frac{\alpha}{2}\right) \tag{4.8}$$

The rotation matrix in terms of the quaternion parameters can be obtained directly from the direct cosine expression (4.6):

$$R_b^e = \begin{pmatrix} q_0^2+q_1^2-q_2^2-q_3^2 & 2\cdot(q_1q_2-q_0q_3) & 2\cdot(q_1q_3+q_0q_2) \\ 2\cdot(q_1q_2+q_0q_3) & q_0^2-q_1^2+q_2^2-q_3^2 & 2\cdot(q_2q_3-q_0q_1) \\ 2\cdot(q_1q_3-q_0q_2) & 2\cdot(q_2q_3+q_0q_1) & q_0^2-q_1^2-q_2^2+q_3^2 \end{pmatrix} \tag{4.9}$$

To be complete, it is necessary to establish the inverse problem, which is how to determine the quaternion parameters as well as the angle values from the rotation matrix. Examining (4.9), it is obvious that:

$$\begin{aligned} q_0 &= 0.5\cdot\sqrt{1+R_b^e(1,1)+R_b^e(2,2)+R_b^e(3,3)} \\ q_1 &= 0.25\cdot\frac{R_b^e(3,2)-R_b^e(2,3)}{q_0} \\ q_2 &= 0.25\cdot\frac{R_b^e(1,3)-R_b^e(3,1)}{q_0} \\ q_3 &= 0.25\cdot\frac{R_b^e(2,1)-R_b^e(1,2)}{q_0} \end{aligned} \tag{4.10}$$

(4.10) is useful to compute the initial value of the quaternion knowing the initial value of the rotation matrix. Similarly, looking at (4.4), the pitch, roll and yaw angles can be retrieved as the following example shows:

$$\begin{aligned} y &= \text{arctg}\left(\frac{R_b^e(2,1)}{R_b^e(1,1)}\right) \\ r &= \text{arctg}\left(\frac{R_b^e(3,2)}{R_b^e(3,3)}\right) \\ p &= \arcsin\left(-R_b^e(3,1)\right) \end{aligned} \tag{4.11}$$

The quadrant uncertainty has still to be solved. Many programming languages provide the *arctan2 (y/x)* function, which tests the sign of both parameters of the division to determine the quadrant of concern. Using *arcsin*, respectively *arccos*, the computation of the cosine, respectively sine, of the angle is needed to set the right quadrant.

4.3. Quaternion update

The critical point remains how to update the quaternion with respect to the last measured angular rate. The quaternion describing the body attitude is related to the angular rate by the following linear differential equations [Grub70]:

$$\dot{\mathbf{q}} = \frac{1}{2} \cdot \mathbf{q} \times \omega \tag{4.12}$$

which is often expressed using the skew-matrix either of the angular rate $\dot{\Theta}$ or of the quaternion:

$$\begin{aligned} \dot{\mathbf{q}} = \frac{1}{2} \cdot \dot{\Theta} \cdot \mathbf{q} &= \frac{1}{2} \cdot \begin{pmatrix} 0 & -\dot{r} & -\dot{p} & -\dot{y} \\ \dot{r} & 0 & \dot{y} & -\dot{p} \\ \dot{p} & -\dot{y} & 0 & \dot{r} \\ \dot{y} & \dot{p} & -\dot{r} & 0 \end{pmatrix} \cdot \begin{pmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} \\ &= \frac{1}{2} \cdot \begin{pmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{pmatrix} \cdot \begin{pmatrix} \dot{r} \\ \dot{p} \\ \dot{y} \end{pmatrix} \end{aligned} \tag{4.13}$$

Assuming ω constant over a small Δt, the solution of the differential equation can be written as:

$$\begin{aligned} \mathbf{q}_{k+1} &= e^{0.5 \cdot \dot{\Theta} \cdot \Delta t} \mathbf{q}_k \\ &= \left(I + \frac{1}{2} \cdot \dot{\Theta} \cdot \Delta t + \frac{1}{4 \cdot 2!} \cdot \dot{\Theta}^2 \cdot \Delta t^2 + \frac{1}{8 \cdot 3!} \cdot \dot{\Theta}^3 \cdot \Delta t^3 + \cdots \right) \cdot \mathbf{q}_k \end{aligned} \tag{4.14}$$

In [Wong88], terms with even power are separated from the ones with odd power, making the Taylor series of the sine and cosine functions visible:

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \left[\left(\cos\left(\frac{\lambda}{2}\right) - 1 \right) \cdot I + \frac{1}{\lambda} \cdot \sin\left(\frac{\lambda}{2}\right) \cdot \Theta \right] \cdot \mathbf{q}_k \tag{4.15}$$

where $\Theta = \int \dot{\Theta} \cdot \Delta t$ and $\lambda = \text{norm}(\omega) \cdot \Delta t$.

In [Edwa71] and [Vaga93], the solution given in (4.14), truncated to the 3rd order, is considered but a corrective term is added to take into account the change of the rotation axis:

$$\mathbf{q}_{k+1} = \left[1 + \frac{\theta}{2} + \frac{1}{2!}\cdot\left(\frac{\theta}{2}\right)^2 + \frac{1}{3!}\cdot\left(\frac{\theta}{3}\right)^3 + \frac{1}{48}\cdot\left(\theta^*\cdot\theta - \theta\cdot\theta^*\right)\right]\cdot\mathbf{q}_k \qquad (4.16)$$

where $\theta = \omega\cdot\Delta t$ and θ^* refers to the angle corresponding to the former angular rate vector. Mathematical rules for operations between quaternions and vectors can be found in [Kant89]. Both methods have been implemented and identical results have been obtained.

The whole attitude computation process is now defined. It consists of:

- computing the transformation matrix knowing the initial roll, pitch, yaw angles (4.4),
- deducing the initial quaternions (4.10),
- updating the new quaternions each time the gyros output data (4.15) or (4.16),
- computing the new rotation matrix with respect to the components of the quaternions (4.9),
- deducing the roll, pitch, yaw angles from the elements of the transformation matrix (4.11).

4.4. Reduced configuration of sensors

The most general strapdown algorithm, by instrumenting a set of 6 sensors (see § 4.1), allows determining the trajectory of an object, which has all possible degrees of freedom (3 linear and 3 angular motion components). In the case of land navigation, motions are obviously more limited. Therefore, it is worthwhile to study the characteristics of such motions in order to set the influence on the kind and number of sensors required.

A perfect 2D motion is completely defined by the yaw angle and the distance travelled along the longitudinal axis. This kind of motion is encountered in indoor robotics applications without obstacles. In this case, odometers are widely used to provide both angle and distance [Bore96]. Similarly to the principle of odometry where the sole longitudinal displacement is measured, only one accelerometer placed along the longitudinal axis x_b has been considered and a modified strapdown algorithm has been developed to deal with this new configuration. In the proposed case, a_{xb} is measured but no

information is available on the centrifugal information a_{yb}. Therefore, it is not possible to apply the frame transformation matrix to the acceleration vector since the 2nd component of body acceleration is totally unknown. Nevertheless, the velocity vector can be fully determined. Since it is always tangential to the trajectory, only its first component is non-trivial and is obtained by integrating a_{xb}:

$$\begin{aligned} v_{xb} &= \int a_{xb} \cdot dt \\ \mathbf{v}_b &= (v_{xb} \;\; 0 \;\; 0) \end{aligned} \tag{4.17}$$

Thereafter, the velocity in the earth-frame can be computed using the frame transformation matrix:

$$\mathbf{v}_e = R_b^e \cdot \mathbf{v}_b \tag{4.18}$$

In the case of land navigation, the vehicle has also pitch and roll motions due to the non-planarity of the ground. They are taken into account in the frame transformation matrix to correct the linear motion measured in the body frame (see § 4.2). In the configuration using one accelerometer, only the first column of R_b^e is considered because of the null elements of the velocity vector:

$$\mathbf{v}_e = \left(R_b^e(1,1) \cdot v_{xb} \;\; R_b^e(2,1) \cdot v_{yb} \;\; R_b^e(3,1) \cdot v_{zb}\right) \tag{4.19}$$

Looking at (4.4), it can be observed that these three components of the matrix do not include the roll angle. The reason is that roll motion does actually not influence the co-ordinate of a point of the trajectory. Hence, considering only the velocity vector does not allow one to distinguish between two motions with different roll components only. Given that, using only one accelerometer makes the system blind to roll motion. However, roll rate is not important for most land navigation applications. Therefore, the proposed configuration of 3 sensors (1 accelerometer and 2 gyros) may be an interesting alternative to usual systems (2-3 accelerometers and 3 gyros). It induces a reduction of cost and complexity and, as illustrated in Chapter 7, can improve the accuracy of the system.

4.5. Design of the Kalman filter

4.5.1. Complementary filter

As previously mentioned in Chapter 2, inertial systems must be used in conjunction with other navigation instruments, which perform an aiding role. Since the inertial system is completely self-contained, the aiding system provides information that is obviously redundant, besides being often partial and intermittently available. The Kalman filter is used to blend optimally all the information in order to yield the best estimate of the trajectory.

The problem of combining two noisy, continuous measurements of the same signal s(t) can be solved using the *complementary* approach [Brow72], where the filter operating on the second input is constrained to be the complement of the first filter Y(s) (Figure 4.4).

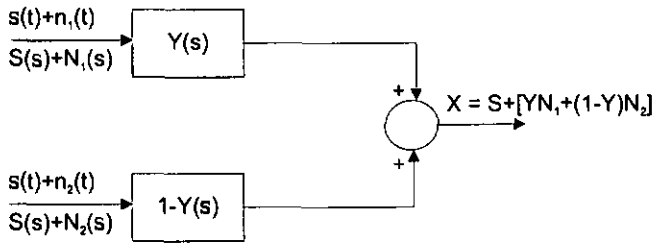

Figure 4.4: Complementary filter.

Figure 4.5 shows an alternate form of this concept, which will be useful to justify the structure of the navigation filter.

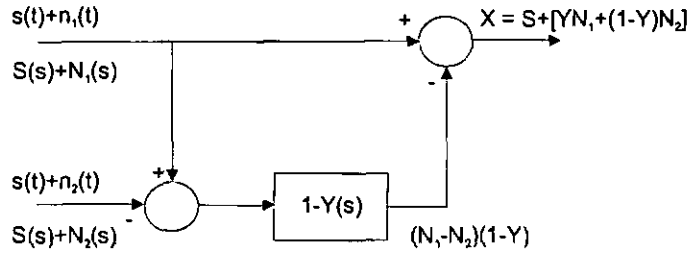

Figure 4.5: Alternate form for the complementary filter.

In the complementary design, it can be shown that no a priori assumptions about the signal are needed, contrary to the former Wiener solution, which requires a statistical model of the signal [Wien49]. In other words, the filter operates only on the noises and the signal passes through the system without any distortion. This arrangement deals with a lot of physical situations where assumptions on the signal are not available, such as navigation applications where it is generally not possible to establish a model of the vehicle motion. Thus, this complementary approach has been widely used in navigation systems. In the variety of mixes which have been developed, the inertial system is most often thought as the reference system, since it is self-contained, and all the other navigational data as augmenting the inertial data (Figure 4.6). The equivalence of Figure 4.6 with Figure 4.5 is straightforward.

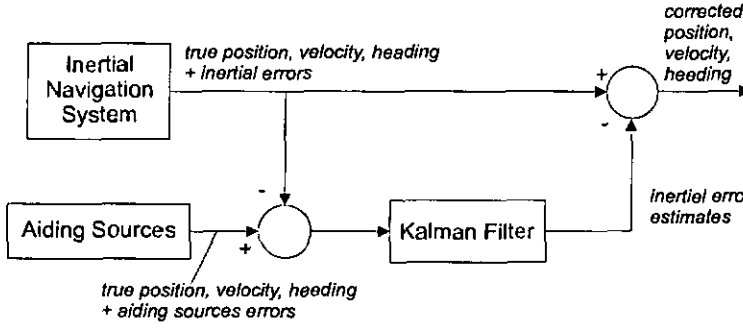

Figure 4.6: Typical integrated navigation system (Feedforward / Linearised form).

4.5.2. Structure of the filter

At that point, the model of the above-described concept to be used in the Kalman filter can be discussed. Refer to Chapter 3 for the notions used below. At first glance, besides the aiding data z_3, the accelerations z_1 and the angular rates z_2 sensed by the inertial system may be considered as system measurements. Doing so, the process model has to contain some external forcing functions, which give rise to the translation and angular motion of the vehicle. As already observed, establishing a statistical model for these functions is not a reasonable assumption in many navigation applications. The solution is therefore to transfer the measurements z_1 and z_2 into the process model. They can indeed be considered as known functions since they are available on a real-time basis. Finally, only the aiding data are considered as system measurements.

The state vector basically contains the system variables to be estimated. These are true position, velocity and angular rate:

$$\mathbf{x} = \begin{pmatrix} x_e & y_e & z_e & v_x & v_y & v_z & p & r & y \end{pmatrix}^T \tag{4.20}$$

Higher order variables are generally not considered for the following two reasons. Firstly, this basic set of variables is often the only one of interest and adequate to describe motions with moderate dynamics. Secondly, the process model, which describes the evolution of the dynamical variables with time, is a first-order equation. Therefore, it is necessary to know how to describe the first derivation of each additional component of the state vector. Since the sensors provide the accelerations and the angular rates, the prediction equations corresponding to the minimal state vector of (4.20) are directly available from (4.1) to (4.3). In return, putting higher order variables in the state vector implies the design of models.

The prediction equations (4.1) to (4.3) are obviously non-linear. Thus, the state matrix is obtained after the linearisation of these equations around the last estimate of the state vector, leading to the extended Kalman filter form (see Chapter 3). In other words, corrections are made at the sensor level rather than at the system output (Figure 4.7). Also called *feedback*

configuration, this approach is usually preferred to the linearised Kalman filter form, also referred to as *feedforward* configuration (Figure 4.6) to avoid divergence with the nominal trajectory, especially in applications where the time mission is long. In return, specific applications, where the trajectory is pre-programmed, would use the linearised form.

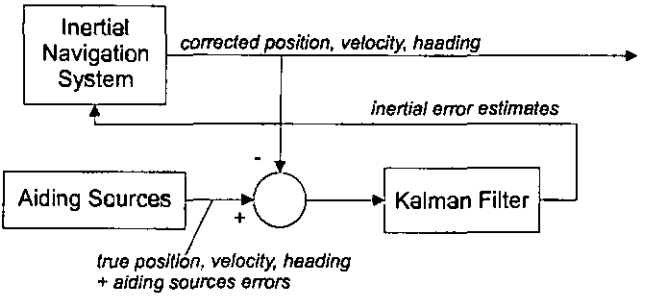

Figure 4.7: Integrated navigation system (Feedback / Extended form)

Appendix B gives detailed structures of the state matrices that have been designed, regarding different configurations (number of sensors and use of quaternions for the computation of the rotation matrix).

The measurement vector contains the aiding data, which most often are directly state variables (position, velocity or orientation). Thus, the measurement matrix contains simply either zeros or ones regarding the variables concerned and the measurement equations are linear. Since aiding sources may arise intermittently, the measurement update is less frequent than the prediction stage.

4.6. Summary

Since inertial sensors are attached to the moving vehicles, their measurements have to be transformed from the body frame to an earth-fixed frame. The frame transformation matrix can be expressed with quaternions to avoid the use of trigonometric functions. In land navigation (2D navigation but taking into account the non-planarity of the ground), the complete

configuration of 6 sensors (3 accelerometers and 3 gyros) is not productive. Alternatively to usual systems (2 accelerometers and 3 gyros), a simpler configuration with only 1 accelerometer and 2 gyros has been demonstrated as being suitable for land motions.

The integration of aiding sources with the inertial systems is based on the complementary filter principle. In this configuration, the state variables are the navigation parameters to be estimated and the state equations are given by the dynamical equations. The measurements of the system are all aiding data that can be connected with the chosen state variables. Due to the non-linearity of the system, an extended Kalman filter is used to yield the best estimate of the trajectory.

4.7. References

[Bore96] J. Borenstein, "Navigating Mobile Robots: Systems and Techniques", AK Peters Ltd, Wellesley, Massachusetts, 1996.

[Brow72] R. G. Brown, "Integrated Navigation Systems and Kalman Filtering: A Perspective", Navigation: Journal of the Institute of Navigation, Vol. 19:4, pp. 355-362, winter 1972-1973.

[Brow92] R. G. Brown, "Introduction to Random Signals and Applied Kalman Filtering", John Wiley & Sons Inc., New York, 1992.

[Edwa71] A. Edwards, "The State of Strapdown Inertial Guidance and Navigation", Navigation: Journal of the Institute of Navigation, Vol. 18:4, pp. 386-401, winter 1971-1972.

[Frie78] B. Friedland, "Analysis Strapdown Navigation Using Quaternions", IEEE Transactions On Aerospace and Electronic Systems, Vol. AES-14:5, pp. 764-768, September 1978.

[Grub70] C. Grubin, "Derivation of the Quaternion Scheme via the Euler Axis and Angle", Engineering Notes, Journal of Spacecraft, Vol. 7:10, pp. 1261-1263, October 1970.

[Kant89] M. Kantor, "Hypercomplex Numbers: An Elementary Introduction to Algebras", Springer-Verlag New-York Inc., 1989.

[Radi91] J. C. Radix, "Répertoire Géodésique en Vue de la Navigation", Éditions Cépadues, Toulouse (F), 1991.

[Vaga93] J. Vaganay, "Conception d'un Système Multisensoriel de Localisation Dynamique 3D pour Robot Mobile", PhD thesis, Université de Montpellier II Sciences et Techniques du Languedoc, Montpellier (F), 1993.

[Wei90] M. Wei, "A Strapdown Inertial Algorithm Using an Earth-Fixed Cartesian Frame", Navigation: Journal of the Institute of Navigation, Vol. 37:2, pp. 153-167, Summer 1990.

[Wien49] N. Wiener, "Extrapolation, Interpolation and Smoothing of Stationary Time Series", John Wiley & Sons Inc., New York, 1949.

[Wong88] R. V. C. Wong, "Development of a RLG Strapdown Inertial Survey System", PhD thesis, University of Calgary, Department of Surveying Engineering, Calgary, Alberta (Ca), 1988.

