Underwater Vehicle Integrated Navigation

author : A.F.Neele

TECHNISCHE UNIVERSITEIT DELFT
Faculteit der Elektrotechniek
Vakgroep Telecommunicatie- en
Verkeersbegeleidingssystemen

Titel : Underwater Vehicle Integrated Navigation

Auteur : A.F. Neele

Aard : Afstudeerverslag

Omvang : 222 pagina's

Datum : 07 april 1993

Hoogleraar : Prof. Dr. Ir. D. van Willigen
Prof. Ir. J.A. Spaans

Mentor(en) : -

Opdrachtnr. : A - 490

Periode : januari 1992 - maart 1993

Most of the research done on navigation nowadays, deals with the optimization of space, air, land and sea-surface navigation. Research into underwater navigation however, is a somewhat neglected area, although there are many submersible vehicles in use, ranging from small unmanned remotely operated vehicles (ROVs) to large nuclear submarines.
Navigators of submarines - these vessels probably representing the largest group of submersibles - are particularly in need of accurate underwater navigation methods. At present, navigators on board submarines of some European countries use a concept called the 'Pool of Errors' (POE) to establish position with associated confidence region. This is a purely graphical method, the accuracy of which is questionable. As this is recognized, a request for an investigation on how to improve underwater navigation on board submarines was forwarded by the Royal Netherlands Navy.

The paper starts with an evaluation of the concept of the POE as currently used, pointing out some important shortcomings in its use. In order to improve underwater navigation and to overcome the shortcomings, a mathematical and a statistical model, based on rigorous formulae, to be used for position fixing and quality control, are proposed. These models are implemented into a computer simulation program which is developed to show the main features of integrated navigation and can serve as a basis for an integrated navigation system to be implemented in a real-time environment on board submarines.

Table of Contents

1. Introduction

Most of the research done on navigation deals with systems that can be used for navigation of vehicles on land, in the air or at the sea surface. There is however a category of users that can only make use of these systems for a limited amount of time. This is the group of navigators on board submersible vehicles, such as submarines. While the vehicle is at the sea surface, use can be made of, for example, satellite and terrestrial navigation systems to obtain an accurate position. Once submerged, information is very much limited to the use of the vessel's bounded sensors or special underwater navigation systems. Acoustic bottom references such as bottom mounted acoustic transponders are impractical because of their limited area of coverage. Perhaps an occasional check-point from a recognizable terrain feature is the best that can be hoped for as a bottom reference. One must therefore fall back on dead reckoning to establish a position.
At present day, navigators on board submarines belonging to navies of some European countries use a concept called 'Pool of Errors' (POE) to establish their position with associating confidence region. This is a purely graphical method which usefulness and accuracy is questionable. Furthermore, it is a time consuming method especially since all factors affecting position accuracy need to be included, leaving no room to identify any errors present in information provided by the systems.

A few reasons may be given to illustrate why a more accurate method of position fixing underwater has to be developed. One of them is the fact that the 'Marineraad'[1] of the Royal Netherlands Navy urged for research on underwater position fixing in one of its verdicts after investigating an accident that had happened to one of the Dutch submarines as a result of misinterpretation of information leading to a wrong position. Another reason is that over the past few years several incidents took place between fishermen and submerged submarines of the Royal Navy (UK) in which the fishermen claimed damage allegedly caused by submarines. In court the evidence of the submarine's position at the time of the accident based on the concept of the 'Pool of Errors', was not admitted as sufficient, resulting in claims by the fishermen being granted. Inaccurate position fixing can also lead to trespassing in territorial waters by any vehicle navigating underwater. This has always been a delicate matter easily escalating into a diplomatic incident.

---

[1] Naval Council

1.1 Aim of the research

Before starting the research, it is important to state what will be investigated and what will be left out of consideration, especially because the variety of sensors and system that can be used for navigation is immense. Furthermore, the group of underwater vehicles comprises the whole range between a small unmanned remotely operated vehicle (ROV) and large nuclear submarines, so a choice has to be made what sort of vehicle and application to focus on.
As the question to investigate how navigation underwater can be improved was posed by the Royal Netherlands Navy, and since the group of submarines probably forms the largest and most important group of submersible vehicles in need for accurate underwater navigation, the research will focus on the way navigation can be improved on board submarines of the Royal Netherlands Navy. This choice automatically leads to the sensors and systems to be considered, namely those present on board the mentioned submarines.

After having chosen platform and range of sensors and systems to be used, the aim of the research has to be described. The aim is threefold :

1. Evaluation of the concept of the POE as it is currently used on board submarines of the Royal Netherlands Navy;

2. Find a suitable mathematical and statistical model that can be used to calculate the Most Probable Position (MPP) by combining information provided by the vessel's bounded and/or unbounded systems and sensors. It must also be able to give figures of position accuracy and perform quality control of information provided.

3. The development of a computer simulation program based on the mathematical and statistical model, that shows the main features of integrated navigation and that can serve as basis for an implementation in a real-time environment (i.e. on board the submarine).

It should be noted that it is not the intention of this research to give the optimum solution for integration of systems, but merely to show how integration of navigation systems using a computer can lead to more reliable position fixing. So each system is regarded as an entity and no special study has been made on integration of a minimum number of systems to reach maximum performance (eg. GPS with an Inertial Navigation System), although when giving suggestions for further research and improvements in the final chapter of this paper, some remarks will be made on this matter.

1.2 Preview

In order to get an impression on the contents of the chapters and their relation with respect to each other, a brief summary of each chapter will be given here.

In chapter 2 the concept of the POE as it is currently used on board submarines will be explained, together with the problems encountered when using it.

After having evaluated the concept of the POE, in chapter 3 the main characteristics of the systems and sensors available at present on board submarines to obtain information for position fixing, are described. The chapter does not deal with the characteristics in great detail since plenty of good textbooks are available. The chapter is merely included to state what is available. Next, Chapter 4 deals with the errors present in the data provided by the systems and sensors described in chapter 3. Although this chapter forms the basis for the statistical models used in calculations, care had to be taken again not to go into too great detail when describing the error sources. Therefore, only of those error sources that contribute significantly to the total error of each system and sensor are briefly described, resulting in an error budget for each observable.

Having dealt with the main characteristics and error sources of the systems and sensors available, the algorithms that can be used to calculate the MPP from observations, need to be looked at. These algorithms, that form the main calculation part of the computer simulation program, are provided in the next two chapters of the paper. In chapter 5 a description is given of the least squares algorithm used to derive an MPP from LOPs obtained from unbounded systems (i.e. external information). The algorithm can only be used when the submarine is at the sea surface. The chapter also contains a section on statistical tests performed to evaluate the quality of the MPP derived by means of least squares. Chapter 6 deals with the Kalman filter algorithm, which is either used to combine the MPP derived using the least squares algorithm with a position based on a ships model and bounded sensor (i.e. internal information) or to predict the probable (future) position of the vessel based on the ships model and bounded sensor information only. The advantage of this algorithm is that it can be used both when the submarine is at the sea surface or submerged.

Chapter 7 will focus on the developed computer simulation program. It will provide an outline of the computer program, a general description of the software, its outputs and a concise user manual. Furthermore some limits of the program will be given.

Finally, chapter 8 contains concluding remarks and suggestions for further research. Here an evaluation of the integrated system - as simulated by the computer program -, will be given. Furthermore some recommendations will be made on the optimum use of already available systems and sensors and possible changes in equipment needed to improve the accuracy of position fixing underwater in particular. Also suggestions for further development of the simulation program and underlying mathematical and statistical models will be given in order to make it ready to be used in a real-time environment.

2. Concept of the Pool of Errors

2.1 Introduction

In this chapter the concept of the Pool of Errors (POE) as it is currently used on board submarines will be discussed. The chapter starts however, with a section containing definitions of important terms that will be used throughout the paper. Since many subject related terms and abbreviations will be used, appendix 1 has been added in which a complete list of definitions, terms and abbreviations used, can be found. In the next section current use of the concept of the POE will be described, using information from existing naval publications on this subject. The third section describes in short how the theory, as discussed in the second section, is actually used on board. In the final section of this chapter some of the most important shortcomings in the use of the concept are listed, giving a good idea why a more rigorous treatment of this subject is needed.

2.2 Definitions

Although appendix 1 will contain a list of definitions, it is important to discuss a few of them here in greater detail before starting to investigate any theories, since a good understanding of their meaning is relevant for a good understanding of the theory and conclusions.

Because a considerable part of the theory of statistics is used in the theory described in this paper, related terms will be explained first. The most important terms to be used are **precision**, **reliability** and **accuracy**. There is generally much confusion on the interpretation of these terms.

- **PRECISION** is the degree of agreement between repeated measurements of the same quantity to each other.

When an instrument is used for observing a quantity, precision gives an indication of the spreading of the measurements when the observation is repeated many times under the same conditions. The need for same conditions is important for the following reasons:

- systematic errors present in the observations will not have direct effect on the precision. However, before the precision of measurements can be evaluated from observations made, it is important that any systematic errors present are the same for all measurements made. A bias due to the systematic error will exist;

- if the conditions under which the measurements are taken change, the precision might change as well.

Precision of a measurement gives information about the magnitude of ***random errors*** present. It is stated as standard error in units of the system (eg. metres when measuring ranges, angles when measuring bearings). In order to be able to validate the accuracy of observations made using different systems, the standard error will, for all systems, be transformed into meters. How this is done is further explained in chapter 4.

- **RELIABILITY** is the measure of ease with which a blunder in a measurement can be detected.

Several quantities have been suggested to give a quantitative measure of reliability [Spaans,1988:2; Cross et al.,1985]. The definition as given above refers to a statistical property of a measurement. It should not be confused with the reliability of a system, which is given as [FRP,1990] :

- **RELIABILITY** is the probability of performing a specified function without failure under given conditions for a specified period of time.

- **ACCURACY** is the degree of conformity between the true value of a quantity and the most probable value derived from a series of measurements (estimate).

A distinction can be made between several types of accuracy. Since the true value of the measured quantity is unknown, it is better to talk about ***predictable accuracy***, when the quantity is derived from measurements, taking into account all predicted errors. Consider for example a situation in which positioning equipment is placed at an already known position. If this position is re-measured over a period of time, using this equipment, a number of calculated positions scattered round the true position will be found. From this 'scatter plot' a measure of ***absolute accuracy*** by which the position could be derived using the given instrument, should be stated.

Since accuracy is a statistical measure of performance, a statement of the accuracy of a measuring device or measured quantity is meaningless unless it includes a statement of uncertainty. This will be derived from the VCV matrix of the measured quantities in which on the main diagonal the variances of the quantities are given. Note that in the linear case, i.e. one measured quantity, this VCV matrix will contain only one element as could be expected. Using these values a number of confidence regions can be defined. For position fixing in 2D the most common ones are the ***standard error ellipse***, the ***a% error ellipse*** - where ***a*** is normally given as 95% or 99.9%[1] - the d_{rms} or the ***CEP*** whereas in 3D these will normally be the ***standard error ellipsoid***, the ***a% error ellipsoid***, or the ***SEP***. In section 5.5 of chapter 5, information concerning the relationships existing between these different quantities will be given.

As a final remark it should be stressed again that accuracy is the degree of conformity with the correct value, while precision is the degree of refinement of a measured value. Therefore, when blunders and/or systematic errors are present in the measurements of a quantity, it can still be determined with high precision but it will have low accuracy. Any systematic errors present will lead to bias. In figure 2.1 this is visualized for a one dimensional quantity being measured. When comparing the two pdf's, p_1 shows less precision than p_2, while on the other hand p_2 shows higher accuracy then p_1.

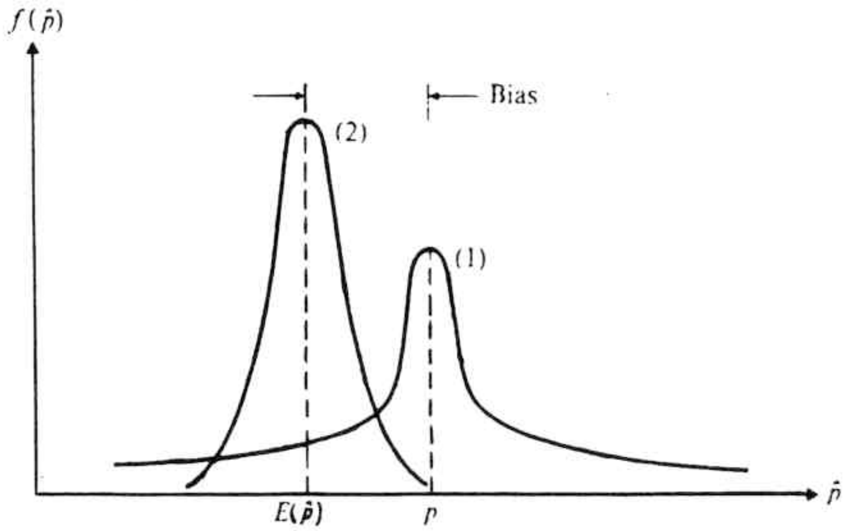


Figure 2.1 Accuracy and precision of a one-dimensional variable p being measured.

---

[1] With respect to aids to navigation the IMO, in its resolution A.529 (13) of 17 November 1983, states that : "The 95% probability figure should be used to describe the accuracy of a system fix" [IALA,1990]

The definitions that are given next are related to errors. In general the following three types of errors can be distinguished :

- **BLUNDERS** : This type of error is difficult to define since a deviating value could be well within the possible values of a p.d.f. A possible definition could therefore be that a blunder is a measurement which differs significantly from the expected value making it very likely that certain external circumstances are present other than the ones that would make it a random error (under normal circumstances) [Spaans,1988:2].

- **SYSTEMATIC ERRORS** : errors that follow some law by which they can be predicted.

Constant and low frequency disturbances (i.e. the DC and low frequency part of the error spectrum) are considered to be systematic errors. A distinction should be made between systematic errors present in the system used to measure a quantity and those introduced by the user of that system. Constant and low frequency errors of a system are obtained from analysis of the error spectrum which provides insight in the dynamical behaviour of the system. In some cases the constant error can be removed from the system by means of calibration, but often corrections need to be applied to the measurements.
Systematic errors introduced by the user are more difficult to recognize. However, experience could lead to assessment of errors introduced after which corrections can be applied to the measurements.

- **RANDOM ERRORS** : These errors are unpredictable in magnitude and/or sign. They are governed by laws of probability which means that they can be characterized by a p.d.f.

The total random error present in a measurement will in most cases consist of a mixture of random errors introduced in different parts of the observation process (due to environment, measurement sensor or system, user etc.). Each of these random errors will have its own pdf. In general, this pdf is assumed to be zero mean Gaussian, but other types should be considered as well. Errors introduced by rounding off can serve as an example. These errors have an uniform pdf.
The **central limit theorem** states that when the number of mutually independent random variables (i.e. the error components) increases, the distribution of their sum gets closer to the Gaussian (normal) distribution. It is therefore not unreasonable

at all to use the normal distribution as the probability model for the total random error in an observation when this error is assumed to have many components.
The random errors present in consecutive measurements obtained from a system or sensor, exhibit an auto-time-correlation which can in most cases be described by a first-order Gauss-Markov process. For a short description see appendix 6. For RPF systems, the correlation time is dependent on carrier frequency used : the higher the frequency, the smaller the correlation time.

The last set of definitions to be given in this section are the ones used to classify position fixes. The following types of fixes will be used when discussing position fixing at sea :

- DEAD RECKONING (DR) POSITION : this is the position based on the most recent fix, updated using information from the bounded sensors that measure the vessels *heading*, *speed through the water* and *attitude* (roll,pitch etc.). When establishing this position, the estimated effects of wind, currents and tidal streams, sea state etc. are also taken into account.
- MOST PROBABLE POSITION (MPP) : this is the best position that can be derived using all information available.

It should be noticed that, given the circumstances for position fixing, each position fix, even a DR position can be regarded as MPP.

2.3 Principles of the 'Pool of Errors'

2.3.1 Philosophy and use

The POE is a concept used on board submarines to provide a graphical indication of the maximum likely errors in the MPP at any time. By using the POE the commanding officer can determine safe course, speed and depth. When the submarine is submerged, the commanding officer has to decide when a new update of the position, using external navigation systems, is required. This decision depends on the size of the POE, the area in which the submarine is operating and the dangers present.

For optimum use of the POE it is very important to make an assessment of the various errors - both systematic and random - present and their effects on the submarine's position accuracy.

This can be very difficult, especially if the submarine is operating in areas where knowledge of external disturbances acting on the vessel is limited. Therefore standard errors of measurements are estimated in such a way that their true values - which are unknown - will be guaranteed to be smaller than those estimated. After observations have been made over a period of time, in which estimated positions are compared against position updates, the standard errors can be redefined in such a way that they agree more with the actual (true) values. One need to be reminded of the fact that in all cases average standard errors are used.

The way the POE is used during sailing consists of three steps:

1. When the submarine is at the sea-surface, its position is fixed by means of external sources. As long as the submarine stays at the surface, the position can be updated continuously. This leads to the normal position fixing and quality control using common mathematical methods such as least squares. The POE will increase in time, but the continuous flow of sensor information makes it possible to reduce the size continuously, keeping it to an acceptable size.

2. As soon as the submarine submerges, use can no longer be made of external sources. Its position is therefore calculated using bounded sensor and current information. Now the POE is introduced as an expanding mathematical figure. If possible, information from external sources such as eg. bottom contours and seabed slopes measured using an echo sounder, will be used to update the position and to decrease the size of the POE.

3. Eventually the dimensions of the POE reach limits, making it unacceptable to stay underwater using DR and expanding POE. The submarine is brought either to periscope depth or surfaced making it possible for the navigator to update the position using information from external sources, resulting in a new MPP and small POE.

The means available for position fixing when the submarine is at the sea surface are terrestrial radio position fixing (RPF) systems - such as Loran-C, Omega and Decca -, satellite position fixing systems and compass bearings, radar distances etc. Once the submarine submerges, use can only be made of its bounded sensors and systems and occasionally of its echo-sounder. All these systems have errors which can be evaluated quite well during trials at the surface and errors will in general remain unchanged when the submarine submerges. The currents however pose a bigger problem as it is difficult to estimate their influence on the movements of the submarine. Their direction and strength can change quite significantly with depth. These changes make it very difficult to estimate their effect on the POE. To be able to use the POE in unknown waters without reaching unacceptable limits quickly, three types of operation have been defined, each giving its own restrictions to the limits of the POE. These limits depend highly on the sort of area and navigational dangers present. The three types of operation are :

1. area operations : the submarine will be manoeuvred in a designated area in open sea. The smaller the size of the area, the more course and/or speed changes will be made. Generally, the movements of the submarine are not limited by seabed features;

2. transit : the journey from one area to another;

3. confined waters operations : operations in those areas where manoeuvring is restricted due to bottom features and/or landmasses present.

Each of the above mentioned operations imposes its own needs for positional precision and accuracy, resulting in different methods for construction of the POE. Beside this it is of course important always to evaluate the general bottom contours and find dangers present at the most likely depth at which the submarine is operating.

2.3.2 Construction of the POE

After having given a broad description of the purpose of the POE, it is also important to give information on the way the POE is actually being constructed. As the navigator is not able to spend much time on plotting positions and constructing a POE, position fixing and plotting in a chart has to be done quickly. Therefore special hand methods have been developed, the results of which are shapes that are relatively easy to draw such as circles and rectangles with rounded corners. Basically, the following methods are used at present :

- expanding circle method : a circle, having the current DR position as its centre, is drawn. The radius of the circle is calculated from

$$R_{ec} = a + \epsilon \times \Delta t \quad \text{[metres]} \tag{2.1}$$

where

R_{ec} = radius of circle
a = the d_{rms} value of the last MPP based on external sources
ϵ = expansion factor
Δt = time interval between last MPP based on external sources, and current DR position

The magnitude of the expansion factor is a combination of the accuracy with which the current direction and speed are known and the standard errors of the bounded systems used to obtain the current DR position. No fixed value for the expansion factor is given, but values between 900 - 1800 metres (1000 - 2000 yards) per hour are generally used. This way, the radius of the circle will expand linearly in time. This method is used when the submarine is operating in deep, open waters where tidal streams are either circular or variable.

- circle segment method : The expanding circle is used again, with ϵ set to a value representing the total error in log and gyro. To allow for uncertainties due to tidal streams and currents, part of a circle segment can be constructed having a depth (d) and width (2a) (see figure 2.2). The size of these measures depends on the present situation. The

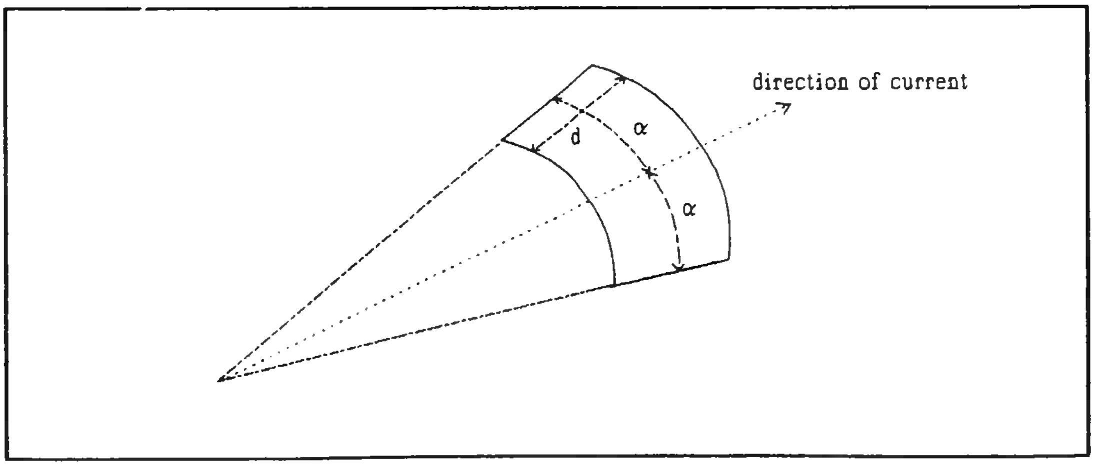


Figure 2.2 The part of the circle segment used to construct the uncertainty in the position fix introduced by current.

resulting area, giving the uncertainty in the position fix due to tidal streams and current, is combined with the expanded circle to give the POE.
This method is used in areas where the general direction and speed of tidal streams and currents are quite well known.

2.4 Problems related to the use of the POE

Although the concept of the POE in itself works really well, there are some problems when deriving a POE and using information found. In this section the most important problems will be discussed.

To start with, there is not sufficient backup of systems or sensors on board the submarines. This leads to the following problems :

1. when no backup of a sensor or system is available at all:

   - when sensors drift (such as gyros used in the gyro compass or the Ships Inertial Navigation System (SINS)), the navigator will have no means to observe this because no reference is available;

   - If a system or sensor fails completely, a primary source for position updating and quality control is lost.

2. when only one backup system or sensor is available for a particular system or sensor :

   - If only one of the two is running at the time and this system or sensor has a total failure, valuable time will be lost during start-up of the backup system, leading to a temporary loss of a primary source for position updating and quality control. Beside that, no information is present on differences between the two instruments, since no comparison had been made between two working systems;

   - If both systems are running and one system drifts from the right value, the navigator will not be able to determine which system is working correctly so basically neither of the systems can give good information to be used for position fixing. Again a primary source for position update and quality control is lost.

On the quality control side of position fixing the following group of problems can be distinguished :

1. Use of 'a priori' standard errors as precision of measurements :

   - A priori standard errors are stated for several systems and sensors, but no statement has been made on how these values were obtained. Furthermore, it is left to the judgement of the commanding officer to change these values if this is thought to lead to better results. No guidelines are given here.

   - How can the user be sure that even if he could estimate a priori standard deviations, these will be the correct ones. It could well be necessary - because of the changing conditions - to scale the a priori standard deviations continuously by σ_0 obtained from initial calculations.

2. The following remarks can be made on the current use of the POE, i.e. purely as a hand method :

   - the method does not give a unique solution for the position fix and size of the POE. There is too much room for personal interpretation of the problem;

   - the expansion of the circle used in both methods to obtain the POE as described in section 2.3.2, is based on the direction and distance between the last fix and current DR position. It is made independent of the actual track between these positions. Especially when the submarine has made many changes in course and/or speed, fix errors tend to be under-estimated;

   - as a submarine is able to move underwater and therefore in 3D space, its position needs to be treated as a 3D position. This means that the POE needs to give 3D information. At present, the POE only provides 2D information;

   - it is not possible to get valid information regarding the quality of the fix, reliability of LOPs or a priori and a posteriori standard errors by using the plot of the POE. Therefore the POE cannot give a true representation of the probability area of the fix;

   - once the navigator has been able to calculate the position of the submarine using external information, the already existing DR position is not integrated with the new position fix to form a new MPP. The DR position is disregarded. Even though the POE of this DR position may be quite large, the position can still be integrated with the new position fix derived from external sources, possibly leading to new information on standard errors to be used.

At the moment there is no integration of systems leading to a direct plot of the POE on for example a plotting table or computer screen. Everything has to be done by hand. To do this properly, taking into account all the errors present in the sensors and systems used, is quite a complicated job.

From the above it should be clear that preferably three independent working sets of each piece of equipment should be used and running. Not only will this provide backup when a systems breaks down, it will also show if a system drifts. Normally, this is not feasible. Therefore, sufficient backup information should be provided by the other working systems as this will provide redundancy and a system cross-checking ability. A consistent mathematical solution to position fixing and quality control should be available to the navigator as this is the most objective way by which results can be obtained, compared and saved for future use. Mathematical models that provide all this will be presented in chapters 5 and 6.

GPS 3-D Positioning Accuracies* (Metres, 95%)

Operating Mode		Dual Frequency		Single Frequency	
SA	A-S	PPS	SPS	PPS	SPS
OFF	OFF	37	37	51	51
ON	OFF	37	170	51	174
OFF	ON	37	NA**	51	51
ON	ON	37	NA**	51	174

* Worldwide over 24 hours; minimum satellite elevation 5°.
** NA indicates not available with A-S on.

GPS Horizontal Positioning Accuracies* (Metres, 95%)

Operating Mode		Dual Frequency		Single Frequency	
SA	A-S	PPS	SPS	PPS	SPS
OFF	OFF	21	21	29	29
ON	OFF	21	98	29	100
OFF	ON	21	NA**	29	29
ON	ON	21	NA**	29	100

* Worldwide over 24 hours; minimum satellite elevation 5°.
** NA indicates not available with A-S on.

GPS Vertical Positioning Accuracies* (Metres, 95%)

Operating Mode		Dual Frequency		Single Frequency	
SA	A-S	PPS	SPS	PPS	SPS
OFF	OFF	34	34	46	46
ON	OFF	34	156	46	159
OFF	ON	34	NA**	46	46
ON	ON	34	NA**	46	159

* Worldwide over 24 hours; minimum satellite elevation 5°.
** NA indicates not available with A-S on.

table 4.2 : Summary of PPS and SPS positioning accuracies under various conditions of SA and A-S [NATO,1991:1].

be able to distinguish between signals received from different satellites, the carrier frequencies are phase modulated with codes known as P- and C/A-codes. While performing measurements on multipath signals, Kranendonk [1992] found indications that cross-interference exists between signals from different satellites. This is considered to be the result of the fact that the C/A-codes for the different satellites are not completely orthogonal to each other, resulting in unwanted peaks in the cross-correlation function.
Spread spectrum interference can simply be modelled the same way as multipath (see above) [Van Nee,1992]. The only difference is that cross-correlation peaks can precede the autocorrelation peak as if they were multipath signals with a negative delay. The fading bandwidth is now equal to the frequency difference of the interfering satellites. Instead of using the SMR, one should use the Signal-to-Interference Ratio (SIR).
Cross-correlation can cause errors of up to a few metres.

9. user dynamics

The effect of user dynamics on the position fix depends very much on the dynamics of the vessel and position of the receiver antenna. A shipborne receiver should be able to accept as input information provided by the ship's attitude and water speed sensors. The heading and water speed input signals can be used to assist in satellite acquisition. If only a poor estimate of the position and time is available locating and locking onto any satellite in view is slowed down. Once the carrier and code tracking loops are locked no position, velocity or time from outside sources is needed. The roll and pitch input signals (if present) can be used to compensate for antenna motion.
The reader is referred to NATO [1991:1 & 1991:2] for more detailed information on this subject.

10. Selective Availability (SA)

Early test results showed that positions obtained by using only C/A code had a much better predictable accuracy than expected. In order to deny unauthorised users the access to this relatively high accuracy, the Block - II SVs have been equipped with SA. SA exists of controlled but to a user unpredictable variations in the C/A-signal, introducing errors in ranges measured. These variations consist of two types :

- a frequency dither introducing errors in the navigation time coded signals;
- offsets in satellite ephemeris data giving an apparent shift of satellite position.

resulting in a wrong pseudo-range. The influence of multipath signals depends on [Van Nee,1992:1 & 1992:2] :

- the Signal-to-Multipath Ratio (SMR) and multipath signal delays;
- fading bandwidth and tracking loop bandwidth (plus pre-detection bandwidth for non-coherent DLL);
- early-late spacing (d) with respect to chip time (T_c);
- the receiver antenna attenuation;
- the measuring technique of the GPS receiver (non-coherent DLL or coherent DLL).

Table 4.1 gives an overview of whether mean code tracking errors must be expected when multipath signals are present or not. For further information, the reader is referred to Van Nee [1992:1 & 1992:2].

	slow fading	fast fading
non-coherent DLL	a non-sinusoidal signal is present, resulting in a mean delay error greater than zero	the resulting 'S-curve' is the summation of the different 'S-curves', leading to a mean delay error greater than zero
coherent DLL	a mean delay error is present	no tracking error is present

Table 4.1 Mean code tracking delay errors in multipath environment.

The influence is restricted to multipath signals with a maximum delay of $T_c + d/2$. The maximum and mean tracking errors are proportional to the early-late spacing and SMR [Brouwer et al., 1989; Van Nee,1992:1 & 1992:2; Kranendonk, 1992]. In order to reduce the chance of getting interference from multipath signals, receiver antennas that suppress signals arriving from below a certain elevation angle are used. This way, interference from signals reflected at the sea-surface is minimised. In order to avoid reception of signals reflected off the ships structure, the antenna position has to be selected with care.

8. GPS signals cross-interference

All satellites use the same carrier frequencies. In order to

The effect of the troposphere on propagation delays can be modelled using measurements of the surface temperature, atmospheric pressure and relative humidity. The correction model can be split into two parts :

- dry component : errors due to a dry (i.e. without water vapour) troposphere can be predicted to a high degree;
- wet component : more difficult to model mainly because the surface measurements of relative humidity do not accurately reflect the distribution of water vapour along the signal path.

The algorithm normally used in GPS receivers is a function of satellite elevation angle only and typically corrects for 90 percent of the tropospheric delay [Braasch,1990] :

$$\delta R_{tropo} = \frac{2.4224}{0.026 + \sin E} e^{-0.13345 H} \quad [\text{metres}] \qquad (4.3)$$

where

H = altitude of receiver above the earth's surface (in km)

E = satellite elevation angle

As the troposphere is non-dispersive in the RF part of the spectrum, models developed for Transit (see below) are applicable.

5. receiver noise and resolution

The performance of state-of-the art GPS receivers is such that receiver measurement errors due to for example quantization resolution or oscillator-phase noise of a digital tracking loop is small compared to the tracking error introduced by thermal noise. For a typical C/N_0 of 48 dB-Hz, the code phase measurement error due to thermal noise only is in the order of 1.5 m (1σ, 4Hz code tracking loop bandwidth) [Braasch, 1991].

6. multipath errors

Multipath errors are caused by satellite signals that were reflected off surfaces and thus arriving at the receiver delayed in time with respect to the line-of-sight (direct) signals. The receiver cannot distinguish between a line-of-sight and a reflected signal since both are coded the same way. Multipath signals present will lead to a shift of the zero-point using the early-late signal correlation technique, the Earth's atmosphere causing errors due to group delay and carrier advance, scintillation and refraction.

The effect of group delay and carrier phase advance due to the ionosphere can be reduced by using dual-frequency observations because the ionosphere is dispersive in the RF part of the spectrum. From measuring pseudo-ranges on two frequencies, a corrected distance can be found :

$$R_c = \frac{R_{p1} - \left(\frac{f_2}{f_1}\right)^2 R_{p2}}{1 - \left(\frac{f_2}{f_1}\right)^2} \quad \text{[metres]} \qquad (4.2)$$

where

R_c = pseudo-range corrected for ionospheric effects
R_{p1} = pseudo-range measured at frequency 1
R_{p2} = pseudo-range measured at frequency 2
f_1 = carrier frequency 1 (L_1)
f_2 = carrier frequency 2 (L_2)

It should be noted that when code correlation techniques are used to determine pseudo-ranges, L_2 can only be used for calculating the ionospheric correction when P-code access is granted. If access is denied a ionospheric delay model has to be used to reduce ionospheric effects. Parameters to be used in this model are provided in the broadcast ephemeris. These coefficients are updated at 10-day intervals, or more often if necessary, to account for seasonal and solar activity changes. After applying this correction algorithm to single-frequency pseudo-ranges, the remaining residual range error is due to short-term ionospheric range errors not accounted for by the model.
A way to gain access to the L_2 signal without having access to the P-code is by using squaring techniques in which the incoming signal is multiplied by itself. The result is a codeless carrier wave. This way a corrected pseudo-range can be measured from actual observations.
Ionospheric scintillation, which consists of a rapid fluctuation of the Total Electron Content (TEC), is unpredictable, correlated with the solar cycle and particularly severe in high latitudes. It results in a variation of amplitude and Doppler shift of the incoming signals which can lead to a loss of phase lock due to a lower SNR and/or a sudden Doppler shift outside the tracking bandwidth of the receiver.

From formula (4.1) it becomes clear which primary error sources cause the measured pseudo-range to be unequal to the true range. These error sources will be discussed briefly in turn.

1. satellite clock error

   The satellite clock is continuously monitored at the MCS. The offset of the clock from the GPS system time is measured, transmitted as data in the broadcast message and allowed for in the receiver. Remaining errors in the satellite clock (eg. from temperature changes) are very small and can not be distinguished from certain components of ephemeris data errors.

2. receiver clock error

   As is the case with the satellite clock, the clock in the receiver will not be synchronised with the GPS System Time. Unlike satellite clock errors, this difference is not monitored and since the receiver clocks are less accurate than satellite clocks, large errors can develop due to drift. In order to reduce the effect of this error, the receiver clock bias is normally taken as one of the unknowns in the position calculation algorithm.

3. ephemeris errors

   At the MCS the position of the satellites in space is calculated, based on information being received from monitoring stations. The satellite ephemeris data is updated and uplinked to the satellites every second orbit which is about once every 24 hours.

   The errors resulting from ephemeris data can be divided into two groups :

   a. errors common to all satellites used for a position fix. This results in an apparent error in the receiver clock and can be compensated for in calculation of the receiver clock offset;
   b. satellite dependent errors.

   The errors in ranges resulting from errors in ephemeris data are small.

4. ionospheric and tropospheric errors

   Part of the path between satellite and receiver goes through computation model used to calculate MPPs, presents the right results. This is done by means of statistical testing based on hypotheses to detect for example blunders (described in chapter 5) and identification of biases by extending the dynamic state vector (described in chapter 6).

4.2 The error budget of EPF systems

4.2.1 Satellite Position Fixing Systems

NAVSTAR / GPS

A minimum of four satellites is needed to determine a 3D position, whereas a 2D position can be obtained using three satellites, provided the antenna height of the receiver with respect to the spheroid is given as a parameter. The velocity of the receiver in ECEF coordinates can be calculated by using receiver velocity relative to the satellites tracked as determined by the carrier tracking loop. Both positional and velocity information are converted to the WGS84 Earth model.

The main observable is a time difference from code measurement, resulting in a 'pseudo-range' to GPS satellites[1]. The pseudo-range from a satellite to a receiver is given by :

$$R_p = R_t + c\,\delta t_p + c(\delta t_r - \delta t_s) + \delta R_e + \delta R_m + \delta R_n + SA \quad \text{[metres]} \qquad (4.1)$$

where

- R_p = measured pseudo-range satellite - receiver
- R_t = true range satellite - receiver
- c = propagation velocity of radio waves
- δt_p = timing error due to propagation delay
- δt_s = timing error due to satellite clock offset from GPS system time
- δt_r = timing error due to receiver clock offset from GPS system time
- δR_e = range error due to errors in ephemeris data
- δR_m = range error due to multipath
- δR_n = range error due to noise
- SA = Selective Availability

[1] Other observables that will not be discussed further here are: integrated Doppler count and carrier phase measurement.

- not converting the position given by the receiver to the geodetic datum of the chart;
- plotting errors.

Except for the rounding off errors, which will be regarded as random errors, all user errors should be regarded as either being systematic errors or blunders.

Apart from the types of errors mentioned above - which contribute to the total error resulting in a figure of precision for a measurement (LOP) -, geometry of transmitting stations with respect to the receiver plays an important role when assessing the accuracy of an MPP obtained by combining LOPs. Dilution of Precision (DOP) has been introduced as dimensionless factors to describe how geometry affects position fix accuracy. Although geometry is inherent to each system, no attention is paid to this subject in this chapter. The least squares algorithm, used to calculate the MPP from observations made, automatically accounts for the effect of stations - receiver geometry when calculating the accuracy of an MPP. This algorithm will be discussed in the next chapter. In this chapter, therefore, the key issue is the assessment of a figure of precision for a single LOP given by each system or sensor. For a short discussion on xDOP see appendix 3.

In the following two sections the systematic and random errors thought to be present in the observables of each of the systems and sensors described in the previous chapter, will be discussed. This leads to an error budget, resulting in a value for standard error as measure of precision of the observable (LOP or position and/or velocity) obtained. The value of standard error given is based on the assumption that no blunders are present and all systematic errors have been removed from the observation only leaving random errors. The standard error can be given in units according to the observable provided by a sensor or system, such as degrees for bearings, metres for ranging systems or centilanes when using hyperbolic position fixing systems. However, when observations of different type are combined to obtain an MPP, it is preferred to have all standard errors stated in the same unit. All values of standard deviation will therefore be converted to metres. The main reasons why this conversion is preferred will be given in the next chapter.

Before starting the discussion on errors present in observations the following remark should be made : **Any systematic error that is not removed from an observation is treated as a random error.** However, errors assumed to be zero mean random errors in nature, that contain in fact a systematic component will lead to biased estimators and therefore to degradation of the predictable accuracy and reliability of the MPP.
Therefore, quality control methods are needed to ensure that the

For each system and sensor the errors that influence the predictable accuracy of its observable can be divided into three groups :

1. System errors

   This group of errors consists all errors inherent to the total system layout, irrespective of the receiver equipment used to obtain information from the system. The group includes errors such as position accuracy of the transmitter stations, stability of transmitter oscillators, susceptibility to sky-wave interference and interference from other signals. But also errors resulting from propagation fluctuations, precipitation, seasonal changes in climate and vegetation. Most of these errors are systematic in nature and corrections to be applied to observations in order to reduce their effect can be found by calibrating the system.

2. Receiver errors

   These errors not only comprise those inherent to the receiver itself such as tracking loop errors (magnitude depending on loop construction, C/N_0 and loop bandwidth), resolution, zero errors and repeater errors, but also those resulting from spatial separation between the receiver and its antenna such as delays caused by leads or radiation noise from power cables. Some of the receiver errors can be allowed for by careful calibration of the system, while others are corrected by applying corrections to the measurements. These corrections are either obtained after long-term measurements using statistics or as a result from mathematical models, and are normally combined with the corrections to allow for system errors as mentioned above.
   When the receiver equipment gives a position as output, errors resulting from imperfections of the algorithms used to convert observations to this position must be regarded as well.

3. User errors

   These errors are those made by the user in deriving a MPP from data given by the receiver. The errors that can be made are numerous and whether they are made or not depends very much on the skill of the operator. The most important errors are:

   - not applying corrections to readings to allow for differences between actual propagation velocity and the mean velocity used to draw lattices on charts;
   - reading off errors;
   - rounding off errors;
   - incorrect receiver settings;
   - not correcting for spatial separation of receiver antennas when combining LOPs from different systems;

4. Error Budgeting

In the previous chapter the main characteristics of several sensors and systems, used for position fixing on board submarines, were outlined. In order to be able to use the information provided by them in the most effective way, it is important to know what sort of errors (their magnitude and characteristics) can be expected under operational conditions. Being able to estimate the errors likely to exist is not only important when deciding whether a system can be used or not, it may even be more important to know them when evaluating the quality of the fix, especially if different systems are integrated to obtain that fix.

In this chapter the most important errors sources for each system and sensor used, will be dealt with. The first section reviews the theory of errors in general, giving an overview of the different types of errors. In the next two sections an error budget for each of the sensors and systems described in chapter 3 will be derived, resulting in figures for standard errors of data obtained - whether this is an LOP or a position fix. For easy reference, the order in which the sensors and systems were discussed in the previous chapter is maintained as much as possible.

The results from this chapter will be used in the Least Squares algorithm and Kalman Filter, which are discussed in chapters 5 and 6 respectively. These algorithms form the basis for calculation of the MPP.

4.1 Types of errors

An error can be defined as *'the difference between a specific value and the correct or standard value'*. We distinguish between the following categories of errors : **blunders, systematic errors** and **random errors**, for which definitions were given in chapter 2. In order to be able to evaluate the predictable accuracy of systems, it is important that blunders and systematic errors are removed from observations made. Therefore, not only error sources introducing random errors will be discussed but also those giving systematic errors along with means to reduce their effect. It will, however, not always be possible to remove the systematic errors completely.
As was stated in chapter 2, the total random error of an observed quantity (range, phase-/time difference etc.) is assumed to have a time correlated Gaussian distribution.

3.4.2 Navigation software

Quite a number of software packages of navigation software is already available on the market and more are developed. The programs range from simple programs performing basic functions such as finding course and distance from one place to another, coordinate transformations from one geodetic datum to another or astro-navigation to very sophisticated special purpose programs. What is important with all software programs used for calculation of an MPP from observations is that quality control of the MPP derived is displayed in some sort, making it possible for the navigator to distinguish between systems and/or sensors providing LOPs or information that are reliable and those that are doubtful. This way the navigator will be able to make a choice which systems and/or sensors to combine in order to obtain the best MPP in a statistical sense.
Furthermore, it must be clear to the user of navigation software what sort of algorithms are used to obtain an MPP and whether data is filtered or not and in what way.

through the water, to form the depth reference plane.

3.3.7 Periscope

The periscope consists of a complex set of lenses and mirrors, giving the navigator the possibility to take bearings and measure distances while the submarine is still below the sea surface.

This instrument is only of tactical importance for submarines and not used on board other submersibles. Yet it can be used to collect information for updating the DR position and therefore reducing the dimensions of the POE.

3.4 System integration and software

3.4.1 System integration

The introduction of small microcomputers has had a major influence in the integration of navigation systems. All systems that have been described in this chapter each have their own advantages and disadvantages. By combining the results obtained from different systems not only redundancy is guaranteed, giving a better possibility to increase the predictable accuracy of a position fix; the reliability of the combination of systems is also increased since drifting or total failure of one of the systems is covered by one or more systems still working, making position fixing still possible.

By considering system integration we need to make a distinction between the following methods [Appleyard et al., 1988] :

1. Integrated position fixing : In this case raw position information (observables) from several systems and/or sensors is fed into a computer. An MPP is derived by combination of the data available at any time in the computer and is based on a mathematical model to achieve an optimum solution.

2. Hybrid position fixing : In this case two (or more) systems are available, each giving an independent position or unbiased observable. These positions or observables are compared with each other (integrity check) and combined to obtain an MPP.

3.3.6 Pressure sensor

Pressure sensors on board submersibles are used to measure the depth with respect to the sea-surface at which the vessel is situated.
The method of obtaining pressure using the pressure sensor is based on measuring the frequency of a precise quartz crystal resonator whose frequency of oscillation varies with pressure induced stress. A quartz crystal temperature signal is provided to thermally compensate the calculated pressure in order to achieve high accuracy over a broad range of temperatures.

The measured frequency (f) is converted into a pressure (P) using the following equation :

$$P = P_d + P_{ref} = \alpha \left| A\left(1 - \frac{f}{f_0}\right) - B\left(1 - \frac{f}{f_0}\right)^2 \right| \quad \text{[Pascal]} \tag{3.2}$$

where

- f_0 = theoretical frequency of crystal in vacuum provided by manufacturer (≈ 40 kHz)
- A,B = sensor dependent calibration coefficients provided by manufacturer (A ≈ 9750 psi, B ≈ 5000 psi)
- α = conversion factor to convert pressure given in psi to pressure in Pa (α = 6894.757)
- P_d = pressure measured due to water column
- P_{ref} = pressure measured due to atmospheric pressure

Using the calculated pressure obtained from equation (3.2), the depth (d) can be calculated using

$$d = \frac{\beta}{\gamma} [P - P_{ref}] \quad \text{[metres]} \tag{3.3}$$

where

- β = conversion factor to convert pressure in Pa to metres water column with a specific weight of 1000 kg / m^3 (β = 101.9716 10^{-6})
- γ = conversion factor to allow for specific weight of seawater (γ ≈ 1.026)

The depth information provided can be used in combination with the depth calculated using the inclination angle (provided as attitude angle from SINS) in combination with measured velocity

- timing mechanism : is used to measure the time between start of transmission of a pulse and reception of received pulse.

- recorder : The recorder has a broad strip of paper moving slowly over a flat metal surface. A belt on which one or more styluses are fastened runs over two pulleys, driven at a constant speed by an electric motor. When the receiver supplies an electric voltage to the stylus, the upper layer of the paper is burnt away, so leaving a depth trace on the paper. In order to be able to use this information later on, a transmission mark is given at the top of the paper. With modern echo sounders, the depth is also represented digitally.

Depending on the environment in which the echo sounder is used, a transmitting frequency has to be chosen. Low frequencies will transmit energy efficiently over long distances because the power will not be rapidly reduced by attenuation. Although high frequencies are prone to greater attenuation and therefore less range, the pulse duration can be shorter making higher resolution possible. Therefore the following classification can be made :

- shallow water E/S : These echo sounders use high frequencies of about 200 kHz. The pulse will not penetrate sediment and has a high resolution. The E/S can be used in waters up to 100 meters.

- medium depth E/S : These echo sounders, which are used in waters of 100 - 1000 meters, use frequencies of about 30 kHz. These frequencies penetrate soft sediment and have medium resolution.

- deep sea E/S : These echo sounders use frequencies of about 10 kHz. They are used in waters deeper then 1000m. At these frequencies, the pulses penetrate sediments quite deeply, giving echoes of underlying layers. Resolution is completely lost.

In all cases the precision of depth measurements is highly dependable on the knowledge of the actual speed of sound in seawater. In shallow waters this can be obtained rather easily by measurements. However, when navigating in deep waters, especially in the oceans, the speed of sound is initially set to 1500 m/s and tables are used to correct measured depth to actual depth.

For a good working order of the gyro compass, it is important to give the compass system correct speed and latitude information, as this is used for counteracting drift and tilt.

3.3.4 Inclinometer

The inclinometer is a device giving the direction of the main axis of the vessel with respect to the horizontal plane. The most reliable way to get information about the inclination of the submarine is to make use of SINS. Attitude angles can be obtained from the synchro packs attached to each of the gimbals. By comparing the already known angle with the synchro value, an error signal can be produced, which can be used to update the known value.

3.3.5 Echo sounder

The echo sounder (E/S) is simply a sonar with a vertical axis. A pulse of acoustic energy is projected from a transmitter to the seabed and reception of the reflected pulse is measured so that range is derived by multiplying the assumed speed of sound in seawater by half the time measured between transmission and reception. This is represented to the user a on paper trace and/or digital equipment.