Chapter 5
Local filter

This chapter presents the problem of taking the microsensor errors into account to improve the accuracy of the results. Only those errors that can be corrected using Kalman filtering are considered.

First, explanations are given on the strategy of completing the complementary design of the Kalman filter (see Chapter 4) in order to deal with the sensor errors. Such methods have been successfully used in the past but at that time, the sensors used were far more accurate. The main feature of these approaches is that the estimation stage is dependent on the availability of the aiding system. Therefore, alternative designs are investigated in this study, in such a way that prediction and estimation of the errors arise at the same time.

For this purpose, the way of separating a signal from a noise through a Kalman filter, using only one measurement of the signal of interest, is studied and illustrated by simulation examples. Moreover, a novel algorithm is presented, which allows estimating the true signal without any assumptions on its shape. The proposed method is used to correct the offset of real gyro signals.

An intermediate stage is the modelling of the drift of the sensor. The whole procedure including hardware as well as software options is also presented in the following.

5.1. Reduction of sensor errors

Fundamentally, accelerometers and gyros provide rate information. Hence, the sensor output must be integrated to provide absolute measurements of orientation, position and velocity. Accordingly, errors accumulate, leading to unacceptable results. An inertial system, whatever its quality, is inherently characterised by position errors that grow with time and distance. Moreover, current inertial microsensors are far less accurate than former mechanical or optical devices. Thus, the problem of dealing with sensor errors becomes a crucial one, whatever the level of accuracy required by the application.

There are two ways of overcoming the unbounded growth of errors. The first one consists of periodically resetting the inertial system with other sensing mechanisms. This approach has been presented in Chapter 4 where the coupling between an INS and an aiding source has been thoroughly explained. The second one is the reduction of the sensor errors. Enhancing the quality of sensor data does not avoid updating the INS but allows extending the duration of use in a stand-alone mode. Besides, the aiding source does not have to fulfil too tight requirements (e.g., high accuracy at high rates). Thus, an error correction system may decrease the complexity and the cost of the overall integrated system.

Inertial sensors suffer from errors due to internal causes (e.g., parasite vibration modes) as well as external causes (e.g., temperature, pressure) (see Chapter 2). The corrective methods depend on the characteristics of the perturbations. If the frequency range of the perturbations is different from the bandwidth of the signal, suitable filtering in the frequency domain eliminates undesired signals. This stage is usually necessary but does not suffice to suppress the sensor errors. In most cases, the key idea is to model the sensor behaviour and thereafter, estimate the error in the measurements with respect to the model.

Some errors are deterministic, then correction can be simply made by applying the known mathematical relationship with respect to the varying parameter (e.g., quadratic dependence of the output of a pressure sensor with respect to

the temperature) [Bert96]. On the contrary, some undesired phenomena are random. Nevertheless, they can also be modelled through a shaping filter (see Chapter 3). Each time-dependent error can be properly modelled with a state space description and can be potentially eliminated using a Kalman filter.

The classical method to reduce such errors while integrating aiding information consists of augmenting the 9-states basic state vector (see Chapter 4) with additional states corresponding to the INS errors. The state matrix is modified accordingly:

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{v} \\ \boldsymbol{\theta} \\ \mathbf{e}_g \\ \mathbf{e}_a \end{pmatrix}_{k+1} = \begin{pmatrix} & \Phi_n & & \Phi_{e,n} & \\ & & & & \\ 0 & 0 & 0 & \Phi_g & 0 \\ 0 & 0 & 0 & 0 & \Phi_a \end{pmatrix} \cdot \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \\ \boldsymbol{\theta} \\ \mathbf{e}_g \\ \mathbf{e}_a \end{pmatrix}_k + \mathbf{G}_k \cdot \mathbf{w}_k \qquad (5.1)$$

where x, v, θ are the position, respectively velocity and orientation states. This subset is the 9-states vector, which is suitable for usual 3D navigation description. $\mathbf{e}_g$ (m x 1) and $\mathbf{e}_a$ (n x 1) are the additional states and refer to the states describing the m gyro, respectively n accelerometer, errors. Φ_n (9 x 9) describes the dynamic equations, which link the navigation parameters, according to the developments of Chapter 4. Φ_g (n x n) and Φ_a (m x m) are derived from the state models of sensor errors. $\Phi_{e,n}$ describes the influence of the sensor errors on the navigation parameters. The augmentation of the state vector is the only change to be done to consider the errors. Except for that, the filter works in the same way as the one of the complementary method. When external information is available, all states, including the errors, are estimated. Provided the state matrix has been correctly defined, the error contribution is deduced from the navigation results and the accuracy increases.

The number of additional state variables depends on the sensors used as well as on the application. As previously mentioned, the errors that are put in the state vector are usually time-dependent random errors, which can be modelled using a shaping filter.

The choice of the state vector as well as the structure of the overall system is unfortunately poorly documented because former research on navigation systems has been concentrated in very confidential circles (see Chapter 1). Nevertheless, the Department of Geomatics Engineering at the University of Calgary (Ca) has been deeply involved in the research on integrated systems and provides many publications on the subject [Schw83]. As a typical example, in [Lapu90], the errors of a mechanical 3D strapdown system are described by a gyro drift $\mathbf{b}_g$ and an accelerometer bias $\mathbf{b}_a$, both modelled by a first-order Gauss-Markov process:

$$\dot{\mathbf{b}}_g = -\varsigma \cdot \mathbf{b}_g + \mathbf{w}_g$$

$$\dot{\mathbf{b}}_a = -\beta \cdot \mathbf{b}_a + \mathbf{w}_a \qquad (5.2)$$

Therefore, 6 state variables are added to the state vector. This model is found to suit the offset behaviour of mechanical inertial sensors. The values for the correlation times ς and β as well as the strength of the white noise are obtained through calibration tests prior to the survey.

Besides the inertial system errors, the errors of the aiding sources are also considered in most integration designs. Two options are usually presented [Schw90] [Schw94]. In the centralised approach, the INS and aiding system errors define a common state vector. In the decentralised approach, the errors of both systems are described in separate Kalman filters. However, the errors of the aiding sources are outside the scope of this research and are not considered in the design of the Kalman filters.

The specific feature of these classical methods (insertion of sensor errors in the navigation state vector) is the update rate of the states: error sensors are estimated only when the measurement source is available. These techniques have been successfully used in the past. However, one has also to keep in mind that, at that time, inertial sensors were mainly of the high precision type and computational load of the algorithms was one of the most important limiting factors [Wagn96]. Nowadays, the complementary approach is still valuable since none of the navigation systems is able to provide a satisfactory navigation solution in a stand-alone mode. It is anyway

worthwhile to design new systems more suited to the current characteristics of the systems (e.g., lower accuracy of the inertial sensors, more powerful processors). Part of this research has been done consequently to this consideration. The key idea was to find a method to estimate the sensors that does not depend on the aiding sources. The approach and the corresponding results will be presented mainly in the following. As a starting-point, the basic way of using a Kalman filter (or the background Wiener theory) in noise filtering is described. Although the Kalman filter is often used in the complementary form, the original goal of the Wiener theory was the separation of a signal from noise, when only a single combination of both was available. The functioning of the filter in this configuration is explained below.

5.2. Use of the Wiener/Kalman theories in noise filtering problems

5.2.1. The Wiener filter problem

The Kalman filter provides an alternative approach to the Wiener filter problem using state space methods (see Chapter 3). Kalman's contribution has been essential in applied work because its solution is readily implemented with modern digital methods and because it does not require the stationarity assumption. However, the theoretical problem was solved through Wiener filter theory, which will now be briefly characterised.

The Wiener filter was originally designed to separate a signal from a noise when both signal and noise are noiselike in character and when there may be an overlap in their spectra. The signal and noise are defined as random processes with known spectral characteristics or equivalently, known auto- and cross-correlation functions. The Wiener filter consists of a linear operation, described by a transfer function, applied on the input combination of noise and signal to yield the best separation by minimising the average squared error. This type of filtering is loosely referred to as *least square filtering*, though

linear minimum mean-square error filtering would be a more descriptive name.

5.2.2. Design of the Kalman filter

Straightforward applications of the Wiener filter to classical noise filtering problems are not as common place as one might expect. One reason may be that it demands that the signal be noiselike in character whereas in most applications signals have at least some deterministic structure. Therefore, the Wiener filter has been used rather in the complementary form (see Chapter 4), where the transfer function affects only the noises without distorting the signal. This is typically an instrumentation application where redundant measurements of the same signal are available and where the problem is to combine all the information in such a way as to minimise the instrumentation errors.

Leaving the complementary design aside, let us now consider how a Kalman filter is designed in order to separate two random processes in the case where an input combination of both signal and noise is available.

The state vector is composed of all the variables describing the models of the random processes to be estimated. Additional state variables are appended to account for either non-white state or measurement noise.

The state transition matrix is directly deduced from the chosen models. Likewise, the state noise covariance matrix is computed from the state equations.

The observation vector is simply the measured signal, which is a mix of the signal and noise. Therefore, the observation matrix is usually a combination of unit matrices and zeros matrices according to the state vector. A more complicated form of the design matrix results when the updating measurement is a linear combination of a number of states.

Once the system is properly designed, the algorithm of the discrete Kalman filter is used to estimate the signal. It is worth noting that the stages of update and prediction arise at the same rate.

5.2.3. Examples

Signal and noise purely noiselike in character

In the following example, the signal is assumed an exponentially correlated process (correlation time α = 0.01). Therefore, a single state variable is used to model it:

$$\mathbf{x}_{k+1} = e^{-\alpha}\mathbf{x}_k + \alpha\sqrt{1 - e^{-2\alpha}}\,\mathbf{w}_k \qquad (5.3)$$

The measurement noise is assumed white of strength 0.01. Of course, the noise could also be coloured (e.g., random walk) and modelled accordingly in the state vector. Here, no additional state variable is needed. The state matrix is reduced to the scalar $e^{-\alpha}$, and the observation matrix equals simply 1. The state noise covariance is directly deduced from (5.3). The measurement noise covariance is given by the variance of the white noise. Figure 5.1 shows that the filter is able to separate the signal from the noise.

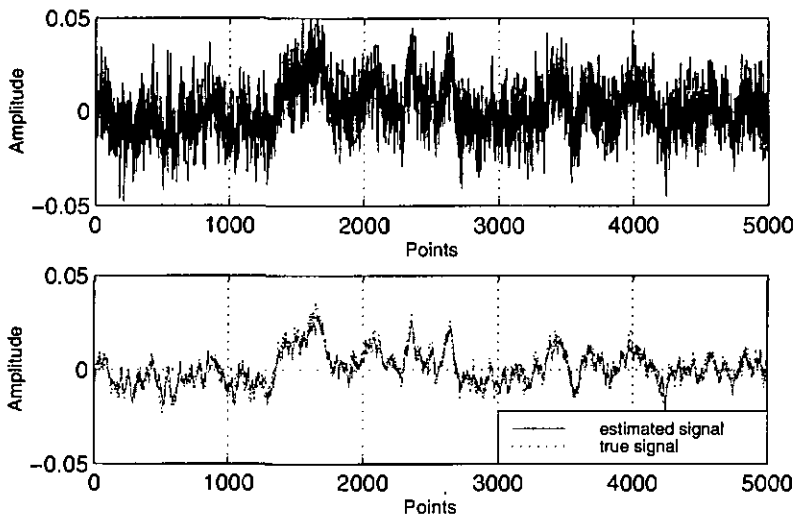

Figure 5.1: Estimation of an exponentially correlated process. Top: measurement – Bottom: estimated signal compared with true signal.

Deterministic structure of the signal

Let us now consider that the signal has a deterministic structure but that some parameters (e.g., slope of a ramp, amplitude of a sine) are unknown. A Kalman filter can be used to estimate them. In the following test, the signal is assumed a ramp. Two state variables are needed to model it. x_1 is the signal itself and x_2 is the slope. Since the signal is assumed purely deterministic, no white noise is added in the state equations:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_k \tag{5.4}$$

The corrupting noise is assumed an exponentially correlated process, though any random model (e.g., ARMA) would be convenient, leading to the addition of variables to the state vector. Here, one state variable is added to the state vector. The state matrix is completed accordingly. Besides, white noise of strength 0.1 is superposed. It is included in the measurement equations. The covariance matrix of the measurement noise is the variance of the white noise. The measurement matrix is:

$$H = (1 \quad 0 \quad 1) \tag{5.5}$$

meaning that the measurement is modelled as the sum of the ramp and the exponentially correlated process. In this example, the value of the slope is unknown. An arbitrary initial value for x_2 of 10 is provided to the filter whereas the true value is actually 0.001. The corresponding term of the initial error covariance matrix is set to a non-trivial value to allow the filter to change this value. Figure 5.2 shows the ability of the filter to separate the ramp from the noise and to recover the true value of the slope although a wrong initial value has been provided to the algorithm.

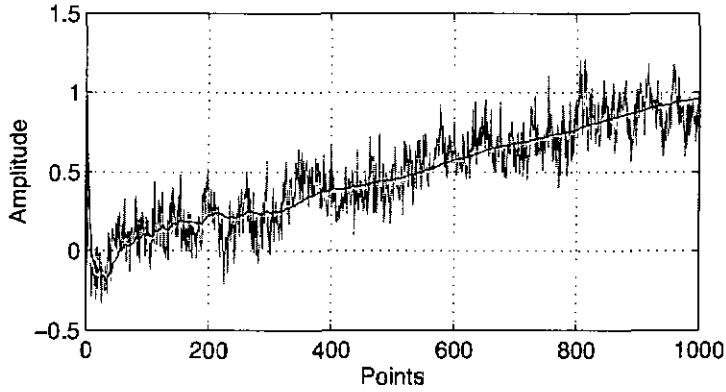

Figure 5.2: Estimation of the slope of a deterministic ramp. Sum of a ramp, an exponentially correlated process and white noise (grey) – Ramp estimated by the filter (black).

5.3. The pseudo-model

5.3.1. Motivation

The above examples have illustrated the correction of a signal using only one measurement, contrary to the complementary concept. The Kalman filter is able to separate a signal from noise, provided suitable models for both signal and noise are defined. The processes can be purely or partially random, stationary or not stationary. The method has therefore an interest in a wide range of applications since many types of signals are concerned. Its interest in sensor correction problems is obvious. The algorithm allows improving the sensor output, whatever the nature of the physical signal and the sensor errors. Let us for example consider an application of vibration damping. A control system uses microaccelerometer measurements to generate an anti-phase signal to cancel vibrations. The quality of the damping depends straightforwardly on the quality of the measurement. The method described in § 5.2.2 can be used to correct the accelerometer output. The vibration can reasonably be

modelled by a sine of which either amplitude or frequency may be unknown. The accelerometer errors can also be modelled. A properly designed filter will be able to track the sine. On one hand, it will separate the sine from other noise, including the accelerometer noise. On the other hand, it will estimate the unknown parameters of the sine.

Despite its interest, this method is seldom used, to the extent of our knowledge. Its main drawback is that it requires a model of the signal whereas in some applications no assumptions about the signal that the sensor measures are available. Especially in navigation applications, this requirement is too restrictive. Hence, systems are usually designed regarding the complementary concept (see Chapter 4). The drawback, in turn, is that errors are estimated only whenever the aiding source is available. Nevertheless, in [Bars93], a system designed for robotic navigation is proposed where the rates of prediction and estimation are the same. The observation vector is composed of the measured angular rates and accelerations. The state vector contains the true values of position, velocity and acceleration and the sensor error models. The acceleration and angular rate signals are modelled with a random constant.

In the following, an alternative design is proposed where the offset and the sensitivity error of a sensor are modelled whereas no a priori information at all is given on the signal to be estimated [Mars97] [Mars98].

5.3.2. Design of the Kalman filter

A reasonable model for the output a_m of a sensor is:

$$a_m = (k_0 + k) \cdot a + b \tag{5.6}$$

where a is the true signal. k_0 is the nominal scale factor given by the manufacturer. k is the sensitivity error. b is the offset variation. In the following, only these two errors are considered.

The state vector is chosen as containing the sensitivity and offset variations as well as the physical signal:

$$\mathbf{x} = (k \;\; b \;\; a)^T \tag{5.7}$$

The number of variables to describe k and b depends on the models found during the identification stage. In the following, the intermediate state variables are ordered in such a way that the first one corresponds to the output of the shaping filter. The main feature of the design is that only one state variable is attributed to the signal, whatever its real shape.

The first rows of the state matrix are composed with respect to the shaping filters of the sensor errors. The state equation corresponding to the signal is obtained from (5.6):

$$a = \frac{a_m - b_1}{k_o + k_1} = f(k_1, b_1) \tag{5.8}$$

Since the function f in (5.8) is non-linear, the last row of the state matrix is given by the partial derivatives of f regarding the state variables:

$$\Phi(n,1) = \frac{\partial f(k_1, b_1)}{\partial k_1} = -\frac{a_m - b_1}{(k_0 + k_1)^2} \tag{5.9}$$

$$\Phi(n, p+1) = \frac{\partial f(k_1, b_1)}{\partial b_1} = -\frac{1}{k_0 + k} \tag{5.10}$$

$$\Phi(n,n) = \frac{\partial f(k_1, b_1)}{\partial \omega} = 0 \tag{5.11}$$

The observation of the system is simply the measurement of the sensor. Therefore, the measurement equation is:

$$a_m = (k_o + k_1) \cdot a + b_1 = h(k_1, b_1, a) \tag{5.12}$$

Due to the non-linearity of (5.12), the elements of the observation matrix are obtained as follows:

$$H(1) = \frac{\partial h(k_1, b_1, a)}{\partial k_1} = a \tag{5.13}$$

$$H(p+1) = \frac{\partial h(k_1, b_1, a)}{\partial b_1} = 1 \tag{5.14}$$

$$H(n) = \frac{\partial h(k_1, b_1, a)}{\partial \omega} = k_o + k_1 \tag{5.15}$$

This design is referred to as *pseudo-model* in the following.

5.3.3. Simulation results

The behaviour of the filter using the pseudo-model has been studied through simulations. Sensor measurements are generated as the sum of a true signal and an offset. In the first case shown, the true signal is a step whereas in the second example, it is a sum of 2 sine waves. The offset is assumed an ARMA noise defined by 2 poles ($p_1 = -1$, $p_2 = 0.5$) and 1 zero ($z_1 = 0.1$). It is described with two state variables (see Chapter 3):

$$\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & 1 \\ -0.5 & 0 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}_k + \begin{pmatrix} 1.1 \\ -0.5 \end{pmatrix} \cdot w_k \qquad (5.16)$$

The sensor offset b is given by the first state variable b_1. The white noise w_k has a variance of $\sigma^2 = 0.01$.

The sensitivity is assumed constant and modelled accordingly with one state variable.

Obviously, different models for the offset or for the scale factor as well would lead to a different set of variables. The physical signal, in turn, is always described with one state variable, whatever its nature (5.8).

According to § 5.3.2, the state vector has 4 elements (2 for the ARMA noise, 1 for the scale factor error model and 1 for the signal pseudo-model).

Figure 5.3 and Figure 5.4 show the ability of the filter to recover the real signal, although no model for the signal has been provided.

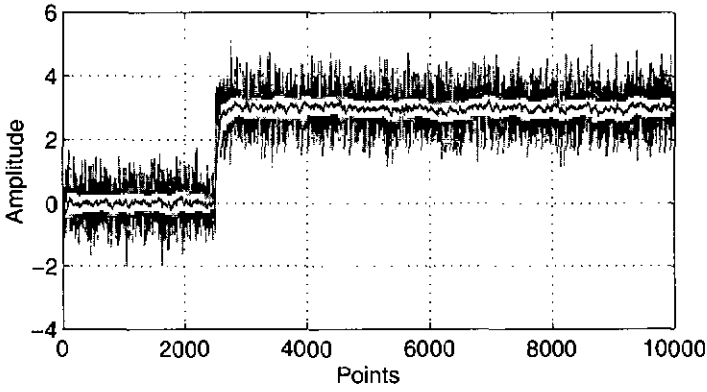

Figure 5.3: Estimation of a step using the pseudo-model. The offset is assumed an ARMA signal.

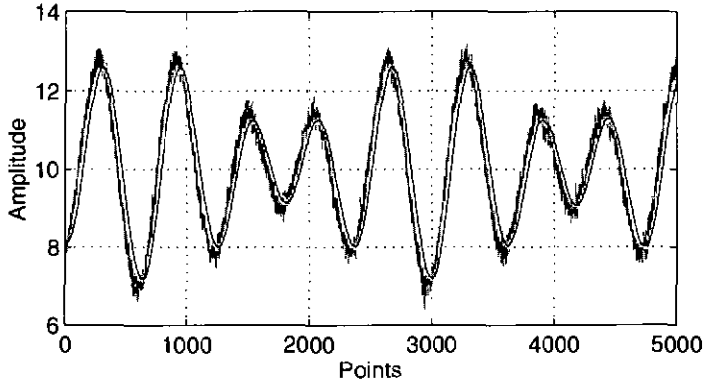

Figure 5.4: Estimation of a sinusoidal signal using the pseudo-model. The offset is assumed an ARMA signal.

5.4. Offset models for inertial microsensors

5.4.1. Methodology

Models of the sensor offsets are determined using the system identification procedure described briefly in § 3.5. These models will be further used in the Kalman filter to reduce the sensor errors. The proposed models are *a posteriori* models since they are derived from empirical data. No decision about the model is made regarding a physical principle. They are also *black-box* models since only the output of the sensor and not the internal structure is considered. The modelling stage in this research is restricted to fix parametric models. Adaptive modelling is not investigated. The models are established during an off-line processing. Once they are validated, they are included in the system description and no more modified, whatever the results of the correction done by the Kalman filter.

The first step of the identification procedure consists of designing an experiment to collect observations. It includes hardware and software design as well as specifications of the type of acquisitions. Secondly, data are carefully examined and a model structure is selected accordingly. Next, the parameters of the model are identified by using a linear regression approach. Finally, the model is validated through comparison with several batches of empirical data. The residual represents misfits between data and model. On one hand, the presence of any information remaining in the residuals is a clue that the model might be insufficiently complex or otherwise inappropriate. On the other hand, a major objective is to obtain a model at least possible complexity in order to reduce the computation load in subsequent processing. Therefore, despite a lack of fit with the data, the model may be accepted, provided the required accuracy is reached.

The *System Identification Toolbox* of *Matlab*, which provides basic routines for identification techniques, is used to determine the models. The most common techniques for computing models are available, including parametric and non-parametric estimation. Functions for evaluating the models are also provided. The software allows using different

representations (e.g., state space, polynomial and poles/zeros). The identification result, including the model structure and its estimated parameters as well as other relevant information, is contained in a packed matrix called *theta matrix*. This is the basic format with the *System Identification Toolbox*. It is used by all parametric identification methods and can be easily transformed into many model representations. The *System Identification Toolbox* is matched directly to the textbook [Ljun87]. However, it is also worth reading [Joha93], which provides a more comprehensive treatment of the system identification theory.

5.4.2. Data acquisition

The sensors are fixed on a gyro test bench (ACUTRONIC AG simple axis position rate table 120) isolated as much as possible from the floor vibrations. When the table is rotating, the spectrum of the signal shows peaks due to mechanical resonances of the table. Therefore, although the sensitivity has been considered in the theoretical derivation of the filter, attempts to model the scale factor error have been forsaken. To model the offset, the sensor is kept immobile whereas varying velocity tests are realised to show the correction of the signal obtained with the Kalman filter. The output voltage of the sensor is recorded using the 16 bit A/D converter of the DSP card (see Appendix C). Figure 5.5 shows the experimental set-up.

The sampling frequency is set to 1280 Hz to avoid aliasing. A real-time decimation (1/5) after suitable low-pass filtering is done to reduce the size of the data files. Finally, during the off-line processing stage, the signals are resampled again from 256 Hz to 32 Hz. The suitable bandwidth for navigation is indeed in a ten Hz range.

There is no absolute rule to set the duration of acquisitions. It just must be long relative to the typical variations of the signal. Spectra obtained with different acquisition lengths have been compared. Trials have shown that an acquisition time of about 10 min is sufficient to model the offset.

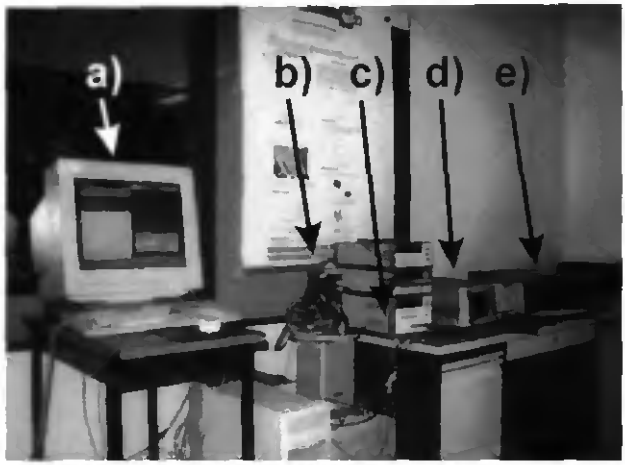

Figure 5.5: View of the experimental set-up with: a) the computer with the DSP card, b) the rotating table with the IMT gyro in a metallic box, c) the power supply, d) the lock-in amplifier and e) the controller for the rate bench.

5.4.3. IMT gyro offset model

The sensor output is examined after a warm-up to limit the influence of the temperature variation. The signals exhibit a low-frequency drift with a superimposed noise of higher frequency (Figure 5.6). The sensitivity of the sensor is 88 μV/deg/s.

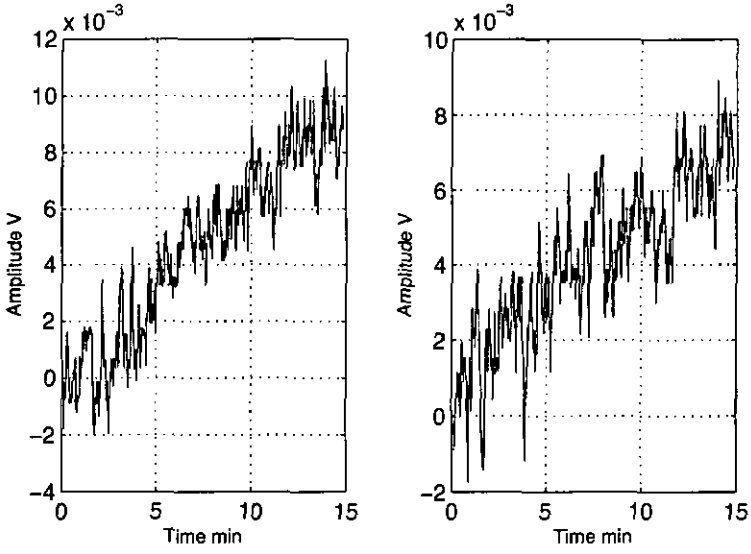

Figure 5.6: Typical offset realisations of the IMT gyro.

It has been decided to separately model both effects since attempts to find a single model to fit the entire signal had failed. First, the offset signal is filtered in order to select only the DC component. Next, the higher frequency residual is simply obtained by subtracting the DC part from the signal and is studied separately.

Model of the higher frequency noise

Once the DC trend of the offset is removed, the model for the residual noise is estimated with an ARMA structure. Providing enough poles as well as zeros allows indeed fitting all peaks and valleys of the spectrum. The number of poles and zeros results from a trade-off between the size and the quality of the model. All details of the spectral content can be modelled, provided a large amount of poles and zeros is considered. Once the size of the model is chosen, the parameters are estimated using the iterative Gauss-Newton algorithm [Joha93]. Figure 5.7 shows the poles and the zeros estimated when a structure of 3 poles and 2 zeros is chosen.

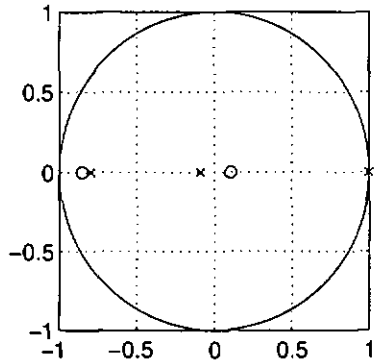

Figure 5.7: Poles and zeros of the ARMA model – 'o' denoting zeros and 'x' poles.

In order to validate the model, the procedure is done on several batches of data. Close results are obtained. An additional check consists of generating an ARMA signal with respect to the model and comparing its PSD with the PSD of the true signal. Figure 5.8 shows the frequency fitting between the model and the signal. Moreover, the autocorrelation of the residuals is computed. The result is a Dirac function, which proves the adequacy of the model (Figure 5.8).

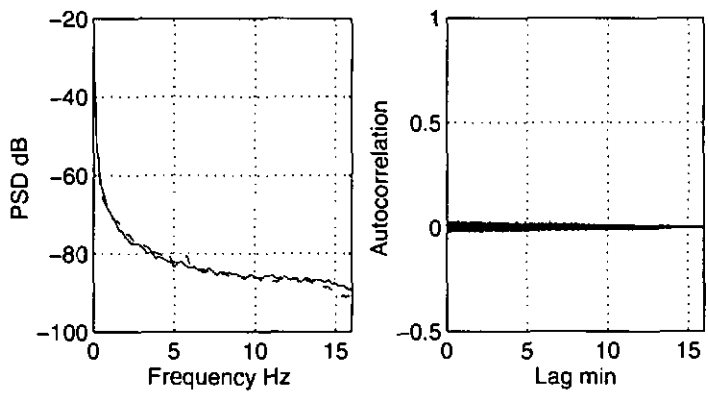

Figure 5.8: Left: Comparison between the PSD of the model (solid line) and the PSD of the measurement (dotted line). Right: Autocorrelation of the residuals.

Model of the low-frequency drift

The IMT gyro exhibits a drift, which is obviously not stationary. This unbounded trend has to be modelled with an unstable system. Experiments have shown that the drift can be described by an integrated random walk, which is the process obtained by integrating twice a white noise. It has 2 poles at +1.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_{k+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_k + \begin{pmatrix} 0 \\ \sigma \end{pmatrix} \cdot w_k \qquad (5.17)$$

5.4.4. Murata gyro offset model

Similarly to the IMT gyro, the Murata gyro can be described by a low-frequency drift and an additional noise of higher-frequency (Figure 5.9). For the sake of brevity, the description of the modelling procedure is not repeated. In this experiment, the sensitivity of the gyro is 30 mV/deg/s.

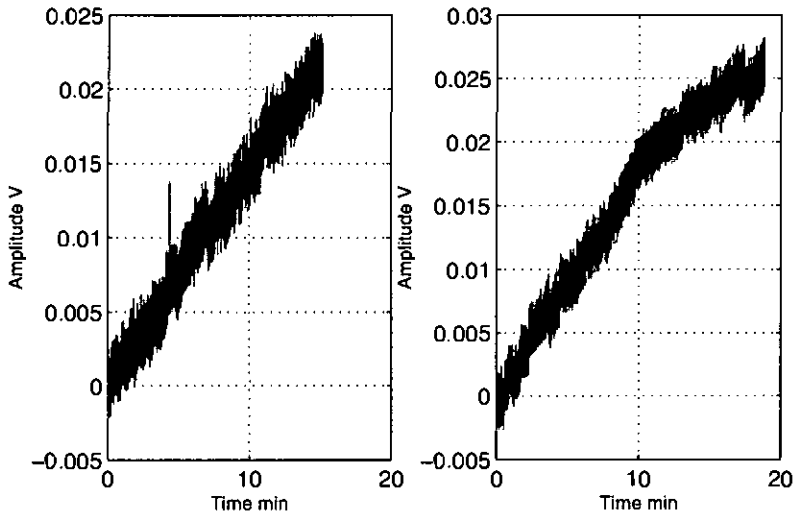

Figure 5.9: Typical offset realisations of the Murata gyro.

5.5. Correction of the sensor offset

5.5.1. IMT gyro correction

The Kalman filter corresponding to the pseudo-model design is now used to correct the output of the IMT gyro during a varying velocity test. The offset is modelled as the sum of an integrated random-walk and an ARMA signal with respect to the model found in § 5.4.3. No model is given for the signal to be estimated. The state vector contains 8 variables altogether (one for the scale factor, 3 for the ARMA component of the offset, 2 for the DC drift of the offset, 1 for the entire offset, 1 for the pseudo-model of the signal). The sensor output is recorded during 35 min (Figure 5.10). In the middle of the experiment, the angular velocity is changed from 0 deg/s to 100 deg/s. The raw output of the sensor exhibits a large offset drift, leading quickly to unacceptable results. After correction, the high-frequency noise is attenuated whereas the DC drift is removed. Thus, the accuracy is drastically increased. However, a low-pass effect appears when the signal changes suddenly. The adaptation of the algorithm lasts for a while. Tests have shown that the choice of the noise measurement matrix has direct influence on this low-pass effect and that a trade-off has to be found between the attenuation obtained and the adaptation of the filter.

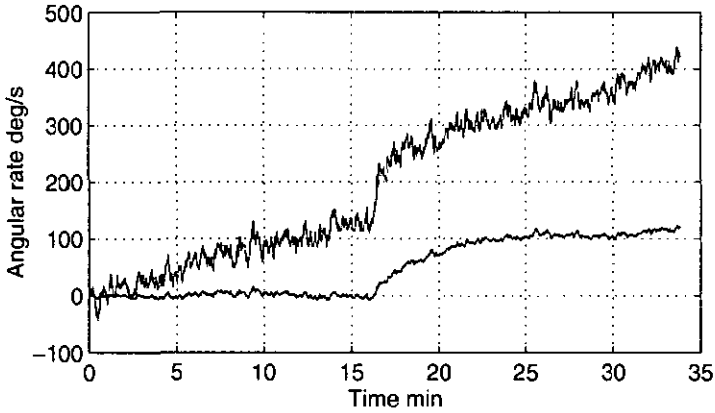

Figure 5.10: Correction of the IMT gyro during a step change of angular velocity.

5.5.2. Murata gyro correction

The procedure described in the previous section is similarly applied to the Murata gyro. A suitable model is defined as the sum of an integrated random walk and an ARMA signal. The parameters of the models are obviously different from those defined for the IMT gyro. Hence, the slope of the ramp is smaller since the Murata gyro is far more accurate (see Chapter 2). Figure 5.11 shows an experiment where the angular velocity is changed from 50 deg/s to 75 deg/s in the middle of the recording period of 35 min. The correction algorithm allows reducing the drift.

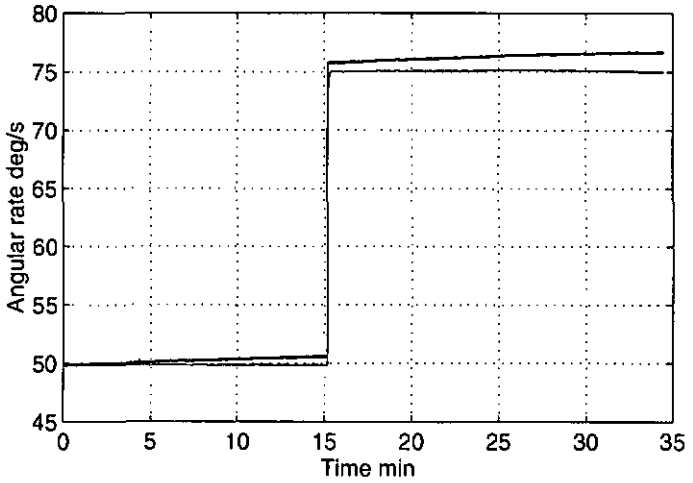

Figure 5.11: Correction of the Murata gyro during a step change of angular velocity.