The main parts of an echo sounder are :

- pulse generator : generates pulses of electrical energy to be transmitted
- switching unit : connects the output from the pulse generator to the transmitter at the right moment. Then switches back to the receiver.
- transmitter : a piezo-electric transducer that converts the electrical power provided by the pulse generator into an acoustic pulse
- receiver : a piezo-electric transducer that receives the reflected pulses and converts them into an electrical signal
- amplifier : The pre-amplifier boosts the very weak signal to the strength needed to activate the recording system. A power amplifier is used to produce enough power to mark reception of the signal on paper

The main part of the log consists of a coil inserted in a watertight flow probe which is fixed underwater on the outside of the vessel's hull. Two types of the probe are available, one is hull mounted ('Flush' mounted log) and the other is retractable through a sea valve. The probe has a streamlined shape in order to minimize the effect of water being dragged with the ship. This has to be done because the ship's speed is measured relative to the water. The water surrounding the flow sensor acts as the conductor forming the loop and the magnetic field in the coil induces a voltage in the water. This voltage is proportional to the speed of water along the coil, which is essentially the same as the speed of the ship through the water. An electrode, fitted on the probe, picks up the induced voltage and passes it to a measuring device. This device transforms the voltage into speed.

3.3.3 Gyro compass

The principle of the gyro compass is based on a fast rotating gyroscope, whose axis of rotation maintains its direction in space. This means that the direction of the spin axis will change with respect to an earth fixed reference frame due to earth rotation. This is given by two parameters : drift, which is the angular rotation of the spin axis round a local vertical axis (i.e. azimuth of axis), and tilt, the rotation of the spin axis around a local horizontal axis (i.e. inclination of axis). Because of these changes of spin axis with respect to the earth fixed reference frame, it is not possible to use the gyroscope on its own as compass. By counteracting the two disturbances, the spin axis is made 'North seeking'.

The main parts of a gyrocompass system are :

- gyroscope : a specially constructed rotor;
- damping system : a construction exerting forces on the gyroscope to counteract the movement of its axis in the earth centred system, therefore making the gyroscope 'North seeking'. Two damping systems are used : horizontal damping and vertical damping;
- gimbals in which the gyroscope is mounted, giving it three degrees of freedom;
- compass housing;
- corrector mechanism : used to eliminate both the damping error and course, latitude and speed error (see section 3.3 of chapter 4)

In figure 3.4a the basic set up of the accelerometers and gyros in shown while in figure 3.4b the basic principle of deriving speed and distance travelled is shown.

The system as described so far will work when the vessel is stationary on a non-rotating earth. In order to be able to use the system on earth a few extra corrections have to be applied to the platform :

- the platform has to be rotated around an axis parallel to the earth rotation axis to compensate for the earth's rotation and ship's motion in E/W direction;
- the platform has to be rotated around the local E/W-axis to compensate for the ship's motion in N/S direction;
- the platform has to be rotated around the local vertical axis to compensate for convergence of the earth meridians.

The corrections that need to be applied are found by transforming velocities obtained into rotation angles.

As opposed to the closed loops of the horizontal channels, the vertical channel is an open loop. This means that any accelerometer errors are unbounded and will increase with time to a square law. Therefore, a displacement in vertical direction given by the vertical axis of SINS is combined with information provided by the pressure sensor (section 3.3.6). These systems are largely complementary to each other. The SINS vertical channel needs to be bounded by external reference (provided by pressure sensor depth) when used over longer periods, but provides direct information about vertical accelerations and good reference for use during short periods of diving or climbing. The pressure sensor on the other hand, provides good depth information when the submarine is sailing at a nearly horizontal level over longer periods, but is less accurate during short dives or climbs.

3.3.2 Electromagnetic log

The working of the electromagnetic (EM) log is based on the Maxwell-Faraday induction law : if a conductor is moved through a magnetic field, an electric force (E) is induced in the conductor, having its direction at right angles to both the magnetic field (B) and velocity (v) :

$$E = v \times B \qquad (3.1)$$

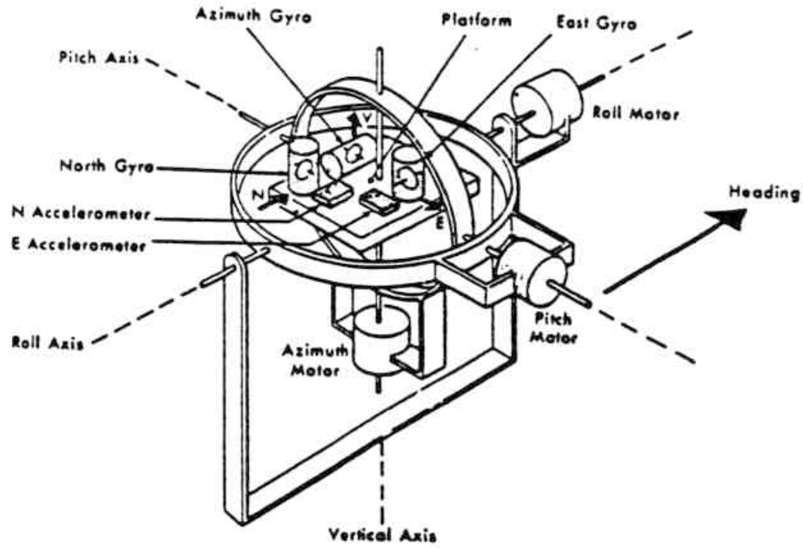


Figure 3.4a SINS platform arrangement.

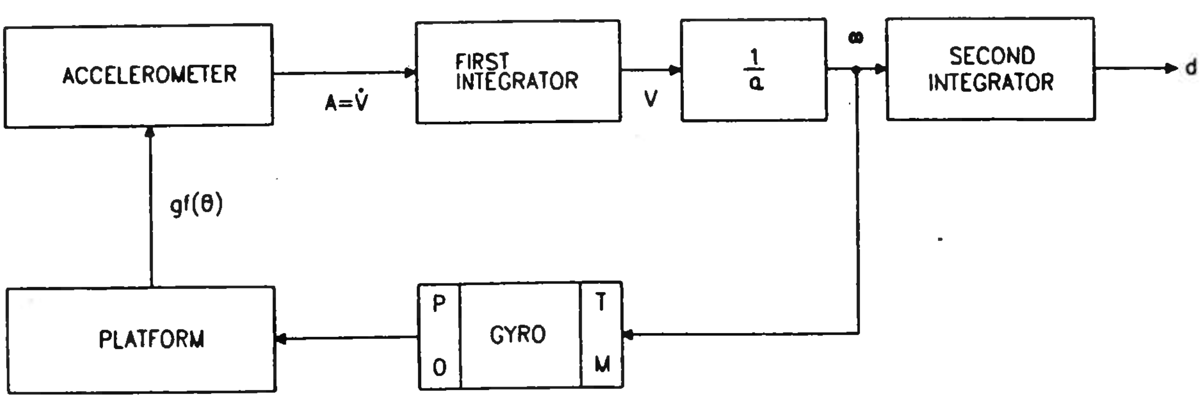


Figure 3.4b Basic principle position calculation using a one channel system.

3.3 Ships bounded and unbounded systems and sensors

In this section all remaining sensors and systems on board a submarine that can be used for position fixing, whether it gives an MPP based on DR or external information, are discussed. We consider two main groups of sensors and systems : the bounded and unbounded. Bounded means that no external equipment providing information is needed for the sensor or system to give its information - i.e. they are self-contained (autonomous) -, whereas unbounded means that the sensor or system needs information from the outside. This classification implies that based on bounded systems and/or sensors only, a DR position will be derived.
All sensors and systems described in this subsection can be used both when the submarine is at the sea surface and when it is submerged.

3.3.1 Ships Inertial Navigation System

The Ships Inertial Navigation System (SINS), is a self contained DR device. It employs an Inertial Measurement Unit (IMU) with three gyros and three accelerometers to provide continuous output of the following data :

- ship's geographic position;
- horizontal an vertical linear velocity components;
- angle magnitude and rate for heading, roll and pitch.

The accelerometers are positioned on a platform in the three main directions so forming an XYZ - 3D coordinate system. By integrating accelerations measured, both speed and displacement in the three main directions can be calculated. In order to be able to use this information to calculate the present position, the orientation of the system of accelerometers needs also to be fixed with respect to a terrestrial reference frame. This is done by aligning one accelerometer with the East/West direction, one with the North/South direction and one in the vertical plane at right angles to the other two.
To yield useful information, the platform on which the accelerometers are mounted, must be constantly maintained in a known orientation, which is chosen to be horizontal, i.e. at right angles to the direction of gravity at any point. This can be achieved by suspending the platform in a gimballed system and using gyros which are also fixed on the platform. Any deviation of the platform from the horizontal will produce an angular rate which is used in conjunction with additional terms provided by the computer, to torque one or more gyros and drive the platform to the horizontal level. This way a closed undamped loop is formed, whose period is equal to the Schuler period.

3.2.3 Radar

Radar is a user-based microwave transmitter/receiver with a rotating antenna transmitting short microwave pulses in all directions sequentially and receiving the echoes from its own pulses reflected by surrounding obstacles. All the reflected signals are processed and displayed on a radar screen with their correct individual range and bearing.
The radar equipment is operated at the user's vessel.

- signal characteristics : The two main frequency bands used for marine radar are 2920 - 3100 MHz and 9320 - 9500 MHz.

- availability : This depends on the position of the vessel with respect to objects that can be used for navigational purposes.

- coverage : is practically limited by the characteristics of the radar installation (transmitted power, radio horizon and receiver sensibility) and is normally up to 45 km (25 nm) for shipborne radars, but can be as high as 75 km (40 nm) under good conditions.

- fix rate and dimensions : Once objects that can be used for navigation are present, radar can be used continuously for position fixing. Normally the antenna makes 20 - 30 turns per minute continuously updating the display. From each object bearing and distance can be measured giving a LOP.

  LOPs obtained from two or more different objects provide a 2D position.

- ambiguity : There is in general no ambiguity. Only objects very close to the transmitter/receiver might give a second echo on the display, but in general this is not the case. Sea clutter and interference might give confusion, but proper processing of the incoming signals reduces this very much.

- accuracy : The accuracy of a position derived from radar information depends on the distance of the object from the transmitter/receiver. For one object the predictable accuracy of a bearing is approximately 1° (1σ) whereas the accuracy of a distance measurement is about 0.01 (1σ) part of the range scale selected. The predictable accuracy of a position fix depends on the geometry of the navigational objects used.

series of the master station is called the Group Repetition Interval (GRI) of the chain. This GRI is made unique for each chain, making chain identification possible.
The system is based upon measurement of the difference in time of arrival (TOA) between pulses from the master station and the secondaries. The measurements of Time Difference (TD) are made by a receiver which achieves high accuracy by comparing the zero crossing of a specified cycle - the so-called 'time reference point' (TRP) - within the pulse transmitted by the master and the TRP in the pulses from the secondary stations within the chain. The TRP is obtained in two stages : first the proper cycle to be used for zero-crossing tracking is identified making use of the pulse envelope (Cycle Identification or coarse measurement). Once the proper cycle has been found the exact moment of zero crossing within the cycle is determined by means of a phase tracking loop (fine measurement), triggering a timing mechanism. The TRP is chosen to be the zero crossing in the third cycle because at that point the signal has adequate amplitude and the pulse will not yet be affected by sky wave interference.

- signal characteristics : All Loran-C stations transmit at 100 kHz using time-division multiplexing.
  Each pulse has a well defined envelope. Because of the short rise-time of the pulses, their transmissions occupy a relatively broad frequency band, 99 percent of their energy being distributed between 90 and 110 kHz.
  Additionally, the pulses transmitted by the stations are phase coded. Master and secondary stations each have a different phase code functions.

- availability : The Loran-C transmitting equipment is very reliable. Redundant transmitting equipment is used to reduce system downtime, making availability better than 99%.

- coverage : as shown in figure 3.3.

- fix rate and dimensions : Once the receiver is in the coverage area, 10 to 20 independent position fixes can be made per second depending on the GRI used. Two or more LOPs can be obtained, giving a 2D fix.

- ambiguity : As with all hyperbolic systems, ambiguity is present. However, because of the design of the coverage area of each chain, the ambiguous fix is at a great distance from the desired fix and therefore easily resolved.

- accuracy : Within the coverage area, Loran-C will provide the user with predictable accuracy of less than 450 m (2 d_{rms}). The accuracy is, however, highly dependent on the GDOP at user's location.

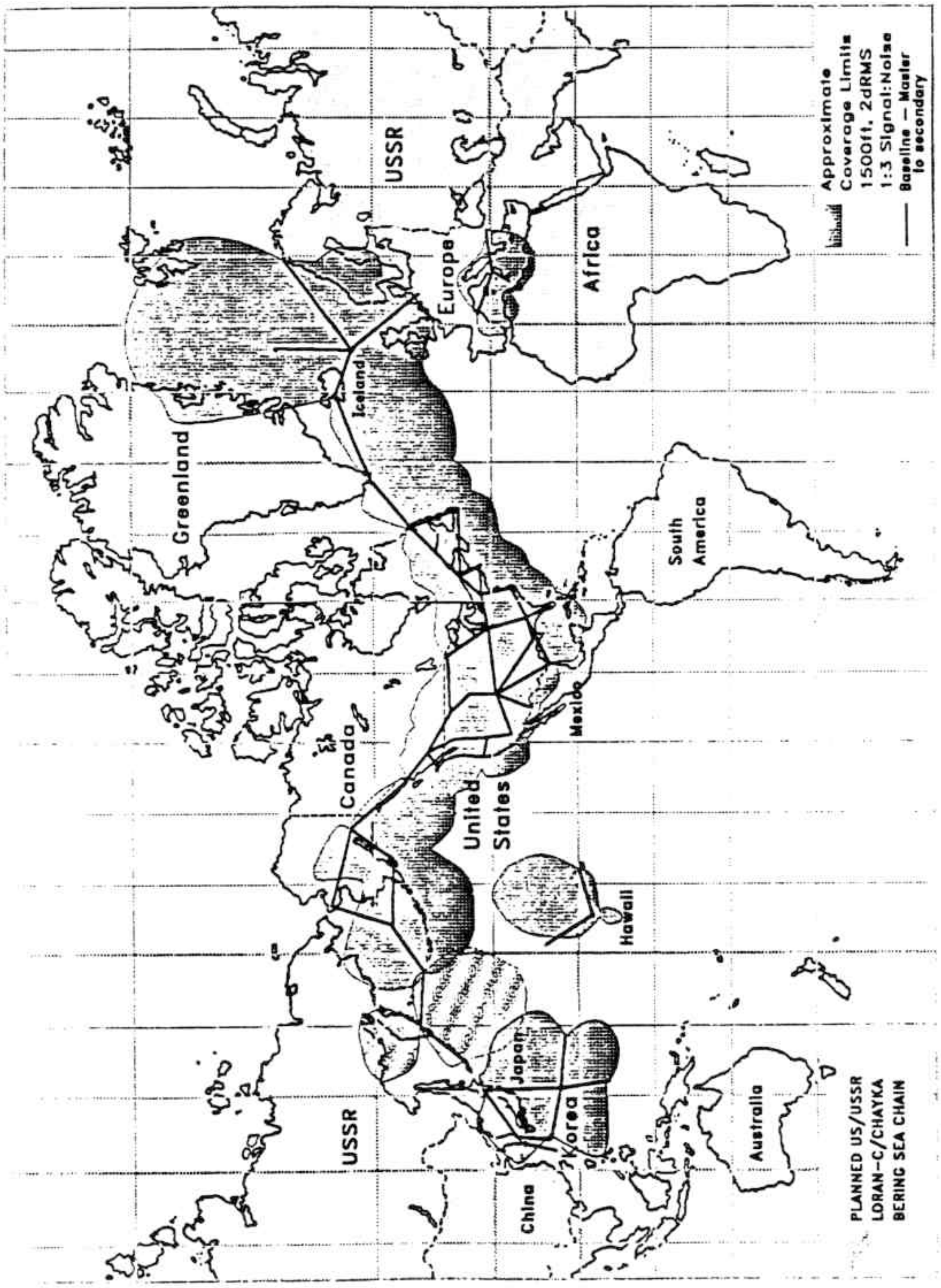


Figure 3.3 Coverage provided by U.S. operated or supported Loran - C stations [FRP,1990]

- signal characteristics : Each stations transmit signals on four frequencies in the following order: 10.2 kHz, 13.6 kHz, 11 1/3 kHz and 11.05 kHz. In addition to these four common frequencies, each station transmits a unique frequency to aid station identification and to enhance receiver performance. A combination of time- and frequency sharing is used to distinguish between signals from the different transmitters. The signal transmission format has a cycle of ten seconds.

- availability : Annual system availability has been greater than 97% with scheduled off-air time included. Equipment redundancy has been designed into nearly all functions of the Omega transmission process, which will contribute to station reliability and availability.

- coverage : Essentially worldwide coverage.

- fix rate and dimensions : Omega provides an independent position fix once every 10 seconds. Two or more LOPs can be obtained giving a 2D position fix.

- ambiguity : When the system is used in hyperbolic mode, cycle ambiguity is present. Single frequency receivers use the 10.2 kHz signal whose lanewidth is about 14.8 km (8 nm) on the baseline between stations. Therefore the EP needs to be known to within 7.4 km (4 nm). Multiple frequency receivers, however, extend the lanewidth for purpose of resolving lane ambiguity.

- accuracy : The accuracy of the Omega system is limited by the accuracy of the propagation corrections that must be applied to the individual lane readings. The system provides a predictable accuracy of 4 - 7 km (2 - 4 nm) (2 d_{rms}). The accuracy depends very much on receiver location, station pairs used, time of day and validity of propagation corrections.

LORAN-C

Each Loran-C chain consists of a master station and up to six secondary stations. Each station transmits a series of energy bursts at a carrier frequency of 100 kHz. The shape of the bursts is well defined. The master station transmits first after which the secondaries each transmit in turn starting after tightly controlled time intervals. This interval for each secondary is known as the Emission Delay (ED). When all stations of a chain have transmitted their series of bursts, the master station transmits again. The time interval between the start of two burst

each of the stations at 10.2 kHz. Since the stations are widely spaced, each station works autonomously. In order to be able to use time-sharing, each station has four Cs-standard atomic clocks making it possible to maintain synchronization. Since the phases are synchronized, the measurements may either be taken in pairs to give hyperbolic LOPs (hyperbolic mode), or may be taken with respect to a precision time source in the receiver, to give circular LOPs (rho-rho mode).
No chains are defined. At any point on the surface of the earth at least five LOPs will be available, allowing the navigator to take advantage of LOP redundancy. For each part of the earth an optimal set of stations is given, based on geometry of stations with respect to receiver and angle of cut of LOPs derived.

Although the standard deviation of an Omega LOP is large (see section 4.2.2) compared to other EPF systems available, making the Omega system less useful as stand alone system or integrated with other navigation systems especially now GPS has become operational, one of its advantages is that underwater reception of signals might be possible because of the very low frequencies involved. However, no information on the way this can be used on board submarines of the Royal Netherlands Navy is given in this paper due to operational classification.

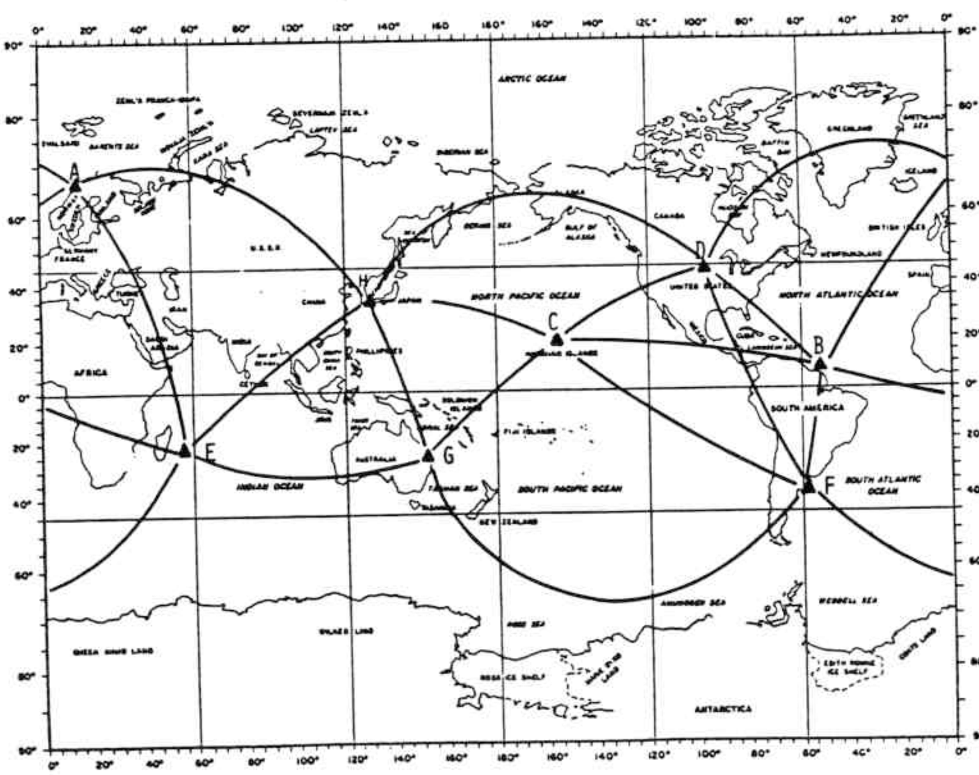


Figure 3.2 Omega System Configuration [IHR,1972]

sky wave interference (1-hop-E layer sky wave distance). Beyond this distance, a fix from MP readings is about 20 % more accurate than the fine pattern fix [Decca,1979]. This is because the MP signal has the important property that it remains stable in phase in the presence of mutual phase shifts in its constituents. As the range further increases, there comes a point at which the fine pattern becomes unreliable and the MP signals are then the sole means of fixing. The limit of night-time coverage is set by phase shifts in the MP constituents due to sky wave interference, introducing unwanted peaks in the MP signals. This is expected to occur at distances beyond 440 km (240 nm) from any station.

- fix rate and dimensions : Decca provides continuously two or three LOPs, making continuous position fixing possible. If Multi Pulse is used for position fixing, only once every 20 seconds a fix can be made.
Having two or three LOPs provides a 2D position fix.

- ambiguity : As phase measurements are performed on carrier waves, cycle ambiguity is present. The area between lines of zero phase difference is called 'lane'. To avoid the fact that the position of the submarine needs to be known to within a few hundreds of meters (since the lanewidth of patterns on the baseline ranges from approximately 350 - 600 m), a coarse pattern is created via MP signals. The lanewidth of this coarse pattern is approximately 10.6 km on the baseline (which equals a zonewidth on the baseline), making it possible to resolve cycle ambiguity within a zone. The position now needs to be known to within half a zone-width.

- accuracy : The predictable accuracy varies from 50 m ($2d_{rms}$) at approximately 80 km from the master to 450 m ($2d_{rms}$) up to 250 km from the master. Due to sky wave interference this accuracy reduces at nighttime by a factor 6 to 8.
Since the predictable accuracy is dependent on many factors, one needs to refer to the Decca Chain Data Sheets to find the most likely value for predictable accuracy at a given position for a certain time and even date.

OMEGA

Omega is a very low frequency (VLF) radio navigation system, comprising eight transmitting stations situated throughout the world. Worldwide coverage was achieved when the station in Australia became operational.
The basic measurement in Omega is the phase of the signal from

3.2.2 Terrestrial LF and VLF Radio Position Fixing Systems

In this section the most frequently used LF and VLF Radio Position Fixing (RPF) systems will be viewed at. They can only be used when the submarine is at periscope depth or at the surface.

Decca Navigator System

The Decca Navigator System (DNS) is a hyperbolic radio navigation system. A Decca chain consists of one master and two or three slaves, designated Red, Green and Purple. LOPs are formed by phase comparison between the signals received from the master and slaves.

- signal characteristics : The DNS utilizes unmodulated signals in the 70 - 130 kHz band. The basic frequency f_0 is about 14.2 kHz. The value of f_0 varies slightly from chain to chain. The master transmits at a frequency of $6f_0$, the slaves at $8f_0$ (Red), $9f_0$ (Green) and $5f_0$ (Purple). Phase comparison between master and slave is done at $24f_0$ (Red), $18f_0$ (Green) and $30f_0$ (Purple), resulting in a fine pattern. The necessary phase lock between the master and slave transmissions is ensured by the control equipment at the slave station. Generally, the phase of the slave signal with respect to the master signal is so adjusted that the baseline extension at the master station has the fraction value zero. At the slave station the baseline extension has the value of the residual lane fraction if the baseline is not equal to a whole number of lanes. Several chains depart from this convention, however.

  Once every 20 seconds lane identification signals, known as Multi Pulse, are transmitted from each station in turn so that the receiver can extract a signal of frequency f_0 from the master and each slave. Comparing the phase of these signals generates a coarse pattern. An additional phase difference meter in the receiver responds to this coarse pattern and gives periodic readings which indicate, in turn, the correct lane of each pattern within a known zone.

- availability : Taking downtime of the chains in consideration, the availability will be approximately 99.8%.

- coverage : At present day (December 1992) 42 chains are in use worldwide. A chain can be used up to 750 km (400 nm) from any station during daylight. At night, the accuracy of phase comparison on the fine pattern degrades at distances over approximately 200 km (110 nm) from any station due to

The main characteristics of the Transit system are as follows:

- signal characteristics : The satellites emit continuously at two frequencies : 399.968 MHz and 149.988 MHz. In practice these frequencies will vary slightly from satellite to satellite and drift with time. Because each satellite has only one oscillator, the lower frequency can be kept always 3/8 times the higher. The carrier frequencies are phase modulated in order to carry the satellite's broadcast ephemeris and to provide a specific time-marker every even numbered minute of UTC.
  The transmitted signals are right-hand circularly polarised.

- availability : the availability is better than 99% when a Transit satellite is in view. Being 'in view' depends on user latitude, antenna mask angle, user manoeuvres during satellite pas, number of operational satellites and satellite configuration.

- coverage : Coverage is worldwide but not continuously due to the relatively low altitude of the Transit satellites and the precession of the satellite orbits.

- fix rate and dimensions : Once a satellite is in view, it is visible for up to 18 minutes which gives ample information for approximately 40 positions during perfect satellite pass. The satellite wait time varies with latitude, theoretically from an average of 110 minutes at the equator to an average of 30 minutes at 80 degrees latitude. Due to non-uniform orbital precession, the Transit satellites are no longer in evenly spaced orbits. Consequently, a user can occasionally expect a period greater than 6 hours between fixes.

  Transit provides 2D positioning.

- ambiguity : there is no ambiguity

- accuracy : The predictable accuracy is highly dependent on the user's knowledge of his velocity and course. On average one gets :

  - dual frequency : 100 - 350m ($2d_{rms}$)
  - single frequency : 200 - 500m ($2d_{rms}$)

NNSS / TRANSIT

Transit is a space-based radio-navigation system consisting of three major segments :

· Space segment

There are 9 satellites, from which only 5 are operational, in approximately 1100 km polar orbits, having a period of orbit of approximately 105 minutes. The spacing of the orbits in azimuth is not uniform due to precessional effects.
Each satellite carries a radio receiver and transmitter along with a small computer, stable oscillator and crystal clock.

· Ground segment

The satellites are tracked by four main stations, forming OPNET. They provide the tracking information necessary to update satellite orbital parameters every 12 hours.
Additionally there is a number of stations all over the world tracking the satellites when in view. They form TRANET. Their tracking results are send to the U.S. DMA, which computes the 'precise ephemeris' for the satellites. A user tracking a certain satellite can use this 'precise ephemeris' for post processing rather than the real-time broadcast ephemeris.

· User segment

The user segment consists of an antenna, preamplifier, a receiver with microprocessor and power supply.
The receiver measures successive Doppler shifts as a satellite passes the user. A typical measurement interval is either approximately 30 seconds or two minutes. The Magnavox 1502 receiver, for example, uses 4 batches of 5 words (approximately 23 sec each) and the two minute marker as time gates.
The microprocessor computes the satellite's position at the beginning and the end of each Doppler count interval using the broadcast ephemeris data. It is also fed with the ship's estimated position, ground course and speed, enabling the receiver to estimate the slant ranges and differences over the time interval. The receiver then calculates the geographical position of the user based on knowledge of the satellite's position and range difference from the measured Doppler count in eg. a least squares position estimation process, either as part of an integrated or stand alone system.

ephemeris data, atmospheric propagation correction data and satellite clock bias information.
The transmitted signals are right-hand circularly polarised.

- availability : expected to approach 100% when fully operational.

- coverage : Fully operational worldwide 3D coverage will be provided. At the moment (December '92) 2D coverage is available 24 hours a day whereas 3D coverage is available over 23 hours per day.

- fix rate and dimensions : When a receiver is switched on, some time will elapse before a first position can be calculated, due to the fact that the receiver needs to acquire the satellite signals and navigation data. This so-called Time to First Fix (TTFF) can range from 30 seconds for a fully operational set to 25 minutes for a cold start using a one-channel receiver. Figure 3.1 gives an overview of TTFF that can be expected under various conditions. Once the receiver is locked on, a position, velocity and time solution will be provided approximately every second, depending on the receiver used.

  GPS provides 3D positioning and velocity fixes as well as accurate time information when fully operational.

- ambiguity : there is no ambiguity

- accuracy : Two levels of navigation are provided by the GPS, the Precise Positioning Service (PPS) and the Standard Positioning service (SPS). The PPS is a highly accurate positioning, velocity and timing service which is only made available to authorized users. The SPS is a less accurate position and timing service which is available to all GPS users.
  When the system is fully operational, the predictable accuracy will be [FRP,1990] :

	SPS	PPS
horizontal (2 d_{rms})	100 m	17.8 m
vertical (2σ)	156 m	27.7 m
time (1σ)	167 ns	100 ns

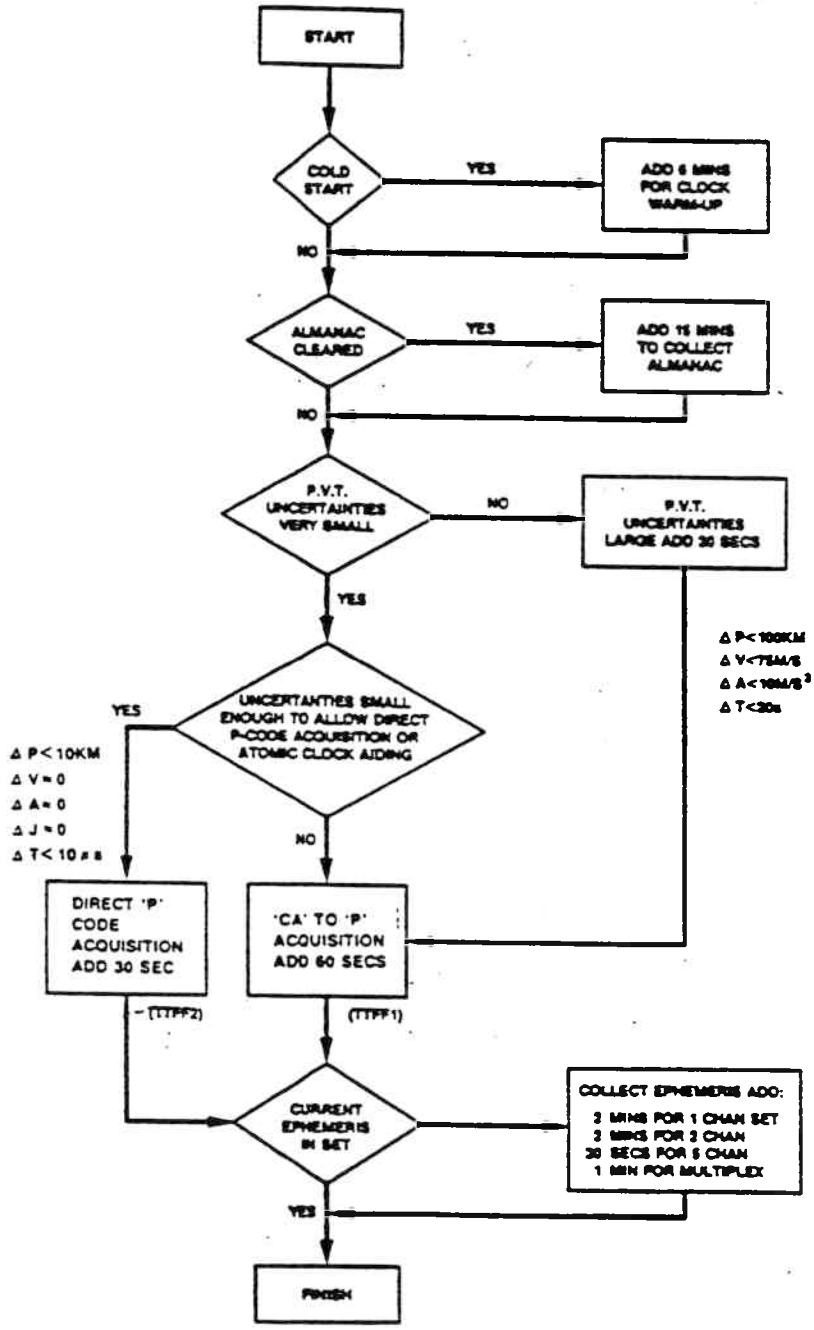


Figure 3.1 Block diagram for estimation of Time To First Fix [NATO,1991:1]

of 20200 km. The orbits have an inclination angle of 55° relative to the equator and are separated by 60° in azimuth. The period of orbit of the satellites is about 12 hours. Spacing of the satellites in their orbit will be in such a way that, when the system is fully operational, the user will have at least 4 satellites in view at any time and any place.

- Control segment

The control segment comprises five monitor stations, three ground antennas and a Master Control Station (MCS). Each monitor station is a remote unmanned station consisting of a GPS receiver, an atomic frequency standard, communication equipment and environmental sensors. The receivers track all satellites in view, accumulating range data. At the same time the clock signals are observed and compared to the station frequency standard. Local meteorological data is also measured. All this information is transmitted to the MCS where it is processed to determine and predict ephemeris data and clock-bias for each satellite. The navigation message of each satellite is updated by transmitting this information to the satellite via the Earth uplink antennas. This is done every second orbit or once about every 24 hours.
In addition to the monitor stations, the GPS satellites are also tracked by a number of semi-permanent tracking stations.

- User segment

The user segment consists of an antenna, receiver, processor and I/O devices. The receiver demodulates the navigation signals to obtain the pseudo-range and delta pseudo-range measurements. The microprocessor converts these measurements to a position, velocity and time.
Various receivers are available commercially, ranging from one channel, single-frequency (C/A-code) receivers to multi-channel, dual-frequency (C/A-code plus P-code) receivers.

The main characteristics of the GPS system are as follows :

- signal characteristics : Each satellite continuously transmits at two L band frequencies : L1 (1575.42 MHz) and L2 (1227.60 MHz). Using Bi-Phase Shift Keying (BPSK), so-called Pseudo Random Noise (PRN) codes are modulated on the carrier frequencies. The L1 carrier is modulated with a precise (P) code plus a coarse/ acquisition (C/A) code whereas L2 is only modulated with the P-code. The resulting frequency spectrum for the carrier, due to BPSK, equals 20 MHz for the P-code and 2 MHz for the C/A code. The carrier frequency is suppressed.
  Both frequencies have a navigation data message superimposed. This navigation message contains satellite

3.2 Ships Electronic Position Fixing Systems

In this section a general description will be given on the electronic position fixing (EPF) systems used on board submarines. These systems can only be used when the submarine itself is at the sea surface or when it is able to have a receiving antenna at the surface. When one or more of these systems are in use, they can in most cases be used continuously for position updating. An important factor from a tactical point of view is of course the time needed for a first fix if the systems are only used for updating the DR position and decreasing the dimensions of the POE.

The group of EPF systems not only comprises the well known terrestrial position fixing systems such as Decca, Loran-C or Omega, but also satellite navigation systems and radar. In evaluating these systems it is important not only to describe the configuration of the system used, but also to look at characteristics such as

- signal acquisition and tracking continuity
- signal integrity
- coverage
- availability

Other important factors to be considered are :

- signal characteristics
- accuracy
- fix rate of independent LOPs or position fixes
- fix dimension (2D or 3D)
- ambiguity

3.2.1 Satellite Position Fixing Systems

<u>NAVSTAR / GPS</u>

GPS is a space-based position, velocity and time system that has three major segments :

- Space segment

When fully operational this segment will consist of 21 satellites plus three active spares, in six orbital planes. Up to date (December '92) 19 satellites are operational. The satellites operate in almost circular orbits at an altitude

3. Sensors, Systems and their Characteristics

3.1 Introduction

In this chapter an overview will be given of the sensors and systems available for collecting information to be used for position fixing and quality control. Since a wide range of sensors and systems is available, only those generally used on board submarines are discussed.

The chapter is divided into three sections, each section describing a particular group of sensors and systems used in one of the three stages of navigation, according to the concept of the POE. In the first section the systems that can only be used at the sea surface will be discussed. The next section gives all systems and sensors that can be used when the submarine is at the sea surface or submerged. Here a distinction is made between bounded and unbounded sensors. In the final section information will be given on the possible integration of the various sensors with each other and with a main computer. Also some remarks on software used for navigation will be made.

It is not the intention to go into great detail about the different systems as there are plenty of good textbooks describing them. The reason for including this chapter in the paper is not only because then the systems that are available will be named, but also because it will make the reader aware of the fact that since a submarine can be navigated either at the sea surface or under water, restrictions are posed on the availability of position fixing systems.

In this chapter only mean values for predictable accuracy will be given. For most systems, however, the standard error of data provided depends on many factors, making the predictable accuracy to vary in time under given conditions. Thus a more detailed analysis of errors contributing to the total standard error of a sensor or system is needed in order to be able to calculate predictable accuracy of a position fix at any time. This analysis will be performed in chapter 4.

considered to be of importance) and interference from other sources transmitting in the 50 - 150 kHz frequency band. Errors due to CWI are normally negligible, especially since the DNS makes use of exclusive frequency allocations and the receivers have a narrow input bandwidth. This narrow bandwidth also makes the DNS relatively immune to atmospheric noise.

3. Sky wave interference

Decca signals can reach the receiver either by the direct path between station and receiver, the so-called **ground** wave, or via reflections from the ionosphere, the **sky wave**. Interference of these two signals results in the Decca reading being subject to a random error. The magnitude of this error is a function of the sky wave signal strength with respect to that of the ground wave (SIR) and phase angle of sky wave relative to phase angle of ground wave:

$$\delta\varphi_{sky} = \arctan\left|\frac{\sin\alpha}{\frac{\hat{g}}{\hat{s}} + \cos\alpha}\right| \qquad (4.11)$$

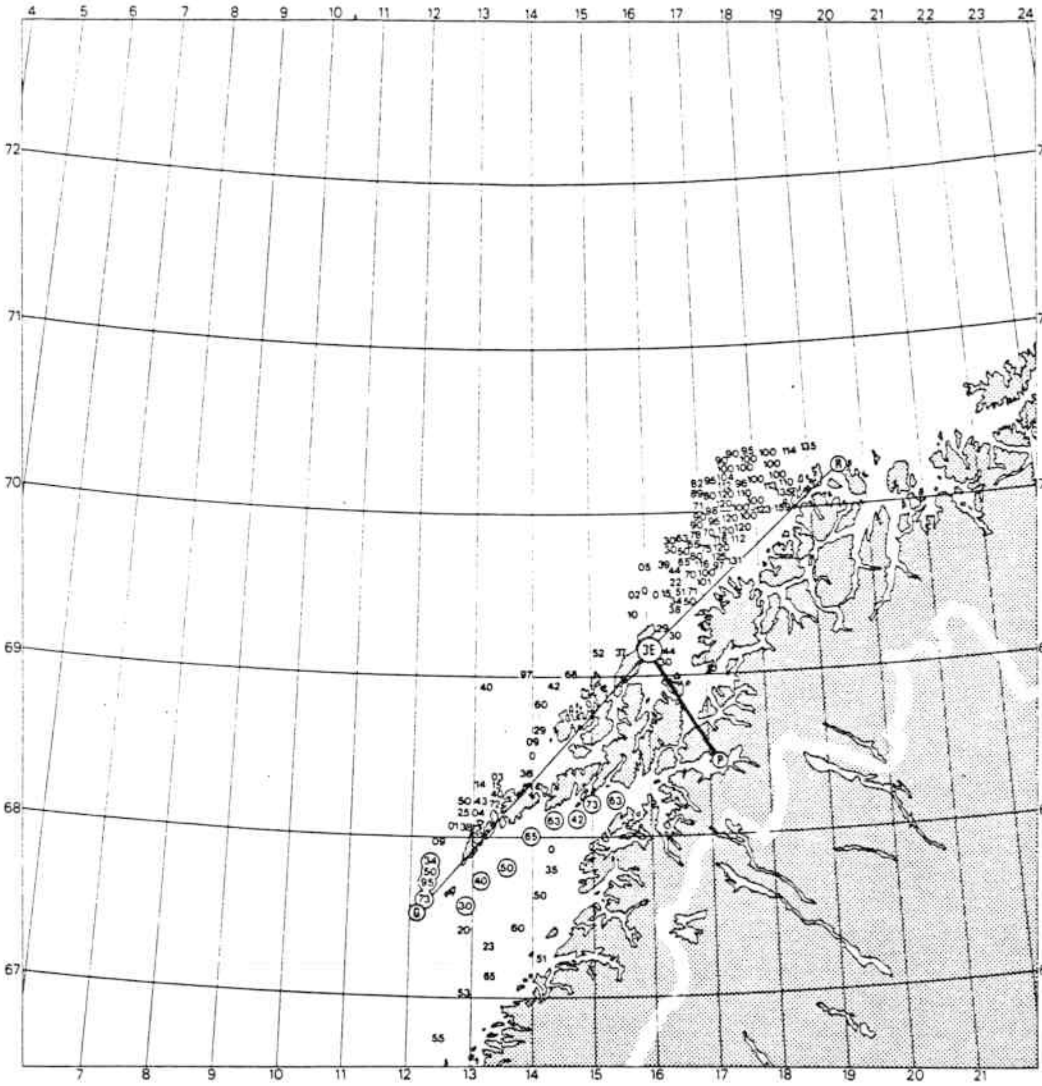
where

$\delta\varphi_{sky}$ = sky wave induced phase error
$\hat{s}$ = sky wave amplitude
$\hat{g}$ = ground wave amplitude
α = relative phase of sky wave with respect to ground wave

A maximum phase disturbance will arise when the two signals are almost in quadrature ($\alpha \approx 90°$) if $\hat{s} \ll \hat{g}$. If $\hat{s}$ becomes larger than $\hat{g}$, laneslips can occur.

The sky wave interference is more pronounced during nighttime than daytime and stronger in winter than in summer. To avoid these interferences from sky waves, the navigator should, if possible, not use a chain at ranges over approximately 750 km (400 nm) during daytime and 200 km (110 nm) at night from any station since sky wave interference is most likely to happen at these ranges and over (1-hop-E layer sky wave distance)[1]. The effects of

[1] The Decca Navigator Company Ltd. gives a SIR of 10 dB as approximate representative value for satisfactory receiver operations [Decca,1979]. Graphs showing the intensity of 100 kHz ground and sky-waves with increasing distance indicate that this SIR value is reached approximately at the ranges stated.



THE DECCA NAVIGATOR SYSTEM — LOFOTEN CHAIN (3E)

PURPLE PATTERN FIXED ERROR CORRECTIONS

CORRECTIONS TO APPLY TO OBSERVED PURPLE DECOMETER READINGS TO OVERCOME FIXED ERRORS
VALUES SHOWN ARE IN HUNDREDTHS OF A LANE UNITS
FIGURES ENCIRCLED SHOULD BE SUBTRACTED
FIGURES NOT ENCIRCLED SHOULD BE ADDED

Figure 4.6 Decca Navigator System - Lofoten Chain Purple pattern fixed error corrections [Decca Data Sheets] sky wave interference can be reduced using MP readings for position fixing in areas where sky wave interference is likely to occur since MP signals are less susceptible to sky wave interference (see section 2.2 of chapter 3).

The magnitude of errors due to sky wave interference are predictable within reasonable limits of confidence, based on statistical analysis of countless observations at fixed monitor stations in differing environments and at various ranges from the transmitters.
The error in phase of the signals under sky wave conditions arriving at the receiver station is considered to be due to the following phase deviations :

- phase error in the signal from Master to Slave ($\delta\varphi_{ms}$);
- synchronization error in the Slave ($\delta\varphi_{sync}$);
- phase error in the signal from Master to Receiver ($\delta\varphi_{m}$);
- phase error in the signal from Slave to Receiver ($\delta\varphi_{s}$).

For each pattern, the four components can be combined to give an expression for the total error of that pattern at the receiver [Decca,1979] :

- Red : $\delta\Delta\varphi_R = 4\delta\varphi_m - 4\delta\varphi_{mR} - 3\delta\varphi_{syncR} - 3\delta\varphi_R$
- Green : $\delta\Delta\varphi_G = 3\delta\varphi_m - 3\delta\varphi_{mG} - 2\delta\varphi_{syncG} - 2\delta\varphi_G$
- Purple : $\delta\Delta\varphi_P = 5\delta\varphi_m - 5\delta\varphi_{mP} - 6\delta\varphi_{syncP} - 6\delta\varphi_P$

This information can be used to calculate the variances and covariances of Decca LOPs. These are given in appendix 4.

B. Receiver errors

Differential phase errors in the signal channels and subsequent sections of the receiver are corrected by a 'reference' facility. The reference input forms a phase datum whereby the measured phase difference should be zero in the absence of errors. When a systematic error is present in the receiver circuits, the user restores the zero reading - and therefore minimizes the receiver error - by adjusting the phase shifter in the appropriate receiver channel. [Decca,1979].
This leaves the receiver with short-term phase changes, which can be considered random of nature. These errors are negligible with respect to the random errors due to sky wave interference and noise.

If the Decca receiver output is a geographic position or if the receiver is part of an integrated system, errors due to reading off and plotting do not exist. Large errors might be present in the position if the fixed error corrections (see below) have not been incorporated in the solution.

C. User errors

Errors that can be introduced by the user were already mentioned at the beginning of this section. Two of those are explained here in slightly more detail in case a receiver like the Mark 21 Decca receiver is used :

- setting the equipment : in this case the zone letter and / or whole lane number are set wrongly. Both errors should be regarded as gross errors.
  Setting the wrong zone letter becomes quite obvious when plotting positions in a chart because of the zonewidth. However, wrong setting of the lane number might not be so obvious, since the lanewidth of the patterns is only a few hundreds of metres.

- applying 'fixed error corrections' : when plotting Decca lattices on charts a mean signal propagation velocity of 299550 km/s is used. This way hyperbolas which are smooth will be plotted. The actual propagation velocity will deviate from this mean value, resulting in irregular lines of equal phase difference. To allow for the discrepancy between plotted lattices and LOPs based on the actual propagation velocity, corrections need to be applied to observed values. The corrections needed are called 'Fixed Error Corrections' and are given in the Decca Data Sheets. These corrections can be as large as several tenth of lanes in coastal areas. Figure 4.6 gives an example of extreme corrections needed on the purple pattern of the Lofoten Chain (3E).

If receiver errors and user errors are removed as described above and the 'Fixed Error Corrections' are applied to the Decca readings, the errors introduced by sky wave interference are the major source of error in a Decca LOP. Based on data provided by the Decca Navigator Company, the following typical values of standard deviations in fractions of lanes can be given[1] for Summer night conditions in temperate latitudes and for a propagation path over sea-water or good soil [Decca,1979] :

[1] The distance to both master and slave is taken to be the range given in the table.

range (km)	Red	Green	Purple
100	0.029	0.021	0.044
200	0.065	0.048	0.101
300	0.120	0.088	0.189
400	0.215	0.157	0.338
500	0.341	0.250	0.534

Table 4.3 Typical lane accuracies (1σ) in lanes

The conversion of standard error of a Decca LOP from lanes (σ_{LOP}) to metres (σ_m) is given by :

$$\sigma_m = \frac{1}{2}\lambda_c \operatorname{cosec}(\frac{1}{2}\gamma)\sqrt{\sigma^2_{LOP}} \quad \text{[metres]} \qquad (4.12)$$

where

γ = angle subtended by master station and slave at receiver position

OMEGA

Similar to the Decca receiver, the Omega receiver is a device that measures phase differences. Because of the long propagation paths between transmitting station and receiver and the low frequencies used, prediction of phase disturbances becomes more difficult. The Omega system error sources can be divided into four categories :

A. Predictable errors in assumed propagation model

This category of errors comprises those that are either fixed or can be forecasted to certain extent. Corrections that need to be applied to measurements to reduce the effect of these errors are stated in the Omega Propagation Correction Tables (OPCT). The wave propagation mathematical model on which the OPCT is based, is revised periodically to account for changes in solar activity and other propagation anomalies. The OPCT are updated every other year on average.