5.6. Summary

A Kalman filter is well suited to eliminate time-dependent sensor errors, which are either partially or totally random, provided adequate state space models could be defined to describe the undesired phenomena. In classical methods, these error models are put in the basic state vector of the navigation filter and estimated whenever the aiding source is available. In this study, an additional filter is designed where the measurement is the sensor output itself. The state vector contains the models of both the signal and the errors. This method allows improvement of the sensor output without external aiding and thus widens the range of applications. However, the need of a model for the signal may be too restrictive for some applications. Therefore, an alternative approach, referred to as pseudo-model, is proposed where the model of the signal to be estimated is derived from the measurement equations. The algorithm is used to improve the IMT gyro and the Murata gyro. The study focuses on the correction of their offsets. The modelling stage shows that the

offset can be described with two components: an integrated random walk and an ARMA signal. Important corrections of velocity increasing step-wise are obtained by reducing the sensor drift.

5.7. References

[Bars93] B. Barshan, "Inertial Navigation Systems for Mobile Robots", University of Oxford, Department of Engineering Science, Robotics Research Group, report n°1998/93, 1993.

[Bert96] C. Berthoud, "Effective Static Response Compensation Suitable for Low-Power ASIC Implementation with an Application to Pressure Sensors", IEEE Instrumentation & Measurement Technology Conference, IMTC/96, Brussels (Be), pp. 1168-1173, June 4-6 1996.

[Joha93] R. Johansson, "System Modeling and Identification", Prentice Hall, Englewood Cliffs, New Jersey, 1993.

[Lapu90] D. Lapucha, "Precise GPS/INS Positioning for a Highway Inventory System", PhD thesis, University of Calgary, Department of Surveying Engineering, Calgary (Ca), 1990.

[Ljun87] L. Ljung, "System Identification: Theory for the User", Prentice Hall, Upper Saddle River, New Jersey, 1987.

[Mars97] C. Marselli, "Error Modelling of a Silicon Angular Rate Sensor", Symposium Gyro Technology, Stuttgart (D), pp. 4.0-4.9, September 16-17 1997.

[Mars98] C. Marselli, "Application de Filtres de Kalman à la Réduction de Bruits de Microcapteurs", Colloque C2I 98, Cachan (F), November 18-19 1998, to be published.

[Schw83] K. P. Schwarz, "Inertial Surveying and Geodesy", Reviews of Geophysics and Space Physics, Vol. 21:4, pp. 878-890, May 1983.

[Schw90] K. P. Schwarz, "Testing a Decentralised filter for GPS/INS Integration", IEEE Position Location and Navigation Symposium, Las Vegas, pp. 429-435, March 20-23 1990.

[Schw94] K. P. Schwarz, "Aided versus Embedded: A Comparison of two approaches to GPS/INS Integration", IEEE Position Location and Navigation Symposium, Las Vegas, pp. 314-322, April 11-15 1994.

[Wagn96] J. Wagner, "Aspects of Combining Satellite Navigation and Low Cost Inertial Sensors", Symposium Gyro Technology, Stuttgart (D), pp. 11.0-11.20, September 1996.

Chapter 6
GPS and position representation

In Chapter 4, means of fully determining the position and the heading of a vehicle with a set of accelerometers and angular rate sensors have been presented. They involve strapdown mechanisation equations and integration of an aiding source through the use of a Kalman filter.

In this study, these methods are applied to microsensors and illustrated in real-field tests where microaccelerometers and microgyros are used to compute the trajectory of a car while a GPS receiver provides updating fixes.

Practical issues of the experiment are given in Chapter 7. However, for the sake of clearness, this chapter gathers basic elements of a navigation application. It concerns parts of the work that are required by the navigation problem but without relation with inertial sensors.

First, different features of the GPS (Global Positioning System) are reviewed. Next, the problem of position co-ordinates is addressed.

6.1. Global Positioning System

6.1.1. Principle

The *NAVSTAR-GPS* (*NAVigation System by Timing And Ranging-Global Positioning System*) is a satellite-based radio navigation system providing means of determining position, velocity and time around the globe [Hofm94]. Initially designed by the *U.S. Department of Defense* (*DoD*), it guarantees also civilian access through an agreement between the DoD and the *Department of Transportation* (*DoT*).

The goal is to get a fix on a position by receiving information from satellites. The basic principle lies on *trilateration*: the receiver's position is determined by using geometry of triangles using only distance measurements. Knowing the position of a satellite, the measurement of the distance from a satellite to the receiver determines a sphere. Provided three satellites do not lie on a straight line, the intersection of the three spheres gives two points, where one can be discarded and the other is the desired point (Figure 6.1).

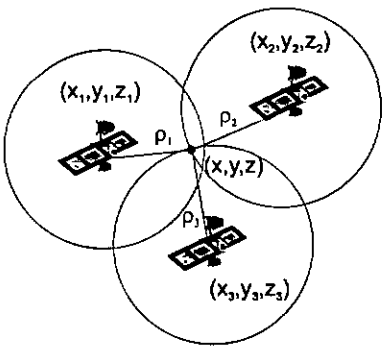

Figure 6.1: 3D Positioning principle.

The measurement of the satellite-receiver distance is based on the principle of time of arrival ranging. The time interval taken for the signal transmitted from the satellite to reach the receiver is measured and multiplied by the speed of light giving

the transmitter-receiver range. However, due to the clock offset between satellite and receiver, the actual range cannot be known exactly and the measured range is referred to as *pseudo-range*. Therefore, a fourth satellite is needed to calculate both the clock offset and the receiver's position. The corresponding equations are:

$$\begin{aligned}
\rho_1 &= \sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}+c\cdot\Delta t\\
\rho_2 &= \sqrt{(x-x_2)^2+(y-y_2)^2+(z-z_2)^2}+c\cdot\Delta t\\
\rho_3 &= \sqrt{(x-x_3)^2+(y-y_3)^2+(z-z_3)^2}+c\cdot\Delta t\\
\rho_4 &= \sqrt{(x-x_4)^2+(y-y_4)^2+(z-z_4)^2}+c\cdot\Delta t
\end{aligned} \qquad (6.1)$$

where x_i, y_i, z_i are the known satellite positions, ρ_i are the measured pseudo-ranges, c is the speed of light, Δt is the unknown receiver clock offset from GPS time and x, y, z are the unknown co-ordinates of the receiver's position.

Besides position, the radial velocity of a vehicle can also be determined with the GPS. Because of the relative motion of the GPS satellites with respect to a moving vehicle, the frequency of the signal broadcast by the satellite is shifted when received at the vehicle. This measurable Doppler shift is proportional to the relative radial velocity. Since the radial velocity of the satellites is known, the radial velocity of the moving vehicle can be deduced from the Doppler observations.

6.1.2. GPS signals

Like most radio signals, GPS signals are modulated carrier signals. Two carriers have been chosen in the L band (L1: 1575 MHz, L2: 1227.6 MHz) to satisfy a trade-off between atmospheric absorption and bandwidth. The two modulation signals are *pseudorandom noise* (*PRN*) codes. PRN codes allow the use of small signal levels suitable for the small antennae receiver and attributing a different code to each satellite. The civilian signal called *C/A code* has a pulse wavelength of 300 m, whereas the military signal (*P code*) has a pulse wavelength of 30 m, leading to an intrinsic higher precision. In addition to the PRN codes, a navigation message is broadcast

on both carriers at the slow rate of 50 bits/s. It contains a large amount of information to assist the receiver while computing the position, among others:

- precise ephemeris data of the satellite orbit, which will be used to compute the satellite position,
- the *almanac* which contains rough ephemeris data of other satellites to predict when new satellites will become available to the receiver.

6.1.3. GPS system components

The system is composed of three main parts: the *space segment* (satellites), the *user segment* (receiver) and the *ground segment* (management).

The space segment consists of a constellation of 24 satellites arranged in six orbital planes, each inclined at 55 ° to the equatorial plane. Orbits are nearly circular, 20000 km in altitude and repeat exactly twice per sidereal day (≈ 12 h). The constellation is designed to ensure a worldwide coverage 24 hours a day. However, the reception requires a clear line of sight since the signal is distorted by foliage, buildings and other.

The user segment refers to the receivers. The function of the receiver is to internally generate the same PRN codes as the satellites and to perform the correlation between these internal codes and those externally received from the satellites. Hence, they measure the propagation time and deduce the pseudo-range. They also do compensation of known errors (atmospheric delays) and finally implement the solution to the navigation equations (6.1) using a least squares algorithm or a Kalman filter algorithm [Rnet96]. They can compute the positioning fix at a typical rate of 1 Hz.

The ground segment consists of five ground stations, which maintain the overall system time and update the orbital parameters of the satellites.

6.1.4. Differential GPS

Using the C/A code the physically achievable 40 m precision is deliberately degraded to 120 m (2σ rms) by the *selective availability* (*SA*), which is the introduction of errors into the broadcast ephemeris and other. Additional physical errors are atmospheric delays or multi-path effects. Moreover, the effect of these errors can increase due to a bad geometry of the satellites. Thus, the accuracy of the standard system is far too low for many applications.

The technique of *differential GPS* (*DGPS*) allows dramatically improving the system accuracy. DGPS is based on a simple idea. A 'home' receiver, sometimes referred to as a *rover*, whose position is exactly known compares the measured pseudo-ranges with the actual distances and transmits the computed errors to receivers that are reasonably close (less than 100 km), so they can correct their own pseudo-ranges by that same amount.

In Switzerland, a DGPS service provided through a collaboration of the *Federal Office of Topography* and *Swisscom* is available since January 96 (Figure 6.2). The reference GPS antenna is placed in Zimmerwald, near Bern. Corrections are broadcast over a network of 6 FM radio stations (see Appendix D), using the *RDS*, which is a European standard for distributing digital data on broadcast FM. The coverage attains 60 % of the country but will be extended in a near future.

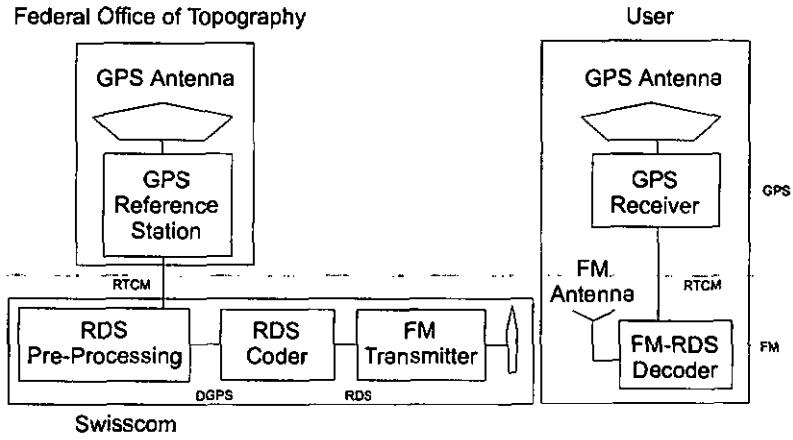

Figure 6.2: DGPS system. RTCM SC-104 is the format used for DGPS data transmission. It is maintained by the Radio Technical Commission for Maritime Service.

6.1.5. NMEA standard

To send positioning information to other instruments, GPS receivers use the *NMEA (National Marine Electronics Association)* standard, which defines the electrical interface and data protocol for communications between marine instruments. The NMEA data is transmitted as plain ASCII text at 480 characters/s in a sequence of sentences up to 80 characters in length each. A sentence begins with the $ mark, followed by the type of device, the sentence type and the relevant information. The end is an optional checksum that can be used to check for corrupted sentences. For example, in the case of the GPS receiver, a possible sentence structure is:

$GPRMC, 225446, A, 4659.72, N, 00656.60, E, 000.5, 054.7, 191197, 000.8, E*68

where GP stands for GPS, 225446 is the UTC time, 4659.72, N means latitude 46°59.72' north, 00656.50, E means longitude 6°56.50' east, 0.5 is the velocity (knots), 54.7 is the heading (degrees), 191197 is the date, 0.8 is the magnetic variation (degrees), *68 is the checksum.

6.2. Position co-ordinates

Various co-ordinate systems are used around the world. Thus, maps with different references cannot be blindly mixed. In this study, geographical transformations are needed in order to blend GPS data with local inertial position. This section mentions the main steps of the procedure as well as basic geographical concepts. The complete algorithm can be found in Appendix E.

6.2.1. Earth representation: ellipsoid, geoid and datum

In geodesy, a reasonable approximation to the earth is an *ellipsoid* of revolution slightly flattened at the poles. The *geoid* is the surface defined physically as the surface along which the gravity potential is constant at all points and for which the gravity vector is normal to all points. It coincides with mean sea level, passing through all landmasses as well as the sea areas. The geoid is often visualised as a contour map of height above or below the reference ellipsoid (Figure 6.3) [Stra97].

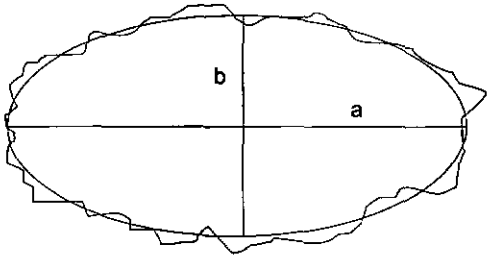

Figure 6.3: Geoid and ellipsoid. a and b are respectively the semi-major axis and the semi-minor axis of the ellipsoid.

It is desirable that the rotational ellipsoid and the geoid match as closely as possible within the area under consideration. Thus, many regional authorities have adopted their own ellipsoid. The set of parameters defining the shape of the ellipsoid as well as its position with respect to the geoid is

referred to as a *datum*. The co-ordinate reference used by the GPS is *WGS-84*, which stands for World Geodetic System 1984, whereas the datum *CH-1903* is used in Switzerland (Table 6.1).

	a [m]	b [m]
WGS-84	6 378 137	6 356 752
CH-1903	6 377 397	6 356 079

Table 6.1: Ellipsoid axis of the global and local datum.

The transformation between two ellipsoids is expressed using the Helmert transformation [Hofm94]:

$$\mathbf{x}_2 = \mathbf{d} + \mu \cdot \mathrm{R} \cdot \mathbf{x}_1 \tag{6.2}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the cartesian co-ordinates in the first and the second ellipsoid respectively, $\mathbf{d}$ is a translation (shift) vector, μ is a scale factor and R is a rotation matrix:

$$\mathrm{R} = \begin{pmatrix} c\beta \cdot c\gamma & s\alpha \cdot s\beta \cdot c\gamma + c\alpha \cdot s\gamma & s\alpha \cdot s\gamma - s\beta \cdot c\alpha \cdot c\gamma \\ -s\gamma \cdot c\beta & c\alpha \cdot c\gamma - s\alpha \cdot s\beta \cdot s\gamma & s\alpha \cdot c\gamma + s\gamma \cdot c\alpha \cdot s\beta \\ s\beta & -c\beta \cdot s\alpha & c\alpha \cdot c\beta \end{pmatrix} \tag{6.3}$$

For converting CH-1903 into WGS-84, these parameters are set to [Topo2]:

$$d_x = 660.077 \text{ m}, \; d_y = 13.551 \text{ m}, \; d_z = 369.344 \text{ m}$$

$$\alpha = 2.484 \text{ cc}, \; \beta = 1.783 \text{ cc}, \; \gamma = 2.939 \text{ cc}$$

$$\mu = 1.00000566 \tag{6.4}$$

For converting WGS-84 into CH-1903, (6.2) is simply inverted to obtain $\mathbf{x}_1$ with respect to $\mathbf{x}_2$.

6.2.2. Co-ordinates

Ellipsoidal co-ordinates

A point is usually characterised by its *geodetic* or *ellipsoidal* co-ordinates. The *ellipsoidal height* h is the perpendicular distance from the point to the ellipsoid. The *geodetic latitude* ϕ is the angle between that perpendicular and the equatorial plane.

The *geodetic longitude* λ is the angle between the meridional plane (through the polar axis of the ellipsoid) and a plane perpendicular to the equatorial plane and containing the normal (Figure 6.4).

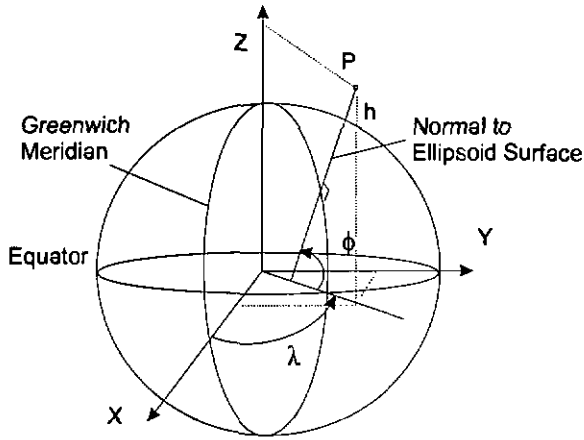

Figure 6.4: Geodetic and cartesian co-ordinate system.

Any point in a geodetic ellipsoidal co-ordinate system can be translated to a point in cartesian co-ordinates with:

$$\begin{aligned} x &= (N+h)\cdot\cos(\phi)\cdot\cos(\lambda) \\ y &= (N+h)\cdot\cos(\phi)\cdot\sin(\lambda) \\ z &= \left(\frac{b^2}{a^2}\cdot N + h\right)\cdot\sin(\lambda) \end{aligned} \tag{6.5}$$

where:

$$N = \frac{a^2}{\sqrt{a^2\cdot\cos^2(\lambda)+b^2\cdot\sin^2(\lambda)}} \tag{6.6}$$

The inverse formulae to obtain ellipsoidal co-ordinates from cartesian co-ordinates are not straightforward and the solution is either an exact iterative procedure or a direct approximated equation [Hofm94].

Map projection

Representation of the 3D earth on a 2D-mapping plane requires a projection. The projections in geodesic applications preserve angles. The mapping used by the Swiss system is the cylindrical projection around a transverse axe with the origin at the old observatory of Bern, whose co-ordinates are:

$$(X_0, Y_0) = (200000\ \text{m}, 600000\ \text{m})$$

$$(\phi_0, \lambda_0) = (46°57', 7°26') \qquad (6.7)$$

The exact procedure is given in [Topo1].

Direct and inverse conversion between WGS-84 and CH-1903

The GPS receiver provides ellipsoidal co-ordinates using the global datum WGS-84, whereas maps in Switzerland are based on the local datum CH-1903 with plane cartesian co-ordinates.