1. Diurnal ionospheric changes

   When a receiver is placed stationary at a point on the Earth surface, changes in phase can be observed when the incoming signals are observed. These changes are caused by variations in the height of the ionosphere during a 24 hour period, causing variations in propagation speed of radio waves, hence introducing the observed phase shifts. These diurnal changes are quite well known and as such listed as corrections to LOPs.

2. Geography

   Because of the large distances, waves paths between transmitter and receiver lie over a mixture of land and sea stretches, resulting in deviations from assumed propagation velocity (which is normally based on sea path only). The corrections can either be obtained by calibration or by calculation using mathematical models that make use of ground conductivity maps.

3. Icebound regions

   The Omega signals are severely attenuated in icebound regions, especially in the Arctic regions (eg. Greenland, parts of Iceland). The corrections given in the OPCT allow for propagation in these regions. Little data is available for these areas, making even the best estimates uncertain. In particular rather rapid phase changes with position may occur as one passes in the 'shadow' of the Greenland icecap [DMA,1981].

4. Modal interference

   Waves having low frequencies as used with the Omega system, propagate in a space confined by the Earths surface and the ionosphere. This space acts as a waveguide. In this waveguide different propagation modes exist, each having its own speed of propagation and attenuation rate. The different modes interfere with each other causing changes in phase angles that are very difficult to predict. This interference takes place at different parts of the transmission path between transmitter and receiver:

   - within 450 nm from the transmitter near field interference can be observed, causing such disturbances on the signal that it should not be used for position fixing;
   - when the path of the signal from transmitter to receiver passes through both day and night conditions, the twilight area will cause interference;
   - when signals cross the Earth magnetic equator, in

particular when the signals travel from East to West the interference is pronounced over a wide area.

B. Irregular errors in assumed propagation model

Forecasting based on long term observations only (on which the OPCT is based), gives the general aspect of phenomena with associating errors that can be expected and is therefore lacking in refinement, since certain sudden irregularities cannot be predicted sufficiently. Consequently, these irregularities lead to degradation of fix accuracy since no corrections can be promulgated in order to reduce their effect.

1. Diurnal ionospheric changes

Propagation conditions during daytime and nighttime are relatively stable, making predictions possible to great extend (< 0.05 lane for daytime and < 0.10 lanes for nighttime). Transition periods (sunrise and sunset) however, are of intermediate stability and present complications in prediction, resulting in considerable departures from predicted corrections, particularly near the end of the sunrise transitions.
It should be noted that even when day or night conditions are present at the receiver station, transition conditions can be present at some part of the propagation path between transmitter and receiver, still influencing the observations.

2. 'Long path' signal reception

Omega signals propagating from the transmitting stations to the magnetic east are attenuated less severely than signals propagating to the magnetic West. Because of this characteristic, an Omega receiver located well west of a station may receive a signal which has come the 'wrong way' round, that is over the long instead of the short path around the Earth. Since Omega lattices on charts and algorithms for position calculation are based on reception of signals using the shortest path from station to receiver, large errors can be introduced.

3. Solar propagation anomalies

Several Omega propagation anomalies result from solar activity. The main sources are :

- Sudden Ionospheric Disturbances (SID) caused by X-ray emitting solar flares;

- Polar Cap Absorption (PCA) caused by proton bombardment of the magnetic polar regions, leading to a concentration of high-energy particles in the region of the magnetic pole with the result that normal VLF transmission is disrupted. The effect of PCA may be to shift an LOP 11 - 15 km (6 - 8 nm) for a period of several days.

C. Receiver errors

For receivers used nowadays, the phase-tracking error is in the order of 0.01λ (depending on C/N_0 and loop bandwidth). The phase-tracking errors can increase due to course and/or speed changes.
If a receiver performs automatic transmitter identification, which is based on pattern matching, wrong identification of the transmitting stations will lead to laneslips.

D. User errors

Errors similar to those existing with the Decca system can be observed when using Omega. An extra error inherent to the Omega system is misidentification of stations. A wrong station identification when the receiver is manually synchronized with the signal transmission format will result in laneslips.

It can be concluded that total error in Omega LOPs obtained is mainly dependent on the validity of corrections promulgated (accuracy depending on propagation models available) and the means to react adequately on sudden disturbances (PCA and SID), and not due to synchronization of transmissions, noise or receiver performance.
Because of the unpredictable nature of errors present, no definite value for the total standard error of LOPs can be given. Under normal conditions (no PCA or SID present) the following approximated values can be used [Pierce,1965;Draaisma,1982] :

	σ_{LOP}
day path	4.9 μsec
transition path	8.7 μsec
night path	8.9 μsec

These values are transformed into standard errors in metres:

$$\sigma_n = (2 f_c \sigma_{LOP}) \frac{1}{2}\lambda_c \operatorname{cosec}(\frac{1}{2}\gamma) \quad \text{[metres]} \qquad (4.13)$$

In the Federal Radio Navigation Plan [1990] it is stated that "*in most cases the predictable accuracy is consistent with the 2 - 4 nm ($2d_{rms}$) system design goal*".
Field observations show that an error of 2 nm ($2d_{rms}$) during daytime and 4 - 5 nm ($2d_{rms}$) at night are good working values. Propagation anomalies can, however, cause large errors up to 10 nm lasting for days at a time.

A way to increase accuracy is by using Differential Omega. This system is based on continuously monitored Omega signals at monitor stations. By comparing at the station the received signal with a predicted signal, corrections can be calculated. These corrections can then be transmitted to users in the area. Using this technique, Omega predictable accuracy ($2d_{rms}$) can be increased to 450 m (0.25 nm) for receivers within 100 km (50 nm) from the monitor station and to 2 km (1 nm) at daytime and 3 km (1.5 nm) at nighttime when the receiver is 550 km (300 nm) from the monitor station.

LORAN-C

This system is based on measuring the difference in time of arrival (TOA) between pulses from the master- and a secondary station. At the receiver the measured time difference is given by :

$$\Delta t = \left| \frac{D_{ms}}{c_a} + CD + \frac{D_s}{c_a} \right| - \frac{D_m}{c_a} = \frac{(D_s - D_m)}{c_a} + \left| \frac{D_{ms}}{c_a} + CD \right| \quad \text{[sec]} \qquad (4.14)$$

where

- D_{ms} = distance between master station and secondary
- D_m = distance between master station and receiver
- D_s = distance between secondary station and receiver
- CD = secondary coding delay
- c_a = propagation velocity of radio waves

From this equation it can be seen that lines of equal time difference are geodetic hyperbolas with the master- and secondary station at the focal points.

Several factors are responsible for unwanted phase shifts in the pulses received at the receiver, thus causing errors in measured time differences. The main sources are :

A. System errors

1. 'Man made' interference

The Loran-C system suffers from two types of interfering signals [Beckmann,1992] :

- signals generated by the Loran-C system itself : these originate from a different chain operating close to the chain used for positioning. This type of interference is called Cross Rate Interference (CRI);
- all other man-made signals in the spectrum from 50 to 150 kHz. These signals are called Carrier Wave Interference (CWI).

The mentioned groups of interference signals can be subdivided into the following categories :

- synchronous interference signals;
- near-synchronous signals;
- a-synchronous signals.

Each of these three classes generates different problems for Loran-C receivers. The effect of these types of interference on position accuracy is subject of current research. The reader is suggested to Beckmann [1992] for detailed information on this subject.

It is assumed that the effect of CWI and CRI can be equated to an atmospheric noise level of 61 dBμV/m [Last,1992]. This way existing prediction methods, that take only atmospheric noise into account, can be used to calculate the coverage area for which position accuracy is stated below.

2. Sky wave interference

Apart from the above mentioned man-made interference signals, sky wave interference is also present with the Loran-C system. Depending on the distance between transmitting station and receiver, and on effective ionospheric height, sky waves can arrive at the receiver as early as 35 μsec and as late as 1000 μsec after the ground wave. In the first case, sky wave and ground wave

from the same pulse interfere with each other. To overcome this problem, only the first part of the ground wave pulse is used for measurement of time difference. In the second case, the sky wave of one pulse interferes with the ground wave of the succeeding pulse of the eight pulse transmission sequence. By using phase coding on the pulses, the receiver is able to discriminate between ground wave and sky wave of a composite pulse.
The above mentioned provisions make the Loran-C system essentially immune to sky wave signals at most times over most of its coverage area.

3. Inaccuracies in assumed propagation model

The propagation time needed for a signal to travel from a transmitter to receiver is taken as the time needed if the path had been entirely over sea water plus an 'Additional Secondary Factor' (ASF) to allow for land stretches along the path. The ASF values are published and are either entered manually into the receiver or stored in a ROM in the receiver. This way, the effects of land paths on LOP precision are reduced to the residual inaccuracies of ASF mapping and to the effects of seasonal variation in velocity of propagation.

4. Transmitter synchronization

All transmitting stations are equipped with Cs frequency standards. The high stability and accuracy of these standards permit each station to derive its own time of transmission without reference to another station. The objective for control of a Loran-C chain is to maintain constant the observed time difference (Δt) of each master-secondary pair throughout the coverage area. Frequency offsets in the cesium standards and changes in propagation conditions can cause fluctuations in the observed time differences. Therefore, one or more **Service Area Monitor (SAM)** stations with precise receiving equipment are established in the coverage area to monitor continuously the time differences of the master-secondary pairs. When the observed time difference varies from a control time difference (which is established during chain calibration) by one-half of the prescribed tolerance (typically 200 nsec or better) the SAM directs a change in the timing of the secondary station to remove the error.

B. Receiver errors

1. Filtering in the receiver

Because of its burst-type character, the spectrum of the Loran-C signal takes up frequencies roughly between 50 and 150 kHz. In this frequency band a lot of CWI-signals are present. However, the frequency band from 90 -110 kHz only contains Loran-C transmissions.
In order to get rid of interfering signals a band-pass filter is used in the receiver. The effect of this filter on the received Loran-C signal is that the modulation waveform is delayed (40 - 60 µsec !) and distorted : if a band-pass filter with steep slopes is used, most of the CWI signals are suppressed. However, the rising edge of the burst is lost making the sky wave rejection capability of the receiver difficult if not impossible. On the other hand, if the filter has gentle slopes, distortion of the modulation waveform is less severe leading to good sky wave suppression but allowing more CWI signals to distort the phase.
Good filter design is therefore very important and many research is done in developing and optimizing filters.

2. Cycle identification

This process determines the proper cycle of the Loran-C cycle to be used for time difference measurements. An error of one cycle results in a time range error to a station of 10 µsec.

3. Zero crossing phase tracking

Especially synchronous and near-synchronous interference can cause errors in range measurements as these interferences cannot be distinguished from a frequency shift due to a change in receiver position. The error introduced cannot be detected or removed by analysis and filtering of the tracking data.
Asynchronous signals with very large amplitudes can cause the zero-crossing loop to fall out of lock due to an apparent noise increase caused [Beckmann,1992].

C. User errors :

The same errors as mentioned with the DNS can be observed here.

On the accuracy to be expected using Loran-C for position fixing the following, based on Last [1992], can be said :

Improvements in receiver and chain control techniques have reduced resulting timing uncertainties, making them negligible in comparison with uncertainties due to SNR. Traditionally, the noise had been considered to be atmospheric in origin and random in nature. In NW Europe however, the dominant source of noise is CWI. The limit of coverage is assumed to be reached when the SNR (including CWI in NW Europe) at the sampling point falls to -10 dB, (this is equal to C/N_0 = 10 dB for a typical bandwidth of 50 Hz) resulting in a carrier tracking error due to noise of approximately 0.15 μsec (tracking loop bandwidth of 0.1 Hz).

The precision of a Loran-C LOP can be stated as :

	σ_{LOP}
tracking loop (noise)	0.15 μsec
ASF calculation	0.1 μsec
station synchronization	0 - 0.1 μsec
total error (1σ)	0.18 - 0.21 μsec

These errors can be converted to metres using (4.13).

The Federal Radionavigation Plan [1990] states that within the coverage area of a chain the predictable accuracy of a position has to be smaller than 463 m (0.25 nm; $2d_{rms}$).
At the edge of the area cycle matching is more difficult and can result in cycleslips. Each cycleslip gives a time error of 10 μsec to a station.

Field observations give that within the maximum range of the ground wave coverage (approximately 1500 km) predictable fix accuracies of 50 - 450 m ($2d_{rms}$) can be expected.

4.2.3 Radar

A radar can be used to observe both bearing and distance to an object. In both observations random errors will be present. First an overview of errors in bearings will be given , followed by the overview of errors in a distance measurement.

A. Bearings

The main error sources in a relative radar bearing are :

1. misalignment of heading marker with ships heading

This error can be minimised by carefully comparing the direction given by the heading marker against the ships heading given by a gyro compass. When the heading marker is aligned this way, the systematic error can be reduced to approximately 0.1°, provided only random errors are present in the gyro course.

2. 'own ship' positioning

If the position of the 'own ship' is not centred, an error will be introduced in observed bearings, depending on the direction and magnitude of displacement, and bearing. This error will not be constant in magnitude but change according to :

$$\delta\alpha = \frac{d\sin\alpha}{D - d\cos\alpha} \tag{4.15}$$

where

α = bearing to object relative to direction of displacement
D = distance to object
d = displacement

With most modern radars it is possible to automatically align the position of the 'own ship' with the centre of the screen.

3. Electronic Bearing Line (EBL) not positioned on 'own ship'

When taking bearings, this gives a displacement in the same way as with misalignment of the 'own ship' position.

4. size of object on screen

- the size of an object is increased on the screen as a result of beamwidth. This effect increases with increasing distance
- if the bearing to a physically large object is measured, it is important to know which part of the object is used. This error becomes less important as the distance to the object becomes larger.

5. quantization

   Normally a graduation circle is engraved around the border of the CRT. If this is used, an error can be introduced due to parallax. With modern radars, a digital bearing indicator is part of the radar system. This indicator is connected to the position of the synthetic radar cursor, giving the bearing of the cursor. Apart from wrong positioning of the cursor due to parallax, the indicator gives the bearing of a target to the nearest 0.1°.

6. user rounding off

   In most cases, when a radar bearing is plotted, the bearing is rounded off to the nearest one tenth of a degree.

The above mentioned errors make up the total error for a relative bearing. To obtain the absolute error, the total gyro compass error (see section 4.3.3) has to be included since the largest part of the total error depends on the error of the gyro compass. Considerable errors (up to a few degrees) can arise from many speed and/or course changes made in a small time interval.
The total error of an absolute radar bearing is given by [Draaisma et al.,1982;Lenart,1989] :

total gyro compass errors (1σ)	0.65°
misalignment heading marker	0.1°
cursor setting	0.5°
beam shape errors	0.05°
quantization	0.14°
total angle error (1σ)	0.85°

This error is converted to meters in a direction perpendicular to the bearing line by

$$\sigma_{m} = \sigma \frac{\pi}{180} d \quad [\text{metres}] \qquad (4.16)$$

where

d = distance to object in metres

B. Distances

The main sources, causing errors in distance measurements using a radar are :

1. time base errors

   At the instant the RF energy starts to leave the antenna, the electron beam in the CRT comes under the deflecting influence of a linearly increasing magnetic force. The deflective force is produced by the action of a sawtooth current waveform produced by the **timebase generator** which is synchronized to start the deflecting process at the onset of each transmitted pulse. Video pulses of received echoes are applied to the CRT so that the target appears. Any errors in synchronization will lead to errors is position of the target displayed on the CRT. When the radar circuitry is well adjusted, errors due to synchronization are negligible with respect to errors described below.

2. errors in Variable Range Marker (VRM)

   - accuracy of VRM : IALA [1990] states that the 95 percent radial error should be better than 1.5 percent of the maximum range of the scale in use or 70 metres whichever is greater.
   - correct setting of VRM on the contact since the size of the contact is given enlarged on the radar screen, due to pulse length

   Although these errors will be the same in magnitude on the radar screen, their absolute magnitude depends on the range scale used.

3. elevation of the object with respect to antenna height

   If the target is elevated with respect to antenna height, a slant range is measured instead of the horizontal range plotted or used in calculations. The error introduced is generally negligible, especially if the target is at great distance.

4. pulse shape

   - the size of an object is increased on the screen as a result of pulse length. This effect decreases on screen with larger range scale used
   - if the distance to a physically large object is measured, it is important that the point of the echo closest to the centre of the screen is used for measurement.

5. quantization

   The digital distance indicator is connected to VRM, giving the distance a target to the nearest 0.01 nm.

6. user round off error of distance to 0.1 nm.

The error budget of a radar distance is given by [Draaisma et al.,1982; Lenart,1989] :

accuracy VRM	10 - 35 m
positioning VRM	10 - 80 m
quantization	5 m
pulse shape errors	20 m
reading rounding error	55 m
total distance error (1σ)	60 - 105 m

4.3 The error budget of bounded and unbounded systems

4.3.1 Ships Inertial Navigation System

The correct working of the SINS is depending on the following factors introducing errors in the measured accelerations and therefore in the calculated velocities and distance travelled :

1. Timekeeping errors

   In order to obtain the distance travelled, acceleration have to be integrated twice with respect to time. The result of these integrations is also used to calculate the rotation angles needed to correct for rotations of the reference platform due to the ships's movement in N-S and/or E-W direction.

2. Dislevelment of the platform

   It is quite a difficult task to keep the platform at right angles to the local vertical at any time. Transient accelerations introduced by course and/or speed changes, roll, pitch and irregularities in the earth gravitational field can cause platform oscillations. If the platform is not completely horizontal, accelerations are introduced due to gravity. Since the accelerometers cannot make a distinction between accelerations resulting from ship movement and components of the Earth's gravity, both are combined in the calculations of the ship's speed and distance travelled. In order to reduce susceptibility to erroneous accelerations, the reference platform is 'Schüler-tuned'. The result is that errors do not increase in time but oscillate round a mean value.

3. Coriolis force

When the ship is under way, the platform is subject to the Coriolis force, introducing extra accelerations, which have to be corrected for. The magnitude and direction of the acceleration vector can be computed and errors introduced in calculation of velocity and distance travelled can be allowed for.

4. Accelerometer errors

The most important errors in modern inertial quality accelerometers are :

- bias : an output when no acceleration is applied;
- non-linearity : deviations from the least squares straight line for input-output relationships;
- threshold : minimum detectable change in accelerometer output;
- misalignment with respect to direction in which the accelerations are measured.

5. Gyro errors

Drift of the gyro's are caused by internal torques caused by mechanical wear, friction and mass imbalance will lead to a wrong 'horizontal' position of the reference frame.

Misalignment of gyros will lead to wrong stabilization signals, resulting dislevelment of the platform.

It can be shown [Draaisma,1986] that for a single channel a constant error in a measurement made by an accelerometer, an incorrect correction applied to allow for Coriolis force or an error in the gravitational force each will lead to a displacement error (δ_m) of

$$\delta_{\blacksquare}(t) = \frac{R}{g}\left[1 - \cos\left(\sqrt{\frac{g}{R}}\, t\right)\right]\delta a \qquad [\text{metres}] \tag{4.17}$$

where

R = earth radius
g = acceleration due to gravity
δa = error in acceleration measurement

whereas a constant gyro drift or an incorrect correction applied to allow for earth rotation or ship's speed in E/W direction will each lead to an error of

$$\delta_{\blacksquare}(t) = \left[t - \sqrt{\frac{R}{g}}\,\sin\left(\sqrt{\frac{g}{R}}\, t\right)\right]\frac{\pi R}{180}\,\delta_\omega \qquad [\text{metres}] \tag{4.18}$$

where

$\delta\omega$ = error in rotation angle correction

4.3.2 Electromagnetic Log

The EM log measures the ship's speed through the water within the hydrodynamic influence of the ship's hull. This will lead to the following errors in speed measurements :

1. boundary layer

   The velocity of the ship is measured with respect to a small volume of water in the direct vicinity of the ship's hull. Due to the viscosity of water a friction boundary layer is carried along with the hull. The speed is measured with respect to this boundary layer, which differs from the actual speed with respect to the surrounding water.

2. shallow water effect

   When the hull of the vessel is close to the bottom, the water flow velocity distribution changes due to restriction of the region in which the water can flow around the hull, causing an increase in the speed reading of a bottom-mounted sensor.

Since the log is rigidly attached to the hull and not in the centre of gravity of the ship, it undergoes displacements (both translations and rotations) proportional to the distance from the centre of gravity. The errors resulting from this can be divided into two categories :

1. errors which result from the dynamic orientation being different from the designed sensor orientation due to

   - trim
   - instantaneous pitch, roll and yaw angles
   - drift component due to wind

   These sources introduce errors since the log measures the actual speed through the water multiplied by the cosine of the angle

2. manoeuvring errors caused by controlled motions of the vessel, such as

   - turns : introduce an angle between the heading and the actual water velocity vector at the sensor.

Finally there are errors introduced when

- a misalignment exists between the ships fore-and-aft axis and the sensor axis;

- calibration of the log has not been performed correctly, resulting in erroneous settings

The total log error can be divided into three parts : a constant part, a part which is proportional with speed and a random part. If the vessel sails with constant speed, the error proportional to speed can be seen as part of the systematic error.
For an EM log that has been calibrated and correctly aligned, the standard deviation of the vessel's random error can be approximated as [Draaisma et al.,1982] :

$$\sigma = 0.02\ v \qquad (4.19)$$

where
v = speed of vessel through the water.

4.3.3 Gyrocompass

The gyrocompass is subject to several errors, some of which can be eliminated in the design of the compass, while others require manual adjustment. The main error sources are :

1. speed error

If the vessel is moving in another direction than due East or West, the compass is, in effect, being carried in a direction perpendicular to the resultant of the sum of the ship's velocity vector and the Earth rotation velocity vector. This error angle introduced is well known and the error angle (δ_v) can be approximated by [Draaisma et al.,1986] :

$$\delta_v \approx 0.1232 \; V \cos(C) \sec(\varphi) \; [\text{degrees}] \qquad (4.20)$$

where

V = ship's velocity in m/s
C = ship's heading
φ = latitude

In modern gyrocompasses this error is corrected mechanically. Speed and latitude are set by hand and the cosine of the course is introduced automatically. Wrong setting of the latitude and/or speed will still lead to an incorrect velocity correction which is considered to have a uniform pdf with a standard deviation of o = 0.2°

2. latitude error

When the gyro compass is equipped with a vertical damping system, a difference exists between the direction of steady state of the rotation axis of the gyroscope and the direction of the local meridian. The error is called the latitude error and has a magnitude

$$\delta_\varphi \approx 57.3 \; f \tan(\varphi) \quad [\text{degrees}] \qquad (4.21)$$

where

f = ratio between magnitude of horizontal and vertical damping forces

This error has a uniform pdf with a standard deviation of approximately 0.1° (latitude set to the nearest 10° and f = 0.04).

Gyro compasses equipped with a horizontal damping system do not experience a latitude error as such. This error is still introduced, however, when an incorrect latitude is set to correct for speed error.

3. error due to ship's motion

When the North-South component of the ship's speed changes, an accelerating force acts on the gyroscope which results in a precessing force in the horizontal plane, introducing a temporary error in the readings of magnitude :

$$\delta = \alpha \, \Delta(V \cos(C)) \quad [\text{degrees}] \tag{4.22}$$

where

α = multiplication factor depending on gyro compass type used ($\alpha \leq 0.1$)

This error will be systematic in nature and can be as large as a few degrees after manoeuvres.

4. ballistic damping error

A temporary oscillatory error of a gyrocompass introduced during changes of course and/or speed as a result of the means used to damp the oscillations of the spin axis. Provisions are made to counteract this effect when rates of change of course and/or speed exceed certain limits.

5. mechanical wear

Mechanical wear of the compass will lead to precessions different from those that are corrected for. This will result in a deviation of the heading.

6. errors in construction

This group of errors includes

- shift in centre of gravity of gyroscope, resulting in an additional precession;
- errors in the compass corrector-circuits
- misalignment of compass housing with respect to ship's fore-and-aft axis;
- follow-up error of repeater.

For a good working compass a maximum error due to mechanical wear and construction of less than 1° can be expected [Draaisma et al.,1986]. This error will be considered to have a uniform pdf with standard deviation of $\sigma = 0.6°$.

It has to be remembered that temporary disturbances like course and/or speed changes can influence the compass deviation over a long period since the oscillation period is approximately 84 minutes and per oscillation the deviation is reduced to 1/3 of its value. When many changes in course and/or speed are done in a short period of time a large deviation (up to a few degrees) can be the result.

The total gyro compass error can be divided into two parts : a systematic error introduced by course and/or speed changes and a random part. If the vessel sails with constant course and speed, the standard deviation of the random part can be approximately given as $\sigma = 0.65°$.

If a compass repeater is used to take bearings to objects, the following additional errors are introduced :

1. error in the repeater
2. round off error due to rounding to the nearest 0.5°

The total error of a bearing line is given by

error in gyro tc	0.65°
follow-up error of repeater	0° - 0.5°
reading rounding error	0.14°
total bearing error (1σ)	0.7° - 0.8°

This error is converted to meters in a direction perpendicular to the bearing line by using equation (4.16).

4.3.4 Inclinometer

During this research, no information was found on accuracy of attitude angles provided by SINS.

4.3.5 Echo sounder

When a depth is measured using an echo sounder, a combination of factors influence the accuracy and precision of the depth obtained. These factors are partly dependent on the echo sounder used but also on the bottom profile. When the depth obtained is used as bathymetric LOP and as such compared with depths given in charts, the whole process of chart compilation has to be considered as well. The compilation of a complete error budget, including all factors is outside the scope of this paper. Here, only the important factors are given. Nanninga [1985] and

Alper et al. [1985] have made an extensive analysis of precision of echo soundings obtained during surveys. The Royal Navy (U.K.) hydrographic department has also published a professional paper in which precision of soundings is assessed [MODUK,1990]. The description of errors as presented in this subsection is mainly based on these papers.

1. assumed speed of sound in water

   The speed of sound in water is dependent on the water temperature, salinity and water pressure. These quantities change with depth and position. In order to be able to measure depth accurately, sound velocity needs to be known over the whole water column. In most cases, only the velocity near the transducers is measured, introducing errors. As a guideline for errors in sound velocity, the following values can be used :

   - temperature : an error of 1°C results in an approximate velocity error of 3.6 m/s
   - salinity : an error of 1 ppt results in an approximate velocity error of 1.5 m/s
   - depth : an error of 100 m results in an approximate velocity error of 1.5 m/s

   The best way to establish sound velocity over the water column is by using a bathy-thermograph (BT) or a sound velocity probe. This way the sound velocity can be measured accurately with a standard deviation of σ_v = 1 m/s, leading to a standard deviation for a depth measured of:

$$\sigma = \frac{d}{v}\sigma_v \quad \text{[metres]} \tag{4.23}$$

   where
   v = assumed sound velocity
   d = measured depth

2. timing accuracy

   The error in depth measurements is directly proportional to the accuracy with which the time interval between transmitted and received pulse can be measured. The error is normally negligible with respect to errors resulting from using a wrong propagation velocity and bottom profile.

3. bottom profile and composition

   When a depth is measured, this is assumed to be the distance between transducer and top of seabed. It depends however on the frequency used and composition of the top layer of the bottom, whether this is the case or not. Another factor that plays an important role is sediment transportation as this can change the bottom profile considerably.

When the depth is measured, it is assumed that this is the depth directly under the transducer. However, the combination of transducer beamwidth and seabed slope cause errors in assumed position and depth. If a slope of α° exists and the beamwidth of is 10° used the following errors are introduced:

displacement : $\delta = d \sin(5^\circ) = 0.09\ d$ [metres](4.24a)
depth measured : $\delta = 0.09\ d \tan(\alpha^\circ)$ [metres](4.24b)

4. ship motions due to sea and swell

In general, ship motions like roll and pitch and swell can have an effect on the depth measurement of over a metre in bad weather conditions. Under normal circumstances, the effect will only be in the order of a few decimetres. The effect is reduced when the submarine is submerged and negligible at depths greater then 0.5 times the wavelength of the surface waves.

5. squat and settlement

Squat is the change in trim of a vessel under way with respect to that of the vessel stopped.
Settlement is the lowering of a vessel in the water due to the interaction between the hull and seabed. It occurs only in waters of depth less than approximately six times the vessel's draught.

6. survey accuracy and plotting of depth in charts

This comprises both the accuracy and precision of the depth measurement as well as positional accuracy of the depth measured.

The standard for precision of depths (combining both observation and position accuracy) as set from January 1991 by the International Hydrographic Organization is (1σ) :

$$\sigma = 0.25 + 0.0045d \quad \text{[metres]} \qquad (4.25)$$

This value can also be used as approximate value for precision of a depth measured on board the submarine.

To use a measured depth in combination with a depth contour given in a chart to obtain a 'horizontal' LOP, precision of both observations need to be combined. In order to give a value of precision for this bathymetric LOP in a horizontal direction, bottom topography, especially the bottom slope, and survey accuracy, especially in post processing of data, play an important role. This needs further investigation which falls outside the scope of this paper. Some results from current investigation have been published [Kielland et al.,1992; Kielland et al.,1992; Velberg,1992], although no definite values are given yet. Therefore no value will be presented here.[1]

4.3.6 Pressure sensors

The following error sources can be distinguished when using the pressure sensor to obtain a depth :

1. measurement precision

   The measurement precision is stated by the manufacturer as 0.02 percent of the frequency cycle time measured. Regular calibration of the sensor will guarantee that this is achieved.
   The measurement precision is effective on both the measurement of the reference frequency at the sea surface and the measured frequency at depth.

2. sea and swell

   Since a pressure based on the water column above the sensor is measured, water motions at the sea surface such as sea and swell cause fluctuations in the measurements. In order to reduce this effect, depth measurements have to be filtered resulting in a mean value.

3. reference pressure

   Before the submarine submerges, the reference frequency is determined. This is based on the present atmospheric pressure. Any changes in atmospheric pressure will lead to an error in the depth measurement. This error is directly proportional to the difference in surface air pressure.

---

[1] A horizontal precision of 1 - 10 km (1σ) is currently used as value in the Royal Netherlands Navy. This is mainly based on sparse field observations.

4. specific weight of seawater

When the measured frequency is converted to depth, one of the multiplication factors is to allow for specific weight with respect to water with a density of 1000 kg / m^3. The error is directly proportional to the depth.

The error budget of a depth obtained by using the pressure sensor can be given as

reference frequency	1.00 m
frequency at depth	1.00 m
sea and swell	0.25 m
surface pressure	0.05 m
specific weight seawater	0.10 m
total error in depth (1σ)	1.45 m

4.3.7 Periscope

The information on the accuracy with which bearings can be taken using a periscope when the submarine is at periscope depth is regarded as classified information, so no specific figures can be given.
The errors present in a bearing can, however, be divided into two categories :

A. errors present in the gyrocompass system as described in section 4.3.3

B. errors inherent to the periscope itself, such as :

- position and alignment with submarine's fore-and-aft axis;
- optics;
- follow-up errors of the periscope synchros;
- magnification used.

The total error of a bearing line is given by

total gyrocompass error	0.65°
total periscope error	0.25°*
rounding off error	0.14°
total bearing error (1σ)	0.7°

* Although the magnitude of the errors of the periscope itself are classified, a the theoretical value (specs.) is given.

This error is converted to meters in a direction perpendicular to the bearing line by using equation (4.16).

5. Least Squares

5.1 Introduction

In chapters 3 and 4 an overview was given of the sensors and systems available for collecting data to be used for position fixing, on board a submarine. In this chapter a mathematical model that combines the information provided by the unbounded systems and sensors to obtain an MPP will be developed. The model is based on the theory of least squares.
In the first section a short explanation is given why the method of least squares is the preferred method to combine observations into an MPP. Next the mathematical and statistical models on which the least squares method is based are discussed. The least squares algorithm provides unbiased minimum variance estimates of the parameters assuming the mathematical model reflects physical reality and the observations only contain random errors. In real-life this is not always the case. For this reason, statistical tests are performed on the results to check the validity of these assumptions. These statistical tests are described in section four. This section is followed by a section on position confidence regions. The chapter ends with some concluding remarks on how to expand the least squares to incorporate optimal parameter estimation of bounded sensor observables.

5.2 Justification for least squares

When a redundancy in measurements exists, an adjustment is necessary in order to get a unique solution to the problem at hand. The adjustment according to the principles of least squares provides a general and systematic procedure for applications to all sorts of situations. It states that '*the most probable values of measured quantities are those which make the sum of the weighted squares of the residuals a minimum*'. The method is limited when blunders and/or systematic errors are present. The adjustment does not correct the observations though it may improve them. After the adjustment, the observations should be consistent.

Assume blunders and systematic errors to be removed from the observations, leaving only random errors. From a statistical point of view, the least squares estimates of parameters can be considered as 'best estimates' under the assumption made, which means that the following statistical properties apply [Cross, 1983; Spaans,1988:2] :

1. the estimates are unbiased;

2. the VCV matrix of the estimates is the one having minimum trace;
3. derived quantities have minimum variance.

These properties are independent of the pdf of the observational errors, in particular irrespective of whether or not they are normally distributed. As a result of these properties, the least squares estimates are often referred to as the **Best Linear Unbiased Estimates (BLUE)**.
If the observational errors are normally distributed, then the least squares estimates have the additional property of being **maximum likelihood estimates**.

If, however, systematic and/or random errors are still present in the observations, the estimates will be biased, therefore leading to degradation of predictable accuracy. Statistical tests on the results need to be performed in order to indicate possible presence of non-random errors. If an error is suspected to exist in one or more of the observations, the observation(s) will be removed from the set and estimates will be calculated using the remaining observations. How this is done will be explained in section 4 of this chapter.

Apart from the mentioned statistical arguments, there are also some practical reasons for using the least squares method for calculating an MPP :

1. the method is extremely easy to apply because it yields a linear set of normal equations;
2. different types of observations can be mixed to obtain estimates of the parameters;
3. there is no upper limit to observations that can be incorporated in the calculation of parameters; the lower limit is determined by the dimension of the problem (i.e. the number of parameters to be estimated) and statistical tests to be performed;
4. it is flexible with respect to the number of observations being used for calculations, i.e. adding observations to or deleting them from the set of observations is easy;
5. it gives a unique solution;
6. it is, generally speaking, 'unobjectionable' - it is very difficult to form an argument against least squares and in favour of some other method;
7. the method leads to an easy assessment of quality via the a posteriori VCV matrix $C_{\hat{x}}$.

5.3 The mathematical model

In order to be able to use least squares, a mathematical model has to be constructed. This model is often thought of as being composed of two parts [Mikhail, 1976] :

1. functional model : describes the deterministic properties of the physical situation or event under consideration,

2. stochastic model : designates and describes the non-deterministic properties - i.e. the totality of the assumptions of statistical properties - of the variables involved. It includes all model variables and designates those that are considered fixed (constants) and those that are considered free (parameters).

The usefulness of the least squares technique and results obtained, depends very much on the way the model reflects physical reality.

There are three general methods of deriving most probable values. The functional model, as implemented in the computer program, makes only use of the class of observation equations, leading to the following set of equations [Cross, 1983; Spaans,1988:2] :

1. mathematical model	: $F(x) - 1 = 0$	(5.1a)
2. normal equations	: $A\ \delta x = b + v$	(5.1b)
3. LSE parameter corrections	: $\delta\hat{x} = (A^TWA)^{-1}A^TWb$	(5.1c)
4. LSE parameters	: $\hat{x} = x_0 + \delta\hat{x}$	(5.1d)
5. VCV matrix of parameters	: $C_{\hat{x}} = (A^TWA)^{-1}$	(5.1e)
6. residuals	: $\hat{v} = A\ \delta\hat{x} - b$	(5.1f)
7. corrected observations	: $\hat{1} = 1 + \hat{v}$	(5.1g)

5.3.1 Use of meter as calculating unit

Because different systems and sensors are used to obtain the MPP, each having its own unit of measurement (and standard error given accordingly), relative performance of the systems and sensors is difficult to judge from results. Although it is not necessary at all to transform all units to meters in order to obtain the estimate of parameters, since the least squares method is very well capable of dealing with different units, there are some advantages to favour for the meter as unit for calculations:

- the design matrix A is better balanced since all its elements will be in the same order of magnitude. This will result in better accuracy when performing matrix

number of LOPs in observation set	2D	3D
1 LOP	I	I
2 LOPs	II	I
3 LOPs	III	II
4 LOPs	IV	III
≥ 4 LOPs	IV	IV

explanation of categories :

I. no least squares estimates can be calculated from the set of observations. The observations can only be used to update an MPP calculated by the Kalman filter. No statistical tests can be performed;

II. least squares estimates can be calculated from the set of observations. However, the redundancy is zero so no statistical tests can be performed;

III. least squares estimates can be calculated from the set of observations. The redundancy is one, leading to a limited set of statistical tests that can be performed (only detection of outliers is possible);

IV. least squares estimates can be calculated and all statistical tests can be performed. (both detection and identification of outliers is possible).

Table 5.1 Overview showing the possibility to calculate a least squares solution and to perform statistical analysis for observation sets of different size.

operations, especially when inverting matrices;

- the a-priori VCV matrix of the observations (C_l) gives a better overview of the relative weights of the observations with respect to each other, leading to a better and easier evaluation of the results to be expected;

- interpretation of the residuals is easier when given in meters, especially when hyperbolic RPF systems which are sustaining from an LEF, or bearings are used;

- interpretation is easier on the whole since the unit of meter is well known as a unit.

5.3.2 Number of observations

In order to be able to calculate a least squares solution of a position from observations and to perform statistical tests, a minimum number of observations is needed. Table 5.1 gives for observation sets of different size (number of LOPs available) whether or not a least squares solution can be found and whether or not statistical analysis can be performed. A distinction is made between 2D and 3D position fixing.
If the number of observations available is not enough to calculate a least squares estimate of the MPP, the observations will be combined directly with the position based on the dynamic model and bounded sensors provided by the navigation filter (described in chapter 6). This is also done by using the least squares algorithm, where the navigation filter provides two LOPs (latitude and longitude).

5.3.3 The functional model

<u>The design matrix and observation vector</u>

The design matrix **A** contains for each observation the partial derivatives of $F(\mathbf{x},\mathbf{l})$ in x and y (and also z and $c\Delta t$ when GPS pseudo-ranges are observed). Each row of **A** represents one observation. The associated observation vector b contains for each observation the result of the observed minus calculated (O - C) value, calculated using provisional parameter values. For the different types of observation available in the simulation program, the functional model and its derivatives are given in appendix 5.

A fixed order in which the observations are placed in the matrix is used. This means that, in the event observations need to be discarded - based on eg. the outcome of statistical tests

or user choice -, no row-shifts need to be performed. This again leads to an easier interpretation of the matrices. LOPs are in- or excluded from calculations by adjusting the a-priori VCV matrix of the observations and the weight matrix (see section 5.3.4).
The top part of the design matrix contains the observations from GPS (measured pseudo-ranges) and Decca (measured lane fraction in combination with the whole lane number and zone given by the navigator) - the only systems implemented so far.
The bottom part of the design matrix contains one or more bearings, distances and/or radar (BRR) observations. At present the computer program allows up to a maximum of 5 observations to be processed at one time. This suffices under normal conditions since the process of taking more than one or two BRR observations is normally too slow compared with calculation time needed for a complete least squares calculation cycle. This means that the elements in this part of the matrix will only be calculated occasionally when observations are made. The order in which the observations were fed into the computer is kept throughout calculations. Once the MPP is calculated and a new cycle starts, all previous BRR observations are discarded, i.e. the elements in this part of the matrix are set to zero.
The design matrix has therefore the following general form :

$$A = \begin{vmatrix} \text{GPS observation(s)} \\ \text{-------------} \\ \text{Decca observation(s)} \\ \text{-------------} \\ \text{BRR observation(s)} \end{vmatrix} \qquad (5.2)$$

The observation vector b is adjusted according to the design matrix.

The depths observed by the pressure sensor and echo sounder are not incorporated in the least squares algorithm. They are taken directly as measurements in the measurement model of the Kalman filter, which will be described in the next chapter. The reason for doing this is based on the ease with which the navigation filter can be adapted to the different situations distinguished in submarine navigation. These situations will be given below. This approach does not affect the final accuracy of the MPP.

In the mathematical model both the design matrix and observation vector can be expanded easily by adding an appropriate number of rows via adjusting dimensions of matrix and vector. This way implementation of new system observations is an easy task.

Calculating surface

A choice had to be made whether calculations should be performed on the geodetic datum or on a plane projection. Since Omega and Loran-C have not been implemented into the model yet, all calculations can be performed without problems on a plane surface. The reason to use a plane in favour of the geodetic datum is merely based on time available and ease of graphics implementation. This choice does not influence the final results. It is, however, assumed that observations have already been reduced to the geodetic datum when given as input into the computer program. So far, no provisions have been made in the program in order to be able to reduce the observations from the Earth surface to the datum (spheroid). No systematic errors are introduced in the transformations of observations from spheroid to the grid since all observations are reduced to the grid using rigorous formulae.
Use is made of the Transverse Mercator (TM) projection in conjunction with a grid having the following parameters :

1. Central Meridian (CM) : longitude of calculated MPP rounded to the nearest degree;

2. grid origin : latitude = 0°
   longitude = CM

   The geodetic position (0°,CM) has the grid values (0,0);

3. grid scale constant on CM : k_0 = 1.0000;

At startup is the user will choose a geodetic datum[1]. Once the program is running, this datum can be changed if desired. The coordinates of fixed stations such as the transmitting stations of RPF systems will automatically be transformed to the new datum by the program. If coordinates of stations are requested as input by the program, these are assumed to be geodetic coordinates on the datum currently in use. All grid positional results are transformed back to geodetic coordinates on the datum currently

---

[1] The combination of a spheroid (axis and flattening) and origin (a position where the geoid-spheroid separation and deflection of the vertical are defined) constitutes a datum. For example, the International spheroid (a and f) located by the Potsdam 1950 origin (ξ_0,η_0,N_0) constitutes the European 1950 Datum (ED50).

used by means of rigorous formulae, and as such presented to the user.

Parameters

The parameters to be optimized are : latitude, longitude and depth[1]. These quantities are optimized indirectly since all calculations are performed on a plane projection, leading to the parameters x,y and z to be optimized. Once the latter are optimized, the transformation to geodetic coordinates results in optimized values for the first parameters. Three situations can be distinguished :

a. submarine at sea surface

   In this case the model is automatically reduced to a 2D problem, taking depth as being equal to zero and having no error (i.e. a standard deviation equal to zero). The z coordinate obtained from GPS pseudo-ranges is taken as a value for the geoid-spheroid separation.
   In this situation use can be made of all EPF systems and all sensors the submarine has available on board.

b. submarine submerged

   In this case the submarine is below the sea-surface so no use can be made of the RPF systems, radar, periscope etc. In this situation no least squares estimate of the MPP will be obtained. The information from the depth sensor and echo sounder are directly incorporated into the measurement model of the Kalman filter.

c. submarine at periscope depth

   In this situation it is possible to obtain one or more bearings with the periscope which can be used to update the horizontal position parameters. If sufficient observations are available, an estimate of the MPP is calculated. This situation is also considered to be two dimensional (only x and y are estimated).

In the situation described under a., the z coordinate is the sum of geoid-spheroid separation and tidal height. In the simulation program developed, tide has been assumed non-existent, and thus not corrected for. In a real-time situation, tidal height can be incorporated using a tidal prediction program.

---

[1] In the position calculation, mean sea level (MSL) is taken to be zero height. MSL is considered to coincide with the geoid.

In the computer program optimization of assumed propagation velocity of radio waves by introducing for example scale factors is not considered. The signal propagation velocity of terrestrial EPF systems is assumed to be equal to those velocities used to draw lattices on charts. However, provisions are made for input of 'fixed error corrections' (DNS) or ASF corrections (Loran-C) in order to be able to correct for (local) propagation velocity anomalies due to land path, when known. Field test should show whether this approach is sufficient for surface navigation or not.

Provisional coordinates

At startup of the program, the user has to give a DR position in combination with depth, which is taken as initial MPP. This position provides the initial provisional coordinates for the least squares calculation cycle. When the program is running, the coordinates of the last MPP derived are used to provide the provisional values of parameters for the following cycle. The user has, however, the possibility to input a new DR position and to use this position as new starting point.

If the approximate values used to obtain the numerical values of the elements of the A matrix are not close to the final least squares estimates, then the normal equations will not be a true linearization of the functional model. In such case it is necessary to iterate using the least squares estimates from the i-th computation as approximate values for the (i+1)-th computation. The iteration is stopped when the vectors of the parameter corrections and residuals change by insignificant amounts. In the computer program the change in value of the provisional coordinates between iterations is checked : if the correction ($\delta\hat{x}$) to each element of the parameter vector ($\hat{x}$) is less than the default value of one metre - which is considered to be sufficient for navigational purposes -, iteration is stopped. It is possible to change the value to any desired value.

5.3.4 The stochastic model

The a-priori VCV matrix

The a-priori VCV matrix of the observations (C_l) contains the variances and covariances of the observations based on prior knowledge. The ways to obtain values for these elements are many. The a priori VCV matrix used in the computer model is based on the error budgets given in chapter 4. These values are a combination of values given by manufacturers and found in literature. The values are either theoretical or derived from many observations made in the field under various conditions. It

is therefore assumed that these values reflect the actual standard errors of the observations quite well.
The layout of the VCV matrix is associated with the design matrix A and has the following form :

$$C_l = \begin{vmatrix} C_{GPS} & 0 & 0 \\ 0 & C_{Decca} & 0 \\ 0 & 0 & C_{obs} \end{vmatrix} \qquad (5.3)$$

where

- C_{GPS} an n x n VCV matrix belonging to the GPS pseudo-ranges
- C_{Decca} a 3 x 3 matrix formed by using the formula (A4.2)
- C_{Obs} a 5 x 5 diagonal matrix containing the variances for bearings, distances or radar observations as given by the error budgets and equation (4.16)

It is assumed that there does not exist any mutual correlation between the observations presented by different systems, bearings, distances or radar observations.

It is very important to have the variances of observations from different sources scaled correctly with respect to each other, i.e. given the correct **relative weight**. By having analyzed the error sources for data presented by the systems and other observations, resulting in error budgets as given in chapter 4 and converted to metres where necessary, this relative scaling of variances between systems and sensors should be correct. Future field tests have to prove whether these are chosen correctly, since relative scaling may well depend on the actual physical conditions affecting the performance of some systems, but not all. No provisions have been made yet for the user to adjust any variance or covariance values for the systems or other observations since estimating these has to be done as part of the system and model calibration using statistical analysis. It is very likely that input of the wrong values will lead to unwanted results. **Absolute scaling** of variances will be looked at in section 4.3 of this chapter.

In order to be able to leave observations out while retaining the correct order of observations in the various matrices, but still being able to calculate the inverse of the matrix, values of the a-priori VCV matrix are manipulated by the computer program as follows :

- if an observation has to be left out from the least squares calculations, its variance is set to one[1] while all its associated covariances are set to zero;
- the weight matrix (see below) is adjusted accordingly.

Expanding the a-priori VCV matrix

As it is possible to expand the A matrix, allowing new systems and/or more observations to be incorporated, the a-priori VCV matrix needs to be expanded accordingly. This is done simply by adding the VCV matrix of the new system as a sub-matrix at the appropriate position in the matrix or by increasing the sub-matrix C_{Obs}.

The Weight Matrix

The weight matrix (W) of the observations is defined as :

$$W = C_l^{-1} \qquad (5.4)$$

If the a-priori VCV matrix has been modified in order to leave one or more observations out of the calculations, the weight matrix needs to be modified as well. This is done simply by setting the appropriate diagonal element of the weight matrix to zero (in order to avoid rounding errors in the computer), thus effectively introducing a zero weight for the observation under consideration.