In this study, the navigation algorithms include conversion between both sets of co-ordinates. On one hand, WGS-84 ellipsoidal co-ordinates of the GPS need to be transformed into CH-1903 cartesian co-ordinates to be blend with the trajectory computed with the inertial sensors. On the other hand, the resulting INS-GPS position given in CH-1903 cartesian co-ordinates must be transformed into WGS-84 ellipsoidal co-ordinates for display needs (see Chapter 7).

According to the previous explanations, the procedure involves several steps, which are:

- conversion from ellipsoidal to cartesian co-ordinates and vice versa,
- datum conversion from WGS-84 to CH-1903 and vice versa,
- projection of the Swiss ellipsoid.

The whole conversion is summarised in Figure 6.5. The formulae are given in Appendix E.

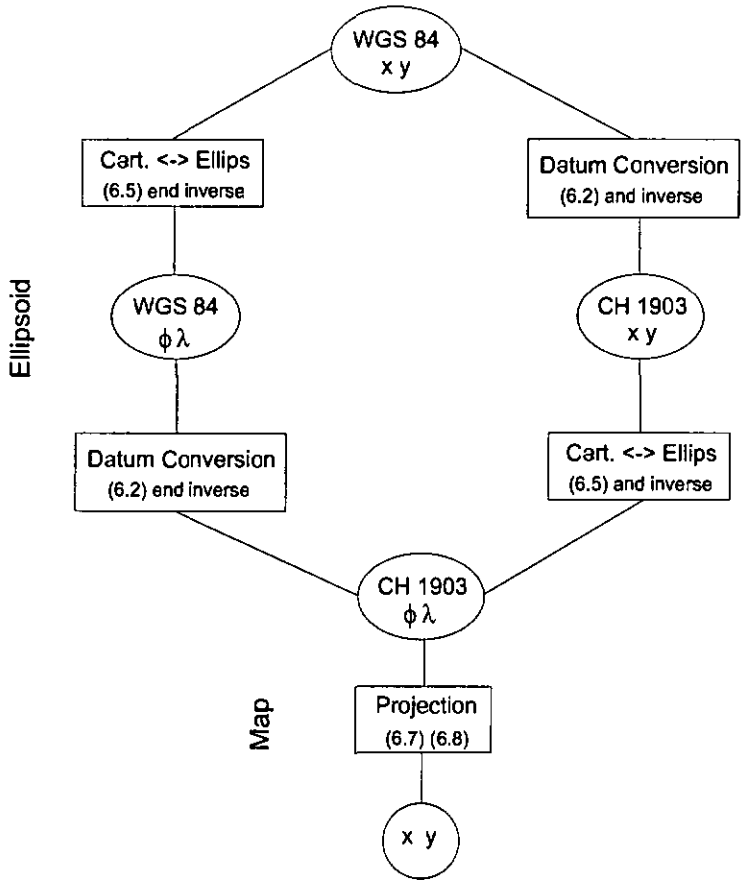

Figure 6.5: Transformation from WGS-84 into CH-1903 and vice versa.

6.3. Summary

The environment of a navigation application, except the inertial sensors, has been presented. The GPS is a satellite-based navigation system, which provides absolutes fixes of the position at 1 Hz and with a basic accuracy of 100 m (2σ rms). In Switzerland, a DGPS service allows any subscribed receiver to improve the accuracy up to 2 m. A GPS receiver can be

interfaced with a PC through the serial port using the NMEA protocol.

GPS fixes are given in the WGS-84 co-ordinate system, whereas Swiss maps are based on the CH-1903 system. Conversion of co-ordinates includes transformation between different ellipsoids describing the earth, projection of the sphere into plane co-ordinates as well as transformation between ellipsoidal co-ordinates (latitude, longitude) and cartesian co-ordinates.

6.4. References

[Hofm94] B. Hofmann-Wellenhof, "Global Positioning System: Theory and Practice", Springer-Verlag, New York, 1994.

[Ruet95] A. Rüetschi, "Etude d'Algorithmes de Correction d'Erreurs d'un Signal GPS", student semester project, ass. C. Marselli, University of Neuchâtel, Institute of Microtechnology, Electronics and Signal Processing Laboratory, Neuchâtel (CH), Winter 1995-1996.

[Stra97] G. Strang, "Linear Algebra, Geodesy, and GPS", Wellesley Cambridge Press, Wellesley (MA), 1997.

[Topo1] Federal Office of Topography, "Schweizerisches Projektionssytem: Formeln für die Umrechnung von Landeskoordinaten in Geographische Koordinaten und umgekehrt", Bern (CH).

[Topo2] Federal Office of Topography, "Transformation von Landeskoordinaten CH-1903 in WGS-84 Koordinaten", Bern (CH), 1990.

Chapter 7
Positioning tests

This chapter reports on positioning test results. Measurements show how microsensor non-idealities influence the trajectory accuracy and how the navigation filter developed in Chapter 4 is able to correct the position errors.

Two series of experiments have been carried out. Firstly, a scale train is used. Position and orientation references are provided by means of optical sensors. Secondly, car navigation tests are realised. A GPS receiver provides position references. The detailed functioning of the car navigation system is given along with a result analysis.

As in the rest of the study, only the effects of bias and sensitivity errors as well as of the number of sensors are considered.

7.1. Scale train navigation

7.1.1. Experimental set-up

First positioning tests were done using a scale train. Thus, the reference trajectory is repetitive and perfectly known. Besides, tests are performed in-door with relatively light means.

The train is running on an eight-shaped track (to avoid wires become tangled!). Due to space constraints, only two

sensors are fixed on the train (Figure 7.1). According to the reduced configuration for a 2D motion (see Chapter 4), an accelerometer placed along the train's longitudinal axis and a gyro measuring the yaw suffice to compute the trajectory.

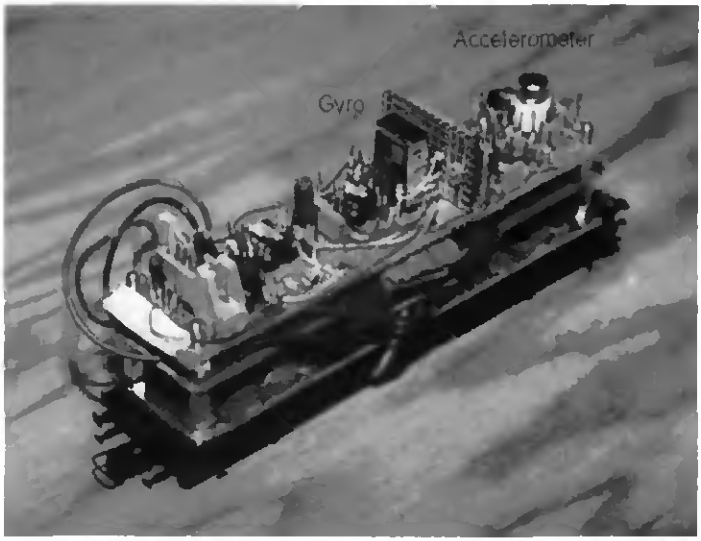

Figure 7.1: Scale train carrying the microsensors.

Absolute references are provided by optical means. Eleven small mirrors are placed regularly along the tracks at known positions. An IR emitter-receiver fixed under the wagon allows detecting when the train has reached a reference position.

Inertial and optical sensor data are recorded using two multiplexed or three individual channels of the DSP board. Data are processed either in real-time or off-line with *Matlab*. In both cases, the display is done with *Matlab*.

7.1.2. Raw trajectory

Figure 7.2 shows typical sensor readouts. The test lasts 60 s. The train is first accelerated and then kept at a constant speed. At the beginning, the train stands at the intersection of the two loops. Then, it runs clock-wise the right-side loop.

The acceleration graph shows clearly the starting acceleration of 0.17 g. The rest of the signal is perturbations, including sensor noise and shocks sensed due to imperfect rail joining.

The angular rate graph exhibits repetitive shapes corresponding to the loop runs. Because of the eight-shaped design, the train runs clock-wise and thereafter counter-clock-wise, changing the sign of the angular rate. The two loops of the circuit are not equal. The first half is a perfect circle, the second half is not. Thus, the angular velocity exhibits either a constant level or a non-regular shape, depending on the loop of concern.

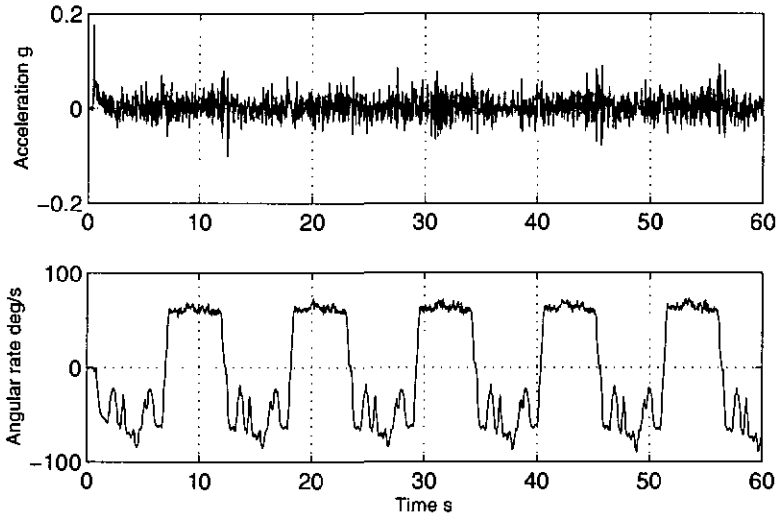

Figure 7.2: Train acceleration and angular rate.

Figure 7.3 shows the trajectory computed with the raw sensor data. The bold curve in the upper right corner is the ideal trajectory whereas the normal curve is the measured trajectory. The computed trajectory and the reference trajectory are equal at the start of the experiment but the computed trajectory exhibits rapidly important drifts both in orientation and loop size.

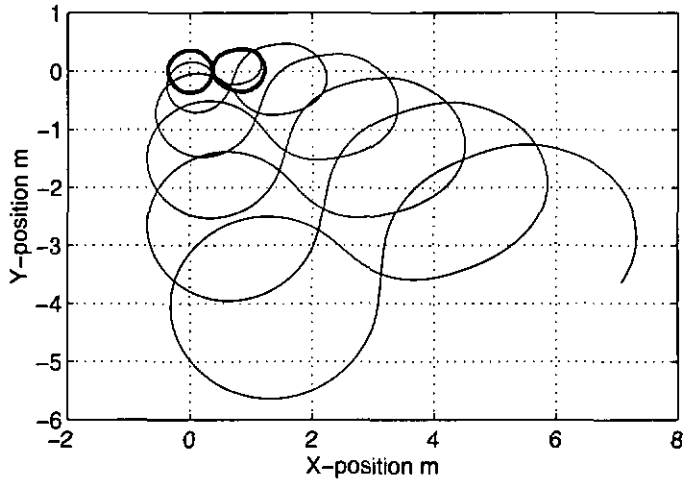

Figure 7.3: Raw computed trajectory.

The gradual extension of the loop size is due to the accelerometer error. It is shown in Figure 7.2 that, after the peak due to initial acceleration, the accelerometer signal exhibits a non-zero signal. This noise is biased at approximately $e_a = 0.03\ m/s^2$. Like real acceleration, it is twice integrated and the resulting distance error Δd is:

$$\Delta d = \frac{1}{2} \cdot e_a \cdot t^2 \tag{7.1}$$

Looking at the left side of the trajectory in Figure 7.3, one sees that the circle radius of a given loop is approximately 0.3 m longer than the radius of the previous loop. This error corresponds well to (7.1), when putting $t = 4.6$ s, which is the duration of the left loop.

To isolate the accelerometer bias effect, the acceleration is now forced to zero once the constant velocity has been reached. The resulting trajectory is shown in Figure 7.4.

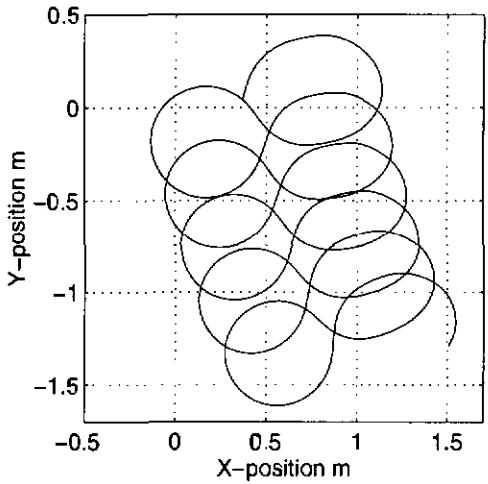

Figure 7.4: Trajectory with acceleration artificially forced to zero.

As expected, the loops do not widen anymore. However, the whole trajectory exhibits a drift due to the fact that no loop is closed before the direction of the angular velocity changes.

Tests have shown that changing the accelerometer gain does not eliminate this error. This phenomenon is due to the gyro scale factor, which is too weak compared to the required sensitivity. Here, the gyro gain is set to 32 mV/deg/s, according to the gain specified by the manufacturer and the gain of the front-end electronics (see Chapter 2). Reducing this value to 29.8 mV/deg/s decreases the orientation shift because the inflexion points of the curve at the beginning and at the end of a loop become closer than in the previous case (Figure 7.5).

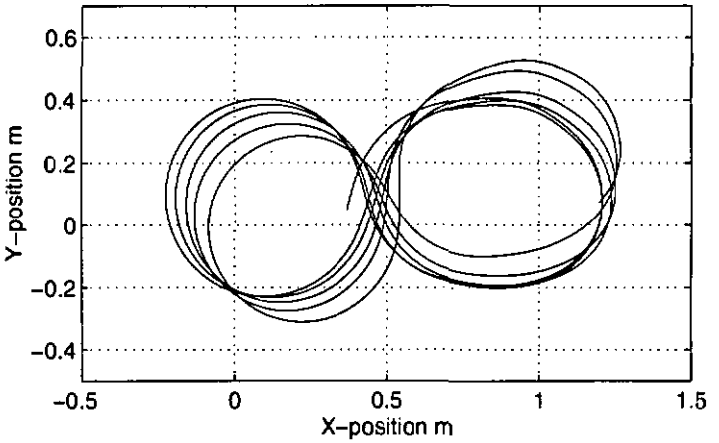

Figure 7.5: Trajectory after the change of the gyro scale factor.

The effect of the gyro scale factor error Δk on the angle θ, can be expressed as:

$$\Delta\theta = \frac{\Delta k}{k} \cdot \omega \cdot t \tag{7.2}$$

Considering the right-side loop, t is about 6 s, ω is about 60 deg/s, thus $\Delta\theta$ is 26.5 deg. This angle corresponds well to the deviation measured on Figure 7.4 between the two starting and ending inflexion points of the right-side loop.

It is worth noting that the scale factor error of the accelerometer is of less importance in this experiment, since it induces errors only when the train is accelerating. Similarly, the gyro bias effect is reduced compared to the gyro sensitivity error since the train undergoes always some rotation.

7.1.3. Use of aiding sources

The above results have shown that the computed trajectory is highly sensitive to sensor non-idealities. As expected, small errors in bias and scale factor lead rapidly to unacceptable results and an absolute reference must be considered to limit the error growth. Figure 7.6 shows the trajectory corrected by means of the optical references. The used Kalman filter

algorithm is described in Chapter 4. The observation vector contains the x/y-position, and the yaw updates. Since the aiding sources are known to be exact, the measurement covariance matrix R is set to zero. The state vector contains the x/y position, the tangential velocity and the yaw. The tuning of the state noise covariance matrix is essential for trajectory accuracy. As shown below, the corrected trajectory is relatively smooth. Due to the state space description, corrections are extended even to state variables that are not measured, such as velocity. However, a part of the right side of the trajectory exhibits a worse accuracy, due to the variation of the gyro scale factor, which is not modelled in the system description and thus not corrected by the filter.

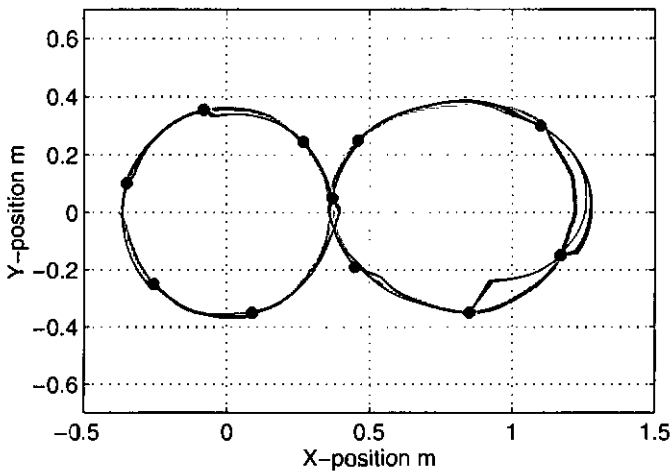

Figure 7.6: Integration of optical references. Bullets represent the position of the mirrors. The ideal trajectory is in grey while the trajectory computed with the Kalman filter is in black.

7.2. Car navigation

The set-up of the car navigation experiment is relatively complex. Especially, the programming is intricate because of the multi-sensor and real-time constraints.

7.2.1. Hardware set-up

Computer

The PC is the host for the DSP board used to sample sensor outputs and to process real-time navigation (see Appendix C). It also makes possible receiving GPS data in the NMEA format through its serial ports. It allows user interaction and trajectory display. For the purpose of navigation tests, it must be easy to transport. *BSI* (*Broadax Systems Inc.*) provides *transportable computers* that are compact, relatively small (41.4 x 24.9 x 28.7 cm^3) and light (13 kg) whereas offering the full power of a desktop model and the ability to use regular desktop add-on cards in up to 8 expansion slots (Figure 7.7).

Microsensors

A 3D inertial platform has been built. Three boards containing each one accelerometer, one angular rate sensor and their front-end electronics (see Appendix A) are orthogonally mounted on three sides of a cubic structure. The construction is rigidly attached in the car boot (Figure 7.7) and connected through shielded wires to the DSP board.

DGPS-GPS

Many low-cost GPS receivers are nowadays available. The one used in this work is *Garmin GPS 12XL Personal Navigator*. It is a small handheld GPS receiver (5.3 x 14.7 x 6.5 cm^3 / 269 g), which can track up to twelve satellites and supports the NMEA

and RTCM-104 interface formats. It has its own display of the navigation information and provides many functions to make the user guidance easy. It can be supplied with batteries or any DC voltage from 10 to 40 V.

For enhancing the reception of GPS signals, an external antenna placed on the car body is used instead of the built-in receiver antenna. An additional FM antenna is required to receive the DGPS correction sent in the RDS format (see Chapter 6). Tests have shown that both antennas must be kept away as far as possible from each other to avoid disturbances on the reception.

In the car, the differential GPS receiver *AZTEC RDS3000v3* decodes the RDS signal and supplies the Garmin GPS receiver with differential correction data according to the RTCM SC 104 Version 2 standard. The Garmin receiver is connected to one of the PC serial ports. The antenna location with respect to the sensor location is of no importance since the car is considered as a solid. The antenna and the sensors are assumed having the same motion. The initial position of the car is given by the average of acquired DGPS positions before the drive, the sensors measure then the relative motion.

Voltage supply

Power is pulled from the vehicle battery. The 12 V is brought to the GPS-DGPS assembly, to the active FM antenna as well as to the sensor block. A DC to AC converter allows supplying the PC with alternative 230 V. The additional –12 V required by the sensor front-end electronic is pulled from the PC.

Figure 7.7: Car set-up. Top left: sensors mounted in the car boot – Right: test car (probably the only Fiat Uno with a navigation system). Bottom left: transportable computer – Right: GPS and FM antenna.

7.2.2. Programming

Real-time implementation

The complete real-time processing includes reading the inertial sensor data, computing the trajectory, updating it whenever a GPS position is available and displaying the path. It involves four programs running simultaneously. The data flow is sketched in Figure 7.8. The program tasks can be summarised as follows:

- reading GPS data: using an interrupt which indicates when data is available at the serial port, this program fetches all characters sent and checks them in order to detect the beginning and the end of the NMEA sentence of interest. Once the complete NMEA sentence is recovered, the program picks up the latitude and longitude, according to the sentence structure. Since this position is given in WGS-84 ellipsoidal co-ordinates, it has to be transformed into CH-1903 plane co-ordinates (see Chapter 6). Finally, the X-Y positions are transferred to the DPRAM with a flag set to indicate that a GPS update occurred. For off-line processing purposes, the time of GPS events is picked up and recorded along with the GPS position in files,
- display program: to get a display of the trajectory according to the map of the place, an off-the-shelf software *Fugawi* is used. It allows displaying in real-time a GPS position on a digital map, provided NMEA sentences are available. The digital map can also be obtained by scanning a suitable paper map. Provided co-ordinates of three points of the site are known, the software calibrates the rest of the map. When a GPS NMEA sentence is available at the serial port, the co-ordinates are displayed on the selected map. Since the automotive navigation tests were realised downtown Neuchâtel, a map of scale 1:5'000 was bought from the cadastre and digitised. However, since co-ordinates are not indicated on such maps, an additional map (scale: 1:25'000)

was used to determine the co-ordinates on three points easy to locate that were reported to the 1:5'000 map,

- DSP program: using interrupts, data sampled from the 6 input channels are acquired according to the chosen sampling frequency. They are converted into physical units and processed using the navigation algorithm (see Chapter 4). When a GPS update is notified on the DPRAM, it is read and processed through the Kalman filter. The results of the navigation process given in CH-1903 co-ordinates as well as the sensor data are put on the DPRAM to be available to the PC. A flag is used to indicate when the DPRAM is full and has to be emptied by the PC,
- PC program: this program transfers some parameters (e.g., sampling frequency, acquisition length, calibration value) to the DSP before processing. Then, it fetches the results given by the DSP when the DPRAM is full and releases the flag. These results are recorded in files and are displayed by the *Fugawi* software, thanks to the following tricky transformation. First, the Swiss plane INS-GPS positions given by the DSP are transformed into WGS-84 ellipsoidal co-ordinates and encoded according to the NMEA standard as if they were real GPS data. Time and date are picked up from PC clock and the checksum is accordingly computed. The resulting sentence is sent to a serial port connected through a null-modem cable to the serial port that the *Fugawi* software is reading.

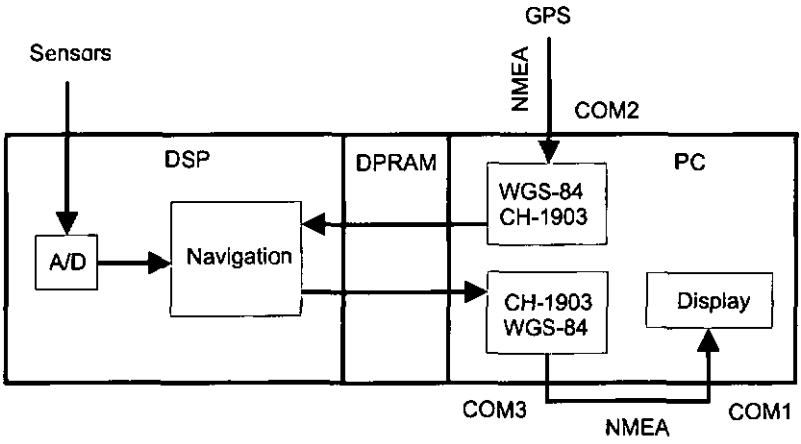

Figure 7.8: Data flow.

Off-line processing

For off-line processing purposes, the DSP program and PC program are simplified to do just the sensor data acquisition and recording whereas the GPS reading program does not communicate with the DSP anymore. Once the acquisition is completed, the off-line processing is done with *Matlab* programs. Since the time of GPS events as well as the starting time and duration of the sensor acquisition have been recorded along with the data, an a posteriori synchronisation of both signals is realised before they are processed with the Kalman navigation filter. The display is done either using *Matlab* or the *Fugawi* software.

7.2.3. Change of body co-ordinate frame

The body frame used up to now had to be modified to suit geographical maps. In Chapter 4, the x-axis and y-axis of the measurement frame were defined towards the front, respectively left side of the moving body. According to a right-handed co-ordinate frame, the z-axis was pointing upwards. However, the car trajectory is expressed using Swiss co-

ordinates, which are defined with an x-axis and y-axis pointing towards north and east, respectively. This corresponds implicitly to a downwards pointing z-axis, according to usual sign rules. Thus, the frame transformation matrix developed in Chapter 4 will no longer be valid.

Instead of computing a new rotation matrix, it has been decided to change the body frame by putting the x-axis and y-axis pointing towards the front, respectively the right side of the car, and accordingly the z-axis downwards. On the contrary of the scale train navigation, a clock-wise motion will now give rise to both positive yaw rate and normal acceleration whereas a counter clock-wise motion will correspond to both negative yaw rate and normal acceleration.

7.3. Typical car signals

One problem that arose during navigation tests is that the pitch/roll rates and the vertical acceleration of the car are too small with regard to parasitic signals due to bumps in the road and to the car body deformations. Hence, pitch, roll and z-acceleration signals cannot be used since they are far too noisy. Accordingly, only longitudinal, normal accelerations and yaw rate will be considered in the rest of the chapter.

In the following experiment, microsensor measurements were recorded over one minute while the car was driven counter clock-wise around a block of houses. Figure 7.9 shows the corresponding unsigned x-y accelerations and yaw rate. The yaw rate signal, like the normal acceleration, exhibits three peaks, which correspond to the three turns of the trajectory. The longitudinal acceleration signal shows several accelerations and decelerations that occurred during the drive. For example, it is clearly shown that the car reduces speed before turning and accelerates in the middle of a turn. Besides, at time 23 s, there is an additional braking, which corresponds to a crossroad. Finally, during the last third of the drive, the car is driving more slowly than in the beginning due to increasing traffic.

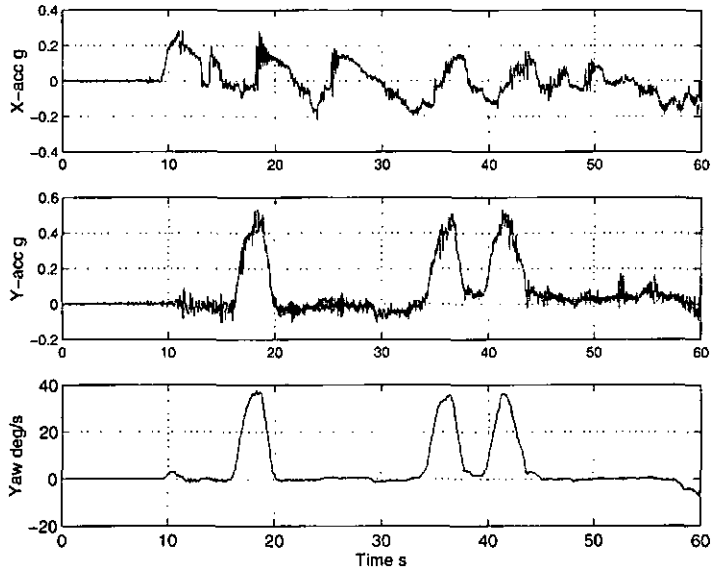

Figure 7.9: Typical unsigned car signals. Top: longitudinal acceleration — Middle: normal acceleration — Bottom: yaw rate. Actually, both yaw rate and normal acceleration are negative.

7.4. Raw trajectory

The resulting trajectory corresponding to the signals of Figure 7.9 is shown on Figure 7.10. The grey trajectory is obtained using the two x/y-accelerometer readouts whereas the black trajectory is computed using only the longitudinal accelerometer, according to the modified algorithm of Chapter 4. When the normal acceleration is taken into account, the trajectory exhibits rapidly important errors. The accuracy is far better when using the reduced set of sensors, although the loop should ideally be closed, according to the true trajectory.

Worse results are due to the noise of the y-accelerometer signal. Even when the car is not turning, the lateral accelerometer output exhibits a noise, which induces the trajectory divergence. By way of proof, a trajectory similar to the one computed with the modified algorithm is obtained by

forcing the lateral acceleration to zero except during the three corners.

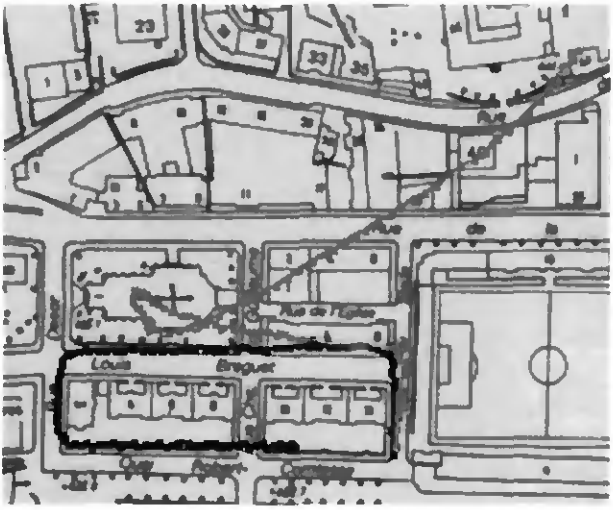

Figure 7.10: Raw car trajectories.

More generally, when examining trajectories computed from raw measurements, it is observed that the gyro scale factor error and the accelerometer bias have the same effect on the position error in car experiments than in the scale train experiments. Besides, in car tests, the accelerometer sensitivity error is of importance, especially during decelerations. If the scale factor is not accurate, the car braking is understood as a car backing and the computed direction becomes the exact opposite of the true one. Hence, this condition is checked, forcing to zero negative tangential velocities to avoid inappropriate change of direction.

7.5. Integration of GPS fixes

Figure 7.11 shows the trajectory computed using the navigation filter to blend INS and DGPS data. Crosses indicate DGPS fixes. The DGPS accuracy is lower than the short-term accuracy of the inertial system. However, provided that no

signal jams occur, the GPS accuracy does not decrease with time. As expected, due to the Kalman filter, the new trajectory is a trade-off between both systems: the shape is closed and the accuracy becomes average (2-10 m).

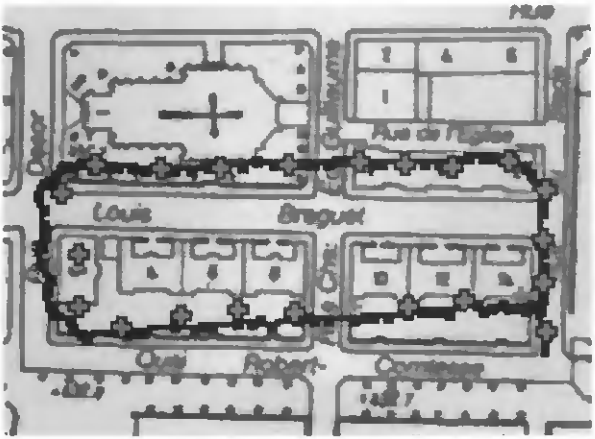

Figure 7.11: Car trajectory obtained by integration of DGPS data.

Figure 7.12 shows GPS fixes recorded over a drive lasting 10 min. During most of the drive, the receiver gave regularly (every two seconds) positions with 5-10 m accuracy. Places, where the car was stopped (crossroads or traffic lights), are indicated by the superposition of GPS fixes. On the lower left side of the graph, a GPS jam can be observed. Due to a loss of satellites, the GPS receiver continues to provide an information computed with an internal processing of estimation but the position accuracy becomes worse and not reliable anymore.

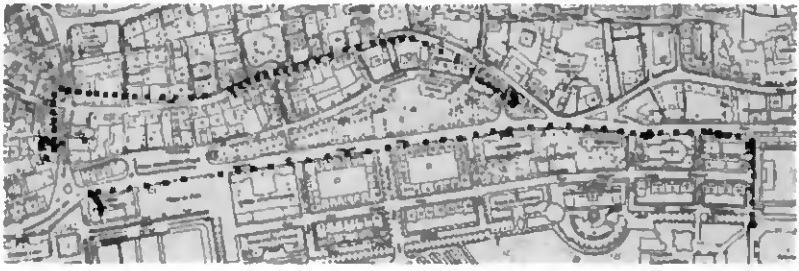

Figure 7.12: GPS fixes.

Figure 7.13 shows the trajectory obtained with the GPS-INS coupling. The change of orientation with respect to Figure 7.12 must not be taken in consideration. Figure 7.13 is oriented according to a vertical North-South axis whereas the map in Figure 7.12 has a different orientation for geographical reasons.

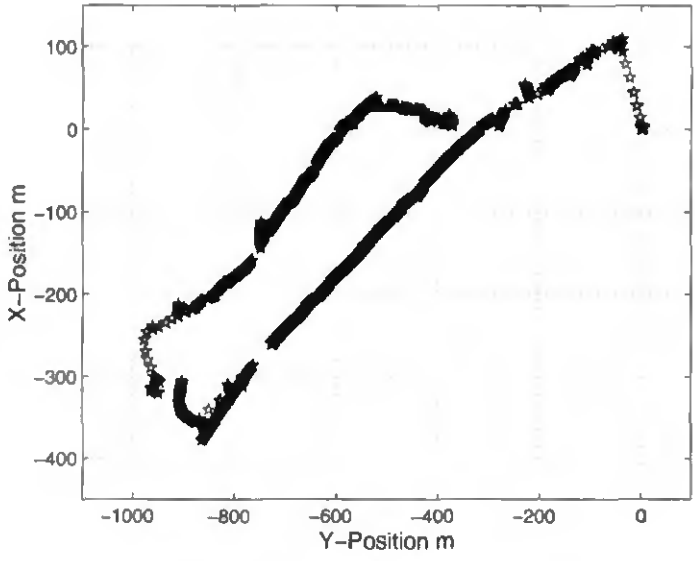

Figure 7.13: Car trajectory obtained by integration of DGPS data.

Figure 7.13 shows the advantages of blending the two systems, due to their complementary characteristics. On one

hand, the regularity of GPS fixes prevents the divergence of the trajectory due to integration of sensor noise, which would have been important since stops were long. On the other hand, the INS provides a 32 Hz navigation information, which allows filling intervals between GPS fixes. It is worth anyway representing the INS-GPS fixes with separate points rather than with a linking curve, which may create the impression that the trajectory is better than it really is. Doing so, little gaps can be observed when the Kalman filter updates abruptly the trajectory. When the GPS quality is good, the measurement covariance matrix is low and the Kalman filter relies on the GPS fixes correcting strongly the inertial information when it is too different from the GPS position.

Let us now examine how the system reacts during the GPS jam. In this part of the drive, due to the bad quality of GPS fixes, the Kalman filter relies mainly on the inertial information. Thus, on the contrary of Figure 7.12, the INS-GPS trajectory is continuous at the corner level, thanks to the gyro information. Nevertheless, an orientation mismatch remains due to the accumulated error in the orientation of the car. The GPS does not provide orientation information and the car heading computed with the gyro readouts becomes wrong with time.

To make the system more robust and accurate, an absolute orientation information should be provided. Means of having orientation fixes with the GPS involves generally the use of two antennas. [Most97] proposes also a method to deduce the orientation of the car, using the last displacement vector. Another improvement of the system would be to make decisions about the treatment of the GPS information by examining the pseudo-ranges rather than the final positions given by the receiver. The *Garmin* receiver computes a position based not only on the received pseudo-ranges but also on the estimation involving additional information (e.g., previous velocity of the car). Thus, it is difficult to estimate whether the GPS position should be considered.

7.5.1. Summary

This chapter gives experimental results of the navigation experiments. The trajectories of both a scale train and a car are computed using inertial microsensor measurements. The effects of the sensor bias and the scale factor error are outlined, especially in the scale train experiments. As expected, integration of sensor errors leads to fictitious displacements. The simple inertial systems, which have been designed, allowed characterising separately the influence of the sensor error. However, the influence of the sensor errors on the trajectory accuracy depends also on the inertial system design. Thus, the reader who is interested in a different configuration should refer to [Brit71], who provides a unified error analysis for terrestrial inertial systems.

Main car results are twofold. First, the reduced sensor configuration (no lateral accelerometer) motivated in Chapter 4 is proved to give best results. Finally, the benefits of coupling the INS with the GPS are presented.

7.5.2. References

[Brit71] K. Britting, "Inertial Navigation Systems Analysis", Wiley-Interscience, John Wiley & Sons, Inc., New York, 1971.