A - posteriori VCV matrices

i. parameters : $C_{\hat{x}} = (A^T W A)^{-1}$ (5.5a)

This matrix is used for evaluation of the precision of the parameters and accuracy of the position derived using a given set of observations;

---

[1] The variance of an observation not used in calculations is technically equal to infinity. Since the observation is made uncorrelated with other observations, its calculated weight will be 1 over infinity. In order to avoid rounding errors, the weight is set to zero manually, which means that any value (except zero) can be chosen in the a priori VCV matrix.

ii. residuals : $C_{\hat{v}} = C_l - AC_{\hat{x}}A^T$ (5.5b)

This matrix is used in calculations of reliability figures for observations (τ) and in the statistical test for detection of outliers.

iii. observations : $C_{\hat{l}} = C_l - C_{\hat{v}} = AC_{\hat{x}}A^T$ (5.5c)

This matrix is normally not used. It shows however a theoretical VCV matrix of the adjusted observations and should ideally be the same as the a-priori VCV matrix of the observations if the latter one was given correctly.

It should be noted that these matrices can be calculated based on the mathematical model only, i.e. no actual observation values are needed. Therefore the matrices could be used in advance to determine the minimum number and geometry - and therefore optimum set - of observations needed to provide results meeting specified precision criteria. In the computer program this analysis part has not been implemented for the simple reasons that generally no time is available for pre-analysis and that the navigator will always use as many observations as possible to calculate the MPP at any given moment.

The a-posteriori VCV matrices are only a measure of the precision of the position fixes. This is not sufficient to describe fully the quality of a fix; it is essential to quote also some measures of reliability and to have some indication of whether or not systematic errors and/or blunders are present in the observations. The way this can be done is discussed in the following section.

5.4 Statistical tests

The method of least squares is used to calculate estimates of random variables (parameters) from samples (set of observations). Related with this estimation is the task of determining accuracy and reliability with which the estimates are obtained (confidence measures) and whether the results of the estimations are in agreement with the initial assumptions (hypothesis). Therefore, once the MPP is calculated using the observations made, statistical tests are performed to validate the observations and final results derived. The general procedure of statistical testing always refers to a **null hypothesis (H_0)** - that is, the set of population parameters (mean, variance etc.) to which the statistics are compared. The result of the test is a statement whether, according to the available evidence, H_0 can be considered acceptable or not. In the developed computer model the following tests are performed:

- test for presence of outliers (blunders and large systematic errors);
- test for sufficient reliability;
- check on correct assumed value of unit variance.

The following type of test for hypotheses is performed in these situations :

$$H_0 \quad : p = p_0 \tag{5.6a}$$
$$H_1 \quad : p \neq p_0 \tag{5.6b}$$

where

p_0 represents a given standard value of the parameter
H_1 is the alternative hypothesis.

The significance level of the tests is set to $\alpha = 0.05$ (5%) and $\alpha = 0.01$ (1%).

Based on the outcome of the tests, a decision is made whether a LOP needs to be discarded from the observation set and the MPP to be recalculated. Additionally an advice is given whether or not the a priori VCV matrix should be scaled.

It is not important that the errors in observations are distributed according to a normal distribution to apply the least squares algorithm to calculate an MPP from LOPs. However, statistical tests as described in this section are based on a multivariate normal distribution. In chapter 2 it was reasoned that, due to the central limit theory, the distributions of the total error of observations (LOPs) are approaching the normal distribution, leading to a multivariate normal distribution of the calculated parameters.

5.4.1 Test for identification of outliers

Outliers in observations can either be blunders (eg. laneslips, multipath, malfunction of a system, sky wave interference, user input errors when for example bearings are included etc.) or large systematic errors (eg. in time changing signal interference, propagation velocity model errors etc.). In order to obtain an accurate position fix, observations sustaining errors like this should not be used. The test described here is used to detect and identify any outliers present.

When testing the observations for outliers, the following statistical test, known as B-method of testing after Baarda who introduced it, is performed [Spaans,1988:2] :

$$H_0 : \hat{l}_i = l_i - \hat{v}_i \tag{5.7a}$$

$$H_1 : \hat{l}_i = l_i + \Delta - \hat{v}_i \tag{5.7b}$$

were

- l = true observed value
- $\hat{l}$ = least squares estimate of observation
- $\hat{v}$ = residual
- Δ = blunder present in the observation.

The test is based on the following reasoning. If the least squares process is repeated without observation l_i, a new solution $\hat{x}'$ with associating $\hat{v}'$ is found. Using this information, a new value $\hat{l}_i'$ is calculated. From this the following quantities can be defined :

$$d_i = \hat{l}_i - \hat{l}_i' \tag{5.8a}$$

$$\sigma_{d_i} = \left(e_i^T W C_{\hat{v}} W e_i \right)^{-1/2} \tag{5.8b}$$

where

e_i = a null vector except for unity at the corresponding observation (i) being tested

The test statistic is consequently defined as

$$w_i = \frac{d_i}{\sigma_{d_i}} = -\left(e_i^T W \hat{v} \right) \sigma_{d_i} \tag{5.8c}$$

If the observation l_i does not contain a blunder, the test statistic will be a normalized statistic, having a normal distribution with zero mean and standard deviation one. If the observation does contain a blunder, w_i will have a normal distribution with a mean (actually a bias) equal to

$$m_w = \frac{\Delta_i}{\sigma_{d_i}} \tag{5.8d}$$

With this information, the hypotheses to be tested as given in (5.7) will be restated as :

$$H_0 \quad : E[w_i] = 0 \qquad (5.9a)$$
$$H_1 \quad : E[w_i] \neq 0 \qquad (5.9b)$$

which is a standard statistical test for sample mean with known standard deviation ($\sigma = 1$). From this it follows that, with significance level α given, H_0 will be accepted if :

$$P(-x < w_i < x) = 1 - \alpha \qquad (5.10a)$$

or when

$$P(abs(w_i) < x) = 1 - \alpha/2 \qquad (5.10b)$$

In the computer program each observation will be tested for given confidence level in the following way :

1. $abs(w_i) \leq 1.960$: H_0 will be accepted for both $\alpha = 0.05$ and $\alpha = 0.01$. The observation is not considered to be an outlier and will therefore always be accepted;

2. $1.960 < abs(w_i) \leq 2.576$: H_0 will be accepted when $\alpha = 0.01$ but rejected when $\alpha = 0.05$. In this case the observation will be accepted, but monitored in order to get additional information;

3. $2.576 < abs(w_i)$: H_0 will be rejected for both $\alpha = 0.05$ and $\alpha = 0.01$. The observation is assumed to be an outlier and will therefore always be rejected.

The values for confidence levels as stated above have been chosen to avoid too many unjust rejections of LOPs ($\alpha = 0.05$ means a chance of 1 in 20 of a LOP being rejected while H_0 is true, whereas $\alpha = 0.01$ gives a chance of 1 in 100), but to give the navigator on the other hand ample warning that limits are being reached due to for example low frequency systematic errors present in one or more LOPs. This makes it possible for the navigator to react adequately by for example de-selecting a LOP from the set of LOPs used for MPP calculation and/or choosing a more optimal set of observations to be used.
For the criteria 2 and 3 as described above, further investigation of the source of the error should be performed. It may well be caused by a constant systematic error not corrected for during system calibration (eg. wrong assumed propagation velocity), by low frequency disturbances resulting in a temporary large value of w_i (eg. due to noise, interference etc.) or by sudden unfavourable situations (eg. loss of receiver phase lock resulting in a laneslip, interference, multipath etc.).

If for one or more observations the value of w_i is larger than 2.576 (criterion 3), the observation having the largest absolute value will be considered to contain a blunder and will therefore be discarded from the set before recalculation of the position is performed. This process is repeated until all observations in the remaining set suffice either criterion 1 or 2 as given above, or when the remaining set of observations becomes too small to perform identification of outliers (see table 5.1).

5.4.2 Check on sufficient reliability

In chapter 2 reliability has been defined as 'the ease with which a blunder in a measurement can be detected'.

To give a measure for reliability, the reliability factor (τ) can be used, which is defined as [Cross,1983; Spaans,1988:2] :

$$\tau_i = (e_i^T C_l e_i e_i^T W C_{\hat{v}} W e_i)^{-0.5} = \left(\frac{\sigma_i^2}{\sigma_{d_i}^2} \right)^{-0.5} = \frac{\sigma_{d_i}}{\sigma_i} \qquad (5.11)$$

The larger the value of τ ($\tau > 1$), the less is the reliability of the LOP under consideration. The reliability factor can be used to calculate the maximum undetectable blunder under H_0. Associated with the reliability factor is the variance factor (VF), which is defined as :

$$VF_i = \frac{e_i^T (C_l - C_{\hat{v}}) e_i}{e_i^T C_l e_i} = \frac{\hat{\sigma}_i^2}{\sigma_i^2} \qquad (5.12)$$

where

σ_i = a priori standard deviation of observation (from C_l)
$\hat{\sigma}_i$ = a posteriori standard deviation (from equation 5.5c)

If VF > 0.9, the observation is considered to be unreliable, i.e. not independently checked which means that a blunder in the observation would remain undetected.

Both quantities are given in the computer simulation program and can help the navigator to decide whether a LOP should be used for position fixing or not.

5.4.3 Unit variance

A statistic known as the standard **error of unit weight** ($\hat{\sigma}_0$) is normally computed after the least squares estimates of the parameters and associating residuals have been obtained. It is a test statistic used to assess the a priori variances and covariance, and is defined as :

$$\hat{\sigma}_0 = \sqrt{\frac{\hat{v}^T W \hat{v}}{n - m}} \qquad (5.13)$$

where

n = number of observations
m = number of parameters estimated

It can be shown [Cross,1983;Spaans,1988:2] that if the correct a priori VCV matrix is used :

$$E[\sigma_0^2] = 1 \qquad (5.14)$$

If $\hat{\sigma}_0^2$ differs significantly from unity and statistical testing has shown that no blunders or large systematic errors are present in the observations, it is normally assumed that the a priori error VCV matrix has on average been underestimated by a factor $1/\hat{\sigma}_0^2$, i.e. absolute scaling of the variances and covariances had been wrong given the current physical conditions. One has to be careful with this assumption as errors in the mathematical model can also lead to a wrong value of $\hat{\sigma}_0^2$.

Using $\hat{\sigma}_0^2$ as scaling factor for the a priori VCV matrix will not have any effect on the least squares estimates of the parameters. It will however affect the a posteriori VCV matrices (equations 5.5a - 5.5d) and therefore the size of the error ellipse / ellipsoid, making assessment of fix quality unreliable.
In order to decide whether $\hat{\sigma}_0^2$ differs significantly from unity, the following hypotheses are tested :

$$H_0 \quad : \sigma_0^2 = 1 \qquad (5.15a)$$

$$H_1 \quad : \sigma_0^2 \neq 1 \qquad (5.15b)$$

The test statistic to be used is given by Mikhail [1976] as :

$$T = \frac{[(n - m) - 1]\, S^2}{\sigma_0^2} \tag{5.16}$$

where

S^2 = sample unit variance, calculated using equation (5.13)
σ_0^2 = a priori unit variance
n = number of observations
m = number of parameters

It can be shown that this test statistic has a chi-square distribution with (n-m)-1 degrees of freedom. The test consists of checking if the value of the sample is within the confidence region given by

$$P\left[\frac{\sigma_0^2}{(n-m)-1}\,\chi^2_{\frac{\alpha}{2},(n-m)-1} < S^2 < \frac{\sigma_0^2}{(n-m)-1}\,\chi^2_{1-\frac{\alpha}{2},(n-m)-1}\right] = 1 - \alpha \tag{5.17}$$

If H_0 is rejected and no blunders are identified, it can be concluded that the a priori VCV matrix had been scaled incorrectly and should be scaled by S^2.

In the computer program, scaling is not automatically performed as this is not desirable because of the small number of redundancy. As long as a least squares MPP is calculated, not only the variance of unit weight based on the sample is calculated using equation (5.13), but also a 'cumulative' unit variance, using the following recursive formula

$$\hat{\sigma}^2_{i+1} = \hat{\sigma}^2_i + \frac{r_{i+1}}{R_{i+1}}\left(S^2_{i+1} - \hat{\sigma}^2_i\right) \qquad i = 0,1,2,\ldots \tag{5.18}$$

where

$\hat{\sigma}^2$ = cumulative unit variance with initial value $\hat{\sigma}_0^2 = 0$
S^2 = sample unit variance calculated using equation (5.13)
r = redundancy of sample
R = cumulative redundancy with initial value $R_0 = 0$

This way a more reliable test can be performed as the cumulative redundancy R increases. As an indication of the number of position fixes needed before a reliable test can be performed, Spaans [1988:2] gives a total redundancy of R = 200, which means that for a 4 LOP - 2D position fix with sample redundancy r = 2, at least 100 consecutive position fixes have to be made. On the other hand, Cross [1983] gives an example of the danger of using S^2 as scaling factor. This means that special techniques have to be developed to decide whether or not the a priori VCV matrix should be scaled. This has not been performed as part of the research. Therefore, **automatic** scaling of the a priori VCV matrix is not performed and it is left to the navigator to decide whether this should be done or not. In the computer simulation program a provision is made to switch automatic scaling on and off.

5.5 Ellipsoids of constant probability

5.5.1 Multidimensional distribution

Although the least squares theory of adjustment does not require a specified distribution, the statistical testing following the adjustment is based on random vectors with elements having a normal distribution, leading to the multivariate density function :

$$f(x_1 \ldots x_n) = f(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C_x}}\, e^{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu}_x)^t C_x^{-1} (\mathbf{x} - \boldsymbol{\mu}_x)} \qquad (5.19)$$

The function

$$h(x) = (\mathbf{x} - \boldsymbol{\mu}_x)^T C_x^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) = k^2 \qquad (5.20)$$

represents a family of hyper-ellipsoids of constant probability when the quadratic form is positive definite [Mikhail,1976].

In this paper only the situations for n = 2 (the ellipse) and n = 3 (the ellipsoid) will be considered since these are used for 2D and 3D position fixing respectively.

It can be shown [Mikhail,1981] that for every symmetric VCV matrix $\mathbf{C}_x^{-1}$ there exists an orthogonal rotation matrix **R** such that $\mathbf{R}^T\mathbf{C}_x^{-1}\mathbf{R}$ is a diagonal matrix. The columns of **R** are the normalized eigenvectors of C_x^{-1} and the elements of the diagonal matrix are the corresponding eigenvalues. From this it can also be shown that the eigenvalues of the VCV matrix C_x represent the squares of the lengths of the primary axes of the ellipsoids.

5.5.2 The error ellipse

In the 2D situation we find for (5.20) :

$$h(x,y) = \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix}^T C_x^{-1} \begin{bmatrix} x - \mu_x \\ y - \mu_y \end{bmatrix} \tag{5.21}$$

$$= \frac{1}{1-\rho^2}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right] = k^2$$

The equation $h(x,y) = k^2$ for a specific value of k is an ellipse which is known as an **ellipse of constant probability**. The value of the probability depends on the value of k. If k = 1, the ellipse is called the **standard ellipse**. The shape of the ellipse is fully determined by σ_x, σ_y and ρ.

There are several ways of calculating the lengths of the semi-major and -minor axes of the ellipse and the angle between its major axis and the y-axis (direction of North) of the local reference system. In the computer program these values are obtained by calculating the eigenvalues and eigenvectors of C_x using the following characteristic equation :

$$\lambda^2 - Tr(C_x)\lambda + \det(C_x) = 0 \tag{5.22}$$

The parameters for the standard ellipse are then given by

$$a = \left[\frac{1}{2}(\sigma_1^2 + \sigma_2^2) + \frac{1}{2}\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + 4\,\sigma_{12}^2} \right]^{1/2} \quad \text{[metres]} \qquad (5.23a)$$

$$b = \left[\frac{1}{2}(\sigma_1^2 + \sigma_2^2) - \frac{1}{2}\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + 4\,\sigma_{12}^2} \right]^{1/2} \quad \text{[metres]} \qquad (5.23b)$$

$$\alpha = \arctan\left[\frac{\sigma_{12}}{\lambda_1 - \sigma_1^2} \right] \quad \text{[degrees]} \qquad (5.23c)$$

Having defined the standard ellipse this way, it is now possible to calculate the probability that the random vector x takes values within or on an ellipse with semi principal axes ka and kb. The general expression is :

$$P\left\{ \chi_2^2 < k^2 \right\} = 1 - e^{-\frac{k^2}{2}} \qquad (5.24)$$

where

χ_2^2 = chi-square distribution with two degrees of freedom

In order to establish confidence regions, the confidence level (γ) is selected and the multiplier k is calculated from :

$$P\left\{ \chi_2^2 < k^2 \right\} = 1 - e^{-\frac{k^2}{2}} = \frac{\gamma}{100} \qquad (5.25)$$

Table 5.2 gives typical values of k for given P.
The computer model will provide the navigator with the dimensions of both the 95% ellipse (IMO recommended) and the 99.9% ellipse since submarine navigation is more hazardous than normal surface navigation.

	k	P
standard ellipse	1.000	0.3935
50% ellipse	1.177	0.5000
95% ellipse	2.447	0.9500
99.9% ellipse	3.717	0.9990

Table 5.2 Typical values of multiplication factor k (2D)

	k	P
standard ellipsoid	1.0000	0.1988
50% ellipsoid	1.5382	0.5000
95% ellipsoid	2.7955	0.9500
99.9% ellipsoid	4.0332	0.9990

Table 5.3 Typical values of multiplication factor k (3D)

5.5.3 The error ellipsoid

In the 3D situation we find for (5.20), defining the ellipsoid of constant probability :

$$h(x,y,z) = \begin{pmatrix} x - \mu_x \\ y - \mu_y \\ z - \mu_z \end{pmatrix}^T C_x^{-1} \begin{pmatrix} x - \mu_x \\ y - \mu_y \\ z - \mu_z \end{pmatrix} = k^2 \qquad (5.26)$$

Again, the value of probability depends on k. If k = 1, the ellipsoid is called the **standard ellipsoid.** As was the case with the error ellipse, several ways are possible to calculate the lengths of the three main axes of the ellipsoid and its orientation in the local reference system. In the computer program the eigenvalues of $C_{\hat{x}}$ are computed from the characteristic equation :

$$\det\left| C_x - \lambda\, I \right| = 0 \qquad (5.27)$$

Since this is a third order polynomial in λ, its roots can be found using rigorous formulae. Once the eigenvalues are calculated, the normalized eigenvectors of $C_{\hat{x}}$ are calculated. Having found these vectors, the orientation of the ellipsoid is found by solving the equation :

$$E = R(\rho, \theta, \varphi) = R_z(\varphi)\, R_y(\theta)\, R_x(\rho) \qquad (5.28)$$

where

- E = matrix containing normalized eigenvectors of $C_{\hat{x}}$
- R = rotation matrix
- ρ = rotation angle around x-axis
- θ = rotation angle around y-axis
- φ = rotation angle around z-axis

Finally, the general expression giving the probability that the random vector x takes values within or on the surface of the ellipsoid with principle axes of length ka, kb and kc is :

$$P(\chi_3^2 < k^2) = \int_0^{k^2} \frac{1}{2\sqrt{2}\,\Gamma\left(\frac{3}{2}\right)} \sqrt{t}\, e^{-t/2}\, dt = \frac{2}{\sqrt{\pi}} \int_0^{\frac{k^2}{2}} \sqrt{p}\, e^{-p}\, dp$$

(5.29)

$$\Gamma\left(\frac{3}{2}\right) = \int_0^{\infty} \sqrt{t}\, e^{-t}\, dt = \frac{1}{2}\sqrt{\pi}$$

Table 5.3 gives typical values of k for given P.
The computer model will provide the navigator with the dimensions of both the 95% and the 99.9% ellipsoid.

5.5.4 Radial standard deviation

The result of the least squares calculation is an MPP with associating error ellipse / ellipsoid based on LOPs available. If the navigator wants to plot the position with associating confidence region in the chart, it is easier to plot a circle than an ellipse. Since the relationship between the standard ellipse / ellipsoid and radial standard deviation is normally not straight forward, except in the case when the standard ellipse or ellipsoid becomes a circle or sphere respectively, curves and tables have been made available to convert from one to the other.

The only radial standard deviation provided by the computer program is the distance root mean squared (d_{rms}), which is defined as :

$$d_{rms} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \qquad (5.30)$$

where

- σ_x = length of semi major axis of standard error ellipse / ellipsoid
- σ_y = length of semi minor axis of standard ellipse or semi 'medium' axis of standard ellipsoid
- σ_z = length of semi minor axis of standard ellipsoid and zero for the ellipse

Care has to be taken when interpreting the d_{rms} value : although the standard ellipse / ellipsoid represents an area of constant probability (39.4% and 19.9% respectively), the probability associated with a given value d_{rms} varies as function of ellipse / ellipsoid eccentricity.

5.6 Concluding remarks

In this chapter, the parameter vector only contained position parameters (and $c\Delta$ if GPS pseudo-ranges are observed). The least squares is however not restricted to the estimation of only these parameters. The parameter vector can easily be expanded to incorporate other parameters such as velocity, heading etc. It is possible to use the least squares to estimate all elements of the submarine's state-vector which will be given in the next chapter. This is done by incorporating the relevant observation equations in the design matrix and observation vector and by including the error VCV matrix of the sensor or system to be added as sub-matrix in $C_{\hat{x}}$. Normally this method is not used. Instead a filtering algorithm is used to estimate these parameters. This algorithm is discussed in the next chapter. The least squares algorithm only provides position parameter estimates to this filtering algorithm.

6. Kalman Filter

6.1 Introduction

The estimation of the position of the submarine (MPP) is derived from information provided by systems and sensors. This information is generally corrupted by additive noise. Using external measurements together with information from bounded sensors in an optimal way, can result in a positional accuracy which is better than obtainable by either external measurements or sensor information alone.
In the previous chapter a description was given how the least squares algorithm is used to derive an MPP based on measurements from EPF systems. But, this way not all available information is used because sensor output (from log, gyro compass, SINS, depth sensor etc.) and past information is not considered. The Kalman filter, which is the subject of this chapter, provides a recursive computational algorithm which 'remembers' past data and uses this in combination with present measurements and sensor information to calculate the best estimate of the present and future state of the submarine. The state vector considered in this chapter only contains the vessel's 3D position, speed, heading and inclination along with estimates of current velocity and biases in log, gyro compass and inclinometer.
Before giving the Kalman filter equations, the first section looks to the justification of such a filter. The following two sections describe the filter algorithms for linear and non-linear systems. This is followed by a section giving the model as it is implemented in the computer program. In the next section some problems encountered when implementing the algorithm into a computer will be discussed, including an indication how to reduce their effect on performance. If the assumed models used were correct, the navigation filter would provide minimum variance unbiased estimates. In a dynamic environment deviations from the assumed models are encountered. Therefore, statistical tests are needed to monitor correct functioning of the filter in order to be able to react adequately to malfunctioning. These test will be discussed in the penultimate section of this chapter. The final section of this chapter will provide a short evaluation of the expected performance of the navigation filter, when used on board the submarine.

6.2 Why a Kalman filter ?

When in a dynamic environment, such as a submarine at sea, the parameters to be estimated change with time because the submarine is moving and because the system and sensor observables show temporal variations. At any moment in time it is possible to use the least squares algorithm as described in the previous chapter, to calculate single MPPs based on the present observations made. It is possible to incorporate past measurements in the least squares algorithm, but this would lead to an ever increasing number of observation equations as a number of observations is added every time new observations become available. Soon, the calculation time to obtain a new MPP would become unacceptable. The Kalman filter improves the way parameter estimation is performed by the least squares algorithm by relating parameters calculated at previous moments in time to the parameters calculated at present time in a recursive way. The Kalman filter is considered the most common optimal filtering technique for estimating the state of linear systems. Before starting the actual discussion of the Kalman filter, a definition of an optimal estimator, should be given because of its fundamental importance for the theory :

> 'An optimal estimator is a computational algorithm that processes measurements to deduce a minimum error estimate of the state of a system by utilising :
>
> i. knowledge of the system dynamics and measurements;
> ii. assumed statistics of the system noise and measurement errors; and
> iii. initial conditions of information.'

This leads to three main types of estimation problem :

1. **filtering** : the time at which an estimate is desired coincides with the last measurement point;
2. smoothing : the time of interest falls within the span of available measurement data;
3. prediction : the time of interest occurs after the last available measurement.

Since the computer model developed is designed to be used to give the navigator on line information on the submarine's current and future position, only the filtering and prediction estimation problems are of interest. Smoothing will therefore not be considered in this paper.

When obtaining a position from measurements, the following has to be kept in mind :

- external measurements contain random errors that may be significant with respect to the errors rising from the use of a bounded-sensor navigation system;
- bounded-sensor navigation system errors are primarily caused by the random, time-varying errors of the sensors

used.

Considering this, the Kalman filter algorithm shows the advantage that it uses all measurement data available, regardless of their errors, plus the prior knowledge about the system and its environment. This leads to the following characteristics [Gelb,1974] :

- it provides useful estimates of all sensor error sources with significant correlation time;
- it can accommodate non-stationary error sources when their statistical behaviour is known;
- configuration changes in the availability of navigation systems and sensors can easily be accounted for;
- it provides for optimal use of any number, combination and sequence of external measurements.

In order to avoid a growing memory filter resulting from storing all past measurement data, the estimate is sought in a linear recursive form so that there is no need to store past measurements for the purpose of computing present estimates. Again, the Kalman filter provides such a recursive algorithm.

Apart from the advantages mentioned above, there are also some important disadvantages to the filtering technique :

- the filter is based on the assumptions that the system is linear and the random errors are described by Gaussian white noise processes;
- it is sensitive to erroneous a-priori models and statistics;
- the computational burden when exact system and error models are to be used.

The first disadvantage can be overcome by linearization of the system equations (described in section 6.4) and use of shaping filters (described in section 6.3.2), whereas the other two disadvantages can be overcome to a fair extent by carefully selecting model- and error descriptions.

6.3 Linear dynamic systems

In this section the algorithm for optimal filtering will be given. It is based on the equations describing the dynamics of time continuous systems which are transformed into discrete-time equations in order to be able to implement the algorithms on a computer. Gaussian white noise is assumed but some consideration will be given on implementing the filter algorithm in the presence of coloured noise. The final subsection deals with prediction.

6.3.1 The discrete-time model filter equations

The expected future state of a linear dynamic system under the influence of (external) disturbances, may be predicted by the equations of motion, represented by the following linear first order vector differential equation :

$$\dot{\mathbf{x}}(t) = F(t)\,\mathbf{x}(t) + G(t)\,\mathbf{w}(t) \tag{6.1}$$

where

$\mathbf{x}(t)$	n dimensional continuous time state vector
$F(t)$	n x n system dynamics matrix
$G(t)$	n x m matrix representing the effect of noise on the state
$\mathbf{w}(t)$	m dimensional Gaussian white noise process vector having the following statistics

$$E[\mathbf{w}(t)] = 0 \qquad \forall\ t$$
$$E[\mathbf{w}(t)\,\mathbf{w}^T(\tau)] = Q(t)\delta(t - \tau)$$

$Q(t)$	m x m symmetric positive semi definite matrix
$\delta(t)$	Dirac delta function

The matrix Q(t) is the spectral density matrix of the system noise. It is used when calculating the error VCV matrix of the predicted state vector. Its contents will be defined in section 5.4 of this chapter.

Given the state vector at a particular point in time, and a description of the system forcing functions from that point in time forward, the state vector can be computed at any other time from the solution of equation (6.1), which is given by :

$$\mathbf{x}(t) = \Phi(t,t_0)\,\mathbf{x}(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\,G(\tau)\mathbf{w}(\tau)\,d\tau \tag{6.2}$$

where

$\Phi(t,t_0)$ n x n transition matrix which is the solution of the matrix differential equation

$$\dot{\Phi}(t,\tau) = F(t)\Phi(t,\tau)$$
$$\Phi(t,t) = I \qquad \forall\ t$$

Since a computer is used to obtain the most probable position of the submarine, its state vector can only be computed at discrete times using numerical solutions to integration and differentiation. Therefore discrete-time dynamic system equations need to be used to describe the motion of the submarine. These discrete-time equations arise from sampling the continuous system described by (6.1) and (6.2). The system vector difference equations allowing the state x_{k+1} of the submarine at time t_{k+1} to be calculated from the state x_k at time t_k is given by :

$$x_{k+1} = \Phi_k x_k + \Gamma_k w_k \tag{6.3a}$$

where

$$\Phi_k = \Phi(t_{k+1}, t_k) \tag{6.3b}$$

$$\Gamma_k w_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) G(\tau) w(\tau) d\tau \tag{6.3c}$$

The stepsize $\Delta t = t_{k+1} - t_k$ will be chosen to ensure that the transition matrix Φ and the variance of model and measurement noise together with their effects on the system can be considered constant.

Observation data is obtained at discrete moments in time. This data is considered to be linearly related to the state of the system according to

$$z_{k+1} = H_{k+1} x_{k+1} + v_{k+1} \tag{6.4}$$

where

- z p dimensional measurement vector
- H p x n observation matrix
- v p dimensional Gaussian white noise process vector having the following statistics

$$E[v_k] = 0 \qquad \forall\ k$$

$$E[w_k w_l^T] = R_k \delta_{kl}$$

- R_k p x p symmetric positive definite VCV matrix

The matrix R_k is the error VCV matrix of the measurement noise. It is used when calculating of the Kalman Gain, updated state vector and error VCV matrix of the updated state vector. Its contents will be defined in section 5.4 of this chapter.

The noise sequences $\{w_k\}$ and $\{v_k\}$ are assumed to be uncorrelated:

$$E[v_k w_j^T] = 0 \qquad \forall\ j,k \tag{6.5}$$

The initial state of the discrete-time system described by (6.3) and (6.4) is defined as :

$$E[x_0] = \hat{x}_0 \tag{6.6a}$$

$$E\left[\tilde{x}_0\tilde{x}_0^T\right] = P_0 \tag{6.6b}$$

where

- P n x n error VCV matrix of the state estimate
- $\tilde{x}$ error in the state vector estimate $\tilde{x} = x - \hat{x}$

Given the linear system described by (6.3) and (6.4) with initial conditions (6.6), the Kalman filter algorithm provides the estimate $\hat{x}_{k|k}$ of the state x_k that is a linear function of all measurement data $(z_1,\ldots\ldots,z_k)$ which satisfies the following conditions :

- $\hat{x}_{k|k}$ is unbiased;
- $\hat{x}_{k|k}$ is the minimum variance estimate;
- $\hat{x}_{k|k}$ is consistent.

The estimate of the state vector is obtained in two steps :

Step 1 : State vector propagation.

This is a 'forecasting' process describing the discontinuous state estimate and error VCV matrix behaviour between measurement times t_k and t_{k+1}. Since system noise has zero mean, an estimate of the state vector and its VCV matrix at time $t_{k+1} > t_k$ is given by [AGARD,1970] :

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k} \tag{6.7a}$$

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \tag{6.7b}$$

Step 2 : State vector update.

This is the discontinuous state vector estimate and error VCV matrix behaviour across a measurement. The difference between the actual and predicted measurements is used as a basis for calculation of corrections to the estimate of the state vector. The algorithm combines the measurement vector z_{k+1} with the prediction of the state vector as derived in step 1 and is given by the following set of equations [AGARD,1970] :

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\left[z_{k+1} - \hat{z}_{k+1|k}\right] \tag{6.8a}$$

$$= \hat{x}_{k+1|k} + K_{k+1}\left[z_{k+1} - H_{k+1}\hat{x}_{k+1|k}\right]$$

$$P_{k+1|k+1} = \left[I - K_{k+1}H_{k+1}\right]P_{k+1|k} \tag{6.8b}$$

$$K_{k+1} = P_{k+1|k}H_{k+1}^{T}\left[H_{k+1}P_{k+1|k}H_{k+1}^{T} + R_{k+1}\right]^{-1} \tag{6.8c}$$

where

K	n x p optimal gain matrix for the unbiased minimum variance filter
$\hat{z}_{k+1\|k}$	predicted measurement vector

Equations (6.7) and (6.8) describe the complete discrete-time Kalman filter algorithm for linear systems.

6.3.2 Coloured noise

In the derivation of the Kalman filter equations, system and measurement noise are considered to be Gaussian white noise processes. However, white noise is physically not realizable. Instead, most system and measurement noise processes exhibit time correlation (coloured noise). In this subsection a method will be given by which coloured noise can be allowed for in the description of a dynamic time continuous system.

The continuous-time equations for a dynamic system and measurements having coloured noise are given by [AGARD,1970] :

$$\dot{x}(t) = F(t)x(t) + Gn_s(t) \tag{6.9a}$$

$$z(t) = H(t)x(t) + n_m(t) \tag{6.9b}$$

where $n_s(t)$ and $n_m(t)$ are coloured noise processes having the following statistical properties

$$\begin{aligned} E[\mathbf{n}_s(t)] &= 0 \\ E\left[\mathbf{n}_s(t)\mathbf{n}_s^T(\tau)\right] &= D_s(t,\tau) \end{aligned} \qquad (6.10a)$$

$$\begin{aligned} E[\mathbf{n}_{\blacksquare}(t)] &= 0 \\ E\left[\mathbf{n}_{\blacksquare}(t)\mathbf{n}_{\blacksquare}^T(\tau)\right] &= D_{\blacksquare}(t,\tau) \end{aligned} \qquad (6.10b)$$

When the coloured noise can be modelled by a shaping filter, having Gaussian white noise as forcing function (see Appendix 6), state vector augmentation can be used to transform equations (6.9a) and (6.9b) as follows :

$$\dot{x} = F(t)x + Gn_s \qquad (6.11a)$$
$$\dot{n}_s = A_s n_s + w \qquad (6.11b)$$
$$\dot{n}_{\blacksquare} = A_{\blacksquare} n_{\blacksquare} + v \qquad (6.11c)$$

$$z = Hx + n_{\blacksquare} \qquad (6.11d)$$

where

$A_{\blacksquare}, A_s$	system transfer function of shaping filters
w, v	Gaussian white noise processes

or in matrix form :

$$\begin{pmatrix} \dot{x} \\ \dot{n}_s \\ \dot{n}_{\blacksquare} \end{pmatrix} = \begin{pmatrix} F & G & 0 \\ 0 & A_s & 0 \\ 0 & 0 & A_{\blacksquare} \end{pmatrix} \begin{pmatrix} x \\ n_s \\ n_{\blacksquare} \end{pmatrix} + \begin{pmatrix} 0 \\ w \\ v \end{pmatrix} \qquad (6.12a)$$

$$z = \begin{pmatrix} H & 0 & I \end{pmatrix} \begin{pmatrix} x \\ n_s \\ n_{\blacksquare} \end{pmatrix} \qquad (6.12b)$$

Equation (6.12a) gives the description of an equivalent system having a 'white noise' driving force and equation (6.12b) describes 'error free' measurements. Before the Kalman filter algorithms (6.7) and (6.8) can be used with these equations, an additional transformation is needed to allow for the 'error free' measurements since the equivalent R matrix is singular thus preventing the Kalman gain matrix K, required for obtaining the optimal state estimator, to be calculated. Next the equations need to be written in discrete form in order to be able to implement them in the computer model. This is done in analogy with the system described in subsection 6.3.1.

In this paper the system and measurement model described by equation (6.9) is not considered. In order to get a good working filter, extensive analysis of system and sensor information is needed to obtain correct correlation times for the shaping filters. This analysis has not been performed as part of the research.

6.3.3 Prediction

Optimal prediction can be thought of in terms of optimal filtering in the absence of measurements. This is equivalent to optimal filtering with arbitrarily large measurement errors ($R^{-1} \to 0$ so $K \to 0$). Therefore, if measurements are unavailable beyond some time t_k, the optimal prediction of $\mathbf{x}_{k+1}$ for $t_{k+1} > t_k$, given $\hat{\mathbf{x}}_{k|k}$ must be

$$\hat{\mathbf{x}}_{k+1|k} = \Phi_k \hat{\mathbf{x}}_{k|k} \qquad (6.13)$$

having an error VCV matrix given by :

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \qquad (6.14)$$

6.4 Discrete-time non-linear estimation

The general Kalman filter algorithm is based on linear differential equations. However, the equations used to describe the motions of the submarine in 3D space are non-linear time continuous differential equations. This means that the equations need to be linearized. Furthermore, since the Kalman filter is to be implemented on a computer, the differential equations need to be converted to discrete-time difference equations.

In this section the linearized, discrete-time Kalman filter equations, based on a time continuous system model and discrete-time measurements are given. They form the basis of the filter to be implemented in the computer model, which will be discussed in section 6.5.

Non-linear systems having continuous dynamics are described by [Gelb,1974] :

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t),t) + G\mathbf{w}(t) \qquad (6.15a)$$

$$\mathbf{z}_{k+1} = \mathbf{h}_{k+1}(\mathbf{x}(t_{k+1})) + \mathbf{v}_{k+1} \qquad (6.15b)$$

Given the minimum variance estimator $\hat{\mathbf{x}}(t_k)$ at time t_k containing all measurement data up to t_k, the best estimate at t_{k+1} is found in two steps :

Step 1 : State vector propagation.

Between measurement times t_k and t_{k+1} no measurements are taken and the state propagates according to (6.15a). On the interval $t_k \leq t \leq t_{k+1}$, the conditional mean of $\mathbf{x}(t)$ is the solution of the equation

$$\dot{\hat{\mathbf{x}}} = \hat{\mathbf{f}}(\mathbf{x}(t),t) \qquad t_k \leq t \leq t_{k+1} \qquad (6.16)$$

In order to be able to compute $\hat{\mathbf{f}}(\mathbf{x},t)$, the pdf $p(\mathbf{x},t)$ needs to be known. To obtain practical estimation algorithms, methods of computing the mean and VCV matrix which do not depend on knowing the pdf are needed. One way to do this is by expanding $\mathbf{f}(\mathbf{x},t)$ in a Taylor series about the current estimate of the state vector:

$$\mathbf{f}(\mathbf{x},t) = \mathbf{f}(\hat{\mathbf{x}},t) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\hat{\mathbf{x}}}(\mathbf{x} - \hat{\mathbf{x}}) + \ldots \qquad (6.17)$$

which leads to

$$\hat{\mathbf{f}}(\mathbf{x},t) = E\left[\mathbf{f}(\mathbf{x},t)\right] = \mathbf{f}(\hat{\mathbf{x}},t) + 0 + \ldots \qquad (6.18)$$

Using (6.18) in (6.16), the first order approximation of the state propagated state vector and its error VCV matrix can be written as [Gelb,1974]

$$\dot{\hat{x}}(t) = f(\hat{x},t) \tag{6.19a}$$

$$\dot{P}(t) = F(\hat{x}(t),t)P + PF^T(\hat{x}(t),t) + G(t)Q(t)G^T(t) \tag{6.19b}$$

where $F(\hat{x}(t),t)$ is the matrix whose elements are given by

$$F_{ij}(\hat{x}(t),t) = \left.\frac{\partial f_i(x(t),t)}{\partial x_j(t)}\right|_{x(t) = \hat{x}(t)} \tag{6.19c}$$

In order to be able to use this propagation of the state vector in the computer model, equations (6.19a) and (6.19b) need to be written in their discrete-time equation equivalent, which are given by :

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}, t_k) \tag{6.20a}$$

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \tag{6.20b}$$

where

$$\Gamma_k w_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau) G(\tau) w(\tau) d\tau \tag{6.20c}$$

$$f(x_k,t_k) = x(t_k) + \int_{t_k}^{t_{k+1}} f(x(t_k),\tau)\, d\tau \tag{6.20d}$$

$$\Phi_k = \Phi(t_{k+1},t_k) = I + \int_{t_k}^{t_{k+1}} \nabla_x \, d\tau \tag{6.20e}$$

$$\nabla_x = \left.\frac{\partial f(x,\tau)}{\partial x}\right|_{x = \hat{x}_{k|k}} \tag{6.20f}$$

Step 2 : State vector update equations.

After the measurement z_{k+1} is taken, the state vector and its error VCV matrix are improved by

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\left[z_{k+1} - h_{k+1}(\hat{x}_{k+1|k})\right] \tag{6.21a}$$

$$P_{k+1|k+1} = \left[I - K_{k+1}H_{k+1}\right]P_{k+1|k} \tag{6.21b}$$

$$K_{k+1} = P_{k+1|k}H_{k+1}^{T}\left[H_{k+1}P_{k+1|k}H_{k+1}^{T} + R_{k+1}\right]^{-1} \tag{6.21c}$$

where H_{k+1} is the matrix whose elements are given by

$$H_{ij} = \left.\frac{\partial h_i(x(t_{k+1}))}{\partial x_j}\right|_{x(t_{k+1}) = \hat{x}_{k+1|k}} \tag{6.21d}$$

The formulae given in this section were based on truncated Taylor series approximations for computing the estimates. By being linearized about $\hat{x}(t)$, the equations show similarity with the Kalman filter equations for linear systems. They are therefore often referred to as the Extended **Kalman filter** propagation equations.
The Extended Kalman filter has been found to yield accurate estimates in a number of practical applications. However, when measurement and dynamic non-linearities become stronger, higher order methods such as the ***Iterated extended Kalman filter*** or a ***second-order filter*** yield better estimates at the expense of greater computational burden.

6.5 Implementation of the Kalman filter

6.5.1 Dynamic model of the submarine

The motions of the submarine are described by the theory of hydrodynamics. To model the ship dynamics a mathematical model, describing the relation between the input (propeller revolutions and rudder angle) and the state of the ship, is used. The

manoeuvring model described by Inoue, which is based on the Newtonian force equations, can be used best as mathematical model for the Kalman filter [Wulder, 1992].
The model implemented here is a very simplified version of the 2D model described by Wulder [1992]; no external forces or moments are considered. As the submarine can also navigate underwater, making its manoeuvring space 3D, the equations are rewritten to allow for this third dimension.

The dynamic manoeuvring model of the submarine will be described by the following set of differential equations :

$$\dot{x}(t) = v(t)\cos(\epsilon(t))\sin(\psi(t)) \quad (6.22a)$$
$$\dot{y}(t) = v(t)\cos(\epsilon(t))\cos(\psi(t)) \quad (6.22b)$$
$$\dot{z}(t) = v(t)\sin(\epsilon(t)) \quad (6.22c)$$
$$\dot{v}(t) = 0 \quad (6.22d)$$
$$\dot{\psi}(t) = 0 \quad (6.22e)$$
$$\dot{\epsilon}(t) = 0 \quad (6.22f)$$

where

x,y,z = position of the submarine
v = longitudinal ship velocity
ψ = submarine's heading
ϵ = inclination angle

According to this set of differential equations, the motion of the submarine is considered to be uniform along a straight line, taking accelerations as disturbances. This initial assumption can be made as long as the submarine can be considered to be a low dynamic system and a relatively high sample rate (approximately once every second) of gyro compass, log and inclinometer can be achieved.

Three main external disturbances, which influence the ship dynamics, act on the submarine :

i. waves : described by their height and direction with respect to the course steered. It acts on all motions of the submarine;
ii. wind : causes a force and moment on the submarine depending on the relative speed and direction of the wind (with respect to the submarine's heading and speed) and on the shape of the submarine;
iii. current : causes the submarines heading and speed to be different from its ground heading and speed.

From these, only the current will be considered, as this is the only disturbance that will influence the submarine's motion both when the submarine is at the sea surface and when submerged. It can be modelled as a velocity vector

$$\mathbf{v}_c(t) = \begin{pmatrix} v_{cx}(t) \\ v_{cy}(t) \\ v_{cz}(t) \end{pmatrix} = \begin{pmatrix} v_c \sin(\psi_c(t)) \\ v_c \cos(\psi_c(t)) \\ 0 \end{pmatrix} \qquad (6.23)$$

where

v_c = current velocity
ψ_c = current direction

that can be added to the submarine's velocity vector in order to obtain the vessel's groundspeed and course.

Since current velocity and direction are not measured, they have to be estimated by the Kalman filter. The initial values must be given by the navigator, based on local knowledge. The current speed vector is assumed to be constant. However, current direction and speed change over a period of approximately 12^h30^m when the tidal motion is semi-diurnal, or 25^h when the tidal motion is diurnal. As an approximation of the uncertainty in the current model, the following can be used [Wulder, 1992]

$$\dot{v}_{cx}(t) = V_{cm}\omega_{cv} + V_{cm}\omega_{c\psi} \qquad (6.24a)$$
$$\dot{v}_{cy}(t) = V_{cm}\omega_{cv} + V_{cm}\omega_{c\psi} \qquad (6.24b)$$

where

V_{cm} = maximum current velocity
ω_{cv} = frequency of the current velocity
$\omega_{c\psi}$ = frequency of the current direction

Combination of the manoeuvring model (6.22) and current disturbance (6.23) makes up the complete system model to be implemented, which is given by the following set of equations

$$\dot{x}(t) = v(t)\cos(\epsilon(t))\sin(\psi(t)) + v_{cx}(t) \qquad (6.25a)$$
$$\dot{y}(t) = v(t)\cos(\epsilon(t))\cos(\psi(t)) + v_{cy}(t) \qquad (6.25b)$$
$$\dot{z}(t) = v(t)\sin(\epsilon(t)) \qquad (6.25c)$$
$$\dot{v}(t) = 0 + w_v(t) \qquad (6.25d)$$
$$\dot{\psi}(t) = 0 + w_\psi(t) \qquad (6.25e)$$
$$\dot{\epsilon}(t) = 0 + w_\epsilon(t) \qquad (6.25f)$$
$$\dot{v}_{cx} = 0 + w_{cx} \qquad (6.25g)$$
$$\dot{v}_{cy} = 0 + w_{cy} \qquad (6.25h)$$

where

w_v = noise signal acting on the longitudinal velocity
w_ψ = noise signal acting on the heading
w_ϵ = noise signal acting on the inclination angle
w_{cx} = noise signal acting on the x-component of the ship's ground velocity caused by uncertainties in the current model
w_{cy} = noise signal acting on the y-component of the ship's ground velocity caused by uncertainties in the current model

The noise signals describe the uncertainty of the assumption that the derivatives (6.25d) - (6.25h) are zero.

6.5.2 Observation equations

Now the system model has been defined, the observables need to be discussed. Measurements are taken at discrete moments in time and are related to the elements of the state vector by observation equations as given by equation (6.15b).

- submarine's horizontal position observable

  The horizontal position of the submarine is determined by using observations from EPF systems, bearings and distances. These observations are combined using the least squares algorithm as discussed in chapter 5. Therefore, the submarine's position observable can be written as

$$\begin{pmatrix} x_{LSE} \\ y_{LSE} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + v \qquad (6.26a)$$

  where
  v is a noise vector having zero mean and error VCV matrix given by $C_{\hat{x}}$.

- submarine's vertical position observable

  The depth of the submarine with respect to the sea surface is observed by the pressure sensor. The relation between the measured depth and true depth is given by

$$z_m = z + v_p \qquad (6.26b)$$

  where

z_m depth given by the pressure sensor

v_p a random noise signal with zero mean and standard deviation σ_p acting on the measurement.

In order to agree with the definition given for zero depth, the depth given by the pressure sensor should also be corrected for tidal height. This can be done using a tide prediction program giving the height of tide above or below MSL. A prediction program has not been incorporated in the simulation program and is a matter for further investigation since tidal ranges of a few metres are quite common, therefore introducing a low frequency systematic error which may be significant.

· submarine's velocity observable

The speed of the submarine through the water is observed by an EM log. The relation between the measured speed and true speed will be given by (see section 3.2 of chapter 4)

$$v_m = v + \Delta v + v_v \tag{6.27}$$

where

v speed given by EM log

Δv systematic deviation of the log

v_v a random noise signal with zero mean and standard deviation σ_v acting on the speed measurement

· submarine's heading observable

The submarine's heading is measured by using a gyro compass, giving the course steered through the water. The relation between the true heading and measured heading will be given by (see section 3.3 of chapter 4)

$$\psi_m = \psi + \Delta\psi + v_\psi \tag{6.28}$$

where

ψ course given by gyro compass

$\Delta\psi$ deviation of the compass caused by unknown part of the speed error, instrument correction etc.

v_ψ a random noise signal with zero mean and standard deviation σ_ψ acting on the heading measurement

· submarine's inclination angle

The inclination of the submarine is measured by an inclinometer. The relation between the measured inclination angle and true angle will be given by

$$\epsilon_m = \epsilon + \Delta\epsilon + v_\epsilon \tag{6.29}$$

where

ϵ	inclination angle given by inclinometer
$\Delta\epsilon$	systematic deviation of the inclination angle caused by trim, ship motion etc.
v_ϵ	a random noise signal with zero mean and standard deviation σ_ϵ acting on the inclination angle measurement

Although the measurement noise of the sensors described above shows time-correlation in real life, it is considered to be zero mean Gaussian white noise for simplicity, since measuring of correlation times in order to be able to use shaping filters falls outside the scope of this thesis.

6.5.3 The State vector and measurement-vector

The state vector contains the variables describing the dynamics of the submarine, which are

- position : x, y, z
- speed : v
- heading : ψ
- inclination angle : ϵ
- current velocities : v_{cx}, v_{cy}

This state vector is extended with parameters used to eliminate the systematic errors in the (measurement) sensors, which are

- systematic error in log : Δv
- systematic error in gyro compass : $\Delta\psi$
- systematic error in inclinometer : $\Delta\epsilon$

The complete state vector to be used now reads :

$$x = (x, y, z, v, \psi, \epsilon, v_{cx}, v_{cy}, \Delta v, \Delta\psi, \Delta\epsilon)^T$$

The measurement vector contains the data given by the sensors available plus the positional data derived from observations using least squares (in the event enough observations were available to do this), which gives

- horizontal position : x_{LSE}, y_{LSE}
- vertical position : z_m
- log speed : v_m
- heading of gyro compass : ψ_m
- inclination angle : ϵ_m

The complete measurement vector therefore reads :

$$z = (x, y, z, v, \psi, \epsilon)^T$$

6.5.4 Discrete-time Kalman filter equations

The discrete-time Kalman filter algorithm is given by equations (6.20) and (6.21). Using the information given in subsections 6.5.1 to 6.5.3, the linearized model can be derived.

1. propagation of the state vector

The propagation of the state vector is derived from the dynamic system model equations (6.25) using equations (6.20a) and (6.20d), resulting in

$$\begin{bmatrix} x \\ y \\ z \\ v \\ \psi \\ \epsilon \\ v_{cx} \\ v_{cy} \\ \Delta v \\ \Delta \psi \\ \Delta \epsilon \end{bmatrix}_{k+1|k} = \begin{bmatrix} x + v\cos(\epsilon)\sin(\psi)\Delta t + v_{cx}\Delta t \\ y + v\cos(\epsilon)\cos(\psi)\Delta t + v_{cy}\Delta t \\ z + v\sin(\epsilon)\Delta t \\ v \\ \psi \\ \epsilon \\ v_{cx} \\ v_{cy} \\ \Delta v \\ \Delta \psi \\ \Delta \epsilon \end{bmatrix}_{k|k} \qquad (6.30)$$

2. transition matrix (Φ) and system disturbance matrix (Γ)

In order to be able to calculate the error VCV matrix of the propagated state, the transition matrix Φ_k and disturbance matrix Γ_k are needed. They are derived from the system model by linearizing $\mathbf{f}(\hat{\mathbf{x}}_{k|k}, t_k)$ using equations (6.20e) and (6.20c) respectively

$$\Phi_k = \begin{bmatrix} 1 & 0 & 0 & F_2\Delta t & vF_1\Delta t & -vF_3\Delta t & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & F_1\Delta t & -vF_2\Delta t & -vF_4\Delta t & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 1 & \sin(\epsilon)\Delta t & 0 & v\cos(\epsilon)\Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{k|k} \tag{6.31}$$

$$\Gamma_k = \begin{bmatrix} \frac{1}{2}F_2\Delta t^2 & \frac{1}{2}vF_1\Delta t^2 & -\frac{1}{2}vF_3\Delta t^2 & \frac{1}{2}\Delta t^2 & 0 & 0 & 0 & 0 \\ \frac{1}{2}F_1\Delta t^2 & -\frac{1}{2}vF_2\Delta t^2 & -\frac{1}{2}vF_4\Delta t^2 & 0 & \frac{1}{2}\Delta t^2 & 0 & 0 & 0 \\ \frac{1}{2}\sin(\epsilon)\Delta t^2 & 0 & \frac{1}{2}v\cos(\epsilon)\Delta t^2 & 0 & 0 & 0 & 0 & 0 \\ \Delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta t \end{bmatrix}_{k|k} \tag{6.32}$$

where

$$F_1 = \cos(\epsilon)\cos(\psi)$$
$$F_2 = \cos(\epsilon)\sin(\psi)$$
$$F_3 = \sin(\epsilon)\sin(\psi)$$
$$F_4 = \sin(\epsilon)\cos(\psi)$$

3. observation matrix

Since the observation equations as given in subsection 6.5.2 are all linear with respect to the state variables, no linearization is needed and the observation matrix is thus given by

$$H_{k+1} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} \qquad (6.33)$$

4. Error VCV matrices Q_k and R_k

In order to be able to use the filter, statistical models of the system and measurement noise processes (**w** and **v** respectively) have to be estimated. This is normally done by using model identification techniques in which output of the model based on measurements up to time t_k is compared with observations at time t_k. Measurement data to be used has to be collected during sea trials.
The way model identification is performed falls outside the scope of this paper, but the reader is referred to Gelb [1974] for a description.

As no sea trials were performed as part of the research, the variances of the different noise processes are estimated to the writer's best knowledge.

The statistics of the system disturbances are given by the error VCV matrix Q_k. As zero mean gaussian white noise processes are assumed, the VCV matrix is diagonal and defined as

$$E[\mathbf{w}_k \mathbf{w}_l^T] = Q_k \delta_{kl} \qquad (6.34)$$

The random process vector $\mathbf{w}_k$ contains all the statistical information of the disturbances acting on the system described by equation (6.25)

$$w_k = (w_v, w_\phi, w_\epsilon, w_{vcx}, w_{vcy}, w_{\Delta v}, w_{\Delta\phi}, w_{\Delta\epsilon})^T$$

The terms $w_{\Delta v}, w_{\Delta\phi}$ and $w_{\Delta\epsilon}$ have been included to describe the uncertainty of the assumption that the derivatives of Δv, $\Delta\phi$ and $\Delta\epsilon$ are zero.

The diagonal elements (variances) of Q_k are estimated as follows:

- acceleration (w_v)

  The pdf for accelerations will be seen as a combination of a discrete and continuous function. If the probability of having no acceleration is p_0 and the probability of acceleration at maximum rate is p_{max}, while the acceleration probability for all other accelerations is described by a uniform pdf, the variance of acceleration is given by [Gelb,1974] :

$$E[w_v^2] = \sigma_a^2 = \frac{a_{max}^2}{3}\left[1 + 4p_{max} - p_0\right] \qquad (6.35)$$

  If a maximum acceleration of $a_{max} = 0.25\ m / s^2$ is assumed and the probability for maximum acceleration and no acceleration are estimated to be 0.01 and 0.75 respectively, the variance is equal to $\sigma_a^2 = 0.006\ m^2 / s^4$.

- rate of turn (w_ϕ, w_ϵ)

  The standard deviation of rate of turn in horizontal and vertical direction is taken to be half the maximum rate of turn.

- rate of change of current vector (w_{vcx}, w_{vcy})

  Uncertainties in the current model are caused by the assumption that the current vector remains unchanged. As this is not the case, the rate of change has to be estimated. As this is difficult to do, only the maximum rate of change is given (eq 6.24).
  The standard deviation is taken to be half of the maximum rate of change.

In the computer simulation model, a semi diurnal tide with maximum velocity of 1 m / s (2 knots) is assumed, leading to the following variances :

x direction : $\sigma^2_{vcx} = 2\ 10^{-8}\ m^2 / s^4$

y direction : $\sigma^2_{vcy} = 2\ 10^{-8}\ m^2 / s^4$

- rate of change of systematic errors ($w_{\Delta v}, w_{\Delta \phi}, w_{\Delta \epsilon}$)

  It is assumed that the rate change of systematic errors of gyro, log and inclinometer are negligible compared to the errors caused by accelerations and drift rate.
  In the computer program it is therefore assumed that the variances are zero.

The statistics of the measurement disturbances are given by the error VCV matrix R_k, which is defined as

$$E\left[v_k v_l^T\right] = R_k = \left[\begin{array}{c|cccc} C_{\hat{x}} & & 0 & & \\ \hline & \sigma_p^2 & 0 & 0 & 0 \\ & 0 & \sigma_v^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\phi^2 & 0 \\ & 0 & 0 & 0 & \sigma_\epsilon^2 \end{array}\right] \qquad (6.36)$$

where $C_{\hat{x}}$ is the error VCV matrix of the position obtained via least squares as given in chapter 5, and as values for σ_p, σ_v, σ_ϕ and σ_ϵ the standard errors of the sensors as given in chapter 4 are used.

5. Initial state

In order to be able to use the Kalman filter, the values of the state vector at k = 0 and its error VCV matrix must be known. They are defined as follows :

· state vector

$$\hat{x}_{0|0} = (x(0),y(0),z(0),v(0),\psi(0),\epsilon(0),v_{cx}(0),v_{cy}(0),0,0,0)^T$$

with

$x(0),y(0),z(0)$ a start position obtained from external information resulting in an initial estimate of the MPP by means of the least squares algorithm. If the least squares estimate of the MPP cannot be calculated, either because not enough LOPs are available or when the submarine is submerged, the initial position given by the navigator is taken as starting position

$$v_{cx}(0) = v_c(0)\,\sin(\psi_c(0))$$
$$v_{cy}(0) = v_c(0)\,\cos(\psi_c(0))$$

$v_c(0)$ current speed given by the navigator
$\psi_c(0)$ current direction given by the navigator

$$\psi(0) = \psi_{GK}(0)$$

· error VCV matrix

$$P_{0|0} = I$$

where
I is the identity matrix.

6. Confidence areas

In section 5 of chapter 5, expressions were given to obtain the error ellipse (2D) and error ellipsoid (3D). These equations can also be used to calculate the confidence regions of the MPP calculated by the navigation filter. In the equations (5.21) and (5.26), $C_{\hat{x}}$ should then be replaced by the top left-hand submatrix of $P_{k+1|k+1}$.

6.6 Implementation considerations

When the Kalman filter is implemented on a computer, some problems affecting the correct working are encountered. These problems can roughly be divided into two groups :

A. problems related to the model

In order to obtain a truly optimal filter, an exact description of the system dynamics, error statistics and measurement process are assumed to be used. This is not always the case either because an exact model would lead to a computational burden too great for the computer used or because exact statistical characteristics are simply not known.
Therefore, the system is normally simplified leading to a sub-optimal filter. As a result, a discrepancy between the filter- and theoretical state, referred to as divergence, can start to exist. To minimize this discrepancy, it is important to carefully evaluate the dynamic model used. This is normally done by sensitivity analysis and error budget calculations for which special computer test programs are developed. This falls outside the scope of this thesis.

B. problems related to the computer

Firstly, there are the constraints imposed by the computer. These are mainly of physical nature such as size, weight, peripherals, memory capacity and processor speed. This will lead to reduction of the complexity of the Kalman filter equations in order to reduce computational burden. Deleting states can be used as a method of simplifying the model - at the cost of introducing problems discussed in point A -, whereas prefiltering can be used when measurements are available more frequently than possible to process them.
Secondly, the finite nature of the computer will lead to truncation errors when approximating integration and differentiation using numerical algorithms - as is the case when using difference equations instead of differential equation -, whereas finite word length of the computer memory will lead to rounding errors. The latter could lead to unsensible results. One should therefore very carefully select algorithms used.

In order to obtain the 'best' Kalman filter for a certain application, the following steps should be part of the development process :

1. obtain a description of the initial model and error statistics;
2. implement the model on the computer;
3. perform sensitivity analysis and error budget calculations;
4. test the filter in real-time environment;
5. evaluate model and use results to modify the model.

Steps 2 through 5 need to be repeated until the chosen model satisfies the real-time situation best. In this thesis only steps 1 and 2 are performed. Steps 3 to 5 are suggested to be part of further investigation to improve the model as given in this chapter.

6.7 Statistical testing

As is the case with the least squares, statistical tests have to be performed to validate the reliability of the Kalman filter. The tests described here are based on Teunissen [1990] and Lu [1992], but parameters used in the reference have been adjusted to agree with those used in chapter 5 of this paper. This way the comparison with the test performed to detect outliers in observations becomes apparent.

The tests described are used to detect outliers originating from model misspecification. Two tests can be distinguished : the Local Overall Model (LOM) test and the Local Slippage (LS) test. Both tests look at only one epoch at a time and are used for detection and identification of outliers. The tests can be expanded to Global tests (GOM,GS), when more epochs are taken into account. The global tests are used to find drifting errors. In the computer simulation program, only the LOM and LS test are implemented.

The LOM and LS test statistics are based on the predicted residual which is given as

$$\hat{v}_{k+1} = z_{k+1} - \hat{z}_{k+1|k} = z_{k+1} - H_{k+1}\hat{x}_{k+1|k} \tag{6.37a}$$

having as error VCV matrix

$$C_{\hat{v}_{k+1}} = H_{k+1}P_{k+1|k}H_{k+1}^{T} + R_{k+1} \tag{6.37b}$$

Local Overall Model test

The Overall Model test is performed to detect whether an outlier is present or not. It is a test on the overall validity of the mathematical model.

If the dynamic model is valid, the predicted residuals should represent a Gaussian zero mean white noise process. If the model is not correct due to misspecification, one or more of the predicted residuals will not have zero mean any more but will show a bias. Therefore the following hypotheses are tested :

$$H_0^{k+1} : \hat{v}_{k+1} \sim N\left(0, C_{\hat{v}_{k+1}}\right) \tag{6.38a}$$

$$H_1^{k+1} : \hat{v}_{k+1} \sim N\left(m_{\hat{v}_{k+1}}, C_{\hat{v}_{k+1}}\right) \tag{6.38b}$$

where

$m_{\hat{v}}$ = mean of predicted residuals in the event of an outlier

The test statistic for local detection is given as

$$T_{LOM}^{k+1} = \frac{\hat{v}_{k+1}^T C_{\hat{v}_{k+1}}^{-1} \hat{v}_{k+1}}{n_{k+1}} \tag{6.39}$$

where

n number of observations in measurement vector z

It can be shown that this test statistic has a chi-square distribution with n_{k+1} degrees of freedom. H_0 is accepted if the test statistic is within the confidence region, i.e. :

$$P\left[\chi^2_{\frac{\alpha}{2}, n_{k+1}} < T_{LOM}^{k+1} < \chi^2_{1-\frac{\alpha}{2}, n_{k+1}}\right] = 1 - \alpha \tag{6.40}$$

If the test statistic falls outside the region, an unspecified local model error is considered to be present.

Local Slippage Test

Once an outlier is suspected to exist as result of the LOM test, the potential source has to be identified. This identification is performed by the Slippage test, which is the next step in the identification process.

To identify the source of the outlier, a new test statistic is defined

$$t_{LS_i}^{k+1} = \frac{e_i^T C_{\hat{v}_{k+1}}^{-1} \hat{v}_{k+1}}{\sqrt{e_i^T C_{\hat{v}_{k+1}}^{-1} e_i}} \qquad (6.41)$$

If the observation i does not contain a blunder (H_0), the test statistic will be a normalized statistic, having a normal distribution with zero mean and standard deviation one. Under H_1 the test statistic will have a normal distribution with a non-zero mean and standard deviation one.
The test statistic is calculated for each observation at t_{k+1} and checked if it is within the confidence region as described in section 5.4.1 where w_i is replaced by t_{LS}. If for one or more observations H_0 is rejected, then the observation having the largest absolute value of l_{LS} is considered to be the observation containing the outlier.

Once an outlier has been detected and identified, the state vector has to be corrected. This will be done by estimation of the magnitude of the error and its associated variance. This can be achieved by using a two-stage Kalman filter [Lu,1992]. This will not be considered further in this paper.

6.8 Concluding remarks

The navigation filter is derived by means of the ship dynamic model and Kalman filter theory. It estimates variables concerning the 3D motions of the submarine, i.e. position, heading and inclination, and some variables concerning the sensor dynamics, i.e. deviations of log, gyro and inclinometer.
The ship's manoeuvring model used is a very simple model in which the submarine is assumed to sail with constant velocity along a straight line. The result of this is

- when the submarine accelerates or decelerates, the predicted

positions based on the dynamic model will lag the submarine's true position;
- when the submarine changes course and/or depth, again accelerations are introduced, this time with the effect that the predicted positions will show overshoot.

When the submarine is at the sea surface, additional data provided by the unbounded sensors is available, which will make the predicted position to agree better with the submarine's true position after some time. When the submarine is submerged, only bounded sensor can be used with sparse additional positional information provided by depths measurements (bathy LOP) obtained using the echo sounder. This means that no corrections can be applied to counteract the above mentioned effects due to accelerations. Errors introduced due to the accelerations will not be detected. This results in accumulative errors, making position accuracy provided by the navigation filter less reliable.
To improve performance, the filter gain (K) could be decreased when changing speed and/or course, putting more confidence in the position obtained from the external sources. This is again only possible when the submarine is at the sea surface. It is therefore important to assess how position accuracy degrades in time due to model disturbances. This depends very much on the system model used and sensor data sampling interval.

It is also important to consider the effect of errors in the initial state vector as this will give an indication of the filter capability to react to sudden changes in the environment.

For submarine navigation, currents are considered to be the most important disturbances. This is because their direction and/or velocity may change significantly with depth in general, or with position when the submarine is operating near coastlines in particular. These sudden disturbances can influence filter performance considerably when not taken into account.
The response of the filter may be improved by choosing large standard deviations of system noise at the beginning of the run and to decrease this later on. This way, the influence of current on position accuracy will be small at an early stage and becomes more important once the values are estimated. The simulation program has a provision to scale the error VCV matrix Q_k. Again trials at sea are necessary to find the correct error VCV matrix and 'filter setting' to be used.

7. The Computer Simulation Program

The third aim of the research was to develop a computer simulation program that shows the main features of integrated navigation and that can serve as basis for an integrated system to be implemented in a real-time environment on board submarines. In the previous chapters an inventory was made of the systems and sensors available for position fixing (chapters 3 and 4) and the mathematical and statistical models to be used for position calculation and quality control (chapters 5 and 6). In this chapter a concise description will be given of the computer simulation program, which is based on the information provided by the previous chapters.
The chapter starts with a section giving technical details about the program itself such as hardware and software requirements plus an overview showing the relationship between the main program and routines used. In the next section the routines used for position fixing and quality control will be described in more detail. This section will however not provide an in-depth description of the software. The final section of this chapter provides the user of the simulation program with a concise manual.

7.1 Technical description

Hardware requirements

The program development had started on an Intel 8086 based PC. Soon it became obvious that the calculations performed to process input data to obtain a position with associating confidence region and to perform the statistical tests, required a more advanced PC. By now, the program has been tested on several PCs, using several types of screens, for correct working. In order to get a realistic picture of performance (i.e. short calculation times in order to get the high sampling rate required), it is advised to use at least a 80286 based computer or equivalent with maths co-processor. The program does not use colours so a monochrome screen will suffice. The program presumes a hard disk to be present.

Software requirements

The computer simulation program is provided on one 3.5 inch floppy disk. This disk contains an installation program, the simulation program, a sensor input data file plus additional data files used by the program. The program runs under the DOS

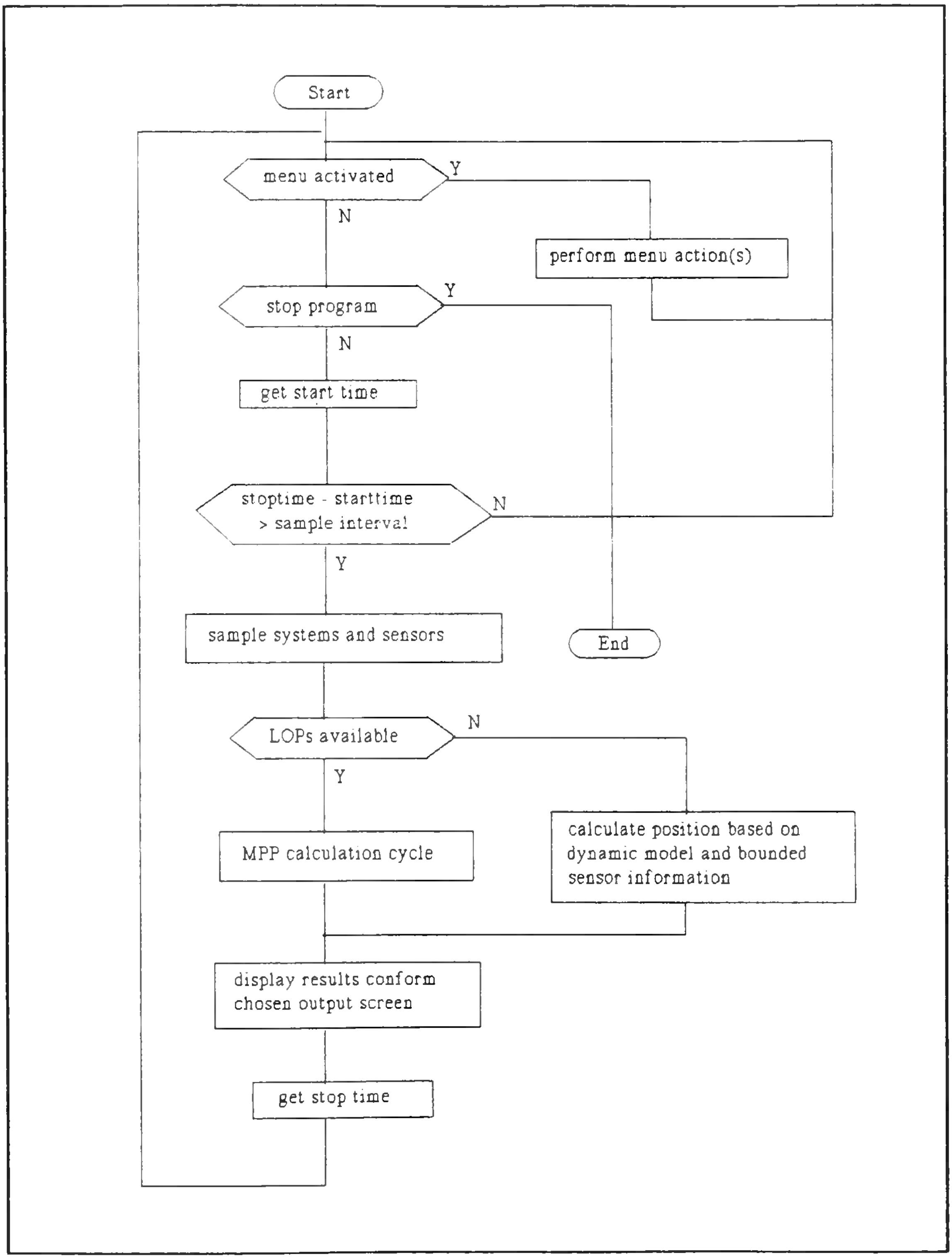


Figure 7.1 Main loop of the simulation program

3.3 Ships bounded and unbounded systems and sensors

In this section all remaining sensors and systems on board a submarine that can be used for position fixing, whether it gives an MPP based on DR or external information, are discussed. We consider two main groups of sensors and systems : the bounded and unbounded. Bounded means that no external equipment providing information is needed for the sensor or system to give its information - i.e. they are self-contained (autonomous) -, whereas unbounded means that the sensor or system needs information from the outside. This classification implies that based on bounded systems and/or sensors only, a DR position will be derived.
All sensors and systems described in this subsection can be used both when the submarine is at the sea surface and when it is submerged.

3.3.1 Ships Inertial Navigation System

The Ships Inertial Navigation System (SINS), is a self contained DR device. It employs an Inertial Measurement Unit (IMU) with three gyros and three accelerometers to provide continuous output of the following data :

- ship's geographic position;
- horizontal an vertical linear velocity components;
- angle magnitude and rate for heading, roll and pitch.

The accelerometers are positioned on a platform in the three main directions so forming an XYZ - 3D coordinate system. By integrating accelerations measured, both speed and displacement in the three main directions can be calculated. In order to be able to use this information to calculate the present position, the orientation of the system of accelerometers needs also to be fixed with respect to a terrestrial reference frame. This is done by aligning one accelerometer with the East/West direction, one with the North/South direction and one in the vertical plane at right angles to the other two.
To yield useful information, the platform on which the accelerometers are mounted, must be constantly maintained in a known orientation, which is chosen to be horizontal, i.e. at right angles to the direction of gravity at any point. This can be achieved by suspending the platform in a gimballed system and using gyros which are also fixed on the platform. Any deviation of the platform from the horizontal will produce an angular rate which is used in conjunction with additional terms provided by the computer, to torque one or more gyros and drive the platform to the horizontal level. This way a closed undamped loop is formed, whose period is equal to the Schuler period.

3.2.3 Radar

Radar is a user-based microwave transmitter/receiver with a rotating antenna transmitting short microwave pulses in all directions sequentially and receiving the echoes from its own pulses reflected by surrounding obstacles. All the reflected signals are processed and displayed on a radar screen with their correct individual range and bearing.
The radar equipment is operated at the user's vessel.

- signal characteristics : The two main frequency bands used for marine radar are 2920 - 3100 MHz and 9320 - 9500 MHz.

- availability : This depends on the position of the vessel with respect to objects that can be used for navigational purposes.

- coverage : is practically limited by the characteristics of the radar installation (transmitted power, radio horizon and receiver sensibility) and is normally up to 45 km (25 nm) for shipborne radars, but can be as high as 75 km (40 nm) under good conditions.

- fix rate and dimensions : Once objects that can be used for navigation are present, radar can be used continuously for position fixing. Normally the antenna makes 20 - 30 turns per minute continuously updating the display. From each object bearing and distance can be measured giving a LOP.

  LOPs obtained from two or more different objects provide a 2D position.

- ambiguity : There is in general no ambiguity. Only objects very close to the transmitter/receiver might give a second echo on the display, but in general this is not the case. Sea clutter and interference might give confusion, but proper processing of the incoming signals reduces this very much.

- accuracy : The accuracy of a position derived from radar information depends on the distance of the object from the transmitter/receiver. For one object the predictable accuracy of a bearing is approximately 1° (1σ) whereas the accuracy of a distance measurement is about 0.01 (1σ) part of the range scale selected. The predictable accuracy of a position fix depends on the geometry of the navigational objects used.

series of the master station is called the Group Repetition Interval (GRI) of the chain. This GRI is made unique for each chain, making chain identification possible.
The system is based upon measurement of the difference in time of arrival (TOA) between pulses from the master station and the secondaries. The measurements of Time Difference (TD) are made by a receiver which achieves high accuracy by comparing the zero crossing of a specified cycle - the so-called **'time reference point'** (TRP) - within the pulse transmitted by the master and the TRP in the pulses from the secondary stations within the chain. The TRP is obtained in two stages : first the proper cycle to be used for zero-crossing tracking is identified making use of the pulse envelope (Cycle Identification or coarse measurement). Once the proper cycle has been found the exact moment of zero crossing within the cycle is determined by means of a phase tracking loop (fine measurement), triggering a timing mechanism. The TRP is chosen to be the zero crossing in the third cycle because at that point the signal has adequate amplitude and the pulse will not yet be affected by sky wave interference.

- signal characteristics : All Loran-C stations transmit at 100 kHz using time-division multiplexing.
  Each pulse has a well defined envelope. Because of the short rise-time of the pulses, their transmissions occupy a relatively broad frequency band, 99 percent of their energy being distributed between 90 and 110 kHz.
  Additionally, the pulses transmitted by the stations are phase coded. Master and secondary stations each have a different phase code functions.

- availability : The Loran-C transmitting equipment is very reliable. Redundant transmitting equipment is used to reduce system downtime, making availability better than 99%.

- coverage : as shown in figure 3.3.

- fix rate and dimensions : Once the receiver is in the coverage area, 10 to 20 independent position fixes can be made per second depending on the GRI used. Two or more LOPs can be obtained, giving a 2D fix.

- ambiguity : As with all hyperbolic systems, ambiguity is present. However, because of the design of the coverage area of each chain, the ambiguous fix is at a great distance from the desired fix and therefore easily resolved.

- accuracy : Within the coverage area, Loran-C will provide the user with predictable accuracy of less than 450 m (2 d_{rms}). The accuracy is, however, highly dependent on the GDOP at user's location.

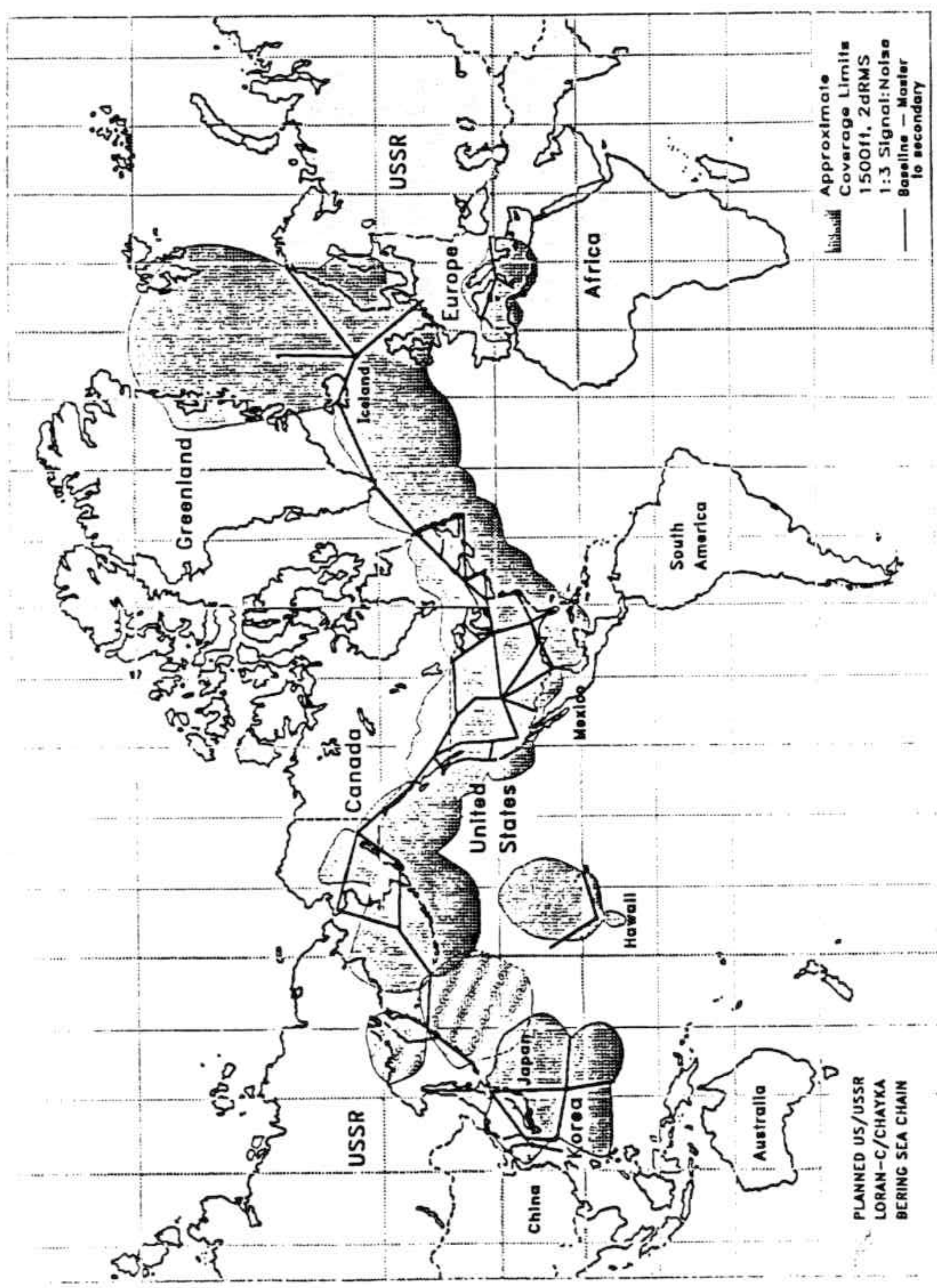


Figure 3.3 Coverage provided by U.S. operated or supported Loran - C stations [FRP,1990]

- signal characteristics : Each stations transmit signals on four frequencies in the following order: 10.2 kHz, 13.6 kHz, 11 1/3 kHz and 11.05 kHz. In addition to these four common frequencies, each station transmits a unique frequency to aid station identification and to enhance receiver performance. A combination of time- and frequency sharing is used to distinguish between signals from the different transmitters. The signal transmission format has a cycle of ten seconds.

- availability : Annual system availability has been greater than 97% with scheduled off-air time included. Equipment redundancy has been designed into nearly all functions of the Omega transmission process, which will contribute to station reliability and availability.

- coverage : Essentially worldwide coverage.

- fix rate and dimensions : Omega provides an independent position fix once every 10 seconds. Two or more LOPs can be obtained giving a 2D position fix.

- ambiguity : When the system is used in hyperbolic mode, cycle ambiguity is present. Single frequency receivers use the 10.2 kHz signal whose lanewidth is about 14.8 km (8 nm) on the baseline between stations. Therefore the EP needs to be known to within 7.4 km (4 nm). Multiple frequency receivers, however, extend the lanewidth for purpose of resolving lane ambiguity.

- accuracy : The accuracy of the Omega system is limited by the accuracy of the propagation corrections that must be applied to the individual lane readings. The system provides a predictable accuracy of 4 - 7 km (2 - 4 nm) (2 d_{rms}). The accuracy depends very much on receiver location, station pairs used, time of day and validity of propagation corrections.

LORAN-C

Each Loran-C chain consists of a master station and up to six secondary stations. Each station transmits a series of energy bursts at a carrier frequency of 100 kHz. The shape of the bursts is well defined. The master station transmits first after which the secondaries each transmit in turn starting after tightly controlled time intervals. This interval for each secondary is known as the Emission Delay (ED). When all stations of a chain have transmitted their series of bursts, the master station transmits again. The time interval between the start of two burst each of the stations at 10.2 kHz. Since the stations are widely spaced, each station works autonomously. In order to be able to use time-sharing, each station has four Cs-standard atomic clocks making it possible to maintain synchronization. Since the phases are synchronized, the measurements may either be taken in pairs to give hyperbolic LOPs (hyperbolic mode), or may be taken with respect to a precision time source in the receiver, to give circular LOPs (rho-rho mode).

No chains are defined. At any point on the surface of the earth at least five LOPs will be available, allowing the navigator to take advantage of LOP redundancy. For each part of the earth an optimal set of stations is given, based on geometry of stations with respect to receiver and angle of cut of LOPs derived.

Although the standard deviation of an Omega LOP is large (see section 4.2.2) compared to other EPF systems available, making the Omega system less useful as stand alone system or integrated with other navigation systems especially now GPS has become operational, one of its advantages is that underwater reception of signals might be possible because of the very low frequencies involved. However, no information on the way this can be used on board submarines of the Royal Netherlands Navy is given in this paper due to operational classification.

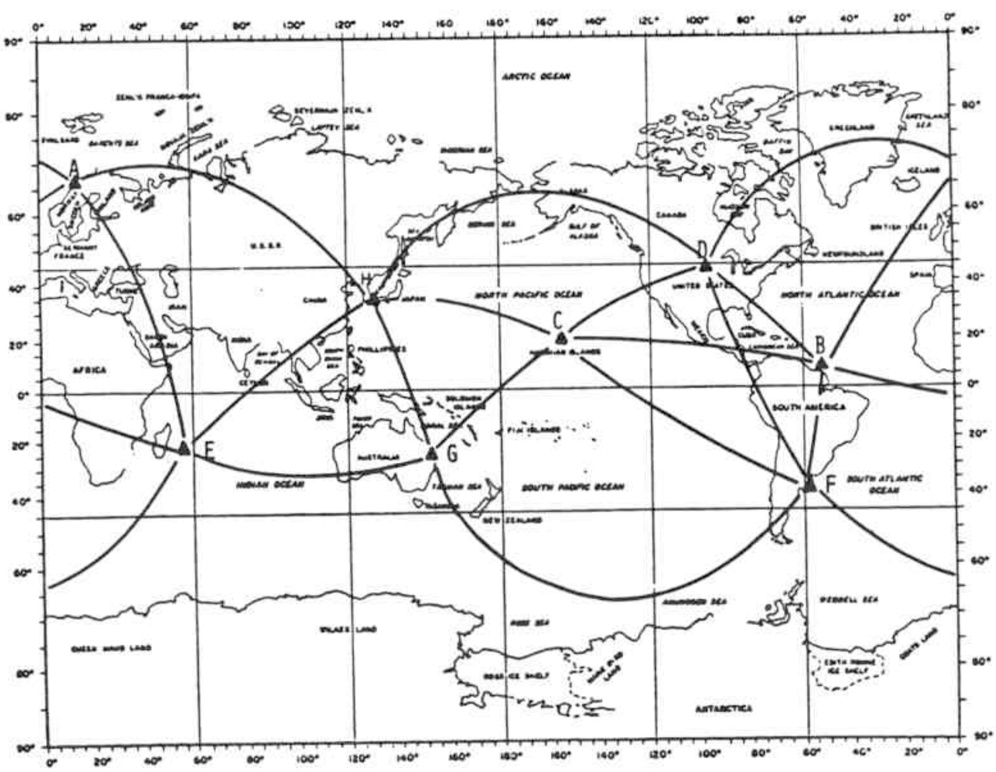


Figure 3.2 Omega System Configuration [IHR,1972]

sky wave interference (1-hop-E layer sky wave distance). Beyond this distance, a fix from MP readings is about 20 % more accurate than the fine pattern fix [Decca,1979]. This is because the MP signal has the important property that it remains stable in phase in the presence of mutual phase shifts in its constituents. As the range further increases, there comes a point at which the fine pattern becomes unreliable and the MP signals are then the sole means of fixing. The limit of night-time coverage is set by phase shifts in the MP constituents due to sky wave interference, introducing unwanted peaks in the MP signals. This is expected to occur at distances beyond 440 km (240 nm) from any station.

- fix rate and dimensions : Decca provides continuously two or three LOPs, making continuous position fixing possible. If Multi Pulse is used for position fixing, only once every 20 seconds a fix can be made.
  Having two or three LOPs provides a 2D position fix.

- ambiguity : As phase measurements are performed on carrier waves, cycle ambiguity is present. The area between lines of zero phase difference is called 'lane'. To avoid the fact that the position of the submarine needs to be known to within a few hundreds of meters (since the lanewidth of patterns on the baseline ranges from approximately 350 - 600 m), a coarse pattern is created via MP signals. The lanewidth of this coarse pattern is approximately 10.6 km on the baseline (which equals a zonewidth on the baseline), making it possible to resolve cycle ambiguity within a zone. The position now needs to be known to within half a zone-width.

- accuracy : The predictable accuracy varies from 50 m ($2d_{rms}$) at approximately 80 km from the master to 450 m ($2d_{rms}$) up to 250 km from the master. Due to sky wave interference this accuracy reduces at nighttime by a factor 6 to 8.
  Since the predictable accuracy is dependent on many factors, one needs to refer to the Decca Chain Data Sheets to find the most likely value for predictable accuracy at a given position for a certain time and even date.

OMEGA

Omega is a very low frequency (VLF) radio navigation system, comprising eight transmitting stations situated throughout the world. Worldwide coverage was achieved when the station in Australia became operational.
The basic measurement in Omega is the phase of the signal from

3.2.2 Terrestrial LF and VLF Radio Position Fixing Systems

In this section the most frequently used LF and VLF Radio Position Fixing (RPF) systems will be viewed at. They can only be used when the submarine is at periscope depth or at the surface.

Decca Navigator System

The Decca Navigator System (DNS) is a hyperbolic radio navigation system. A Decca chain consists of one master and two or three slaves, designated Red, Green and Purple. LOPs are formed by phase comparison between the signals received from the master and slaves.

- signal characteristics : The DNS utilizes unmodulated signals in the 70 - 130 kHz band. The basic frequency f_0 is about 14.2 kHz. The value of f_0 varies slightly from chain to chain. The master transmits at a frequency of $6f_0$, the slaves at $8f_0$ (Red), $9f_0$ (Green) and $5f_0$ (Purple). Phase comparison between master and slave is done at $24f_0$ (Red), $18f_0$ (Green) and $30f_0$ (Purple), resulting in a fine pattern. The necessary phase lock between the master and slave transmissions is ensured by the control equipment at the slave station. Generally, the phase of the slave signal with respect to the master signal is so adjusted that the baseline extension at the master station has the fraction value zero. At the slave station the baseline extension has the value of the residual lane fraction if the baseline is not equal to a whole number of lanes. Several chains depart from this convention, however.

  Once every 20 seconds lane identification signals, known as Multi Pulse, are transmitted from each station in turn so that the receiver can extract a signal of frequency f_0 from the master and each slave. Comparing the phase of these signals generates a coarse pattern. An additional phase difference meter in the receiver responds to this coarse pattern and gives periodic readings which indicate, in turn, the correct lane of each pattern within a known zone.

- availability : Taking downtime of the chains in consideration, the availability will be approximately 99.8%.

- coverage : At present day (December 1992) 42 chains are in use worldwide. A chain can be used up to 750 km (400 nm) from any station during daylight. At night, the accuracy of phase comparison on the fine pattern degrades at distances over approximately 200 km (110 nm) from any station due to

The main characteristics of the Transit system are as follows:

- signal characteristics : The satellites emit continuously at two frequencies : 399.968 MHz and 149.988 MHz. In practice these frequencies will vary slightly from satellite to satellite and drift with time. Because each satellite has only one oscillator, the lower frequency can be kept always 3/8 times the higher. The carrier frequencies are phase modulated in order to carry the satellite's broadcast ephemeris and to provide a specific time-marker every even numbered minute of UTC.
  The transmitted signals are right-hand circularly polarised.

- availability : the availability is better than 99% when a Transit satellite is in view. Being 'in view' depends on user latitude, antenna mask angle, user manoeuvres during satellite pas, number of operational satellites and satellite configuration.

- coverage : Coverage is worldwide but not continuously due to the relatively low altitude of the Transit satellites and the precession of the satellite orbits.

- fix rate and dimensions : Once a satellite is in view, it is visible for up to 18 minutes which gives ample information for approximately 40 positions during perfect satellite pass. The satellite wait time varies with latitude, theoretically from an average of 110 minutes at the equator to an average of 30 minutes at 80 degrees latitude. Due to non-uniform orbital precession, the Transit satellites are no longer in evenly spaced orbits. Consequently, a user can occasionally expect a period greater than 6 hours between fixes.

  Transit provides 2D positioning.

- ambiguity : there is no ambiguity

- accuracy : The predictable accuracy is highly dependent on the user's knowledge of his velocity and course. On average one gets :

  - dual frequency : 100 - 350m ($2d_{rms}$)
  - single frequency : 200 - 500m ($2d_{rms}$)

NNSS / TRANSIT

Transit is a space-based radio-navigation system consisting of three major segments :

· Space segment

There are 9 satellites, from which only 5 are operational, in approximately 1100 km polar orbits, having a period of orbit of approximately 105 minutes. The spacing of the orbits in azimuth is not uniform due to precessional effects.
Each satellite carries a radio receiver and transmitter along with a small computer, stable oscillator and crystal clock.

· Ground segment

The satellites are tracked by four main stations, forming OPNET. They provide the tracking information necessary to update satellite orbital parameters every 12 hours.
Additionally there is a number of stations all over the world tracking the satellites when in view. They form TRANET. Their tracking results are send to the U.S. DMA, which computes the 'precise ephemeris' for the satellites. A user tracking a certain satellite can use this 'precise ephemeris' for post processing rather than the real-time broadcast ephemeris.

· User segment

The user segment consists of an antenna, preamplifier, a receiver with microprocessor and power supply.
The receiver measures successive Doppler shifts as a satellite passes the user. A typical measurement interval is either approximately 30 seconds or two minutes. The Magnavox 1502 receiver, for example, uses 4 batches of 5 words (approximately 23 sec each) and the two minute marker as time gates.
The microprocessor computes the satellite's position at the beginning and the end of each Doppler count interval using the broadcast ephemeris data. It is also fed with the ship's estimated position, ground course and speed, enabling the receiver to estimate the slant ranges and differences over the time interval. The receiver then calculates the geographical position of the user based on knowledge of the satellite's position and range difference from the measured Doppler count in eg. a least squares position estimation process, either as part of an integrated or stand alone system.

ephemeris data, atmospheric propagation correction data and satellite clock bias information.
The transmitted signals are right-hand circularly polarised.

- availability : expected to approach 100% when fully operational.

- coverage : Fully operational worldwide 3D coverage will be provided. At the moment (December '92) 2D coverage is available 24 hours a day whereas 3D coverage is available over 23 hours per day.

- fix rate and dimensions : When a receiver is switched on, some time will elapse before a first position can be calculated, due to the fact that the receiver needs to acquire the satellite signals and navigation data. This so-called Time to First Fix (TTFF) can range from 30 seconds for a fully operational set to 25 minutes for a cold start using a one-channel receiver. Figure 3.1 gives an overview of TTFF that can be expected under various conditions. Once the receiver is locked on, a position, velocity and time solution will be provided approximately every second, depending on the receiver used.

  GPS provides 3D positioning and velocity fixes as well as accurate time information when fully operational.

- ambiguity : there is no ambiguity

- accuracy : Two levels of navigation are provided by the GPS, the Precise Positioning Service (PPS) and the Standard Positioning service (SPS). The PPS is a highly accurate positioning, velocity and timing service which is only made available to authorized users. The SPS is a less accurate position and timing service which is available to all GPS users.
  When the system is fully operational, the predictable accuracy will be [FRP,1990] :

	SPS	PPS
horizontal (2 d_{rms})	100 m	17.8 m
vertical (2σ)	156 m	27.7 m
time (1σ)	167 ns	100 ns

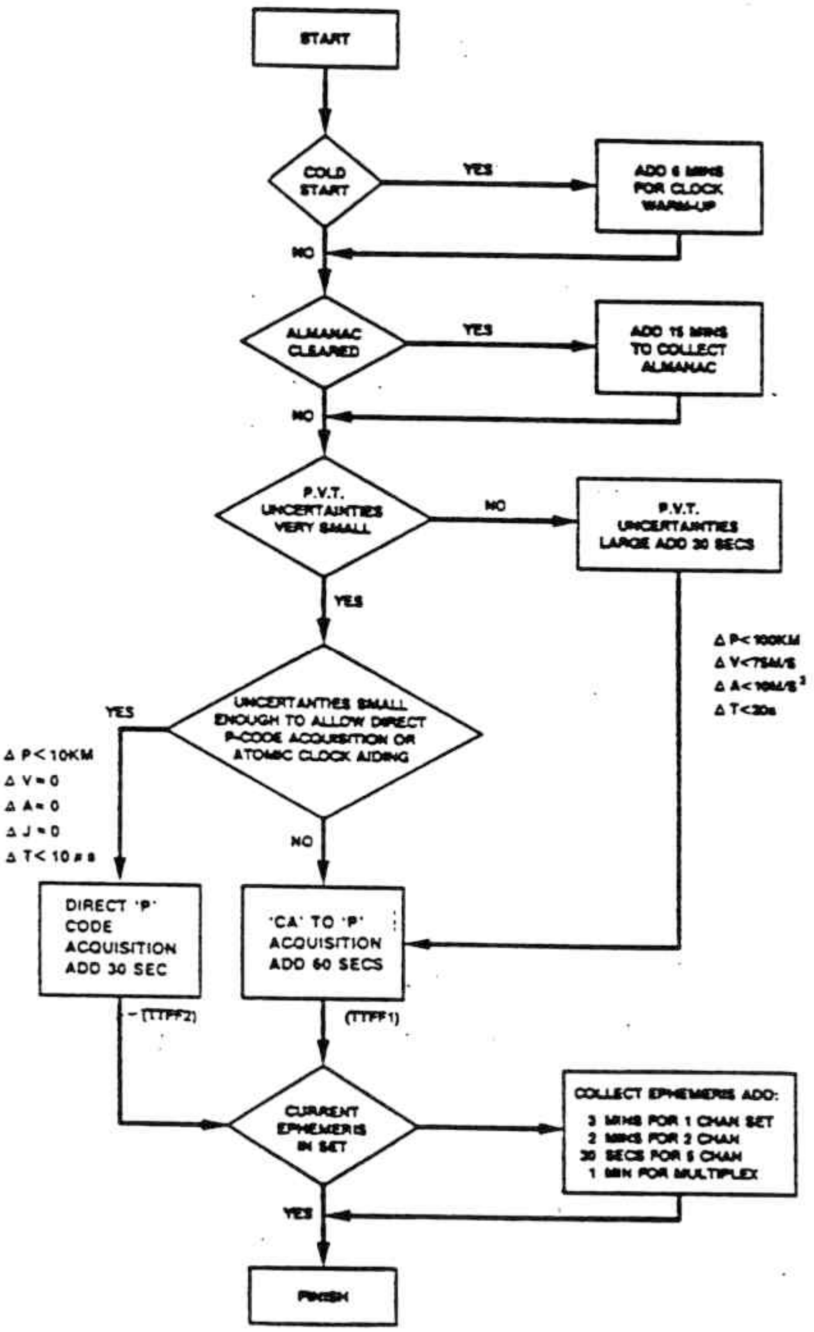


Figure 3.1 Block diagram for estimation of Time To First Fix [NATO,1991:1]

of 20200 km. The orbits have an inclination angle of 55° relative to the equator and are separated by 60° in azimuth. The period of orbit of the satellites is about 12 hours. Spacing of the satellites in their orbit will be in such a way that, when the system is fully operational, the user will have at least 4 satellites in view at any time and any place.