[Most97] K. Mostov, "Initial Attitude Determination and Correction of Gyro-Free INS Angular Orientation on the Basis of GPS Linear Navigation Parameters", Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, 1997.

Chapter 8
Conclusion

The main motivation of this research project was to determine some of the limits of microtechnologies and microsystems through the design of an application example: a navigation system based on inertial microsensors.

Nowadays, because of technology developments, the sensor has to be considered as one component of a system rather than as the key item of the application. In fact, there has been a transfer from the device level to the system level, where the relatively low accuracy of current microsensors can potentially be improved through suitable electronics and signal processing.

Thus, the question of the limits of the microsensors has to be answered considering not only the limits of the sensor itself but also the limits due to the configuration of the system as well as those of the signal processing.

Following the above considerations, the main results of this research are summarised below along with additional comments including possible improvements and further work.

Limits of the sensor

Ideally, the sensor output should be a direct representation of the signal to be measured. As a matter of fact, many other factors influence the sensor behaviour, which may nevertheless be described through a model. Thus, the meaning of *sensor accuracy* is twofold. On one hand, the sensor may perform well

since its design is mature. On the other hand, the sensor accuracy may be poor, but the non-idealities can be modelled and taken into account to improve the measurement. In the latter case, the characterisation of the sensor requires to conduct a series of complex tests using appropriate equipment. Once experiments are carried out, models, either deterministic or random, are identified to describe the phenomena of interest. In any case, one has to speak about the accuracy of a given type of sensors rather than the limit of sensors in general.

In this research, off-the-shelf sensors have been purchased. The selection criterion was that they should suit automotive navigation signals ($a = 1$ g, $\omega = 50$ deg/s) while being reasonably priced for consumer applications. Thus, after a market study (as of 1995), a microaccelerometer and an angular rate microsensor with medium accuracy and relatively low price were chosen. Resolution, sensitivity and dead band of these sensors have been measured. Their accuracy does not fulfil classical navigation specifications. The gyro is less noisy than the accelerometer but its resolution should still be improved four times in order to reach the navigation requirements (0.1 deg/s instead of 0.4 deg/s). Besides, both exhibit a random in-run drift.

Sensor errors can be modelled and reduced through a suitable signal processing (see below). Nevertheless, if the accuracy of the sensors can be increased through a suitable modelling, the choice of the sensors should be made adequately: a good balance between the sensor accuracy and the model sophistication should be found.

Limits of the signal processing

The improvement of the accuracy of any inertial navigation system is basically achieved through two approaches. Firstly, sensor errors have to be reduced. Secondly, regular updates of the trajectory must be provided by an aiding system.

In this study, Kalman filtering forms the heart of the correction data processing. Two Kalman filters have been implemented. The first filter, referred to as *local filter,* was designed to reduce bias and sensitivity error, independently of any aiding source. It has been validated through simulation

examples and through the correction of the in-run drifts of the gyros. Since Kalman filtering is the main theoretical tool of this study, correction has been restricted to random phenomena. Moreover, going further in the behavioural modelling would have required more sophisticated equipment, not available at that time. In order to complete the study, further work could be oriented towards the investigation of other phenomena, which may reasonably be modelled through deterministic models (e.g., temperature dependency).

The second filter, referred to as *navigation filter*, is designed to blend inertial information with aiding information. Scale train and car positioning tests show how it is able to prevent the trajectory from diverging. In the particular case of INS-GPS integration, the INS allows filling intervals between GPS jams. Experiments have shown that, provided the INS is correctly calibrated and the starting position is known accurately, the INS is usable during approximately 40 s.

In these experiments, emphasis is put on the inertial system rather than on the aiding system. Further work could focus on the aiding system, either to improve its accuracy or to modify the types of updating measurements (e.g., orientation update in the car experiments).

Note however that the limits of a signal processing based on Kalman filter are not due to the filter, which is a well defined algorithm, but merely to the quality of the models of the system.

Finally, let us mention that inertial systems may have constraints in size and power consumption. Thus, the DSP implementation realised in this research should be further optimised to suit a physical implementation. In this case, the optimisation at the algorithm level may be an additional limit of the signal processing.

Limits due to the system configuration

The sensor arrangement is obviously application-dependant and inertial navigation systems belong probably to the most complex ones. Firstly, inertial navigation requires two different kinds of sensors. Thus, the limit of the resulting system depends on the combination of the limits of both types of

sensors. Secondly, the sensor readouts are integrated and the accuracy varies with time and position. Thus, the study of the limits of the navigation system with respect to the limits of individual sensor is clearly a difficult task.

In this study, navigation algorithms in the general 6-sensors case were presented. However, practical experiments focused on simpler configurations using only two accelerometers and one gyro. In this way, the effects of sensor bias and sensitivity errors could be separated in the final result. Nevertheless, further work could include the design of new test set-ups enabling the analysis of a 6-sensors positioning system.

Experiments have also shown that the set of sensors used must be carefully chosen. As an example, an alternative system, where the lateral accelerometer was removed and the navigation equations accordingly modified, has been claimed to give better results in land navigation applications.

From sensors to system or from system to sensors?

Current inertial microsensors are not accurate enough to provide a robust solution for navigation. Effort has still to be made to reduce the noise level as well as other errors. To this aim, new designs have to be investigated and techniques of modelisation have to be further developed. However, improvements should be done regarding the targeted applications. It makes probably little sense to design a sensor for general purpose. A design should be optimised regarding the future use of the sensor. As an extreme example, one should ask whether linear sensors are required for an automotive navigation system. As a matter of fact, small rotation changes, corresponding for example to a change of lane, are not of interest whereas bigger angular rates corresponding to a turn are important to compute the trajectory. Thus, it could be more useful to have sensors with a response, which could be logarithmic rather than linear.

In general, the design of a sensor dedicated to a navigation system could also be addressed the other way around as follows. First, the set of possible trajectories should be defined along with the required accuracy. Next, simulations should be

carried on to determine the sensor behaviour required. Starting from a precise definition of the applications needs, the sensor features can be better determined.

What is true for the sensor design is also true for the system design. The signal processing and the sensor mounting have to be optimised regarding the application. In automotive navigation, a fusion algorithm has for sure to be used. However, it should be completed with intelligent processing, which involves models of car behaviour.

Final remarks

The design of any sensor-based system is more than ever an interdisciplinary problem. The aim of the PICS project, which included the present work, was exactly to emphasise this point.

Unfortunately, the schedules for developments in electronics and signal processing and microstructures are quiet different. Therefore, the system aspects had to be developed using existing, commercial sensors.

Nevertheless, due to the contacts established, both microtechnology specialists and electronics and signal processing engineers gained a good insight in each other's problems: from the beginning, the partners have to collaborate in order to finally get a complete, optimised system.

Hopefully, the upcoming modelling and simulation tools will help to fill the gap between the two domains.

Finally, it can be emphasised that, thanks to this international collaboration, the PICS thesis realised in the University of Neuchâtel are validated by the University of Franche-Comté and vice-versa.

Acknowledgements

Being doctor is commonly thought as advantageous. Hence, I would like to sincerely thank the many people who have been associated to this period of my life.

First, I am very grateful to my supervisor Professor Fausto Pellandini, who put his trust in me, giving me the green light for joining his lab and realising this PhD work. He gave me freedom and infrastructure to develop my research work and the opportunity to learn a lot in many various fields. During all our discussions, I really appreciated his clear reasoning and his sensible way to approach problems. The first work is something very important for many reasons and I am glad to have done it at the IMT.

I also would like to thank Dr. Hans-Peter Amann, who was one of the leaders of the PICS project. He was very helpful in attending all meetings, proof-reading the papers as well as the PhD report. He also gave me an invaluable hand for the practical part of the project, for which I had a lot to learn (agence "La Bricole", as we were used to say in french). Thanks to his efforts, the co-diploma procedure between the University of Neuchâtel and the University of Franche-Comté (F) was achieved and this PhD work was also validated by the French University.

Special thanks to Professor Nico de Rooij and Professor Daniel Hauden for their interest in my work and for co-examining my thesis. Two thesis of the PICS project were realised in their lab by Gerold Schröpfer (LPMO) and Florence Grétillat (IMT). I really appreciated working with them. I also

will not forget the friendship token that Florence and her husband Marc gave to me.

I would like to express my gratitude to the students who worked with me. I enjoyed all the moments we shared together. The work done by Andrea Rüetschi was particularly useful.

This research has been funded by the Swiss Foundation for Research in Microtechnology (FSRM). I specially thank his director, Dr Marcel Ecabert, who initiated the project and co-ordinated some of the public presentations.

More generally, I would like to thank all the people who kindly accepted to share their knowledge with me. Among them, I would like to mention the Geodetic Laboratory – EPFL, headed by Professor Bertrand Merminod, as well as the team of the Federal Office of Topography, who both gave me an invaluable help regarding geodesic principles. Besides, I am really grateful to the people, who invited us to present our work during seminars, lectures and conferences.

The working environment at the IMT was very pleasant, partly due to the good ambience between assistants and more generally with all the technical and administrative staff. All were available for professional discussions and help. We spent also private time together, even during holidays! For that, special thanks to all my IMT's colleagues, among them: Christophe Berthoud, Heinz Burri, Giuseppina Biundo, Alain Dufaux, Louisa Grisoni, Olivier Guenat, Catherine Lehnherr, Vincent Pasquier, Sylviane Pochon, Ruud Riem-Vis, Christian Schütz, Dequn Sun, Steve Tanner, François Tièche and especially:

Dominique Daudet (from Valais): since we shared the same office for three years, we had a lot technical discussions, in which he provided much help in discussing different options and avoiding deadlocks. His friendship is also very precious to me.

Vincent Moser: he devoted much time and attention to proof-read the PhD report, in spite of not being involved at all in the PICS project. He gave me great suggestions and a lot of encouragement during the writing.

Sara Grassi: as one of my best friends, she brought me a lot of support in various fields. I really appreciate her generous temper and the time we spend together.

Alexis Boegli: I thank him for all our discussions and the great aquatic souvenirs together, such as the famous participation of the Neuchâtel triathlon.

Christian Robert and Patrick Stadelmann: they shared with me their high computer skills and a lot of jokes. I enjoy a lot their company.

However, the world is obviously not limited to the IMT and I would like to give hearty thanks to all my family and friends:

Merci tout particulièrement à Alain qui m'a tant apporté depuis trois ans.

Merci à ma grand-mère, à ma géniale petite soeur Béatrice et à Christophe. Merci surtout à mes parents non seulement pour ce qu'ils ont fait pour moi mais surtout pour ce qu'ils sont. J'admire les exemples d'humanité, d'intelligence et de vitalité qu'ils me donnent depuis toujours.

Merci à la Suisse, devenue mon pays de coeur depuis dix ans, pour m'avoir donné la possibilité de faire mes études et de travailler.

Merci et bonne route à vous tous, qui me faites la joie de partager la mienne...

Biography

Catherine Marselli was born on January 9, 1971 in Annemasse, France. She received her electrical engineer degree from the Swiss Federal Institute of Technology – Lausanne (EPFL) in March 1994. After a training job in Alcatel Business System – Strasbourg (F), she joined in October 1994 the Institute of Microtechnology (IMT) at the University of Neuchâtel. In December 1998, she defended successfully her PhD thesis entitled 'Data Processing of a Navigation Microsystem'.

Main publications and presentations
(arranged by type and date)

Conference papers

C. Marselli, H. P. Amann, F. Pellandini, "Automotive Navigation based on Inertial Microsensors", accepted for ISATA'99, International Symposium on Automotive Technology and Automation, Vienna, Austria, June 14-18, 1999.

B. de Heij, C. Marselli, J. Hess, "A Smart Inhaler", accepted for ISAM, 12th Int. Congress of the International Society for Aerosols in Medicine, Vienna, Austria, June 12-16, 1999.

C. Marselli, D. Daudet, H. P. Amann, F. Pellandini, "Application of Kalman filtering to Noise Reduction in Microsensor Signals", C2I'98, Colloque Interdisciplinaire en

Instrumentation, Ecole Normale Supérieure de Cachan, Paris, France, November 18-19, 1998, p. 443-450.

C. Marselli, H. P. Amann, F. Pellandini, F. Paoletti, M. A. Grétillat, N. F. de Rooij, "Error Modelling of a Silicon Angular Rate Sensor", Symposium Gyro Technology 1997, Stuttgart, Germany, September 16-17, 1997, p. 4.0-4.9.

Y. Ansel, Ph. Lerch, Ph. Renaud, F. Paoletti, M. A. Grétillat, N. F. de Rooij, G. Schröpfer, S. Ballandras, M. de Labachelerie, C. Marselli, H. P. Amann, F. Pellandini, "Simulation and Design of a Three-axis Navigation Microsystem based on Micromachined Sensors", Microsim'97, Lausanne, Switzerland, September 17-19, 1997, p. 107-116.

C. Marselli, H. P. Amann, F. Pellandini, "Error Correction Applied to Microsensors in a Navigation Application", 1st Europe-Asia Congress on Mechatronics, Besançon, France, October 1-3, 1996, p. 650-655.

Seminars - Lectures

Journées de la microelectronique de l'Association du Centre Microswiss de Suisse Occidentale (ACMSO), Neuchâtel, Switzerland, November 19-20, 1998.

Séminaire Microswiss, "GPS, Système de Positionnement par Satellites", Fribourg, Switzerland, May 27, 1998, see newspaper 'La Liberté', 3rd June 1998.

Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA, November 1997.

Others

C. Marselli, "Traitement de l'Information d'un Microsystème de Navigation", Revue "En Direct", Université de Franche-Comté, Vol. 127, p. 1-2, Besançon, France, March 1999.

E. Heinzelmann, C. Marselli, "Mini-Sensoren Steuern Autos", german version in "Horizonte" (information journal of the Swiss National Fund of Science), Vol. 38, September 1998, p. 24-25, (french version: "Voiture: le pilote est une puce", in "Horizons", Vol. 38, September 1998, p. 24-25).

Appendix A
Design of the front-end electronics

A.1. Purpose and description

Analog front-end electronics has been designed to enable measuring the sensor output regarding the application needs and the devices used. The goals of the circuit are as follows:

- to provide sensors with a stable power supply. The board is supplied with voltages of ±12 V DC. Additional voltages of ±5 V and +2.5 V required by the sensors are locally generated using voltage regulators,
- to amplify the sensor output. The initial sensitivity of the *Murata* gyro is 0.8 mV/deg/s and its maximum range is 90 deg/s. Since the DSP card allows an ±3 V analog input swing, a gain of 40 has been chosen.
  The sensitivity and the full range of the *Analog Devices* accelerometer can be customised. The sensitivity is tuned in two stages. First, resistors are placed on the sensor according to the specifications to set the sensitivity to 200 mV/g. Then, an additional amplifier is added to increase the sensitivity to 3.88 V/g. Due to the DSP constraints, it allows measuring a maximum acceleration of 0.75 g,
- to provide suitable filtering. The motion frequency, which is of interest, ranges from DC to 10 Hz. Low-pass filtering

limits the bandwidth of the accelerometer and the gyro to 10 Hz, and 4 Hz respectively. Thus, it removes high-frequency noise components.

The circuit design is shown in Figure A.1.

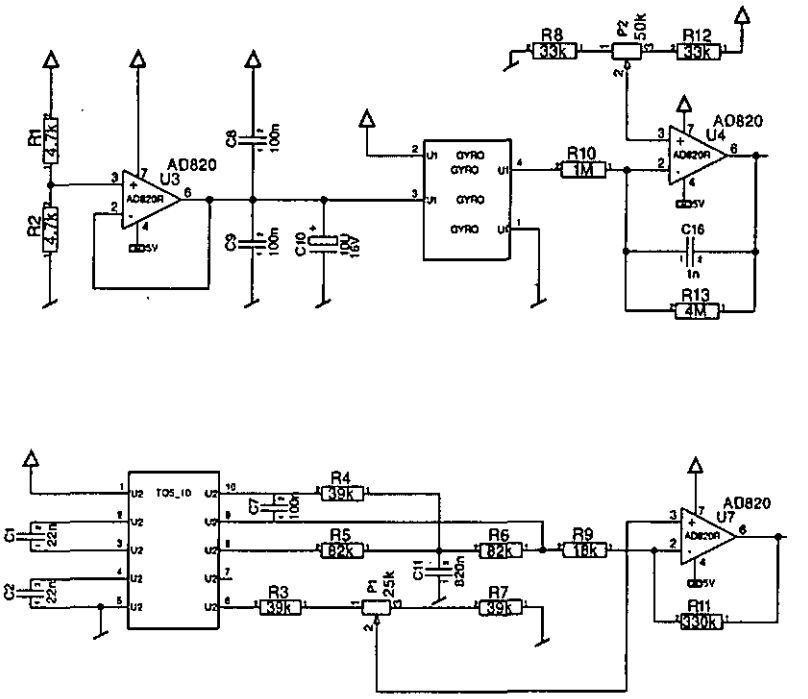

Figure A.1: Electronic design – Top: Murata ENC-05S gyro – Bottom: Analog Devices ADXL05 accelerometer.

The circuit has been wrapped on experimental boards for laboratory tests. PCBs have been realised and mounted on a cubic mechanical structure used for the car navigation tests (Figure A.2).

Figure A.2: Inertial platform for car navigation tests.

A.2. References

[Anal95] Analog Devices, "±1 g to ±5 g Single Chip Accelerometer with Signal Conditioning – ADXL05", Norwood, Massachusetts, 1995.

[Mura93] Murata, "Angular Velocity Sensor Gyrostar ENC Series", Cat. No. S38E, 1993.

Appendix B
Linearised state matrices for the navigation filter

B.1. Configuration with 1 accelerometer and 1 gyro (yaw) – Pure 2D motion

According to Chapter 4, the state vector is chosen as:

$$X = (x_e \;\; y_e \;\; v_b \;\; y)^T \tag{B.1}$$

and the corresponding state matrix is:

I (2 x 2)	$\cos(y)\cdot dt \quad -v\cdot\sin(y)\cdot dt$ $\sin(y)\cdot dt \quad v\cdot\cos(y)\cdot dt$
0 (2 x 2)	I (2 x 2)

(B.2)

B.2. Configuration with 1 accelerometer and 2 gyros (pitch, yaw) – 2D motion with attitude change

According to Chapter 4, the state vector is chosen as:

$$X = (x_e \;\; y_e \;\; z_e \;\; v_{ex} \;\; v_{ye} \;\; v_{ze} \;\; v_b \;\; p \;\; y)^T \tag{B.3}$$

and the corresponding state matrix is:

I (3 x 3)	I (3 x 3).dt	0 (3 x 3)	
0 (3 x 3)	0 (3 x 3)	$\begin{matrix} R_{11} & -v\cdot sp\cdot cy & -v\cdot cp\cdot sy \\ R_{21} & -v\cdot sy\cdot cp & v\cdot cy\cdot cp \\ R_{31} & -v\cdot cp & 0 \end{matrix}$	
0 (3 x 3)	0 (3 x 3)	I (3 x 3)	(B.4)

B.3. Configuration with 1 accelerometer and 2 gyros (pitch, yaw) using quaternions - 2D motion with attitude change

According to Chapter 4, the state vector is chosen as:

$$X = \begin{pmatrix} x_e & y_e & z_e & v_{ex} & v_{ey} & v_{ez} & v & q_0 & q_1 & q_2 & q_3 \end{pmatrix}^T \qquad \text{(B.5)}$$

and the corresponding state matrix is:

I (3 x 3)	I (3 x 3).dt	0 (3 x 3)	0 (3 x 3)
0 (3 x 3)	0 (3 x 3)	$\begin{matrix} R_b^e(1,1) \\ R_b^e(2,1) \\ R_b^e(3,1) \end{matrix}$	$\frac{dv}{dq}$
0 (4 x 4)	0 (4 x 4)	0 (4 x 4)	dq

(B.6)

where:

$$\frac{dv}{dq} = \begin{pmatrix} 2q_0v & 2q_1v & -2q_2v & -2q_3v \\ 2q_3v & 2q_2v & 2q_1v & 2q_0v \\ -q_2v & 2q_3v & -2q_0v & 2q_2v \end{pmatrix} \qquad \text{(B.7)}$$

and:

$$dq = \begin{pmatrix} 1+a & -\dot{r}\cdot b & -\dot{p}\cdot b & -\dot{y}\cdot b \\ b\cdot\dot{r} & 1+a & b\cdot\dot{y} & -b\cdot\dot{p} \\ b\cdot\dot{p} & -b\cdot\dot{y} & 1+a & b\cdot\dot{r} \\ b\cdot\dot{p} & -b\cdot\dot{y} & -b\cdot\dot{r} & 1+a \end{pmatrix} \qquad \text{(B.8)}$$

with:

$$a = \cos(\frac{\theta}{2}) - 1$$

$$b = \frac{1}{\theta} \cdot \sin(\frac{\theta}{2}) \tag{B.9}$$

B.4. Configuration with 3 accelerometers and 3 gyros – 3D motion

According to Chapter 4, the state vector is chosen as:

$$X = (x_e \;\; y_e \;\; z_e \;\; v_{ex} \;\; v_{ey} \;\; v_{ez} \;\; p \;\; r \;\; y) \tag{B.10}$$

Due to space constraints, the whole matrix is not given. Let us just mention that the lower-right (3 x 3) sub-matrix is deduced from the partial derivatives regarding pitch, roll, yaw angles of the equations below (see Chapter 3 and 4):

$$a_x = c(p) \cdot c(y) \cdot a_{bx} + (s(p) \cdot s(r) \cdot c(y) - c(r) \cdot s(y)) \cdot a_{by} + (s(p) \cdot c(r) \cdot c(y) + s(r) \cdot s(y)) \cdot a_{bz} \tag{B.11}$$

$$a_y = s(y) \cdot c(p) \cdot a_{bx} + (s(p) \cdot s(r) \cdot s(y) + c(r) \cdot c(y)) \cdot a_{by} + (s(p) \cdot c(r) \cdot s(y) - s(r) \cdot c(y)) \cdot a_{bz} \tag{B.12}$$

$$a_z = -s(p) \cdot a_{bx} + c(p) \cdot s(r) \cdot a_{by} + c(p) \cdot c(r) \cdot a_{bz} \tag{B.13}$$

Appendix C
DSP Processor

C.1. Description

A navigation application requires real-time computation of the results. Hence, the main part of the algorithms has been implemented on a DSP processor [Laps94]. Moreover, the A/D converters of the DSP board are also used for data acquisition and off-line processing.

The DSP board is a *Loughborough Sound Images* (*LSI*) product. The *PC/C32 TMS320C32* board includes the *Texas Instruments' TMS320C32* processor running at 50 MHz with 32 bit floating point and 50 MFLOPS peak throughput [Loug96].

The on-board memory is divided into two banks of *SRAM* (*static RAM*) 32K x 32 and 512K x 32, used for data storage and fast code execution. The *DPRAM* (*Dual-Port RAM*) of 2K x 16 provides a means to exchange data between PC and DSP.

I/0 is enabled through daughter modules. The main board allows accommodating up to two daughter modules in on-board sites. The interface between the host processor and the daughter module is provided by LSI's standard ASIC *AMELIA*. Additional daughter modules can be plugged on a separate slave board. The off-board access to DSP processor board is enabled through the *DSPLINK2* parallel interface.

In this study, the daughter modules are of type *AM/D16SA* with 2 input and 2 output channels using 16 bit successive approximation data converters. Active anti-aliasing filters are provided for the inputs. Sample rate generation is from a source on the carrier board or from an externally supplied source. Both analog inputs have a maximum input swing of ±3 V, leading to a maximum resolution of 0.18 mV/bit. The extension to 6 I/O channels (corresponding to the maximum set of 6 sensors) is realised through the adjunction of the *PC/DMCB2* carrier board. The link with the DSP processor board is enabled through the DSPLINK. The carrier board can operate using either the same interrupt than the main board or its own interrupt. Choosing the first option allows keeping a fixed order of channel acquisition. To this aim, both boards must be synchronised[1].

The DSP processor has been programmed in C language using the *Texas Instruments' TMS320 Floating-point DSP optimizing C Compiler*. The PC programs are also written in C language (*Microsoft Visual C++, V. 1.52*) and interfaced to the DSP programs using the LSI Interface Library.

C.2. References

[Laps94] P. Lapsley, "DSP Processor Fundamentals: Architectures and Features", 1994-1996 Berkeley Design Technology Inc., 1994.

[Loug96] Loughborough Sound Images, "PC/C32 Real Time Application Platform", Loughborough, Leicestershire (E), 1996.

[1] Surprisingly, LSI does not provide a suitable synchronisation connector. The user has to build its own cable!

Appendix D
Swiss DGPS

The DGPS service broadcasts corrections from the GPS reference antenna, placed in Zimmerwald (Bern), over a network of 6 FM radio stations, which are described in Table D.1.

Radio station	La Dôle	Chasseral	Niederhorn	Rigi	Säntis	St. Christona
Frequency MHz	100.1	105.3	105.8	103.8	105.6	103.6
Reception area	S-West Geneva-Lausanne	Neuchâtel Fribourg Berne	Berner Oberland	Middle Lucerne	North Zurich Bodensee	Around Basel

Table D.1: Main RDS/DGPS transmitters in Switzerland.

The current configuration covers 60 % of the country but will be extended in a near future.

Two levels of correction accuracy are available depending on the subscription price:

- PREMIUM service – 1-2 m accuracy – 800 CHF/year,
- BASIC service – 5-10 m accuracy – 200 CHF/year

The lower accuracy service was chosen in this study.

Further information on the Swiss DGPS service is available at www.swisstopo.ch.

Appendix E
Co-ordinate transformations

E.1. Geographical data

$\pi = 3.14159265...$
$\rho = \pi / 180 = 0.01745329...$

WGS-84 ellipsoid

$a_w = 6378137.000$ m

$b_w = 6356752.314$ m

$e_w^2 = 1 - (b_w - a_w)^2 = 0.00669438...$

CH-1903 ellipsoid

$a_b = 6377397.155$ m

$b_b = 6356078.963$ m

$e_b^2 = 0.00667437...$

Bern Old Observatory ellipsoidal co-ordinates

$\phi_0 = 46.952405556° = 46°57'08.66''$

$\lambda_0 = 7.439583333° = 7°26'22.50''$

Earth radius

$$r = a_b \cdot \frac{\sqrt{(1-e_b^2)}}{(1-e_b^2 \cdot \sin^2(\phi_0 \cdot \rho)} = 6378815.9036 \text{ m}$$

Miscellaneous

$$\alpha = \sqrt{(1+\frac{e_b^2}{(1-e_b^2)} \cdot \cos^4(\phi_0 \cdot \rho))} = 1.00072914...$$

$$r_k = \ln(\tan(\frac{\pi}{4} + b_0 \cdot \frac{\rho}{2})) - \alpha \cdot \ln(\tan(\frac{\pi}{4} + \phi_0 \cdot \frac{\rho}{2}))$$

$$+ \alpha \cdot \frac{e_b}{2} \cdot \ln \frac{(1+e_b \cdot \sin(\phi_0 \cdot \rho))}{(1-e_b \cdot \sin(\phi_0 \cdot \rho)} = 0.00306673...$$

$$b_0 = a\tan(\frac{\sin(\phi_0 \cdot \rho)}{\alpha \cdot \sqrt{1-\left(\frac{\sin(\phi_0 \cdot \rho)}{\alpha}\right)^2}}) \cdot \frac{1}{\rho}$$

E.2. WGS-84 ellipsoidal co-ordinates into CH-1903 cartesian co-ordinates

Helmert parameters

Translation

dx = -660.077 m
dy = -13.551 m
dz = -369.344 m
m = -5.66000 ppm
tx = -2.484568 cc
ty = -1.783951 cc
tz = -2.938272 cc

Rotation matrix

1.00000000	-0.00000462	0.00000280
0.00000462	1.00000000	-0.00000390
-0.00000280	0.00000390	1.00000000

WGS-84 ellipsoidal co-ordinates. into WGS-84 cartesian co-ordinates

$$R = \frac{a_w}{\sqrt{1 - e_w^2 \cdot \sin^2(\phi_w \cdot \rho)}}$$

$$x = (R + h_w) \cdot \cos(\phi_w \cdot \rho) \cdot \cos(\lambda_w \cdot \rho)$$

$$y = (R + h_w) \cdot \cos(\phi_w \cdot \rho) \cdot \sin(\lambda_w \cdot \rho)$$

$$z = (R \cdot (1 - e_w^2) + h_w) \cdot \sin(\phi_w \cdot \rho)$$

WGS-84 cartesian co-ordinates into CH-1903 cartesian co-ordinates

$$\begin{pmatrix} x_b \\ y_b \\ z_b \end{pmatrix} = (1 + \frac{m}{1000000}) \cdot Rot \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + T$$

CH-1903 cartesian co-ordinates into CH-1903 ellipsoidal co-ordinates

$$\phi_0 = a\sin(\frac{z_b}{a_b})$$

$$\lambda_{br} = a\tan(\frac{y_b}{x_b})$$

$$\phi_n = \phi_{n-1} - (\frac{x_b}{\cos(\lambda_{br})} \cdot \tan(\phi_{n-1}) - \frac{a_b \cdot e_b^2 \cdot \sin(\phi_{n-1})}{\sqrt{1 - e_b^2 \cdot \sin^2(\phi_{n-1})}} - z_b)/A$$

where:

$$A = \frac{x_b}{\cos(\lambda_{br}) \cdot \cos^2(\phi_{n-1})} - a_b \cdot e_b^2 \cdot \cos(\phi_{n-1}) \cdot (\frac{\sin^2(\phi_{n-1}) \cdot e_b^2}{\sqrt{1 - e_b^2 \cdot \sin^2(\phi_{n-1})^3}}$$

$$+ \frac{1}{\sqrt{1 - e_b^2 \cdot \sin^2(\phi_{n-1})}})$$

$$\phi_b = \frac{\phi_n}{\rho}$$

$$\lambda_b = \frac{\lambda_{br}}{\rho}$$

$$h = \frac{x_b}{\cos(\lambda_{br}) \cdot \cos(\phi_{br})} - \frac{a_b}{\sqrt{1 - e_b^2 \cdot \sin^2(\phi_{br})}}$$

CH-1903 ellipsoidal co-ordinates into CH-1903 plane co-ordinates

Ellipsoid to sphere

$$b_k = 2 \cdot (a\tan(\exp(\alpha \cdot \ln(\tan(\frac{\pi}{4} + \frac{\phi_{br}}{2})) + r_k - \alpha \cdot \frac{e^b}{2} \cdot \ln(\frac{1 + e_b \cdot \sin(\phi_{br})}{1 - e_b \cdot \sin(\phi_{br})}) - \frac{1}{2}$$

$$l_k = \alpha \cdot (\lambda_{br} - \lambda_0 \cdot \rho)$$

Sphere to map

$$b_q = a\tan$$

$$(\frac{\cos(b_0 \cdot \rho) \cdot \sin(b_k) - \sin(b_0 \cdot \rho) \cdot \cos(b_k) \cdot \cos(l_k)}{\sqrt{(\sin(b_0 \cdot \rho) \cdot \sin(b_k) + \cos(b_0 \cdot \rho) \cdot \cos(b_k) \cdot \cos(l_k))^2 + (\cos(b_k) \cdot \sin(b_k))^2}}$$

$$l_q = a\tan(\frac{(\cos(b_k) \cdot \sin(l_k))}{\sin(b_0 \cdot \rho) \cdot \sin(b_k) + \cos(b_0 \cdot \rho) \cdot \cos(b_k) \cdot \cos(l_k)})$$

$$y_b = r \cdot lq + 600000$$

$$x_b = \frac{r}{2} \cdot \ln(\frac{1 + \sin(b_q)}{1 - \sin(b_q)}) + 200000$$

E.3. CH-1903 cartesian co-ordinates into WGS-84 ellipsoidal co-ordinates

Helmert parameters

Translation

dx = 660.080 m
dy = 13.550 m
dz = 369.348 m
m = 5.66003 ppm
tx = 2.484559 cc
ty = 1.783963 cc
tz = 2.938265 cc

Rotation matrix

1.00000000	0.00000462	-0.00000280
-0.00000462	1.00000000	0.00000390
0.00000280	-0.00000390	1.00000000

CH-1903 plane co-ordinates into CH-1903 ellipsoidal co-ordinates

Map to sphere

$$b_q = 2 \cdot a\tan(\exp(\frac{x - 200000}{r})) - \frac{\pi}{2}$$

$$l_q = \frac{y - 600000}{r}$$

Sphere to sphere

$$b_k = a\tan(\frac{\cos(b_0 \cdot \rho) \cdot \sin(b_q) + \sin(b_0 \cdot \rho) \cdot \cos(b_q) \cdot \cos(l_q)}{\sqrt{1 - (\cos(b_0 \cdot \rho) \cdot \sin(b_q) + \sin(b_0 \cdot \rho) \cdot \cos(b_q) \cdot \cos(l_q))^2}})$$

$$l_k = a\tan(\frac{\cos(b_q) \cdot \sin(l_q)}{\cos(b_0 \cdot \rho) \cdot \cos(b_q) \cdot \cos(l_q) - \sin(b_0 \cdot \rho) \cdot \sin(b_q)})$$

Sphere to ellipsoid

$$\phi_1 = b_k$$

$$\phi_n = 2 \cdot \text{arctg}(e^B) - \frac{\pi}{2}$$

where:

$$B = \frac{\ln(\tan(\frac{\pi}{4} + \frac{b_k}{2}))}{\alpha} + \frac{1}{2} \cdot \ln(\frac{1 + e_b \cdot \sin(\phi_n)}{1 - e_b \cdot \sin(\phi_n)}) \cdot e_b - \frac{r_k}{\alpha}$$

$$\lambda_r = \lambda_0 \cdot \rho + \frac{l_k}{\alpha}$$

$$\phi = \frac{\phi_r}{\rho}$$

$$\lambda = \frac{\lambda_r}{\rho}$$

CH-1903 ellipsoidal co-ordinates into CH-1903 cartesian co-ordinates

$$R = \frac{a_b}{\sqrt{1 - e_b^2 \cdot \sin^2(\lambda \cdot \rho)}}$$

$$x = (R + h) \cdot \cos(\phi \cdot \rho) \cdot \cos(\lambda \cdot \rho)$$

$$y = (R + h) \cdot \cos(\phi \cdot \rho) \cdot \sin(\lambda \cdot \rho)$$

$$z = (R \cdot (1 - e^2) + h) \cdot \sin(\phi \cdot \rho)$$

CH-1903 cartesian co-ordinates into WGS-84 cartesian co-ordinates

$$\begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} = (1 + \frac{m}{1000000}) \cdot Rot \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} + T$$

WGS-84 cartesian co-ordinates into WGS-84 ellipsoidal co-ordinates

$$\lambda_0 = a\sin(\frac{z_w}{a_w})$$

$$\lambda_{wr} = a\tan(\frac{y_w}{x_w})$$

$$\phi_n = \phi_{n-1} - (\frac{x_w}{\cos(\lambda_{wr})} \cdot \tan(\phi_{n-1}) - \frac{a_w \cdot e_w^2}{\sqrt{1 - e_w^2 \cdot \sin^2(\phi_{n-1})}} \cdot \sin(\phi_{n-1}) - z_w)/A$$

where:

$$A = \frac{x_w}{\cos(\lambda_{wr}) \cdot \cos^2(\phi_{n-1})}$$

$$- a_w \cdot e_w^2 \cdot \cos(\phi_{n-1}) \cdot (\frac{\sin^2(\phi_{n-1}) \cdot e_w^2}{\sqrt{1 - e_w^2 \cdot \sin^2(\phi_{n-1})}^3} + \frac{1}{\sqrt{1 - e_w^2 \cdot \sin^2(\phi_{n-1})}})$$

$$\phi_w = \frac{\phi_{wr}}{\rho}$$

$$\lambda_w = \frac{\lambda_{wr}}{\rho}$$

$$h = \frac{x_w}{\cos(\lambda_{wr}) \cdot \cos(\phi_{wr})} - \frac{a_w}{\sqrt{1 - e_w^2 \cdot \sin^2(\phi_{wr})}}$$