· Control segment

The control segment comprises five monitor stations, three ground antennas and a Master Control Station (MCS). Each monitor station is a remote unmanned station consisting of a GPS receiver, an atomic frequency standard, communication equipment and environmental sensors. The receivers track all satellites in view, accumulating range data. At the same time the clock signals are observed and compared to the station frequency standard. Local meteorological data is also measured. All this information is transmitted to the MCS where it is processed to determine and predict ephemeris data and clock-bias for each satellite. The navigation message of each satellite is updated by transmitting this information to the satellite via the Earth uplink antennas. This is done every second orbit or once about every 24 hours.
In addition to the monitor stations, the GPS satellites are also tracked by a number of semi-permanent tracking stations.

· User segment

The user segment consists of an antenna, receiver, processor and I/O devices. The receiver demodulates the navigation signals to obtain the pseudo-range and delta pseudo-range measurements. The microprocessor converts these measurements to a position, velocity and time.
Various receivers are available commercially, ranging from one channel, single-frequency (C/A-code) receivers to multi-channel, dual-frequency (C/A-code plus P-code) receivers.

The main characteristics of the GPS system are as follows :

- signal characteristics : Each satellite continuously transmits at two L band frequencies : L1 (1575.42 MHz) and L2 (1227.60 MHz). Using Bi-Phase Shift Keying (BPSK), so-called Pseudo Random Noise (PRN) codes are modulated on the carrier frequencies. The L1 carrier is modulated with a precise (P) code plus a coarse/ acquisition (C/A) code whereas L2 is only modulated with the P-code. The resulting frequency spectrum for the carrier, due to BPSK, equals 20 MHz for the P-code and 2 MHz for the C/A code. The carrier frequency is suppressed.
  Both frequencies have a navigation data message superimposed. This navigation message contains satellite

3.2 Ships Electronic Position Fixing Systems

In this section a general description will be given on the electronic position fixing (EPF) systems used on board submarines. These systems can only be used when the submarine itself is at the sea surface or when it is able to have a receiving antenna at the surface. When one or more of these systems are in use, they can in most cases be used continuously for position updating. An important factor from a tactical point of view is of course the time needed for a first fix if the systems are only used for updating the DR position and decreasing the dimensions of the POE.

The group of EPF systems not only comprises the well known terrestrial position fixing systems such as Decca, Loran-C or Omega, but also satellite navigation systems and radar. In evaluating these systems it is important not only to describe the configuration of the system used, but also to look at characteristics such as

- signal acquisition and tracking continuity
- signal integrity
- coverage
- availability

Other important factors to be considered are :

- signal characteristics
- accuracy
- fix rate of independent LOPs or position fixes
- fix dimension (2D or 3D)
- ambiguity

3.2.1 Satellite Position Fixing Systems

NAVSTAR / GPS

GPS is a space-based position, velocity and time system that has three major segments :

- Space segment

When fully operational this segment will consist of 21 satellites plus three active spares, in six orbital planes. Up to date (December '92) 19 satellites are operational. The satellites operate in almost circular orbits at an altitude

3. Sensors, Systems and their Characteristics

3.1 Introduction

In this chapter an overview will be given of the sensors and systems available for collecting information to be used for position fixing and quality control. Since a wide range of sensors and systems is available, only those generally used on board submarines are discussed.

The chapter is divided into three sections, each section describing a particular group of sensors and systems used in one of the three stages of navigation, according to the concept of the POE. In the first section the systems that can only be used at the sea surface will be discussed. The next section gives all systems and sensors that can be used when the submarine is at the sea surface or submerged. Here a distinction is made between bounded and unbounded sensors. In the final section information will be given on the possible integration of the various sensors with each other and with a main computer. Also some remarks on software used for navigation will be made.

It is not the intention to go into great detail about the different systems as there are plenty of good textbooks describing them. The reason for including this chapter in the paper is not only because then the systems that are available will be named, but also because it will make the reader aware of the fact that since a submarine can be navigated either at the sea surface or under water, restrictions are posed on the availability of position fixing systems.

In this chapter only mean values for predictable accuracy will be given. For most systems, however, the standard error of data provided depends on many factors, making the predictable accuracy to vary in time under given conditions. Thus a more detailed analysis of errors contributing to the total standard error of a sensor or system is needed in order to be able to calculate predictable accuracy of a position fix at any time. This analysis will be performed in chapter 4.

At the moment there is no integration of systems leading to a direct plot of the POE on for example a plotting table or computer screen. Everything has to be done by hand. To do this properly, taking into account all the errors present in the sensors and systems used, is quite a complicated job.

From the above it should be clear that preferably three independent working sets of each piece of equipment should be used and running. Not only will this provide backup when a systems breaks down, it will also show if a system drifts. Normally, this is not feasible. Therefore, sufficient backup information should be provided by the other working systems as this will provide redundancy and a system cross-checking ability. A consistent mathematical solution to position fixing and quality control should be available to the navigator as this is the most objective way by which results can be obtained, compared and saved for future use. Mathematical models that provide all this will be presented in chapters 5 and 6.

1. Use of 'a priori' standard errors as precision of measurements :

   - A priori standard errors are stated for several systems and sensors, but no statement has been made on how these values were obtained. Furthermore, it is left to the judgement of the commanding officer to change these values if this is thought to lead to better results. No guidelines are given here.

   - How can the user be sure that even if he could estimate a priori standard deviations, these will be the correct ones. It could well be necessary - because of the changing conditions - to scale the a priori standard deviations continuously by σ_0 obtained from initial calculations.

2. The following remarks can be made on the current use of the POE, i.e. purely as a hand method :

   - the method does not give a unique solution for the position fix and size of the POE. There is too much room for personal interpretation of the problem;

   - the expansion of the circle used in both methods to obtain the POE as described in section 2.3.2, is based on the direction and distance between the last fix and current DR position. It is made independent of the actual track between these positions. Especially when the submarine has made many changes in course and/or speed, fix errors tend to be under-estimated;

   - as a submarine is able to move underwater and therefore in 3D space, its position needs to be treated as a 3D position. This means that the POE needs to give 3D information. At present, the POE only provides 2D information;

   - it is not possible to get valid information regarding the quality of the fix, reliability of LOPs or a priori and a posteriori standard errors by using the plot of the POE. Therefore the POE cannot give a true representation of the probability area of the fix;

   - once the navigator has been able to calculate the position of the submarine using external information, the already existing DR position is not integrated with the new position fix to form a new MPP. The DR position is disregarded. Even though the POE of this DR position may be quite large, the position can still be integrated with the new position fix derived from external sources, possibly leading to new information on standard errors to be used.

resulting area, giving the uncertainty in the position fix due to tidal streams and current, is combined with the expanded circle to give the POE.
This method is used in areas where the general direction and speed of tidal streams and currents are quite well known.

2.4 Problems related to the use of the POE

Although the concept of the POE in itself works really well, there are some problems when deriving a POE and using information found. In this section the most important problems will be discussed.

To start with, there is not sufficient backup of systems or sensors on board the submarines. This leads to the following problems :

1. when no backup of a sensor or system is available at all:

   - when sensors drift (such as gyros used in the gyro compass or the Ships Inertial Navigation System (SINS)), the navigator will have no means to observe this because no reference is available;

   - If a system or sensor fails completely, a primary source for position updating and quality control is lost.

2. when only one backup system or sensor is available for a particular system or sensor :

   - If only one of the two is running at the time and this system or sensor has a total failure, valuable time will be lost during start-up of the backup system, leading to a temporary loss of a primary source for position updating and quality control. Beside that, no information is present on differences between the two instruments, since no comparison had been made between two working systems;

   - If both systems are running and one system drifts from the right value, the navigator will not be able to determine which system is working correctly so basically neither of the systems can give good information to be used for position fixing. Again a primary source for position update and quality control is lost.

On the quality control side of position fixing the following group of problems can be distinguished :

$$R_{ec} = a + \epsilon \times \Delta t \quad \text{[metres]} \tag{2.1}$$

where

R_{ec} = radius of circle
a = the d_{rms} value of the last MPP based on external sources
ϵ = expansion factor
Δt = time interval between last MPP based on external sources, and current DR position

The magnitude of the expansion factor is a combination of the accuracy with which the current direction and speed are known and the standard errors of the bounded systems used to obtain the current DR position. No fixed value for the expansion factor is given, but values between 900 - 1800 metres (1000 - 2000 yards) per hour are generally used. This way, the radius of the circle will expand linearly in time. This method is used when the submarine is operating in deep, open waters where tidal streams are either circular or variable.

- circle segment method : The expanding circle is used again, with ϵ set to a value representing the total error in log and gyro. To allow for uncertainties due to tidal streams and currents, part of a circle segment can be constructed having a depth (d) and width (2a) (see figure 2.2). The size of these measures depends on the present situation. The

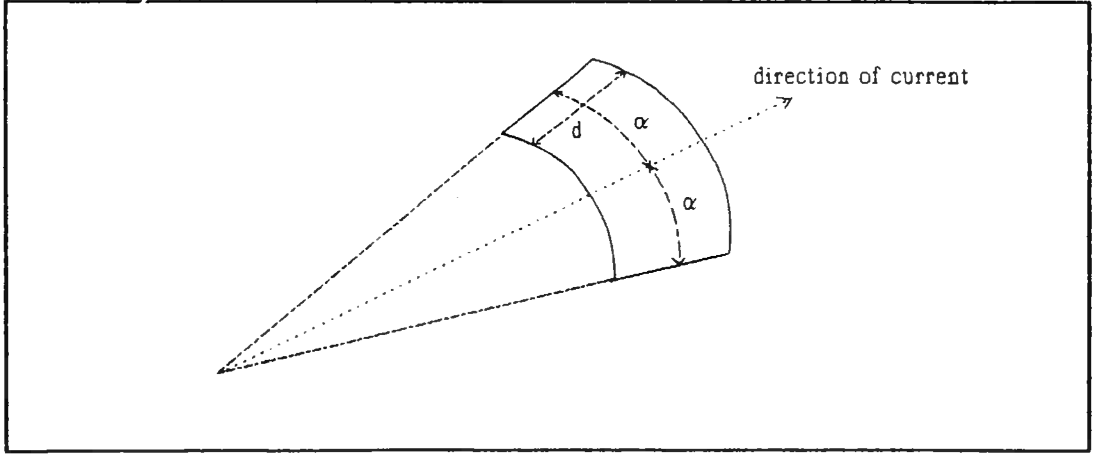


Figure 2.2 The part of the circle segment used to construct the uncertainty in the position fix introduced by current.

a bigger problem as it is difficult to estimate their influence on the movements of the submarine. Their direction and strength can change quite significantly with depth. These changes make it very difficult to estimate their effect on the POE. To be able to use the POE in unknown waters without reaching unacceptable limits quickly, three types of operation have been defined, each giving its own restrictions to the limits of the POE. These limits depend highly on the sort of area and navigational dangers present. The three types of operation are :

1. area operations : the submarine will be manoeuvred in a designated area in open sea. The smaller the size of the area, the more course and/or speed changes will be made. Generally, the movements of the submarine are not limited by seabed features;
2. transit : the journey from one area to another;
3. confined waters operations : operations in those areas where manoeuvring is restricted due to bottom features and/or landmasses present.

Each of the above mentioned operations imposes its own needs for positional precision and accuracy, resulting in different methods for construction of the POE. Beside this it is of course important always to evaluate the general bottom contours and find dangers present at the most likely depth at which the submarine is operating.

2.3.2 Construction of the POE

After having given a broad description of the purpose of the POE, it is also important to give information on the way the POE is actually being constructed. As the navigator is not able to spend much time on plotting positions and constructing a POE, position fixing and plotting in a chart has to be done quickly. Therefore special hand methods have been developed, the results of which are shapes that are relatively easy to draw such as circles and rectangles with rounded corners. Basically, the following methods are used at present :

- expanding circle method : a circle, having the current DR position as its centre, is drawn. The radius of the circle is calculated from

This can be very difficult, especially if the submarine is operating in areas where knowledge of external disturbances acting on the vessel is limited. Therefore standard errors of measurements are estimated in such a way that their true values - which are unknown - will be guaranteed to be smaller than those estimated. After observations have been made over a period of time, in which estimated positions are compared against position updates, the standard errors can be redefined in such a way that they agree more with the actual (true) values. One need to be reminded of the fact that in all cases average standard errors are used.

The way the POE is used during sailing consists of three steps:

1. When the submarine is at the sea-surface, its position is fixed by means of external sources. As long as the submarine stays at the surface, the position can be updated continuously. This leads to the normal position fixing and quality control using common mathematical methods such as least squares. The POE will increase in time, but the continuous flow of sensor information makes it possible to reduce the size continuously, keeping it to an acceptable size.

2. As soon as the submarine submerges, use can no longer be made of external sources. Its position is therefore calculated using bounded sensor and current information. Now the POE is introduced as an expanding mathematical figure. If possible, information from external sources such as eg. bottom contours and seabed slopes measured using an echo sounder, will be used to update the position and to decrease the size of the POE.

3. Eventually the dimensions of the POE reach limits, making it unacceptable to stay underwater using DR and expanding POE. The submarine is brought either to periscope depth or surfaced making it possible for the navigator to update the position using information from external sources, resulting in a new MPP and small POE.

The means available for position fixing when the submarine is at the sea surface are terrestrial radio position fixing (RPF) systems - such as Loran-C, Omega and Decca -, satellite position fixing systems and compass bearings, radar distances etc. Once the submarine submerges, use can only be made of its bounded sensors and systems and occasionally of its echo-sounder. All these systems have errors which can be evaluated quite well during trials at the surface and errors will in general remain unchanged when the submarine submerges. The currents however pose

at all to use the normal distribution as the probability model for the total random error in an observation when this error is assumed to have many components.
The random errors present in consecutive measurements obtained from a system or sensor, exhibit an auto-time-correlation which can in most cases be described by a first-order Gauss-Markov process. For a short description see appendix 6. For RPF systems, the correlation time is dependent on carrier frequency used : the higher the frequency, the smaller the correlation time.

The last set of definitions to be given in this section are the ones used to classify position fixes. The following types of fixes will be used when discussing position fixing at sea :

- DEAD RECKONING (DR) POSITION : this is the position based on the most recent fix, updated using information from the bounded sensors that measure the vessels *heading*, *speed through the water* and *attitude* (roll,pitch etc.). When establishing this position, the estimated effects of wind, currents and tidal streams, sea state etc. are also taken into account.
- MOST PROBABLE POSITION (MPP) : this is the best position that can be derived using all information available.

It should be noticed that, given the circumstances for position fixing, each position fix, even a DR position can be regarded as MPP.

2.3 Principles of the 'Pool of Errors'

2.3.1 Philosophy and use

The POE is a concept used on board submarines to provide a graphical indication of the maximum likely errors in the MPP at any time. By using the POE the commanding officer can determine safe course, speed and depth. When the submarine is submerged, the commanding officer has to decide when a new update of the position, using external navigation systems, is required. This decision depends on the size of the POE, the area in which the submarine is operating and the dangers present.

For optimum use of the POE it is very important to make an assessment of the various errors - both systematic and random - present and their effects on the submarine's position accuracy.

The definitions that are given next are related to errors. In general the following three types of errors can be distinguished :

- **BLUNDERS** : This type of error is difficult to define since a deviating value could be well within the possible values of a p.d.f. A possible definition could therefore be that a blunder is a measurement which differs significantly from the expected value making it very likely that certain external circumstances are present other than the ones that would make it a random error (under normal circumstances) [Spaans,1988:2].

- **SYSTEMATIC ERRORS** : errors that follow some law by which they can be predicted.

Constant and low frequency disturbances (i.e. the DC and low frequency part of the error spectrum) are considered to be systematic errors. A distinction should be made between systematic errors present in the system used to measure a quantity and those introduced by the user of that system. Constant and low frequency errors of a system are obtained from analysis of the error spectrum which provides insight in the dynamical behaviour of the system. In some cases the constant error can be removed from the system by means of calibration, but often corrections need to be applied to the measurements.
Systematic errors introduced by the user are more difficult to recognize. However, experience could lead to assessment of errors introduced after which corrections can be applied to the measurements.

- **RANDOM ERRORS** : These errors are unpredictable in magnitude and/or sign. They are governed by laws of probability which means that they can be characterized by a p.d.f.

The total random error present in a measurement will in most cases consist of a mixture of random errors introduced in different parts of the observation process (due to environment, measurement sensor or system, user etc.). Each of these random errors will have its own pdf. In general, this pdf is assumed to be zero mean Gaussian, but other types should be considered as well. Errors introduced by rounding off can serve as an example. These errors have an uniform pdf.
The **central limit theorem** states that when the number of mutually independent random variables (i.e. the error components) increases, the distribution of their sum gets closer to the Gaussian (normal) distribution. It is therefore not unreasonable

quantity is meaningless unless it includes a statement of uncertainty. This will be derived from the VCV matrix of the measured quantities in which on the main diagonal the variances of the quantities are given. Note that in the linear case, i.e. one measured quantity, this VCV matrix will contain only one element as could be expected. Using these values a number of confidence regions can be defined. For position fixing in 2D the most common ones are the ***standard error ellipse***, the ***a% error ellipse*** - where ***a*** is normally given as 95% or 99.9%[1] - the d_{rms} or the ***CEP*** whereas in 3D these will normally be the ***standard error ellipsoid***, the ***a% error ellipsoid***, or the ***SEP***. In section 5.5 of chapter 5, information concerning the relationships existing between these different quantities will be given.

As a final remark it should be stressed again that accuracy is the degree of conformity with the correct value, while precision is the degree of refinement of a measured value. Therefore, when blunders and/or systematic errors are present in the measurements of a quantity, it can still be determined with high precision but it will have low accuracy. Any systematic errors present will lead to bias. In figure 2.1 this is visualized for a one dimensional quantity being measured. When comparing the two pdf's, p_1 shows less precision than p_2, while on the other hand p_2 shows higher accuracy then p_1.

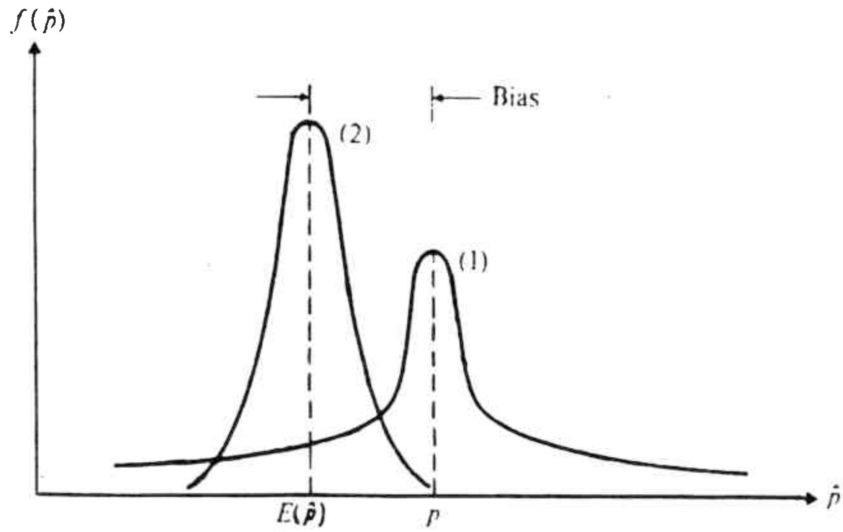


Figure 2.1 Accuracy and precision of a one-dimensional variable p being measured.

---

[1] With respect to aids to navigation the IMO, in its resolution A.529 (13) of 17 November 1983, states that : "The 95% probability figure should be used to describe the accuracy of a system fix" [IALA,1990]

direct effect on the precision. However, before the precision of measurements can be evaluated from observations made, it is important that any systematic errors present are the same for all measurements made. A bias due to the systematic error will exist;

- if the conditions under which the measurements are taken change, the precision might change as well.

Precision of a measurement gives information about the magnitude of ***random errors*** present. It is stated as standard error in units of the system (eg. metres when measuring ranges, angles when measuring bearings). In order to be able to validate the accuracy of observations made using different systems, the standard error will, for all systems, be transformed into meters. How this is done is further explained in chapter 4.

- **RELIABILITY** is the measure of ease with which a blunder in a measurement can be detected.

Several quantities have been suggested to give a quantitative measure of reliability [Spaans,1988:2; Cross et al.,1985]. The definition as given above refers to a statistical property of a measurement. It should not be confused with the reliability of a system, which is given as [FRP,1990] :

- **RELIABILITY** is the probability of performing a specified function without failure under given conditions for a specified period of time.

- **ACCURACY** is the degree of conformity between the true value of a quantity and the most probable value derived from a series of measurements (estimate).

A distinction can be made between several types of accuracy. Since the true value of the measured quantity is unknown, it is better to talk about ***predictable accuracy***, when the quantity is derived from measurements, taking into account all predicted errors. Consider for example a situation in which positioning equipment is placed at an already known position. If this position is re-measured over a period of time, using this equipment, a number of calculated positions scattered round the true position will be found. From this 'scatter plot' a measure of ***absolute accuracy*** by which the position could be derived using the given instrument, should be stated.

Since accuracy is a statistical measure of performance, a statement of the accuracy of a measuring device or measured

1. Diurnal ionospheric changes

   When a receiver is placed stationary at a point on the Earth surface, changes in phase can be observed when the incoming signals are observed. These changes are caused by variations in the height of the ionosphere during a 24 hour period, causing variations in propagation speed of radio waves, hence introducing the observed phase shifts. These diurnal changes are quite well known and as such listed as corrections to LOPs.

2. Geography

   Because of the large distances, waves paths between transmitter and receiver lie over a mixture of land and sea stretches, resulting in deviations from assumed propagation velocity (which is normally based on sea path only). The corrections can either be obtained by calibration or by calculation using mathematical models that make use of ground conductivity maps.

3. Icebound regions

   The Omega signals are severely attenuated in icebound regions, especially in the Arctic regions (eg. Greenland, parts of Iceland). The corrections given in the OPCT allow for propagation in these regions. Little data is available for these areas, making even the best estimates uncertain. In particular rather rapid phase changes with position may occur as one passes in the 'shadow' of the Greenland icecap [DMA,1981].

4. Modal interference

   Waves having low frequencies as used with the Omega system, propagate in a space confined by the Earths surface and the ionosphere. This space acts as a waveguide. In this waveguide different propagation modes exist, each having its own speed of propagation and attenuation rate. The different modes interfere with each other causing changes in phase angles that are very difficult to predict. This interference takes place at different parts of the transmission path between transmitter and receiver:

   - within 450 nm from the transmitter near field interference can be observed, causing such disturbances on the signal that it should not be used for position fixing;
   - when the path of the signal from transmitter to receiver passes through both day and night conditions, the twilight area will cause interference;
   - when signals cross the Earth magnetic equator, in

range (km)	Red	Green	Purple
100	0.029	0.021	0.044
200	0.065	0.048	0.101
300	0.120	0.088	0.189
400	0.215	0.157	0.338
500	0.341	0.250	0.534

Table 4.3 Typical lane accuracies (1σ) in lanes

The conversion of standard error of a Decca LOP from lanes (σ_{LOP}) to metres (σ_m) is given by :

$$\sigma_m = \frac{1}{2}\lambda_c \operatorname{cosec}\left(\frac{1}{2}\gamma\right)\sqrt{\sigma^2_{LOP}} \quad \text{[metres]} \qquad (4.12)$$

where

γ = angle subtended by master station and slave at receiver position

<u>OMEGA</u>

Similar to the Decca receiver, the Omega receiver is a device that measures phase differences. Because of the long propagation paths between transmitting station and receiver and the low frequencies used, prediction of phase disturbances becomes more difficult. The Omega system error sources can be divided into four categories :

A. Predictable errors in assumed propagation model

This category of errors comprises those that are either fixed or can be forecasted to certain extent. Corrections that need to be applied to measurements to reduce the effect of these errors are stated in the Omega Propagation Correction Tables (OPCT). The wave propagation mathematical model on which the OPCT is based, is revised periodically to account for changes in solar activity and other propagation anomalies. The OPCT are updated every other year on average.

present in the position if the fixed error corrections (see below) have not been incorporated in the solution.

C. User errors

Errors that can be introduced by the user were already mentioned at the beginning of this section. Two of those are explained here in slightly more detail in case a receiver like the Mark 21 Decca receiver is used :

- setting the equipment : in this case the zone letter and / or whole lane number are set wrongly. Both errors should be regarded as gross errors.
  Setting the wrong zone letter becomes quite obvious when plotting positions in a chart because of the zonewidth. However, wrong setting of the lane number might not be so obvious, since the lanewidth of the patterns is only a few hundreds of metres.

- applying 'fixed error corrections' : when plotting Decca lattices on charts a mean signal propagation velocity of 299550 km/s is used. This way hyperbolas which are smooth will be plotted. The actual propagation velocity will deviate from this mean value, resulting in irregular lines of equal phase difference. To allow for the discrepancy between plotted lattices and LOPs based on the actual propagation velocity, corrections need to be applied to observed values. The corrections needed are called 'Fixed Error Corrections' and are given in the Decca Data Sheets. These corrections can be as large as several tenth of lanes in coastal areas. Figure 4.6 gives an example of extreme corrections needed on the purple pattern of the Lofoten Chain (3E).

If receiver errors and user errors are removed as described above and the 'Fixed Error Corrections' are applied to the Decca readings, the errors introduced by sky wave interference are the major source of error in a Decca LOP. Based on data provided by the Decca Navigator Company, the following typical values of standard deviations in fractions of lanes can be given[1] for Summer night conditions in temperate latitudes and for a propagation path over sea-water or good soil [Decca,1979] :

---

[1] The distance to both master and slave is taken to be the range given in the table.

sky wave interference can be reduced using MP readings for position fixing in areas where sky wave interference is likely to occur since MP signals are less susceptible to sky wave interference (see section 2.2 of chapter 3).

The magnitude of errors due to sky wave interference are predictable within reasonable limits of confidence, based on statistical analysis of countless observations at fixed monitor stations in differing environments and at various ranges from the transmitters.
The error in phase of the signals under sky wave conditions arriving at the receiver station is considered to be due to the following phase deviations :

- phase error in the signal from Master to Slave ($\delta\varphi_{ms}$);
- synchronization error in the Slave ($\delta\varphi_{sync}$);
- phase error in the signal from Master to Receiver ($\delta\varphi_m$);
- phase error in the signal from Slave to Receiver ($\delta\varphi_s$).

For each pattern, the four components can be combined to give an expression for the total error of that pattern at the receiver [Decca,1979] :

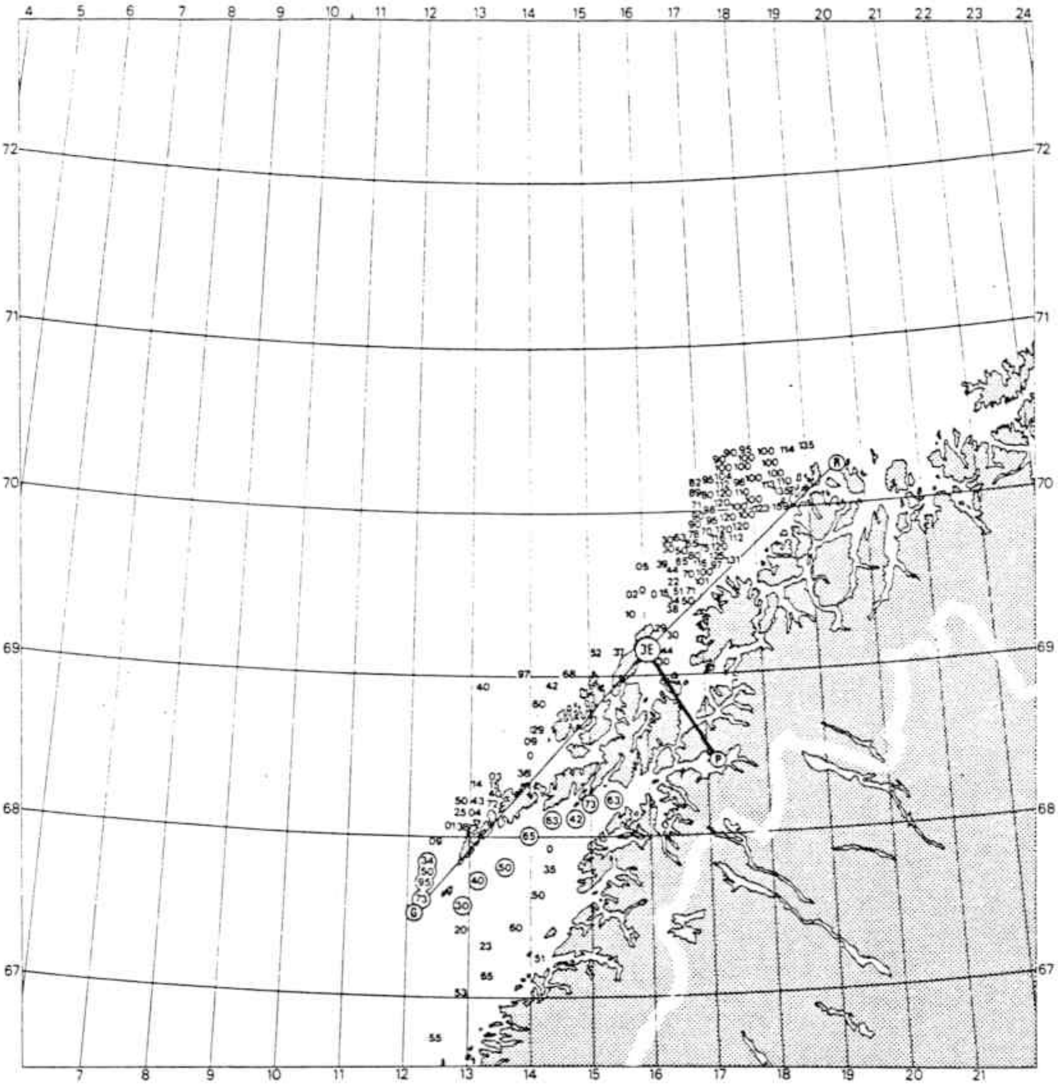
- Red : $\delta\Delta\varphi_R = 4\delta\varphi_m - 4\delta\varphi_{mR} - 3\delta\varphi_{syncR} - 3\delta\varphi_R$
- Green : $\delta\Delta\varphi_G = 3\delta\varphi_m - 3\delta\varphi_{mG} - 2\delta\varphi_{syncG} - 2\delta\varphi_G$
- Purple : $\delta\Delta\varphi_P = 5\delta\varphi_m - 5\delta\varphi_{mP} - 6\delta\varphi_{syncP} - 6\delta\varphi_P$

This information can be used to calculate the variances and covariances of Decca LOPs. These are given in appendix 4.

B. Receiver errors

Differential phase errors in the signal channels and subsequent sections of the receiver are corrected by a 'reference' facility. The reference input forms a phase datum whereby the measured phase difference should be zero in the absence of errors. When a systematic error is present in the receiver circuits, the user restores the zero reading - and therefore minimizes the receiver error - by adjusting the phase shifter in the appropriate receiver channel. [Decca,1979].
This leaves the receiver with short-term phase changes, which can be considered random of nature. These errors are negligible with respect to the random errors due to sky wave interference and noise.

If the Decca receiver output is a geographic position or if the receiver is part of an integrated system, errors due to reading off and plotting do not exist. Large errors might be



THE DECCA NAVIGATOR SYSTEM – LOFOTEN CHAIN (3E)

PURPLE PATTERN FIXED ERROR CORRECTIONS

CORRECTIONS TO APPLY TO OBSERVED PURPLE DECOMETER READINGS TO OVERCOME FIXED ERRORS
VALUES SHOWN ARE IN HUNDREDTHS OF A LANE UNITS
FIGURES ENCIRCLED SHOULD BE SUBTRACTED
FIGURES NOT ENCIRCLED SHOULD BE ADDED

Figure 4.6 Decca Navigator System - Lofoten Chain Purple pattern fixed error corrections [Decca Data Sheets] considered to be of importance) and interference from other sources transmitting in the 50 - 150 kHz frequency band. Errors due to CWI are normally negligible, especially since the DNS makes use of exclusive frequency allocations and the receivers have a narrow input bandwidth. This narrow bandwidth also makes the DNS relatively immune to atmospheric noise.

3. Sky wave interference

Decca signals can reach the receiver either by the direct path between station and receiver, the so-called ground wave, or via reflections from the ionosphere, the **sky wave**. Interference of these two signals results in the Decca reading being subject to a random error. The magnitude of this error is a function of the sky wave signal strength with respect to that of the ground wave (SIR) and phase angle of sky wave relative to phase angle of ground wave:

$$\delta\varphi_{sky} = \arctan\left[\frac{\sin\alpha}{\frac{\hat{g}}{\hat{s}} + \cos\alpha}\right] \qquad (4.11)$$

where

$\delta\varphi_{sky}$ = sky wave induced phase error
$\hat{s}$ = sky wave amplitude
$\hat{g}$ = ground wave amplitude
α = relative phase of sky wave with respect to ground wave

A maximum phase disturbance will arise when the two signals are almost in quadrature ($\alpha \approx 90^\circ$) if $\hat{s} \ll \hat{g}$. If $\hat{s}$ becomes larger than $\hat{g}$, laneslips can occur.

The sky wave interference is more pronounced during nighttime than daytime and stronger in winter than in summer. To avoid these interferences from sky waves, the navigator should, if possible, not use a chain at ranges over approximately 750 km (400 nm) during daytime and 200 km (110 nm) at night from any station since sky wave interference is most likely to happen at these ranges and over (1-hop-E layer sky wave distance)[1]. The effects of

[1] The Decca Navigator Company Ltd. gives a SIR of 10 dB as approximate representative value for satisfactory receiver operations [Decca,1979]. Graphs showing the intensity of 100 kHz ground and sky-waves with increasing distance indicate that this SIR value is reached approximately at the ranges stated.

This phase difference can only be measured in the interval 0 to 2π, resulting in the fraction :

$$L = \frac{\Delta\phi}{2\pi} \pmod{2\pi} = \text{fract}\left| \frac{D_R - D_S + B}{\lambda_c} + PVC + k \right| \quad \text{[lanes]} \qquad (4.10)$$

where

L = fraction of lane
fract[·] = fractional part of expression

In order to remove ambiguity, the user must give the correct zone letter and lane number as input to the receiver. For the correct lane number, use can be made of coarse pattern provided by the MP signals as discussed in section 2.2 of chapter 3.

Errors in the Decca LOP are the result of errors in the measured phase difference. The sources of these errors can be grouped as follows :

A. System errors

1. Inaccuracies in assumed propagation model

   A mean wave propagation velocity, based on an 'all sea' propagation path between transmitting station and receiver, is used in calculations of chart lattices or position. To allow for land stretches along the path, leading to variations in wave propagation velocity, a correction needs to be applied to the observations. For a given path between transmitting station and receiver the value of the correction due to land path will be fairly constant in time. For this reason, the corrections were named 'Fixed Error Corrections' by the Decca Navigator Company and are stated in the Decca Data Sheets when known. If observations are not corrected, a bias that can have values of more than one lanewidth in extreme cases, will arise in the observation.
   Additional temporary variations result for example from changing meteorological conditions (such as rain, snow, fog) affecting the ground conductivity along the given path.

2. Noise

   The main noise sources are atmospheric noise (only that resulting from the equatorial thunderstorm belt is

- the way synchronization between the transmitting stations is maintained;
- accuracy of prediction model for radio wave propagation velocity (predictable and sudden fluctuations);
- receiver accuracy (SNR, thermal noise)
- lane expansion factor (LEF) for hyperbolic LOPs
- susceptibility to interfering signals (sky wave, other signals in same frequency band)

Laneslips :

Any break in the continuation of receiving signals, caused by equipment failure or severe interference may cause a **laneslip**, i.e. the tracking loop may miss or jump one or more lanes. Errors of this kind are unpredictable but are most likely to occur when the receiver is near the edges or outside the limits of the accepted coverage range. In these regions, sky wave interference can be sufficiently large to take over temporarily the control of the phase of the resultant signal so that laneslip occurs. Generally, this situation is most likely to occur at twilight and is worsened by the presence of heavy atmospheric disturbances or interfering transmissions from other stations.

Decca Navigator System

The Decca LOP at a receiver station is obtained by observing the difference in phase angle of two incoming unmodulated signals (one from the master station and one from a secondary station) on a given frequency (f_c). This phase difference is given by :

$$\Delta\varphi = \varphi_s - \varphi_m = 2\pi\frac{(D_m - D_s + B)}{\lambda_c} + PVC + k \qquad (4.9)$$

where

$\Delta\varphi$ = phase difference between master and slave
φ_m = phase angle of signal from master station
φ_s = phase angle of signal from slave
D_m = spheroidal distance master station - receiver
D_s = spheroidal distance slave - receiver
B = spheroidal length baseline master station - slave
λ_c = wavelength signal of comparison frequency
PVC = phase velocity correction
k = minimum lane count

A possibility of reducing errors due to uncertainties in the ground track is by using the Omega RPF or Loran-C RPF system to calculate the ship's ground track and speed. Although the absolute accuracy of these systems itself is not very good, the system can well be used to measure differences, i.e. displacements to a high degree of accuracy. The data obtained by the Omega or Loran-C receiver is compared with the ship's course and speed input, thus improving accuracy.

A few additional remarks on Transit position fixes should be made :

A satellite will during its pass never be longer than 18 minutes visible to an observer on the ground. In this time interval a total of 45 23 second Dopplercounts can be measured. This is an apparently well over determined system in four unknowns (dX, dY, dZ and df). However, all measurements are made in the plane defined by the satellite's orbit and receiver station. This means that although the receiver position is well defined in the mentioned plane, it is weakly determined in the out-of-plane direction. If the antenna height is given and the ship's velocity vector is taken into account a one pass position result will suffice for navigational purposes. A horizontal position accuracy in the order of 100 - 350 m ($2d_{rms}$) for a dual frequency receiver and 200 - 500 m ($2d_{rms}$) for a single frequency receiver can be expected.

Figures 4.3 and 4.4 show that for good results, a satellite should not have a maximum elevation over approximately 70°. In order to keep errors due to ionosphere and troposphere within acceptable limits, the minimum elevation of a satellite should not be below 15°.

4.2.2 Terrestrial LF and VLF RPF systems

The accuracy of a fix provided by a terrestrial LF or VLF RPF system depends on the accuracy and precision with which each LOP used to obtain the fix can be determined, the number of LOPs and the angle of cut between the LOPs. The accuracy and precision with which each LOP can be obtained depends mainly on the following factors :

- the accuracy with which either time differences or phase differences between received signals can be measured;

5. measurement noise

The satellite signal propagates in a direct line of sight. It may be absorbed, reflected or refracted causing all sorts of interference, if there are objects between the satellite and receiver. Noise can also result from interference of other signals in the 150 MHz or 400 MHz range.

6. time errors

The 2 minute intervals of the broadcast ephemeris are based on UTC and if no corrections are made for variations of UT1-UTC a longitude bias will be introduced in the derived position.

After having corrected for the above given error sources, a total standard error, based on random errors present, can be estimated for each calculated range differences [SMS,1988] :

geopotential model	10 - 20 m
polar motion and time errors	0 - 10 m
predicted ephemeris data	5 - 15 m
ionosphere	7 - 20 m
troposphere	3 - 5 m
instrumentation & measurement noise	3 - 6 m
receiver height error	10 m
total error (1σ)	17.8 - 35 m

This standard error is for a single pass, dual frequency, static receiver.

B. Errors in vessel's ground track

Uncertainties in the vessel's ground course and ground speed during satellite pass introduce additional errors in the position fix. These uncertainties can be the result of unknown influences of wind and current and errors in log and gyro. Figures 4.4 and 4.5 show the errors caused by a one knot ground track error North and East respectively as function of the satellite's maximum angle of elevation.

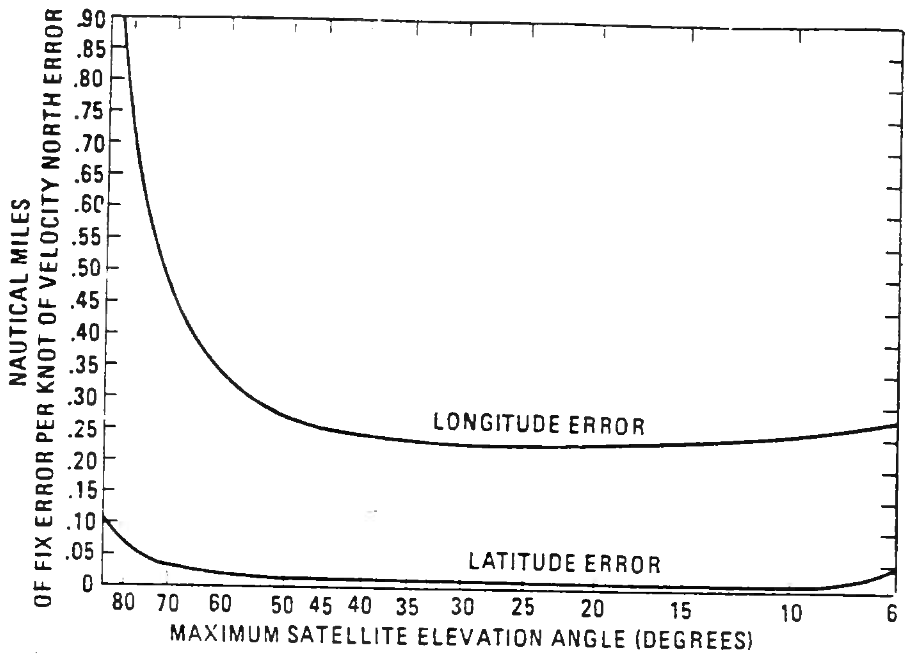


Figure 4.4 Sensitivity of the satellite fix to one knot of ground track North error.

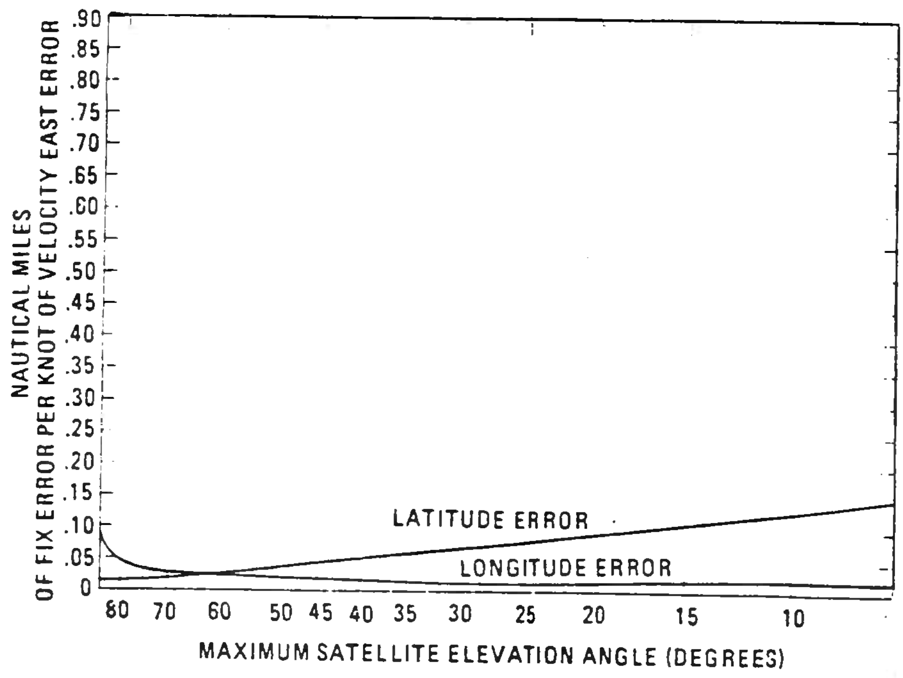


Figure 4.5 Sensitivity of the satellite fix to one knot of ground track East error.

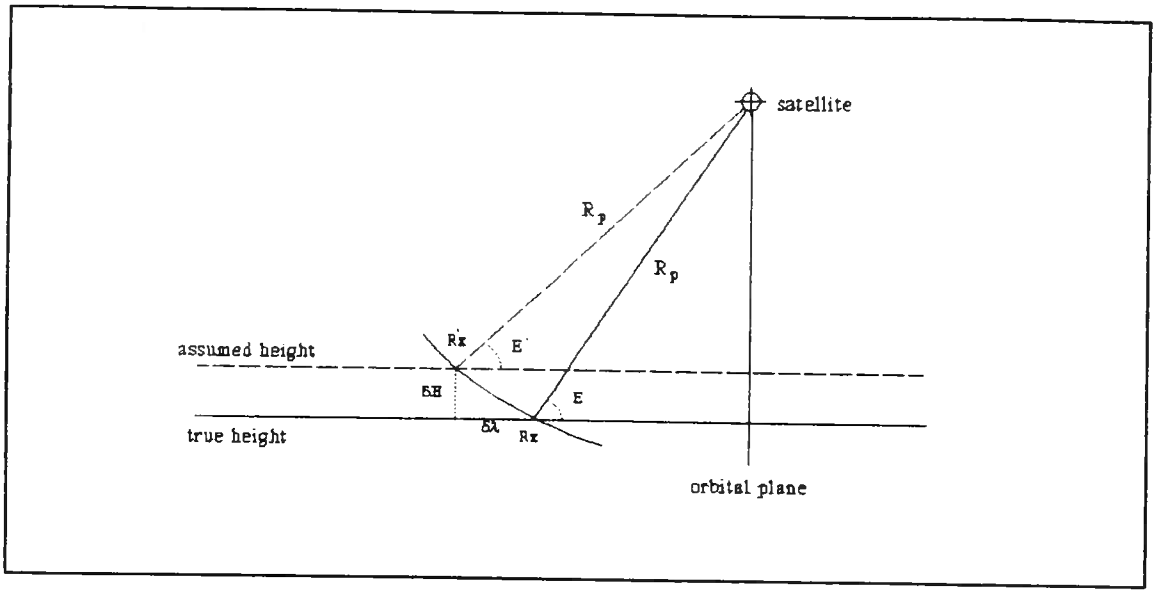


Figure 4.2 Longitude error in relation to height error and satellite elevation

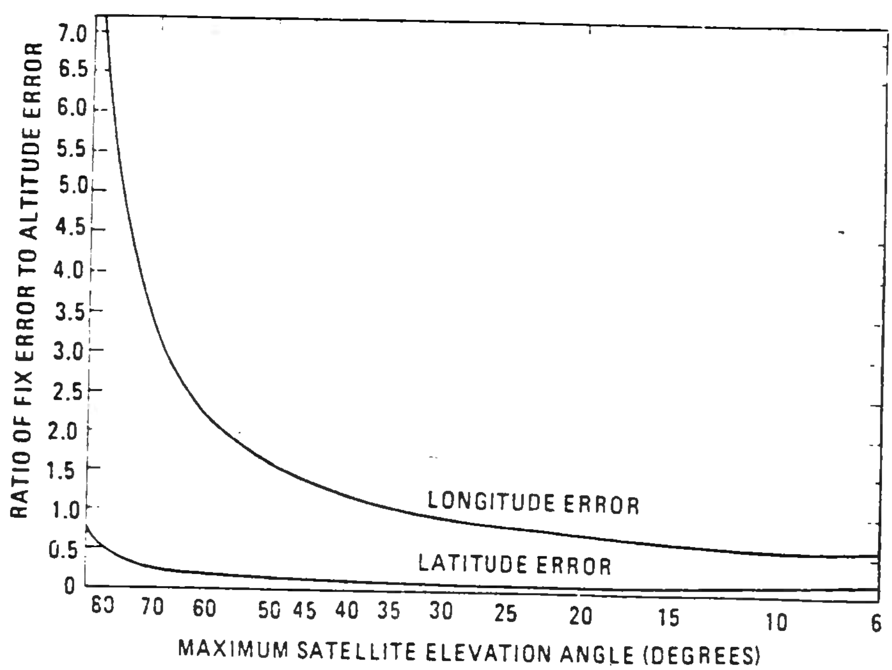
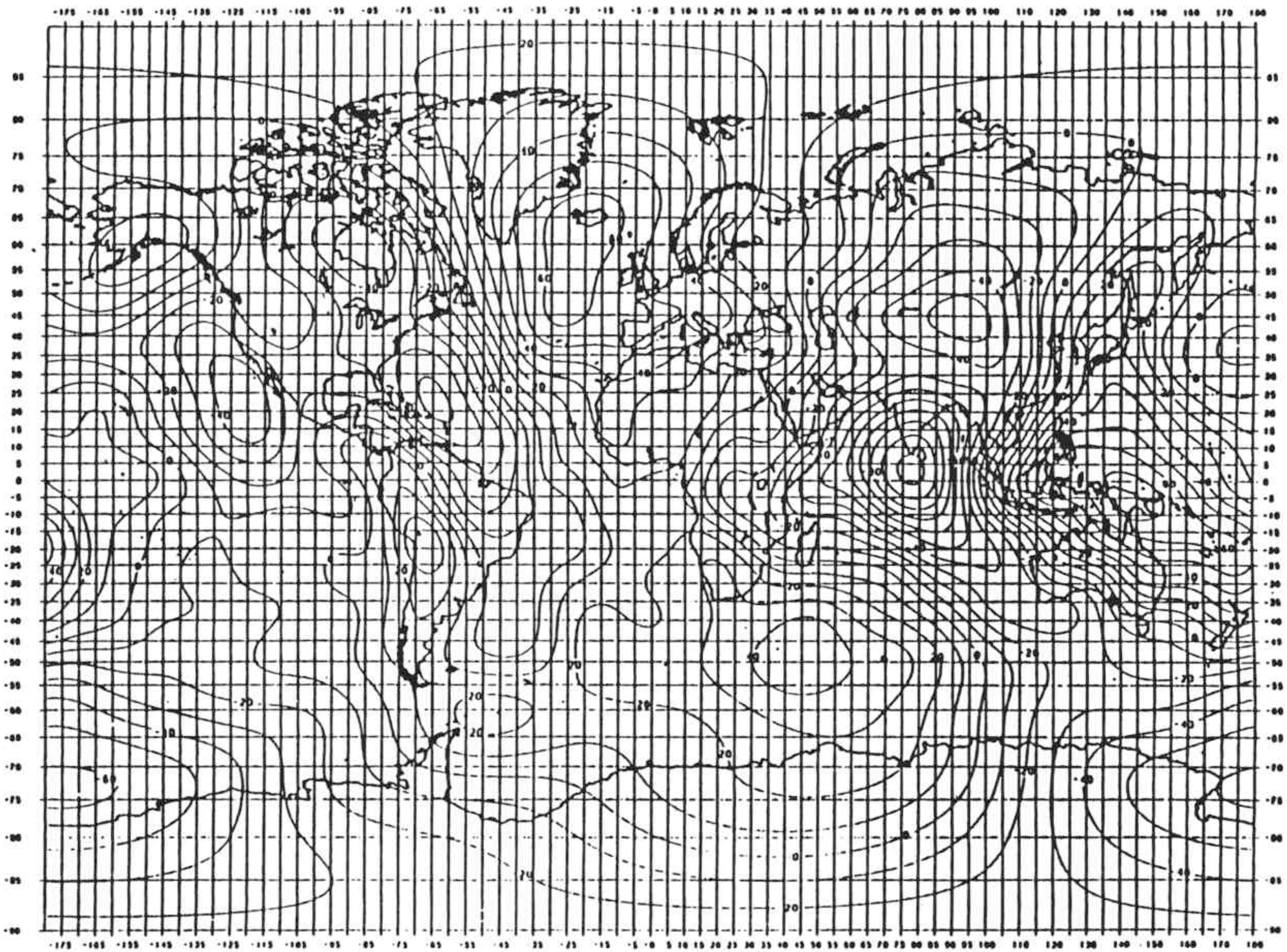


Figure 4.3 Sensitivity of fix error to altitude estimation error



WGS 84 Geoid (n = m = 18 Truncation) Referenced to WGS 84 Ellipsoid ($\overline{\delta C}_{4,0}$, $\delta C_{6,0}$, $\delta C_{8,0}$, and $\overline{\delta C}_{10,0} \neq 0$; Units = Meters).

Figure 4.1 A worldwide WGS84 geoid height contour chart

- height of geoid above the spheroid (called the geoid-spheroid separation);
- height of antenna electrical centre above the geoid (called the orthometric height)

The height of the geoid above or below the spheroid can be derived from models. Figure 4.1 shows contour lines of the geoid height above WGS84, the reference datum used by Transit. For some receivers the geoidal height need not be given since a geoid-spheroid height difference model is implemented. However, uncertainties in the model still give rise to errors. Since in this paper only the use on board submarines will be regarded, the height of the antenna above the geoid can be set to the height above mean sea level (MSL). The error resulting from height errors is mainly in longitude due to the fact that the satellite orbits are almost polar orbits. Figure 4.2 shows the effect of an altitude error (δH) combined with satellite elevation angle (E) on longitude ($\delta\lambda$), whereas figure 4.3 gives the ratio of the fix error to altitude error as a function of the satellites maximum elevation.

3. satellite orbit model

The satellite's orbital parameters are transmitted to each satellite every 12 hours by OPNET. The accuracy of these parameters have a direct influence on the accuracy of the derived receiver position. Irregularities in the Earth's gravity model, causing fluctuations in the predicted satellite orbit, is one of the main error sources. Over the last few years major improvements have been made. Another important error results from miscalculation of the satellite's position which is affected by atmospheric drag due mainly to failure of the adopted atmospheric density model to depict the actual physical situation when the satellite ephemeris is generated.

4. receiver oscillator frequency error

If the frequency of the receiver oscillator drifts, a position longitude error is introduced, depending on the drift-rate and maximum elevation angle of the satellite. The effect of any offset in the receiver oscillator frequency, causing errors in the position fix, can be minimised by taking the unknown offset as one of the parameters to be solved for in the position estimation process.

The effect of the ionosphere on the Dopplercount is a function of the carrier frequency of the signals. By observing the Dopplercount on two different frequencies, a first order ionospheric correction can be calculated for one of the frequencies. This correction is given by :

$$\delta N_{400}^{ion} = \frac{24N_{150}^{o} - 9N_{400}^{o}}{55} \tag{4.7}$$

where

δN_{400}^{ion} = ionospheric correction for 400 MHz signal Dopplercount

N_{400}^{o} = observed Dopplercount on 400 MHz signal

N_{150}^{o} = observed Dopplercount on 150 MHz signal

In the derived case, this correction has to be applied to the 400 MHz signal Doppler-count, giving :

$$N_{400}^{T} = N_{400}^{o} - \delta N_{400}^{ion} \tag{4.8}$$

where

N_{400}^{T} = Dopplercount on 400 MHz signal corrected for ionospheric disturbances

An alternative that can be used with either sort of receivers is to make use of a monitor station giving for a particular area the Dopplercount corrections for atmospheric conditions. This means however, that position data has to be post-processed.

To correct for errors due to the troposphere, models have been developed. These models are partly empirical and partly based on theoretical considerations. The most commonly used model is a simplified Hopfield model in which the atmosphere is regarded as consisting of horizontal layers. Some receivers, such as the Magnavox 1502, use the meteorological data at the receiver station to calculate a correction.

2. altitude error

In order to obtain accurate fixes from satellite systems, the altitude of the receiver antenna above or below the reference datum needs to be known. This height consists of two parts :

As the number of cycles transmitted by the satellite between t_i and t_{i+1} must be equal to the received number of cycles between t'_i and t'_{i+1}, equation (4.4) can be rewritten as

$$N_i = \int_{t'_i}^{t'_{i+1}} f_g \, dt - \int_{t_i}^{t_{i+1}} f_s \, dt \qquad (4.5)$$

where

f_s = frequency of signal transmitted by satellite

Assuming f_g and f_s constant over the measurement interval, (4.5) can be reduced to

$$N_i = f_g (t'_{i+1} - t'_i) - f_s (t_{i+1} - t_i) \qquad (4.6a)$$

or

$$N_i = (f_g - f_s)(t'_{i+1} - t'_i) + \frac{f_s}{c}(R_{i+1} - R_i) \qquad (4.6b)$$

where

R_i = range receiver - satellite at time t'_i
R_{i+1} = range receiver - satellite at time t'_{i+1}

giving a direct relationship between Doppler count and range difference during the given time interval.

The error sources influencing the Dopplercount and derived range difference can be grouped as follows :

A. Transit system errors

1. atmospheric errors

As is the case with GPS, part of the Transit signal path goes through the Earth's atmosphere.

fix obtained.

GPS receivers calculate velocity by measuring the rate-of-change in pseudo-range, known as **delta range**. A 0.2 m/sec per axis (2σ) is achievable for PPS receivers. SPS velocity accuracy is the same as PPS when SA is off. When SA is on, SPS velocity is degraded. The amount of degradation of the velocity is classified [NATO,1991:1].

NNSS / TRANSIT

Each Transit satellite, having a known orbit is broadcasting a phase modulated carrier CW signal at given frequency (f_s). When one of the satellites is in view a receiver can pick up its signal. This received signal has a frequency f_r, which differs from f_s due to the relative motion of the satellite with respect to the receiver. In the receiver a reference signal is generated, having frequency f_g. As long as the satellite is in view, the number of cycles of the beat frequency $f_g - f_r$ can be counted over given time intervals to give the so-called **'Dopplercount'**. This Dopplercount is the basic observed quantity and is a measure of the difference between the satellite - receiver range at the beginning and at the end of the measurement interval, giving a hyperboloidal LOP.
Since the measurement interval is chosen to be much smaller than the time the satellite is in view, the measurement can be repeated, resulting in a set of LOPs by which the MPP of the receiver can be calculated.

The number of counts in a specific time interval is given by :

$$N_i = \int_{t'_i}^{t'_{i+1}} (f_g - f_r)\, dt \qquad (4.4)$$

where

N_i = Dopplercount over measurement interval
t'_i = start time of measurement interval (in receiver timescale)
t'_{i+1} = end time of measurement interval (in receiver timescale)
f_r = frequency of received radio signal
f_g = frequency of receiver reference signal

SA is controlled in such a way that in peacetime the horizontal position predictable accuracy is 100 m ($2d_{rms}$) and the vertical accuracy 156 m (2σ) (see table 4.2). The SPS peacetime velocity degradation due to SA is classified. The value stated for horizontal accuracy does not include errors due to multipath! System accuracy degradation can be increased if it is necessary to do so. The SA position error distribution resembles a Gaussian distribution with a long-term zero mean [NATO,1991:1].
In order to reduce the effect of SA use can be made of Differential GPS (DGPS), which involves the use of correction data provided by a monitor station in the vicinity of the user's receiver.

The errors discussed result in a difference between true and measured range between satellite and receiver. After applied corrections a typical error budget for a static measured range using a dual frequency P-code or single frequency C/A-code receiver without SA is given as [NATO,1991:1;Teunissen,1991; Kranendonk,1992] :

	P-code	C/A - code
satellite clock and subsystem stability	3.0	3.0
prediction of satellite perturbations	1.0	1.0
ephemeris / geopotential model	4.2	4.2
ionospheric delay	2.3	5.0 - 10
tropospheric delay	2.0	2.0
receiver noise & resolution	1.5	1.5
multipath & satellite interference	1.2 - 10*	1.2 - 10*
other	1.1	1.1
UERE (1σ)	6.5 - 11.8	7.9 - 15.3

all errors are stated in metres
* this is a moderate value, the actual situation can be far more worse.

The User Equivalent Range Error (UERE) is a measure of the error in the range measurement to each satellite as seen by the receiver. This value tends to be different for each satellite and tends to be at a minimum following an upload. Apart from the error in each range measurement, geometry of the satellites with respect to the receiver also affects the accuracy of the position

GPS 3-D Positioning Accuracies* (Metres, 95%)

Operating Mode		Dual Frequency		Single Frequency	
SA	A-S	PPS	SPS	PPS	SPS
OFF	OFF	37	37	51	51
ON	OFF	37	170	51	174
OFF	ON	37	NA**	51	51
ON	ON	37	NA**	51	174

* Worldwide over 24 hours; minimum satellite elevation 5°.
** NA indicates not available with A-S on.

GPS Horizontal Positioning Accuracies* (Metres, 95%)

Operating Mode		Dual Frequency		Single Frequency	
SA	A-S	PPS	SPS	PPS	SPS
OFF	OFF	21	21	29	29
ON	OFF	21	98	29	100
OFF	ON	21	NA**	29	29
ON	ON	21	NA**	29	100

* Worldwide over 24 hours; minimum satellite elevation 5°.
** NA indicates not available with A-S on.

GPS Vertical Positioning Accuracies* (Metres, 95%)

Operating Mode		Dual Frequency		Single Frequency	
SA	A-S	PPS	SPS	PPS	SPS
OFF	OFF	34	34	46	46
ON	OFF	34	156	46	159
OFF	ON	34	NA**	46	46
ON	ON	34	NA**	46	159

* Worldwide over 24 hours; minimum satellite elevation 5°.
** NA indicates not available with A-S on.

table 4.2 : Summary of PPS and SPS positioning accuracies under various conditions of SA and A-S [NATO,1991:1].

be able to distinguish between signals received from different satellites, the carrier frequencies are phase modulated with codes known as P- and C/A-codes. While performing measurements on multipath signals, Kranendonk [1992] found indications that cross-interference exists between signals from different satellites. This is considered to be the result of the fact that the C/A-codes for the different satellites are not completely orthogonal to each other, resulting in unwanted peaks in the cross-correlation function.
Spread spectrum interference can simply be modelled the same way as multipath (see above) [Van Nee,1992]. The only difference is that cross-correlation peaks can precede the autocorrelation peak as if they were multipath signals with a negative delay. The fading bandwidth is now equal to the frequency difference of the interfering satellites. Instead of using the SMR, one should use the Signal-to-Interference Ratio (SIR).
Cross-correlation can cause errors of up to a few metres.

9. user dynamics

The effect of user dynamics on the position fix depends very much on the dynamics of the vessel and position of the receiver antenna. A shipborne receiver should be able to accept as input information provided by the ship's attitude and water speed sensors. The heading and water speed input signals can be used to assist in satellite acquisition. If only a poor estimate of the position and time is available locating and locking onto any satellite in view is slowed down. Once the carrier and code tracking loops are locked no position, velocity or time from outside sources is needed. The roll and pitch input signals (if present) can be used to compensate for antenna motion.
The reader is referred to NATO [1991:1 & 1991:2] for more detailed information on this subject.

10. Selective Availability (SA)

Early test results showed that positions obtained by using only C/A code had a much better predictable accuracy than expected. In order to deny unauthorised users the access to this relatively high accuracy, the Block - II SVs have been equipped with SA. SA exists of controlled but to a user unpredictable variations in the C/A-signal, introducing errors in ranges measured. These variations consist of two types :

- a frequency dither introducing errors in the navigation time coded signals;
- offsets in satellite ephemeris data giving an apparent shift of satellite position.

resulting in a wrong pseudo-range. The influence of multipath signals depends on [Van Nee,1992:1 & 1992:2] :

- the Signal-to-Multipath Ratio (SMR) and multipath signal delays;
- fading bandwidth and tracking loop bandwidth (plus pre-detection bandwidth for non-coherent DLL);
- early-late spacing (d) with respect to chip time (T_c);
- the receiver antenna attenuation;
- the measuring technique of the GPS receiver (non-coherent DLL or coherent DLL).

Table 4.1 gives an overview of whether mean code tracking errors must be expected when multipath signals are present or not. For further information, the reader is referred to Van Nee [1992:1 & 1992:2].

	slow fading	fast fading
non-coherent DLL	a non-sinusoidal signal is present, resulting in a mean delay error greater than zero	the resulting 'S-curve' is the summation of the different 'S-curves', leading to a mean delay error greater than zero
coherent DLL	a mean delay error is present	no tracking error is present

Table 4.1 Mean code tracking delay errors in multipath environment.

The influence is restricted to multipath signals with a maximum delay of T_c + d/2. The maximum and mean tracking errors are proportional to the early-late spacing and SMR [Brouwer et al., 1989; Van Nee,1992:1 & 1992:2; Kranendonk, 1992]. In order to reduce the chance of getting interference from multipath signals, receiver antennas that suppress signals arriving from below a certain elevation angle are used. This way, interference from signals reflected at the sea-surface is minimised. In order to avoid reception of signals reflected off the ships structure, the antenna position has to be selected with care.

8. GPS signals cross-interference

All satellites use the same carrier frequencies. In order to

The effect of the troposphere on propagation delays can be modelled using measurements of the surface temperature, atmospheric pressure and relative humidity. The correction model can be split into two parts :

- dry component : errors due to a dry (i.e. without water vapour) troposphere can be predicted to a high degree;
- wet component : more difficult to model mainly because the surface measurements of relative humidity do not accurately reflect the distribution of water vapour along the signal path.

The algorithm normally used in GPS receivers is a function of satellite elevation angle only and typically corrects for 90 percent of the tropospheric delay [Braasch,1990] :

$$\delta R_{tropo} = \frac{2.4224}{0.026 + \sin E} e^{-0.13345 H} \quad \text{[metres]} \qquad (4.3)$$

where
H = altitude of receiver above the earth's surface (in km)
E = satellite elevation angle

As the troposphere is non-dispersive in the RF part of the spectrum, models developed for Transit (see below) are applicable.

5. receiver noise and resolution

The performance of state-of-the art GPS receivers is such that receiver measurement errors due to for example quantization resolution or oscillator-phase noise of a digital tracking loop is small compared to the tracking error introduced by thermal noise. For a typical C/N_0 of 48 dB-Hz, the code phase measurement error due to thermal noise only is in the order of 1.5 m (1σ, 4Hz code tracking loop bandwidth) [Braasch, 1991].

6. multipath errors

Multipath errors are caused by satellite signals that were reflected off surfaces and thus arriving at the receiver delayed in time with respect to the line-of-sight (direct) signals. The receiver cannot distinguish between a line-of-sight and a reflected signal since both are coded the same way. Multipath signals present will lead to a shift of the zero-point using the early-late signal correlation technique, the Earth's atmosphere causing errors due to group delay and carrier advance, scintillation and refraction.

The effect of group delay and carrier phase advance due to the ionosphere can be reduced by using dual-frequency observations because the ionosphere is dispersive in the RF part of the spectrum. From measuring pseudo-ranges on two frequencies, a corrected distance can be found :

$$R_c = \frac{R_{p1} - \left(\frac{f_2}{f_1}\right)^2 R_{p2}}{1 - \left(\frac{f_2}{f_1}\right)^2} \quad \text{[metres]} \qquad (4.2)$$

where

R_c = pseudo-range corrected for ionospheric effects
R_{p1} = pseudo-range measured at frequency 1
R_{p2} = pseudo-range measured at frequency 2
f_1 = carrier frequency 1 (L_1)
f_2 = carrier frequency 2 (L_2)

It should be noted that when code correlation techniques are used to determine pseudo-ranges, L_2 can only be used for calculating the ionospheric correction when P-code access is granted. If access is denied a ionospheric delay model has to be used to reduce ionospheric effects. Parameters to be used in this model are provided in the broadcast ephemeris. These coefficients are updated at 10-day intervals, or more often if necessary, to account for seasonal and solar activity changes. After applying this correction algorithm to single-frequency pseudo-ranges, the remaining residual range error is due to short-term ionospheric range errors not accounted for by the model.
A way to gain access to the L_2 signal without having access to the P-code is by using squaring techniques in which the incoming signal is multiplied by itself. The result is a codeless carrier wave. This way a corrected pseudo-range can be measured from actual observations.
Ionospheric scintillation, which consists of a rapid fluctuation of the Total Electron Content (TEC), is unpredictable, correlated with the solar cycle and particularly severe in high latitudes. It results in a variation of amplitude and Doppler shift of the incoming signals which can lead to a loss of phase lock due to a lower SNR and/or a sudden Doppler shift outside the tracking bandwidth of the receiver.

From formula (4.1) it becomes clear which primary error sources cause the measured pseudo-range to be unequal to the true range. These error sources will be discussed briefly in turn.

1. satellite clock error

   The satellite clock is continuously monitored at the MCS. The offset of the clock from the GPS system time is measured, transmitted as data in the broadcast message and allowed for in the receiver. Remaining errors in the satellite clock (eg. from temperature changes) are very small and can not be distinguished from certain components of ephemeris data errors.

2. receiver clock error

   As is the case with the satellite clock, the clock in the receiver will not be synchronised with the GPS System Time. Unlike satellite clock errors, this difference is not monitored and since the receiver clocks are less accurate than satellite clocks, large errors can develop due to drift. In order to reduce the effect of this error, the receiver clock bias is normally taken as one of the unknowns in the position calculation algorithm.

3. ephemeris errors

   At the MCS the position of the satellites in space is calculated, based on information being received from monitoring stations. The satellite ephemeris data is updated and uplinked to the satellites every second orbit which is about once every 24 hours.

   The errors resulting from ephemeris data can be divided into two groups :

   a. errors common to all satellites used for a position fix. This results in an apparent error in the receiver clock and can be compensated for in calculation of the receiver clock offset;
   b. satellite dependent errors.

   The errors in ranges resulting from errors in ephemeris data are small.

4. ionospheric and tropospheric errors

   Part of the path between satellite and receiver goes through computation model used to calculate MPPs, presents the right results. This is done by means of statistical testing based on hypotheses to detect for example blunders (described in chapter 5) and identification of biases by extending the dynamic state vector (described in chapter 6).

4.2 The error budget of EPF systems

4.2.1 Satellite Position Fixing Systems

NAVSTAR / GPS

A minimum of four satellites is needed to determine a 3D position, whereas a 2D position can be obtained using three satellites, provided the antenna height of the receiver with respect to the spheroid is given as a parameter. The velocity of the receiver in ECEF coordinates can be calculated by using receiver velocity relative to the satellites tracked as determined by the carrier tracking loop. Both positional and velocity information are converted to the WGS84 Earth model.

The main observable is a time difference from code measurement, resulting in a 'pseudo-range' to GPS satellites[1]. The pseudo-range from a satellite to a receiver is given by :

$$R_p = R_t + c\,\delta t_p + c(\delta t_r - \delta t_s) + \delta R_e + \delta R_m + \delta R_n + SA \quad \text{[metres]} \quad (4.1)$$

where

- R_p = measured pseudo-range satellite - receiver
- R_t = true range satellite - receiver
- c = propagation velocity of radio waves
- δt_p = timing error due to propagation delay
- δt_s = timing error due to satellite clock offset from GPS system time
- δt_r = timing error due to receiver clock offset from GPS system time
- δR_e = range error due to errors in ephemeris data
- δR_m = range error due to multipath
- δR_n = range error due to noise
- SA = Selective Availability

[1] Other observables that will not be discussed further here are: integrated Doppler count and carrier phase measurement.

- not converting the position given by the receiver to the geodetic datum of the chart;
- plotting errors.

Except for the rounding off errors, which will be regarded as random errors, all user errors should be regarded as either being systematic errors or blunders.

Apart from the types of errors mentioned above - which contribute to the total error resulting in a figure of precision for a measurement (LOP) -, geometry of transmitting stations with respect to the receiver plays an important role when assessing the accuracy of an MPP obtained by combining LOPs. Dilution of Precision (DOP) has been introduced as dimensionless factors to describe how geometry affects position fix accuracy. Although geometry is inherent to each system, no attention is paid to this subject in this chapter. The least squares algorithm, used to calculate the MPP from observations made, automatically accounts for the effect of stations - receiver geometry when calculating the accuracy of an MPP. This algorithm will be discussed in the next chapter. In this chapter, therefore, the key issue is the assessment of a figure of precision for a single LOP given by each system or sensor. For a short discussion on xDOP see appendix 3.

In the following two sections the systematic and random errors thought to be present in the observables of each of the systems and sensors described in the previous chapter, will be discussed. This leads to an error budget, resulting in a value for standard error as measure of precision of the observable (LOP or position and/or velocity) obtained. The value of standard error given is based on the assumption that no blunders are present and all systematic errors have been removed from the observation only leaving random errors. The standard error can be given in units according to the observable provided by a sensor or system, such as degrees for bearings, metres for ranging systems or centilanes when using hyperbolic position fixing systems. However, when observations of different type are combined to obtain an MPP, it is preferred to have all standard errors stated in the same unit. All values of standard deviation will therefore be converted to metres. The main reasons why this conversion is preferred will be given in the next chapter.

Before starting the discussion on errors present in observations the following remark should be made : **Any systematic error that is not removed from an observation is treated as a random error.** However, errors assumed to be zero mean random errors in nature, that contain in fact a systematic component will lead to biased estimators and therefore to degradation of the predictable accuracy and reliability of the MPP.
Therefore, quality control methods are needed to ensure that the

For each system and sensor the errors that influence the predictable accuracy of its observable can be divided into three groups :

1. System errors

   This group of errors consists all errors inherent to the total system layout, irrespective of the receiver equipment used to obtain information from the system. The group includes errors such as position accuracy of the transmitter stations, stability of transmitter oscillators, susceptibility to sky-wave interference and interference from other signals. But also errors resulting from propagation fluctuations, precipitation, seasonal changes in climate and vegetation. Most of these errors are systematic in nature and corrections to be applied to observations in order to reduce their effect can be found by calibrating the system.

2. Receiver errors

   These errors not only comprise those inherent to the receiver itself such as tracking loop errors (magnitude depending on loop construction, C/N_0 and loop bandwidth), resolution, zero errors and repeater errors, but also those resulting from spatial separation between the receiver and its antenna such as delays caused by leads or radiation noise from power cables. Some of the receiver errors can be allowed for by careful calibration of the system, while others are corrected by applying corrections to the measurements. These corrections are either obtained after long-term measurements using statistics or as a result from mathematical models, and are normally combined with the corrections to allow for system errors as mentioned above.
   When the receiver equipment gives a position as output, errors resulting from imperfections of the algorithms used to convert observations to this position must be regarded as well.

3. User errors

   These errors are those made by the user in deriving a MPP from data given by the receiver. The errors that can be made are numerous and whether they are made or not depends very much on the skill of the operator. The most important errors are:

   - not applying corrections to readings to allow for differences between actual propagation velocity and the mean velocity used to draw lattices on charts;
   - reading off errors;
   - rounding off errors;
   - incorrect receiver settings;
   - not correcting for spatial separation of receiver antennas when combining LOPs from different systems;

4. Error Budgeting

In the previous chapter the main characteristics of several sensors and systems, used for position fixing on board submarines, were outlined. In order to be able to use the information provided by them in the most effective way, it is important to know what sort of errors (their magnitude and characteristics) can be expected under operational conditions. Being able to estimate the errors likely to exist is not only important when deciding whether a system can be used or not, it may even be more important to know them when evaluating the quality of the fix, especially if different systems are integrated to obtain that fix.

In this chapter the most important errors sources for each system and sensor used, will be dealt with. The first section reviews the theory of errors in general, giving an overview of the different types of errors. In the next two sections an error budget for each of the sensors and systems described in chapter 3 will be derived, resulting in figures for standard errors of data obtained - whether this is an LOP or a position fix. For easy reference, the order in which the sensors and systems were discussed in the previous chapter is maintained as much as possible.

The results from this chapter will be used in the Least Squares algorithm and Kalman Filter, which are discussed in chapters 5 and 6 respectively. These algorithms form the basis for calculation of the MPP.

4.1 Types of errors

An error can be defined as '*the difference between a specific value and the correct or standard value*'. We distinguish between the following categories of errors : blunders, systematic errors and random errors, for which definitions were given in chapter 2. In order to be able to evaluate the predictable accuracy of systems, it is important that blunders and systematic errors are removed from observations made. Therefore, not only error sources introducing random errors will be discussed but also those giving systematic errors along with means to reduce their effect. It will, however, not always be possible to remove the systematic errors completely.

As was stated in chapter 2, the total random error of an observed quantity (range, phase-/time difference etc.) is assumed to have a time correlated Gaussian distribution.

3.4.2 Navigation software

Quite a number of software packages of navigation software is already available on the market and more are developed. The programs range from simple programs performing basic functions such as finding course and distance from one place to another, coordinate transformations from one geodetic datum to another or astro-navigation to very sophisticated special purpose programs. What is important with all software programs used for calculation of an MPP from observations is that quality control of the MPP derived is displayed in some sort, making it possible for the navigator to distinguish between systems and/or sensors providing LOPs or information that are reliable and those that are doubtful. This way the navigator will be able to make a choice which systems and/or sensors to combine in order to obtain the best MPP in a statistical sense.
Furthermore, it must be clear to the user of navigation software what sort of algorithms are used to obtain an MPP and whether data is filtered or not and in what way.

through the water, to form the depth reference plane.

3.3.7 Periscope

The periscope consists of a complex set of lenses and mirrors, giving the navigator the possibility to take bearings and measure distances while the submarine is still below the sea surface.

This instrument is only of tactical importance for submarines and not used on board other submersibles. Yet it can be used to collect information for updating the DR position and therefore reducing the dimensions of the POE.

3.4 System integration and software

3.4.1 System integration

The introduction of small microcomputers has had a major influence in the integration of navigation systems. All systems that have been described in this chapter each have their own advantages and disadvantages. By combining the results obtained from different systems not only redundancy is guaranteed, giving a better possibility to increase the predictable accuracy of a position fix; the reliability of the combination of systems is also increased since drifting or total failure of one of the systems is covered by one or more systems still working, making position fixing still possible.

By considering system integration we need to make a distinction between the following methods [Appleyard et al., 1988] :

1. **Integrated position fixing** : In this case raw position information (observables) from several systems and/or sensors is fed into a computer. An MPP is derived by combination of the data available at any time in the computer and is based on a mathematical model to achieve an optimum solution.

2. **Hybrid position fixing** : In this case two (or more) systems are available, each giving an independent position or unbiased observable. These positions or observables are compared with each other (integrity check) and combined to obtain an MPP.

3.3.6 Pressure sensor

Pressure sensors on board submersibles are used to measure the depth with respect to the sea-surface at which the vessel is situated.
The method of obtaining pressure using the pressure sensor is based on measuring the frequency of a precise quartz crystal resonator whose frequency of oscillation varies with pressure induced stress. A quartz crystal temperature signal is provided to thermally compensate the calculated pressure in order to achieve high accuracy over a broad range of temperatures.

The measured frequency (f) is converted into a pressure (P) using the following equation :

$$P = P_d + P_{ref} = \alpha \left| A \left(1 - \frac{f}{f_0} \right) - B \left(1 - \frac{f}{f_0} \right)^2 \right| \quad \text{[Pascal]} \qquad (3.2)$$

where

- f_0 = theoretical frequency of crystal in vacuum provided by manufacturer (≈ 40 kHz)
- A,B = sensor dependent calibration coefficients provided by manufacturer (A ≈ 9750 psi, B ≈ 5000 psi)
- α = conversion factor to convert pressure given in psi to pressure in Pa (α = 6894.757)
- P_d = pressure measured due to water column
- P_{ref} = pressure measured due to atmospheric pressure

Using the calculated pressure obtained from equation (3.2), the depth (d) can be calculated using

$$d = \frac{\beta}{\gamma} [P - P_{ref}] \quad \text{[metres]} \qquad (3.3)$$

where

- β = conversion factor to convert pressure in Pa to metres water column with a specific weight of 1000 kg / m^3 ($\beta = 101.9716 \; 10^{-6}$)
- γ = conversion factor to allow for specific weight of seawater ($\gamma \approx 1.026$)

The depth information provided can be used in combination with the depth calculated using the inclination angle (provided as attitude angle from SINS) in combination with measured velocity

- timing mechanism : is used to measure the time between start of transmission of a pulse and reception of received pulse.

- recorder : The recorder has a broad strip of paper moving slowly over a flat metal surface. A belt on which one or more styluses are fastened runs over two pulleys, driven at a constant speed by an electric motor. When the receiver supplies an electric voltage to the stylus, the upper layer of the paper is burnt away, so leaving a depth trace on the paper. In order to be able to use this information later on, a transmission mark is given at the top of the paper. With modern echo sounders, the depth is also represented digitally.

Depending on the environment in which the echo sounder is used, a transmitting frequency has to be chosen. Low frequencies will transmit energy efficiently over long distances because the power will not be rapidly reduced by attenuation. Although high frequencies are prone to greater attenuation and therefore less range, the pulse duration can be shorter making higher resolution possible. Therefore the following classification can be made :

- shallow water E/S : These echo sounders use high frequencies of about 200 kHz. The pulse will not penetrate sediment and has a high resolution. The E/S can be used in waters up to 100 meters.

- medium depth E/S : These echo sounders, which are used in waters of 100 - 1000 meters, use frequencies of about 30 kHz. These frequencies penetrate soft sediment and have medium resolution.

- deep sea E/S : These echo sounders use frequencies of about 10 kHz. They are used in waters deeper then 1000m. At these frequencies, the pulses penetrate sediments quite deeply, giving echoes of underlying layers. Resolution is completely lost.

In all cases the precision of depth measurements is highly dependable on the knowledge of the actual speed of sound in seawater. In shallow waters this can be obtained rather easily by measurements. However, when navigating in deep waters, especially in the oceans, the speed of sound is initially set to 1500 m/s and tables are used to correct measured depth to actual depth.

For a good working order of the gyro compass, it is important to give the compass system correct speed and latitude information, as this is used for counteracting drift and tilt.

3.3.4 Inclinometer

The inclinometer is a device giving the direction of the main axis of the vessel with respect to the horizontal plane. The most reliable way to get information about the inclination of the submarine is to make use of SINS. Attitude angles can be obtained from the synchro packs attached to each of the gimbals. By comparing the already known angle with the synchro value, an error signal can be produced, which can be used to update the known value.

3.3.5 Echo sounder

The echo sounder (E/S) is simply a sonar with a vertical axis. A pulse of acoustic energy is projected from a transmitter to the seabed and reception of the reflected pulse is measured so that range is derived by multiplying the assumed speed of sound in seawater by half the time measured between transmission and reception. This is represented to the user a on paper trace and/or digital equipment.

The main parts of an echo sounder are :

- pulse generator : generates pulses of electrical energy to be transmitted
- switching unit : connects the output from the pulse generator to the transmitter at the right moment. Then switches back to the receiver.
- transmitter : a piezo-electric transducer that converts the electrical power provided by the pulse generator into an acoustic pulse
- receiver : a piezo-electric transducer that receives the reflected pulses and converts them into an electrical signal
- amplifier : The pre-amplifier boosts the very weak signal to the strength needed to activate the recording system. A power amplifier is used to produce enough power to mark reception of the signal on paper

The main part of the log consists of a coil inserted in a watertight flow probe which is fixed underwater on the outside of the vessel's hull. Two types of the probe are available, one is hull mounted ('Flush' mounted log) and the other is retractable through a sea valve. The probe has a streamlined shape in order to minimize the effect of water being dragged with the ship. This has to be done because the ship's speed is measured relative to the water. The water surrounding the flow sensor acts as the conductor forming the loop and the magnetic field in the coil induces a voltage in the water. This voltage is proportional to the speed of water along the coil, which is essentially the same as the speed of the ship through the water. An electrode, fitted on the probe, picks up the induced voltage and passes it to a measuring device. This device transforms the voltage into speed.

3.3.3 Gyro compass

The principle of the gyro compass is based on a fast rotating gyroscope, whose axis of rotation maintains its direction in space. This means that the direction of the spin axis will change with respect to an earth fixed reference frame due to earth rotation. This is given by two parameters : drift, which is the angular rotation of the spin axis round a local vertical axis (i.e. azimuth of axis), and tilt, the rotation of the spin axis around a local horizontal axis (i.e. inclination of axis). Because of these changes of spin axis with respect to the earth fixed reference frame, it is not possible to use the gyroscope on its own as compass. By counteracting the two disturbances, the spin axis is made 'North seeking'.

The main parts of a gyrocompass system are :

- gyroscope : a specially constructed rotor;
- damping system : a construction exerting forces on the gyroscope to counteract the movement of its axis in the earth centred system, therefore making the gyroscope 'North seeking'. Two damping systems are used : horizontal damping and vertical damping;
- gimbals in which the gyroscope is mounted, giving it three degrees of freedom;
- compass housing;
- corrector mechanism : used to eliminate both the damping error and course, latitude and speed error (see section 3.3 of chapter 4)

In figure 3.4a the basic set up of the accelerometers and gyros in shown while in figure 3.4b the basic principle of deriving speed and distance travelled is shown.

The system as described so far will work when the vessel is stationary on a non-rotating earth. In order to be able to use the system on earth a few extra corrections have to be applied to the platform :

- the platform has to be rotated around an axis parallel to the earth rotation axis to compensate for the earth's rotation and ship's motion in E/W direction;
- the platform has to be rotated around the local E/W-axis to compensate for the ship's motion in N/S direction;
- the platform has to be rotated around the local vertical axis to compensate for convergence of the earth meridians.

The corrections that need to be applied are found by transforming velocities obtained into rotation angles.

As opposed to the closed loops of the horizontal channels, the vertical channel is an open loop. This means that any accelerometer errors are unbounded and will increase with time to a square law. Therefore, a displacement in vertical direction given by the vertical axis of SINS is combined with information provided by the pressure sensor (section 3.3.6). These systems are largely complementary to each other. The SINS vertical channel needs to be bounded by external reference (provided by pressure sensor depth) when used over longer periods, but provides direct information about vertical accelerations and good reference for use during short periods of diving or climbing. The pressure sensor on the other hand, provides good depth information when the submarine is sailing at a nearly horizontal level over longer periods, but is less accurate during short dives or climbs.

3.3.2 Electromagnetic log

The working of the electromagnetic (EM) log is based on the Maxwell-Faraday induction law : if a conductor is moved through a magnetic field, an electric force (E) is induced in the conductor, having its direction at right angles to both the magnetic field (B) and velocity (v) :

$$E = v \times B \qquad (3.1)$$

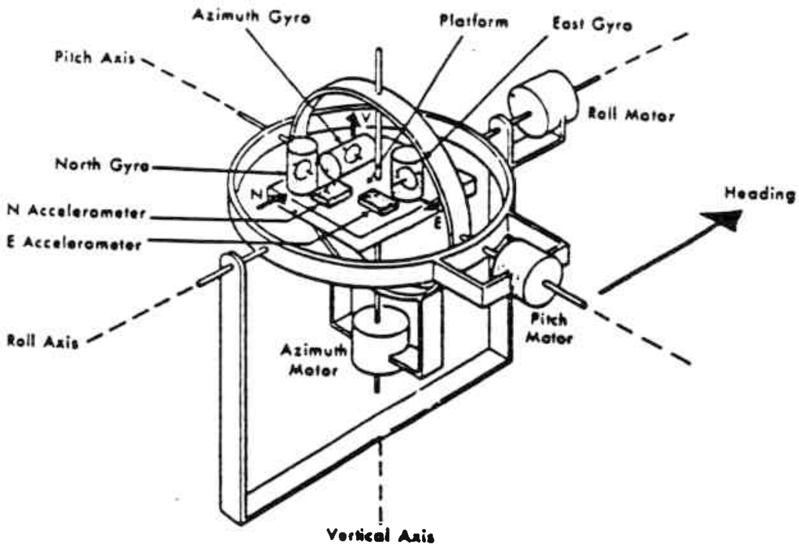


Figure 3.4a SINS platform arrangement.

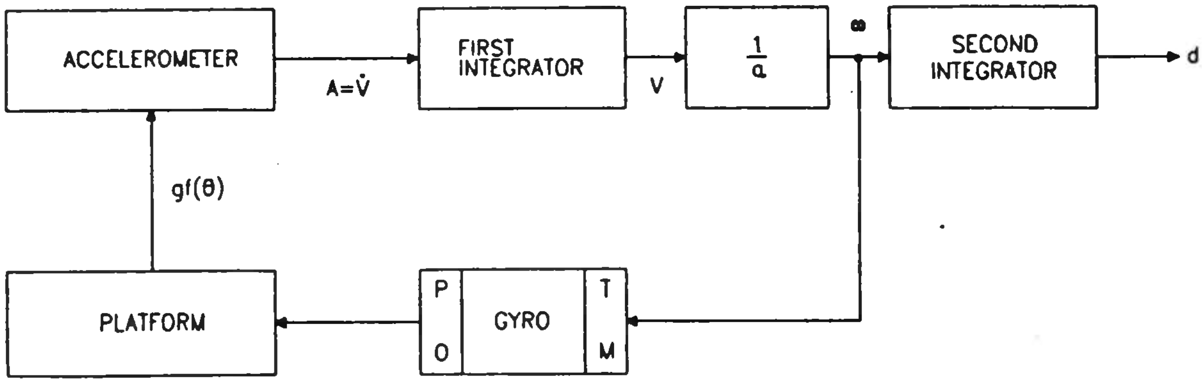


Figure 3.4b Basic principle position calculation using a one channel system.

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 830

Suspended...

Time 08:35:39Z
Systems: G...D
LOPs   :   4
Filter : N/A

CM 005°00.00'E

MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m

0.00040  0.00000  0.00000  0.00000  0.00000
0.00000  0.00040  0.00000  0.00000  0.00000
0.00000  0.00000  0.00110 -0.00024  0.00000
0.00000  0.00000 -0.00024  0.01202  0.00000
0.00000  0.00000  0.00000  0.00000  0.00000

0.00000  0.00000  0.00000  0.00000  0.00000
0.00000  0.00000  0.00000  0.00000  0.00000
0.00000  0.00000  0.00000  0.00000  0.00000
0.00000  0.00000  0.00000  0.00000  0.00000
0.00000  0.00000  0.00000  0.00000  0.00000
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 840

Suspended...

Time 08:35:39Z
Systems: G...D
LOPs   :   4
Filter : N/A

CM 005°00.00'E

MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m

LSE :  935.48      0.92721   0.37455
        76.21     -0.37455   0.92721

       112.00     0.00
         0.00     0.00   -22.00

KF  :

MPP :  935.48      0.92721   0.37455
        76.21     -0.37455   0.92721

       112.00     0.00
         0.00     0.00   -22.00
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 810

Suspended...      | Datum  : ED50
                  | Chain  : 9B
                  | Master :  53°12'12.388" N 007°06'00.590" E
                  | Red    :  55°01'07.110" N 008°41'38.420" E
                  | Green  :  52°35'30.089" N 004°43'47.118" E
Time 08:35:39Z    | Purple :  53°17'07.219" N 009°15'48.858" E
Systems: G...D    |
LOPs   :  4       |   44.862  274.337     1.111
Filter : N/A      |
                  |  -31615.400 5857300.165
CM 005°00.00'E    |  140333.462 5899445.769
                  |  236240.537 6105681.577
MPP               |  -18315.065 5829339.403
Lat 52°50.54'N    |  284292.608 5914986.604
Lon004°31.85'E    |
a      74.87 m    |   177026.82                1.32392
b      21.37 m    |   365220.90   227344.07    0.81675    3.98609    1.07034
Az   112.0°       |    30962.69   173435.71    2.69107    1.58335    5.14909
RMS     6.35 m    |   321036.13   144712.47    1.38371   33.45517    4.49540
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 821

Suspended...      |   0.00000   1.00000       12.45880
                  |   1.00000   0.00000        1.84520
                  |  -0.47409   0.88046       -1.60637
                  |  -0.42885  -0.90336        0.28186
                  |   0.00000   0.00000        0.00000
Time 08:35:39Z    |
Systems: G...D    |  2500.000     0.000     0.000     0.000     0.000
LOPs   :  4       |     0.000  2500.000     0.000     0.000     0.000
Filter : N/A      |     0.000     0.000   916.877    18.050     0.000
                  |     0.000     0.000    18.050    83.585     0.000
CM 005°00.00'E    |     0.000     0.000     0.000     0.000     1.000
                  |
MPP               |
Lat 52°50.54'N    |
Lon004°31.85'E    |
a      74.87 m    |
b      21.37 m    |
Az   112.0°       |
RMS     6.35 m    |
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run              Pg 633

Suspended...          v          σd         w          τ     sei

                    -12.5       52.09      0.26       1.04    *
                     -1.8       60.90      0.04       1.22    *
                      1.6       50.89     -0.09       1.68    *
Time 08:35:39Z       -0.3       44.25      0.17       4.84    *
Systems: G...D        0.0      999.99    999.99     999.99
LOPs    :   4
Filter : N/A       S² :   0.034
                   σ² :   0.021    (4)
CM 005°00.00'E

MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run              Pg 641

Suspended...     Datum      : ED50
                 Projection : TM
                 CM         :  005°00'00.000" E

                               Lat                Lon                x            y
Time 08:35:39Z   LSE :  52°50'32.373" N 004°31'50.958" E  -31615.40   5857300.16
Systems: G...D   KF  :  52°49'59.742" N 004°29'59.447" E  -33709.65   5856305.50
LOPs    :   4    MPP :  52°50'32.373" N 004°31'50.958" E  -31615.40   5857300.16
Filter : N/A
                 a posteriori VCV matrix of MPP :
CM 005°00.00'E
                     814.936  -298.408
MPP                 -298.408   196.755
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m   Standard  :      30.59     8.73      0.00
b      21.37 m   95.0%     :      74.87    21.37      0.00
Az   112.0°      Azimuth   :      112.0
RMS     6.35 m   Elevation :        0.0
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run          Pg 615

Suspended...        Decca statistical data :

                    Chain : 9B

                               Obs     C-O     Calc      LEE       Az
Time 08:35:39Z      Red    :  44.86   -0.05   44.81   3.98609    61.3
Systems: G...D      Green  : 274.34   -0.05  274.29   1.58335   295.0
LOPs   :   4        Purple :   1.11   -0.05    1.01  33.45517   257.6
Filter : N/A
                    a priori VCV matrix (SD) :   916.877    18.050     0.000
CM 005°00.00'E                                    18.050    83.585     0.000
                                                   0.000     0.000     1.000
MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run          Pg 631

Suspended...        Syst   Pat    Read         C-O     Calc            v     Az   Sel
                    GPS    Lat   52°50.55'N   0.00   52°50.54'N    -12.5   90.0   *
                    GPS    Lon 004°31.85'E    0.00  004°31.85'E     -1.8  180.0   *
                    Decca  Red     44.86     -0.05     44.81         1.6   61.3   *
                    Decca  Grn    274.34     -0.05    274.29        -0.3  295.0   *
Time 08:35:39Z      Decca  Prp      1.11     -0.05      1.01              257.6
Systems: G...D
LOPs   :   4
Filter : N/A

CM 005°00.00'E

MPP
Lat 52°50.54'N      Datum : ED50             Proj.: TM     CM : 005°00'00.000" E
Lon004°31.85'E
a      74.87 m      MPP       :  52°50'32.373" N 004°31'50.958" E
b      21.37 m      a/b/c     :   74.87m    21.37m    0.00m
Az   112.0°         Az/elev.  :  112.0°      0.0°
RMS     6.35 m
```

A8.5 Output screen pages layout

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 100
-----------------------------------------------------------------------------
Suspended...    | MPP Datum : ED50                  | LSE Datum : ED50
                |     Lat   :  52°50'32.373" N      |     Lat   :  52°50'32.373" N
                |     Lon   : 004°31'50.958" E      |     Lon   : 004°31'50.958" E
                |     Hg    :    0.0 m              |     Hg    :    0.0 m
GYRO  232.3°    |
LOG    12.7kts  |     RMS   :     6.35m             |     RMS   :     6.35m
RUDR       OFF  |     a / b :   74.87m    21.37m    |     a / b :   74.87m    21.37m
INCL       N/A  |     Az    : 112.0°                |     Az    : 112.0°
PRESS      OFF  |                                   |

GPS              | SINS       | OMEGA      | LORAN-C    | DECCA
                 |            |            |            |
X   3848703.06   |            |            |            | Chain 9B
Y    304893.45   | not        | not        | not        |
Z   5059951.55   | connected  | connected  | connected  | Red     B20.9 *
Lat 52°50.50'N   |            |            |            | Green   F34.3 *
Lon004°31.77'E   |            |            |            | Purple  A51.1
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 611
-----------------------------------------------------------------------------
Suspended...    | GPS statistical data :
                |
                | datum         lat                 lon              Hg
                | WGS84       52°50'30.000" N  004°31'46.200" E    39.5
                | ED50        52°50'32.777" N  004°31'51.052" E    0.00
Time 08:35:39Z  |
Systems: G...D  |         obs             calc
LOPs   :    4   | x :   -31613.554      -31615.400
Filter : N/A    | y :  5857312.625     5857300.165
                |
CM 005°00.00'E  | a priori VCV matrix :  2500.000      0.000
                |                           0.000   2500.000
MPP             |
Lat 52°50.54'N  |
Lon004°31.85'E  |
a      74.87 m  |
b      21.37 m  |
Az   112.0°     |
RMS     6.35 m  |
```

Simulation data file

Date	Time	GPS lat	GPS lon		Decca		Heading	Speed	Rudder
920917	083538	52 50 33.020 N	004 31 51.134 E	.862	.325	.109	232.4	12.7	-.2
920917	083539	52 50 32.781 N	004 31 50.672 E	.862	.337	.111	232.3	12.7	-.2
920917	083540	52 50 32.706 N	004 31 50.497 E	.866	.343	.115	232.2	12.7	-.2
920917	083541	52 50 32.584 N	004 31 50.125 E	.873	.347	.118	232.1	12.7	-.4
920917	083542	52 50 32.509 N	004 31 49.950 E	.877	.353	.118	231.8	12.6	-.5
920917	083543	52 50 32.359 N	004 31 49.663 E	.879	.357	.116	231.7	12.6	.2
920917	083544	52 50 32.269 N	004 31 49.489 E	.872	.361	.112	231.8	12.8	1.2
920917	083545	52 50 32.223 N	004 31 49.289 E	.872	.368	.107	232.3	12.9	1.5
920917	083546	52 50 32.073 N	004 31 48.973 E	.870	.372	.103	232.9	13.0	-.1
920917	083547	52 50 31.997 N	004 31 48.798 E	.869	.376	.104	233.0	13.0	-2.4

etc.

```
NAD 1927         6378206.400  294.9786982     -8.000  160.000  176.000  0.00  0.00  0.00  0.00
NAD 1983         6378137.000  298.257222101    0.000    0.000    0.000  0.00  0.00  0.00  0.00
OSGB 1936        6377563.396  299.3249647    375.000 -111.000  431.000  0.00  0.00  0.00  0.00
TOKYO            6377397.155  299.1528128   -128.000  481.000  664.000  0.00  0.00  0.00  0.00
WGS72            6378135.000  298.26           0.000    0.000    0.000  0.00  0.00  0.00  0.00
WGS72 [DMA]      6378135.000  298.26           0.000    0.000    4.500  0.00  0.00  0.55  0.22
WGS84            6378137.000  298.257223563    0.000    0.000    0.000  0.00  0.00  0.00  0.00
ZANDERIJ         6378388.000  297           -265.000  120.000 -358.000  0.00  0.00  0.00  0.00
*
* End of Data
```

Decca chain data file

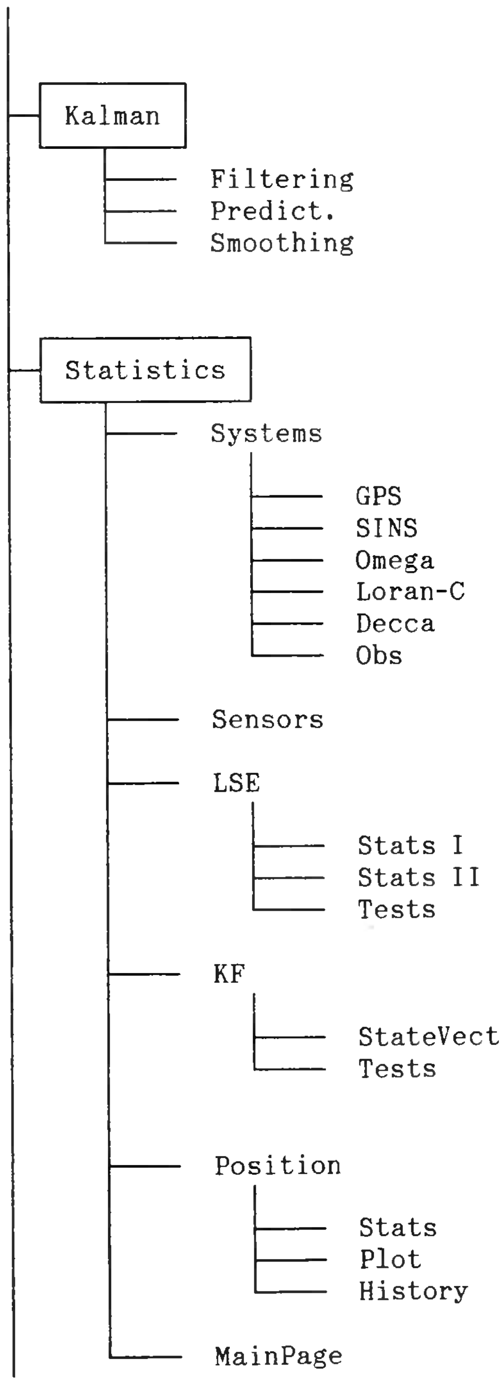
```
* Decca chain data
*
* Frisian Islands Chain (chain 9B)
*
9B
53 12 12.388 N 007 06 00.590 E   85.7200   * Master : Finsterwolde
55 01 07.110 N 008 41 38.420 E  114.2930   * Red    : Hoyer
52 35 30.089 N 004 43 47.118 E  128.5800   * Green  : Heiloo
53 17 07.219 N 009 15 48.858 E   71.4333   * Purple : Zeven
ED50                                       * Geodetic Datum
0                                          * Min. lanecount - Red
0                                          * Min. lanecount - Green
0                                          * Min. Lanecount - Purple
299550.000                                 * Assumed signal prop. speed
*
* Holland Chain (chain 2E)
*
2E
51 36 36.629 N 004 55 36.521 E   84.5500   * Master : Gilze-Rijen
52 35 24.325 N 004 44 37.086 E  112.7333   * Red    : Heiloo
51 13 27.506 N 003 51 41.203 E  126.8250   * Green  : Sas v. Gent
52 11 11.350 N 001 35 46.272 E   70.4583   * Purple : Thorpeness
ED50                                       * Geodetic Datum
-3.920                                     * Min. lanecount - Red
3.205                                      * Min. lanecount - Green
0                                          * Min. lanecount - Purple
299550.000                                 * Assumed signal prop. speed
*
*
* End of Data
```

A8.4 Layout of data files

<u>Geodetic datum data file</u>

```
***********************************************
*
* Parameters of reference spheroids
*
* Update : 29/01/92
*
***********************************************
AIRY 1830                     6377563.396   299.3249646
AIRY MODIFIED                 6377340.189   299.3249646
AUSTRALIAN NATIONAL           6378160       298.25
BESSEL 1841                   6377397.155   299.1528128
BESSEL MODIFIED               6377492.018   299.1528
CLARKE 1858                   6378235.6     294.2606768
CLARKE 1866                   6378206.4     294.9786982
CLARKE 1880                   6378249.145   293.465
CLARKE 1880 MODIFIED          6378249.145   293.4663
EVEREST                       6377276.345   300.8017
EVEREST MODIFIED              6377304.063   300.8017
FISCHER 1960                  6378166       298.3
FISCHER 1968                  6378150       298.3
HAYFORD 1909                  6378388       297
HELMERT 1906                  6378200       298.3
HOUGH                         6378270       297
INTERNATIONAL                 6378388       297
KRASSOVSKY                    6378245       298.3
MADRID 1924                   6378388       297
*
*
***********************************************
*
* Transformation parameters of Datums
*
* Datum ==> WGS84
*
* Update : 08/11/91
*
***********************************************
*
AUSTRALIAN GEODETIC 1966   6378160.000  298.25         -133.000  -48.000   148.000  0.00  0.00  0.00  0.00
AUSTRALIAN GEODETIC 1984   6378160.000  298.25         -134.000  -48.000   149.000  0.00  0.00  0.00  0.00
ED50                       6378388.000  297             -87.000  -98.000  -121.000  0.00  0.00  0.00  0.00
ED50 [DMA]                 6378388.000  297             -87.000  -98.000  -121.000  0.00  0.00  0.00  0.00
ED50 [UKOOA]               6378388.000  297             -86.000  -96.000  -120.000  0.00  0.00  0.00  0.00
ED79                       6378388.000  297             -86.000  -98.000  -119.000  0.00  0.00  0.00  0.00
GRS67                      6378160.000  298.247167427    59.944   -3.843    22.932  0.00  0.00  0.00  0.00
```

- Demo
  - GPS
    - Factor σ
  - Decca
    - PVC
    - Decca CF
  - Params
    - Smple.Int
    - CycleStop
    - S.Weight
  - LSE
    - It.Limit
    - AIM
    - ScaleVCV
  - Plot
    - Size
    - Scale
    - X/Y Fact
  - Recalc LS
- Data
  - Decca
  - A-prio
    - Systems
    - Varcor
  - Weights
  - Ellipsoid
- Run
  - Continue
  - Suspend
  - Start


Kalman
Filtering
Predict.
Smoothing
Statistics
Systems
GPS
SINS
Omega
Loran-C
Decca
Obs
Sensors
LSE
Stats I
Stats II
Tests
KF
StateVect
Tests
Position
Stats
Plot
History
MainPage

A8.3 Menu layout

- Quit
- Edit
  - Prt.Scr.
- Setup
  - Params
    - Datum
    - DR Pos
    - Current
    - Develop
    - Reset
  - Systems
    - GPS
    - SINS
    - Omega
    - Loran-C
    - Decca
  - Sensors
    - Gyro
    - Log
    - Rudder
    - Inclin.
    - Pressure
    - E/S
- Obs
  - Bearing
  - Distance
  - Radar

```
(*------------------------- G P S Datum ------------------------------*)

name := WGS84
axis := 6378137
e2   := 0.00669437999
CM   := 0

(*------------- Default geodetic datum used for calculations ----------*)

name := ED50
axis := 6378388
e2   := 0.00672267002
CM   := 3 / 180 * pi
dX   := -87
dY   := -98
dZ   := -121

(*---------------------------- G P S ---------------------------------*)

PCode            := false         (* C/A code receiver                *)
GPS.SE_SF        := 1             (* error VCV matrix scale factor    *)

(*---------------------------- Decca ---------------------------------*)

Decca.CF         := 64            (* error VCV matrix scale factor    *)
                                  (* Summer's day conditions          *)

(*------------------------ Sensors default ---------------------------*)

s_weight         :=    1.026      (* specific weight of seawater      *)

(*---------------------- LSE variables setup -------------------------*)

epsilon          := 1.0           (* iteration break-off criterion    *)
s02              := 1.0           (* initial value of σ0²             *)
S2               := 0.0           (* initial value of sample variance *)

LOPReject        := 'A'           (* ask if LOPs have to be discarded *)
                                  (* from set after statistical test  *)

(*------------------- statistical variables setup --------------------*)

ReliabInterval   := 5.0           (* reliability interval α = 0.05    *)
ConfRegion       := 95.0          (* confidence level     γ = 95%     *)

ScaleVCV         := false         (* no scaling of a priori VCV matrix *)

(*------------------ Kalman Filter variables setup -------------------*)

DeltaT           := 1             (* sampling interval Δt = 1 sec     *)
```

A8. Computer Simulation Program

A8.1 Simulation program files

```
UVNAV    .EXE    : simulation program executable file

TO_WGS84.DAT     : geodetic datum data file
DECCA    .CHN    : Decca chain data file
UVNAV    .MNU    : menu layout data file
920917_1.RUN     : file with simulation data
```

A8.2 Program parameters default values

```
(*----------------------- Sensors and Systems -------------------------*)

Systems[1]      := 0           (* GPS                                   *)
Systems[2]      := 2           (* SINS                                  *)
Systems[3]      := 2           (* Omega                                 *)
Systems[4]      := 2           (* Loran-C                               *)
Systems[5]      := 0           (* Decca                                 *)

Sensors[ 1]     := 1           (* Gyro                                  *)
Sensors[ 2]     := 1           (* Log                                   *)
Sensors[ 3]     := 0           (* Rudder                                *)
Sensors[ 4]     := 2           (* Inclinometer                          *)
Sensors[ 5]     := 0           (* PressureSensor                        *)
Sensors[ 6]     := 2           (* Echosounder                           *)
Sensors[ 7]     := 2           (* For future use                        *)
Sensors[ 8]     := 2           (* For future use                        *)
Sensors[ 9]     := 2           (* For future use                        *)
Sensors[10]     := 2           (* For future use                        *)

 0  = connected but switched off
 1  = connected switched on
 2  = not connected

(*----------------------- Data files ----------------------------------*)

DatumDataFile   := 'C:\DATA\TO_WGS84.DAT'
DeccaDataFile   := 'C:\DATA\DECCA.CHN'
RunDataFile     := 'C:\DATA\'
```

$$\Phi(t,t_0) = e^{\int_{t_0}^{t}\nabla_x \, d\tau} \approx I + \int_{t_0}^{t}\nabla_x \, d\tau \tag{A7.5}$$

In order to be able to use equation (A7.3) on a computer, it needs to be written in a time discrete form, which can be regarded as sampling of the state, setting $\mathbf{x}_0$ equal to the last best estimate $\hat{\mathbf{x}}_{k|k}$. This leads to

$$\boldsymbol{x}_{k+1} = \Phi_k \boldsymbol{x}_k + \Lambda_k \boldsymbol{u}_k + \Gamma_k \boldsymbol{w}_k \tag{A7.6}$$

where

$$\Lambda_k \boldsymbol{u}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)\left[\boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}) - \nabla_x \hat{\boldsymbol{x}}_{k|k}\right] d\tau$$

$$\Gamma_k \boldsymbol{w}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)\, G(\tau)\, \boldsymbol{w}(\tau)\, d\tau$$

$$\Phi_k = \Phi(t_{k+1},t_k) \approx I + \int_{t_k}^{t_{k+1}} \nabla_x \, d\tau$$

The propagation of the time discrete state vector between two measurements, together with its VCV matrix, is now given by

$$\hat{\boldsymbol{x}}_{k+1|k} = \Phi_k \hat{\boldsymbol{x}}_{k|k} + \Lambda_k \boldsymbol{u}_k \approx \hat{\boldsymbol{x}}_{k|k} + \int_{t_k}^{t_{k+1}} \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k},\tau) + R(\hat{\boldsymbol{x}}_{k|k},\tau) \tag{A7.7a}$$

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \tag{A7.7b}$$

where $R(\hat{\mathbf{x}}_{k|k},\tau)$ is a rest term due to truncation of the Taylor series expansion of $\dot{\mathbf{x}}$ and series expansion of Φ.

A7. Non-linear Filtering

Consider a time continuous system having its state described by the following non-linear vector differential equation

$$\dot{x} = f(x,t) + G(\tau)w(\tau) \tag{A7.1}$$

In order to be able to use this state equation in the Kalman filter algorithm, the equation should be linearized. One method of doing this is by Taylor series expansion of (A7.1), resulting in

$$\dot{x} = f(x_0,t) + \nabla_X\left(x - x_0\right) = \nabla_X x + \left[f(x_0,t) - \nabla_X x_0\right] \tag{A7.2}$$

where

$$\nabla_X = \left.\frac{\partial f}{\partial x}\right|_{x = x_0}$$

The solution of (A7.1) is given by

$$x(t) = \Phi(t,t_0)\,x(t_0) + \int_{t_0}^{t}\Phi(t,\tau)\left[f(x^0,\tau) - \nabla_X x^0\right]d\tau + \int_{t_0}^{t}\Phi(t,\tau)\,G(\tau)\,w(\tau)\,d\tau \tag{A7.3}$$

where $\Phi(t,t_0)$ is the solution of the matrix differential equation

$$\dot{\Phi} = \nabla_X \Phi \qquad \Phi(t,t) = I \quad \forall t \tag{A7.4}$$

resulting in

A first-order Gauss-Markov process, generated by passing an uncorrelated signal w(t) through a linear first-order feedback system - the shaping filter -, is often used to provide an approximation for this band limited signal whose spectral density is flat over a finite bandwidth. The feedback system can be described by the following differential equation

$$\dot{n} = -\beta n + w \qquad \text{(A6.10)}$$

The autocorrelation and spectral density functions of input and output signals for a shaping filter of this kind are then given by

$$R_w(\tau) = 2\beta\sigma^2\delta(\tau) \qquad \text{(A6.11a)}$$

$$\Phi_w(\omega) = 2\beta\sigma^2 \qquad \text{(A6.11b)}$$

$$\Phi_n(\omega) = \frac{2\beta\sigma^2}{\omega^2 + \beta^2} \qquad \text{(A6.11c)}$$

$$R_n(\tau) = \sigma^2 e^{-\beta|\tau|} \qquad \text{(A6.11d)}$$

In the frequency domain the spectral density functions is defined as the Fourier transform of the autocorrelation function

$$\Phi(\omega) = \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau \tag{A6.5}$$

The spectral density functions of the input and output signals are related to each other by

$$\Phi_n(\omega) = |H(j\omega)|^2 \Phi_w(\omega) \tag{A6.6}$$

If the p.d.f. of x(t) and w(t) are restricted to be Gaussian, the process is called a Gauss-Markov process.
The statistics of a stationary first-order Gauss-Markov process are described completely by the following autocorrelation function

$$R_n(\tau) = \sigma^2 e^{-\beta|\tau|} \tag{A6.7}$$

A6.2 Shaping filters

A shaping filter is a linear dynamical system driven by a Gaussian white noise process whose output has the same statistical characteristics as the Gaussian random process $\underline{n}(t)$ that has zero mean and whose correlation is given by

$$E\left[\underline{n}(t)\underline{n}^T(\tau)\right] = D(t,\tau) \tag{A6.8}$$

The autocorrelation of many physical phenomena is well-approximated by

$$R_n(\tau) = \sigma_n^2 e^{-\beta|\tau|} \tag{A6.9}$$

where

β correlation time
σ_n standard error of the noise process.

A6. First-Order Markov Process & Shaping Filters

A6.1 First-Order Markov Process

A continuous process x(t) is a first-order Markov process if for every k and $t_1 \le t_2 \le \ldots \le t_k$ it is true that the distribution function for process $x(t_k)$ is dependent only on the value at one point immediately in the past $x(t_{k-1})$

$$F[x(t_k)|x(t_{k-1}),\ldots,x(t_1)] = F[x(t_k)|x(t_{k-1})] \quad (A6.1)$$

A continuous first-order Markov process can be represented by the following differential equation

$$\dot{x} + \beta(t)x = w(t) \quad (A6.2)$$

The autocorrelation function of input and output are given by

$$R_w(\tau) = E[w(t)w(t+\tau)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_w(\omega)e^{-j\omega\tau}d\omega \quad (A6.3a)$$

$$R_n(\tau) = E[n(t)n(t+\tau)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_n(\omega)e^{-j\omega\tau}d\omega \quad (A6.3b)$$

where

- $\Phi(\omega)$ noise spectral density function
- $R_w(\tau)$ autocorrelation function of input
- $R_n(\tau)$ autocorrelation function of output

They are related to each other by

$$R_n(\tau) = h(-\tau) * h(\tau) * R_w(\tau) \quad (A6.4)$$

where

- $h(\tau)$ time domain impulse response
- $*$ represents the convolution operator.

$$d_0 = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2} \qquad (A5.27)$$

$$\frac{\partial F}{\partial y} = \frac{(y_0 - y_s)}{d_0} \tag{A5.22b}$$

$$\frac{\partial F}{\partial z} = \frac{(z_0 - z_s)}{d_0} \tag{A5.22c}$$

where

$$d_0 = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2 + (z_0 - z_s)^2} \tag{A5.23}$$

<u>azimuth observable</u>

The functional model for an azimuth is written as

$$F(x,l) = \arctan\left(\frac{(x - x_s)}{(y - y_s)}\right) - Az \tag{A5.24}$$

The linearized model Ax = b + v :

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y = Az^0 - \arctan\left(\frac{(x_0 - x_s)}{(y_0 - y_s)}\right) + v \tag{A5.25}$$

Differentials evaluated at $\mathbf{x} = \mathbf{x}_0$ are

$$\frac{\partial F}{\partial x} = \frac{(y_0 - y_s)}{d_0^2} \tag{A5.26a}$$

$$\frac{\partial F}{\partial y} = -\frac{(x_0 - x_s)}{d_0^2} \tag{A5.26b}$$

where

· y coordinate :

functional model

$$F(x,l) = y - y^0 \qquad (A5.17)$$

the linearized model Ax = b + v :

$$\frac{\partial F}{\partial y}\delta y = y^0 - y_0 + v \qquad (A5.18)$$

The differential evaluated at $x = \mathbf{x_0}$ is

$$\frac{\partial F}{\partial y} = 1 \qquad (A5.19)$$

distance observable

The functional model for a measured distance is written as

$$F(x,l) = \sqrt{(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2} - D \qquad (A5.20)$$

The linearized model Ax = b + v :

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z = D^0 - D_0 + v \qquad (A5.21)$$

Differentials evaluated at $x = \mathbf{x_0}$ are

$$\frac{\partial F}{\partial x} = \frac{(x_0 - x_s)}{d_0} \qquad (A5.22a)$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{1}{\lambda_c}\left[\frac{(x_0 - x_m)}{L_m d_{m0}} - \frac{d_{m0}(2x + x_m)}{L_m^2\, 6R^2} - \frac{(x_0 - x_s)}{L_s d_{s0}} + \frac{d_{s0}(2x + x_s)}{L_s^2\, 6R^2}\right]$$

(A5.12a)

$$\frac{\partial F}{\partial y} = \frac{1}{\lambda_c}\left[\frac{(y_0 - y_m)}{L_m d_{m0}} - \frac{(y_0 - y_s)}{L_s d_{s0}}\right] \qquad \text{(A5.12b)}$$

where

$$d_{i0} = \sqrt{(x - x_i)^2 + (y - y_i)^2} \qquad (5.13)$$

<u>Position observable</u>

An observed position in given in grid coordinates leads to two equations :

· <u>x coordinate</u> :

functional model

$$F(x,l) = x - x^0 \qquad \text{(A5.14)}$$

the linearized model $Ax = b + v$:

$$\frac{\partial F}{\partial x}\delta x = x^0 - x_0 + v \qquad \text{(A5.15)}$$

The differential evaluated at $x = x_0$ is

$$\frac{\partial F}{\partial x} = 1 \qquad \text{(A5.16)}$$

$$\frac{\partial F}{\partial N} = -\lambda_s \qquad \text{(A5.7e)}$$

$$\frac{\partial F}{\partial \Delta t'} = \lambda_s (f_g - f_s) \qquad \text{(A5.7f)}$$

where

$$R_1 = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2} \qquad \text{(A5.8a)}$$

$$R_2 = \sqrt{(x_0 - x_2)^2 + (y_0 - y_2)^2 + (z_0 - z_2)^2} \qquad \text{(A5.8b)}$$

Decca observable

The functional model for a Decca lane number is written as

$$F(x,l) = \frac{1}{\lambda_c}\left[\frac{\sqrt{(x - x_m)^2 + (y - y_m)^2}}{L_m} - \frac{\sqrt{(x - x_s)^2 + (y - y_s)^2}}{L_s} + B\right] - N \qquad \text{(A5.9)}$$

with L the line scale factor used to convert grid distances to spheroidal distances, given by

$$L_i = 1 + \frac{x^2 + x x_i + x_i^2}{6 R^2} \qquad \text{(A5.10)}$$

The linearized model $Ax = b + v$:

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y = N^0 - \frac{1}{\lambda_c}\left[\frac{d_{m0}}{L_m} - \frac{d_{s0}}{L_s} + B\right] + v \qquad \text{(A5.11)}$$

where

$$d_0 = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2 + (z_0 - z_s)^2} \qquad (A5.4)$$

Transit observable

The functional model for a range difference based on Dopplercount is written as

$$F(x,l) = (R_2 - R_1) - \lambda_s[N - (f_g - f_s)\Delta t'] \qquad (A5.5)$$

The linearized model Ax + Cv – b = 0 :

$$\left|\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z + \frac{\partial F}{\partial f_s}\delta f_s\right| + \left|\frac{\partial F}{\partial N}v_N + \frac{\partial F}{\partial \Delta t'}v_{\Delta t'}\right| + F(x_0, l^0) = 0 \qquad (A5.6)$$

Differentials evaluated at $\mathbf{x} = \mathbf{x}_0$ are

$$\frac{\partial F}{\partial x} = \frac{(x_0 - x_2)}{R_2} - \frac{(x_0 - x_1)}{R_1} \qquad (A5.7a)$$

$$\frac{\partial F}{\partial y} = \frac{(y_0 - y_2)}{R_2} - \frac{(y_0 - y_1)}{R_1} \qquad (A5.7b)$$

$$\frac{\partial F}{\partial z} = \frac{(z_0 - z_2)}{R_2} - \frac{(z_0 - z_1)}{R_1} \qquad (A5.7c)$$

$$\frac{\partial F}{\partial f_s} = \frac{c[N - f_g\Delta t']}{f_s^2} \qquad (A5.7d)$$

A5. Least Squares Observation Equations

In the least squares calculations, use is made of mathematical models of the observations. In this appendix, the functional model with associating linearized model for the systems as used in the least squares calculations of the simulation program will be given.
In this appendix, the Cartesian coordinates used are coordinates in a TM grid as defined in section 3.3 of chapter 5.

GPS observable

The functional model for a measured pseudo-range is written as

$$F(x,l) = \sqrt{(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2} - c\Delta T - R \qquad (A5.1)$$

The linearized model $Ax = b + v$:

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z + \frac{\partial F}{\partial c\Delta T}\delta c\Delta T = R^0 - (d - c\Delta T)_0 + v \qquad (A5.2)$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{(x_0 - x_s)}{d_0} \qquad (A5.3a)$$

$$\frac{\partial F}{\partial y} = \frac{(y_0 - y_s)}{d_0} \qquad (A5.3b)$$

$$\frac{\partial F}{\partial z} = \frac{(z_0 - z_s)}{d_0} \qquad (A5.3c)$$

$$\frac{\partial F}{\partial c\Delta T} = -1 \qquad (A5.3d)$$

Because transmitting station A is used to obtain both LOPs, correlation will exist in errors present, resulting in a 'full' a priori VCV matrix. The variances and covariances can be calculated as follows :

variance :

$$E[(\delta\Delta\phi_{AB})^2] = \sigma^2_{\Delta AB} = \sigma^2_A + \sigma^2_B + \sigma^2_{syncB} \quad \text{[sq.lanes]} \tag{4.6a}$$

$$E[(\delta\Delta\phi_{AC})^2] = \sigma^2_{\Delta AC} = \sigma^2_A + \sigma^2_C + \sigma^2_{syncC} \quad \text{[sq.lanes]} \tag{4.6b}$$

covariance :

$$E[(\delta\Delta\phi_{AB})(\delta\Delta\phi_{AC})] = \sigma_{\Delta AB\Delta AC} = \sigma^2_A \quad \text{[sq.lanes]} \tag{4.6c}$$

In the above equations, the variances and covariances are given in lanes squared. To get these values in metres squared, the following transformation formula has to be used :

$$\sigma'_{ij} = \frac{1}{4}\sigma_{ij}\lambda^2_c \operatorname{cosec}(\frac{1}{2}\gamma_i)\operatorname{cosec}(\frac{1}{2}\gamma_j) \quad \text{[sq.metres]} \tag{A4.7}$$

In the situation where the two LOPs are obtained using 4 different stations, no correlation exists between the LOPs resulting in a diagonal error VCV matrix (the covariances being equal to zero).

When (A4.3a) and (A4.3b) are used, the vessel is assumed to be positioned well inside the area in which ground wave coverage is maximum and interference by sky waves can be neglected. If this is not the case, unpredictable phase shifts might occur by interference, degrading the accuracy of the Decca LOP.

In the above equations, the variances and covariances are given in lanes squared. To get these values in metres squared, the following transformation formula has to be used :

$$\sigma'_{ij} = \frac{1}{4}\sigma_{ij}\lambda_{c_i}\lambda_{c_j}\mathrm{cosec}(\frac{1}{2}\gamma_i)\mathrm{cosec}(\frac{1}{2}\gamma_j) \quad \text{[sq.metres]} \tag{A4.4}$$

where σ_{ij} = VCV matrix element in sq.lanes

λ_{c_i} = wavelength of comparison frequency for master and slave i

λ_{c_j} = wavelength of comparison frequency for master and slave j

γ_i = at receiver angle subtended between master and slave i

γ_j = at receiver angle subtended between master and slave j

A4.2 Omega error VCV matrix

Consider two Omega LOPs to be used for a position fix. These LOPs are obtained using stations A,B and C. At the receiver phase differences are measured between the signals from A and B (LOP1) and A and C (LOP2). The error in the obtained phase differences are given by :

$$\delta\Delta\varphi_{AB} = \delta\varphi_A - (\delta\varphi_B + \delta\varphi_{syncB}) \tag{A4.5a}$$

$$\delta\Delta\varphi_{AC} = \delta\varphi_A - (\delta\varphi_C + \delta\varphi_{syncC}) \tag{A4.5b}$$

where

$\delta\varphi$ = error in single phase measurement at receiver

$\delta\varphi_{sync}$ = synchronization error of station with respect to station A

$\delta\Delta\varphi$ = error in phase difference measurement

The values for standard deviation of the synchronization errors to be used in equation A4.2 are given by Decca [1979] as

$$3\sigma_{syncR} = 0.012 \quad [\text{lanes}]$$
$$2\sigma_{syncG} = 0.009 \quad [\text{lanes}]$$
$$6\sigma_{syncP} = 0.015 \quad [\text{lanes}]$$

The errors due to sky wave interference have been found to be increasing with covered distance. Table 4.3 given by Decca [1979] shows how standard deviations change with distance. In order to be able to calculate values at intermediate distances, second order polynomials are used to replace the table. These polynomials, calculated using least squares curve fitting are :

$$\sigma = (2.407d^2 - 2.273d + 4.240) * 10^{-3} \quad [\text{lanes}] \qquad \text{(A4.3a)}$$

$$o = (2.843d^2 - 2.597d + 5.000) * 10^{-3} \quad [\text{lanes}] \qquad \text{(A4.3b)}$$

where

d = covered distance from station in 100 km

Formula (A4.3b) should be used to calculate a value for σ_{sP}, whereas formula (A4.3a) has to be used for calculation of all the other variances as given in (A4.2a) and (A4.2b).
The calculated values for standard error using above formulae are valid for summer night conditions over seawater. For a Summer's day these values have to be divided by 8, for a winter's day by 2 and for a winter's night to be multiplied by 1.5.

distance to station	$\sigma_{\blacksquare R}$ / σ_{sR} / σ_{sG} $\sigma_{\blacksquare sR}$ / $\sigma_{\blacksquare sG}$ / $\sigma_{\blacksquare sP}$	$\sigma_{\blacksquare sP}$
100 km	0.0041	0.0050
200 km	0.0100	0.0118
300 km	0.0187	0.0224
400 km	0.0335	0.0400
500 km	0.0532	0.0632

Table A4.1 Standard deviations in transmitted cycles (lanes) for good soil and seawater during Summer nights [Decca,1979].

A4. Decca and Omega Error VCV Matrix

A4.1 Decca error VCV matrix

For each Decca pattern, the phase deviation in the different signals under sky wave conditions, causing errors in the phase measurements at the receiver, can be combined to give an expression for the total error in phase difference of that pattern at the receiver [Decca,1979] :

- Red : $\delta\Delta\varphi_R = 4\delta\varphi_m - 4\delta\varphi_{mR} - 3\delta\varphi_{syncR} - 3\delta\varphi_R$ (A4.1a)
- Green : $\delta\Delta\varphi_G = 3\delta\varphi_m - 3\delta\varphi_{mG} - 2\delta\varphi_{syncG} - 2\delta\varphi_G$ (A4.1b)
- Purple : $\delta\Delta\varphi_P = 5\delta\varphi_m - 5\delta\varphi_{mP} - 6\delta\varphi_{syncP} - 6\delta\varphi_P$ (A4.1c)

Correlation will exist between the phase difference errors of the LOPs because the same master station is used to determine the phase difference for each pattern LOP. This leads to a 'full' a priori VCV matrix, i.e. having both variances and covariances which can be calculated using the following formulae :

variances :

$$E[(\delta\Delta\phi_R)^2] = \sigma^2_{\Delta R} = 16\sigma^2_m + 16\sigma^2_{mR} + 9\sigma^2_R + 9\sigma^2_{syncR} \quad \text{[sq.lanes]} \qquad (A4.2a)$$

$$E[(\delta\Delta\phi_G)^2] = \sigma^2_{\Delta G} = 9\sigma^2_m + 9\sigma^2_{mG} + 4\sigma^2_G + 4\sigma^2_{syncG} \quad \text{[sq.lanes]} \qquad (A4.2b)$$

$$E[(\delta\Delta\phi_P)^2] = \sigma^2_{\Delta P} = 25\sigma^2_R + 25\sigma^2_{mP} + 36\sigma^2_P + 36\sigma^2_{syncP} \quad \text{[sq.lanes]} \qquad (A4.2c)$$

covariances :

$$E[\delta\phi_R\delta\phi_G] = \sigma_{\Delta R\Delta G} = 12\sigma^2_m \quad \text{[sq.lanes]} \qquad (A4.2d)$$

$$E[\delta\phi_R\delta\phi_P] = \sigma_{\Delta R\Delta P} = 20\sigma^2_m \quad \text{[sq.lanes]} \qquad (A4.2e)$$

$$E[\delta\phi_G\delta\phi_P] = \sigma_{\Delta G\Delta P} = 15\sigma^2_m \quad \text{[sq.lanes]} \qquad (A4.2f)$$

- Horizontal Dilution of Precision error (HDOP) - the square root of the sum of the squares of the horizontal components of the position error (σ_x, σ_y). HDOP is sometimes subdivided in its components in the directions East (σ_x) and North (σ_y) which are named EDOP and NDOP respectively.
- Vertical Dilution of Precision (VDOP) - the altitude error (σ_z)
- Time Dilution of Precision (TDOP) - the error in the receiver clock bias, multiplied by the speed of light ($c\sigma_T$).

For submarine navigation, HDOP is most important since GPS can only be used when the submarine is at sea-surface and the vertical component is well known.

A design matrix
$C_{\hat{x}}$ error VCV matrix of parameter estimates

The relationship between the observation errors and parameter estimate errors is a function of satellite geometry only. Therefore, an important consideration in the proper use of GPS is that the four satellites used for the position solution should be arranged in such a geometric relationship that small errors in the pseudo-range measurements (UERE) result in small user position and clock bias errors.
In order to be able to assess the effect of satellite geometry on the position and clock bias errors, it is assumed that each individual pseudo-range measurement has a zero mean error with unit variance, and no correlation exits between the measurements. Under these assumptions, C_l will be the identity matrix, and the error VCV matrix of the parameters will be given by

$$C_{\hat{x}} = \left[A^T A\right]^{-1} \tag{A3.3}$$

The Geometric Dilution of Precision (GDOP) in now defined as the square root of the trace of $C_{\hat{x}}$ when C_l is the identity matrix. Therefore :

$$GDOP = \sqrt{Trace\left|\left(A^T A\right)^{-1}\right|} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_T^2} \tag{A3.4}$$

where

$\sigma_x, \sigma_y, \sigma_z$ standard errors of user position
σ_T standard error of receiver clock bias.

Depending on the user's application, there are five interrelated DOP statistics. In each case they are the amplification factor of pseudo-range measurement errors into user errors due to the effect of satellite geometry with respect to the receiver. Each DOP statistic is a dimensionless number and independent of the coordinate system used.
In a similar way to GDOP, the four other DOPs can be defined as:

- Position Dilution of Precision (PDOP) - the square root of the sum of the squares of the three components of position error $(\sigma_x, \sigma_y, \sigma_z)$;

A3. Dilution of Precision

The GPS exhibits statistical accuracy distributions because of two important parameters :

- User Equivalent Range Error (UERE) :

  A measure of the error in the range measurement to each satellite as seen by the receiver. It tends to be different for each satellite and tends to be at minimum following an upload.

- Geometric Dilution of Precision (GDOP).

  A measure of the error contributed by the geometric relationship of the satellites as seen by the receiver. It varies because the satellites are in constant motion and their geometric relationships are constantly changing.

A3.1 Geometric Dilution of Precision

The radial range measurement equation for one satellite is given by

$$\sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} + c\Delta t - R_i = 0 \qquad (A3.1)$$

where

(x,y,z)	coordinates of receiver
(x_i, y_i, z_i)	coordinates if i-th satellite
Δt	receiver clock bias
c	propagation velocity of radio waves
R_i	pseudo-range measurement to i-th satellite

A minimum of four equations is needed to solve for the four unknowns. A best estimate of the parameters is calculated using the least squares algorithm, which also gives an error VCV matrix of the estimates, based on the a priori VCV matrix of the observations

$$C_{\hat{x}} = \left| A^T C_l^{-1} A \right|^{-1} \qquad (A3.2)$$

where

C_l a priori error VCV matrix of observations

3. for UHF signals (GPS and Transit), signal propagation can be approximated by the laws of optics.
   Vertical profiles of refractivity are estimated, based on measurements at the surface and standard models of the atmosphere. This way corrections for measured times of propagation can be calculated and allowed for. Normally, the corrections are split into two components : a wet and dry component. The dry component is only dependent on vertical pressure and temperature profiles and can be modelled quite well. The vertical water vapour pressure profile is more difficult to model.
   The model most often used for computing the correction for tropospheric refraction is the simplified **Hopfield** model.

A2.4 Propagation in the troposphere

Propagation of radio waves in the troposphere depends on the refractive index. The refractive index of air is a function of height above the earth's surface and is determined by its pressure, temperature and humidity. Variations in refractive index result in two effects : refraction (the angular bending of radio waves) and change in propagation velocity.
Although all signals in the troposphere suffer from tropospheric effects, it becomes most noticeable for frequencies higher then 100 MHz [Van den Berg,1989]. The reason for this is that for higher frequencies the ground wave dampens more quickly, scattering due to objects in the atmosphere is more pronounced and the curvature of the earth plays a more important role.
The navigation systems discussed in this paper can be grouped into three classes, each having its own mechanism describing signal propagation :

1. for VLF signals (Omega), the propagation can best be described by TM modi in a waveguide bounded by the earth's surface and the ionosphere. The propagation velocity depends on the conductivity of the earth's surface and the ionosphere. For Omega, the TM_1 and TM_2 modi play an important role.

2. for the LF signals (Decca and Loran-C), signal propagation is best described by ground wave in combination with a series of signals reflected at the ionosphere, the so-called sky waves.
   The propagation velocity of the ground wave is a function of ground conductivity along the signal path. Normally, the transmission time of a signal from a transmitter to the receiver is considered to consist of two parts : the time needed if the path is an 'all sea' path plus a correction to allow for land path. This correction is called the **Additional Secondary Factor (ASF)**. The ASF for a given position can be obtained in two ways :

   - measure the ASF value in a number of grid points in the area where operating and apply the corrections by interpolation;
   - calculate the ASF values using the method described by Millington. The way this method can be used to correct observations is described by Haagmans [1986].

   When the signals cross coastlines, transient effects occur, resulting in a sudden changes of phase. This can lead to large errors in position fixing when not corrected for.

$$R_c = \frac{f_1^2 R_{p1} - f_2^2 R_{p2}}{f_1^2 - f_2^2} = \frac{R_{p1} - \left(\frac{f_2^2}{f_1^2}\right) R_{p2}}{1 - \left(\frac{f_2^2}{f_1^2}\right)} \quad \text{[metres]} \tag{A2.10}$$

A2.3 Ionospheric correction for TRANSIT signals

The Dopplercount, given by equation (4.6) can be rewritten to allow for ionospheric disturbances using equation (A2.5)

$$N_i = (f_g - f_s)\Delta t + \frac{f_s}{c}\Delta R - \frac{40.25}{c f_s}\Delta TEC \tag{A2.11}$$

If the Dopplercount is measured on two frequencies, both containing a ionospheric disturbance, then for each frequency equation (A2.11) can be formed

$$N_{400}^0 = (f_{g1} - f_{s1})\Delta t + \frac{f_{s1}}{c}\Delta R - \frac{40.25}{c f_{s1}}\Delta TEC \tag{A2.12a}$$

$$N_{150}^0 = (f_{g2} - f_{s2})\Delta t + \frac{f_{s1}}{c}\Delta R - \frac{40.25}{c f_{s2}}\Delta TEC \tag{A2.12b}$$

For Transit, a relation exists between the frequencies of the two transmitted signals, given by $k = f_1 / f_2 = 8/3$. Combining (A2.12a) and (A2.12b) results in

$$N_{400}^0 - k N_{150}^0 = -\frac{40.25}{c f_{s1}}\left[1 - k^2\right]\Delta TEC \tag{A2.13a}$$

or

$$\delta N_{400}^{ion} = -\frac{40.25}{c f_{s1}}\Delta TEC = \frac{N_{400}^0 - k N_{150}^0}{1 - k^2} \tag{A2.13b}$$

$$v_{gr} = n\, c_0 \qquad \text{(A2.6)}$$

Integration of (A2.6), using equations (A2.2) and (A2.3), leads to the expression for group delay :

$$\Delta t' - \Delta t = t' - t = + \frac{40.25}{c_0 f^2}\, TEC \qquad [sec] \qquad \text{(A2.7)}$$

A2.2 Ionospheric correction for GPS signals

The pseudo-range, containing ionospheric disturbances, can be written using equation (A2.7) as

$$c\Delta t = R_p = R_c + \frac{40.25}{f^2}\, TEC \qquad [metres] \qquad \text{(A2.8)}$$

If pseudo-ranges are measured using two frequencies, both containing a ionospheric disturbance, then for each frequency equation (A2.8) can be formed

$$R_{p1} = R_c + \frac{40.25}{f_1^2}\, TEC \qquad [metres] \qquad \text{(A2.9a)}$$

$$R_{p2} = R_c + \frac{40.25}{f_2^2}\, TEC \qquad [metres] \qquad \text{(A2.9b)}$$

From these two equations, the range corrected for ionospheric disturbance can be calculated as

$$n(h) = \sqrt{1 - \frac{f_p^2(h)}{f^2}} = 1 - \frac{1}{2}\frac{f_p^2(h)}{f^2} \qquad \text{(A2.2)}$$

with f_p the plasma frequency which is given by [De Jong,1989] :

$$f_p^2(h) = \frac{N(h)\,e^2}{4\pi^2\epsilon_0 m} = 80.5\,N(h) \qquad \text{(A2.3)}$$

where

h	=	height above the Earth's surface
N	=	electron density
m	=	electron mass
e	=	electron charge
ϵ_0	=	permittivity of vacuum
f	=	signal frequency

Substituting equations (A2.2) and (A2.3) in (A2.1) and defining the Total Electron Content (TEC) along the ray path between transmitting station and receiver as

$$TEC \triangleq \int_{s_0}^{s} N(h)\,dh \qquad \text{(A2.4)}$$

results in a phase advance in time of

$$\Delta t' - \Delta t = t' - t = -\frac{40.25}{c_0 f^2}\,TEC \quad [sec] \qquad \text{(A2.5)}$$

where

R		= total distance travelled
$\Delta t'$	$= t' - t_0$	= total propagation time due to ionosphere
Δt	$= t - t_0$	= total propagation time in free space
s		= upper height of propagation path
s_0		= lower height of propagation path

For modulated signals propagating through the ionosphere an equation similar to equation (A2.5) can be derived using the group velocity (v_{gr}) which is given by

A2. Propagation of Radio waves in the Atmosphere

In this appendix some topics related to propagation of radio waves in the atmosphere will be highlighted briefly. For a more detailed description on this subject, the reader is referred to UKMOD [1984], Haagmans [1986] and De Jong [1989].

A2.1 Propagation in the ionosphere

Because the ionosphere is a frequency dispersive medium, group velocity of a modulated wave will be different from the phase velocity of a monochromatic wave. The group velocity in the ionosphere is smaller than in free space and is referred to as group delay, whereas the phase velocity in the ionosphere is larger than in free space and is referred to as phase advance. Group delay and phase advance are equal in magnitude but opposite in sign. In [Brouwer et al.,1989] it is suggested that this property could be helpful to correct for ionospheric effect on GPS signals when using a C/A code receiver.
In the next part of this section expressions for the group delay and phase advance due to ionosphere will be given.

The propagation time of a monochromatic signal, travelling through the ionosphere can be derived from its equation of propagation velocity, which is given by

$$v_{ph} = \frac{c_0}{n} \tag{A2.1}$$

where

v_{ph} = phase velocity
n = refractive index
c_0 = propagation velocity in vacuum

Direct integration of (A2.1) in order to obtain phase advance is not possible since the refractive index of the ionosphere is dependent on height superscripts

o	observed value
T	transposed of matrix

other

E[.]	mathematical expectation operator
^	estimated value
~	error in estimated value ($\tilde{x} = x - \hat{x}$)
∇	gradient operator
δ	error in measured quantity

$\mathbf{x}$	KF state vector	
$\mathbf{z}$	KF observations vector	
α	significance level	%
α	bearing	°
Γ	time discrete matrix representing the effect of noise on the state	
γ	confidence level	%
γ	angle subtended at receiver by master station and secondary station	°
δN_{400}^{ion}	ionospheric correction for 400 MHz signal Dopplercount of Transit system	
δR_e	range error due to error in ephemeris data	m
δR_m	range error due to multipath	m
δR_n	range error due to noise	m
δR_{tropo}	correction on range for tropospheric effects	m
δt	timing error due to clock offset from GPS time	s
δt_p	timing error due to propagation delay	s
$\delta \varphi_{sync}$	phase angle error due to synchronization	rad
Δt	time difference or time interval	s
Δv	systematic deviation of log	m/s
$\Delta \varphi$	phase difference	rad
$\Delta \psi$	deviation of gyro compass	°
ϵ	inclination angle	°
λ	wavelength	m
λ	longitude	°
σ	standard error	
σ_0^2	unit variance	
τ	reliability factor	
Φ	transition matrix	
φ	phase angle of signal	rad
ψ	submarine's heading	°
ψ_{GK}	gyro course	
ψ_c	direction of current	

subscripts

$k+1\|k$	estimate at time t_{k+1} based on observations $(z_1, \ldots, z_k)$
$k+1\|k+1$	estimate at time t_{k+1} based on observations $(z_1, \ldots, z_{k+1})$
m	master station
o	initial value
r	receiver
s	secondary station
s	satellite

A1.3 Symbols

$\mathbf{A}$	least squares design matrix	m
Az	azimuth	°
B	magnetic field strength	T
B	spheroidal length of baseline	m
$\mathbf{b}$	vector of absolute terms in linearized LS model	m
$\mathbf{C}_{\hat{x}}$	error VCV matrix of Least Squares estimates	m^2
$\mathbf{C}_l$	a-priori error VCV matrix of observations	m^2
c	propagation velocity of EM waves in vacuum	m/s
c_a	propagation velocity of EM waves in air	m/s
D	spheroidal distance	m
d	displacement	m
df	correction on received frequency	Hz
dX,dY,dZ	corrections to Cartesian coordinates in X-, Y- and Z-direction respectively	m
E	electric field strength	N/C
E	elevation	°
f	frequency	Hz
f_g	reference frequency of Transit receiver	Hz
f_s	satellite transmitting frequency	Hz
f_r	frequency of received signal	Hz
$\mathbf{F}$	system dynamics matrix	
$\mathbf{G}$	matrix representing the effect of noise on the state	
$\mathbf{H}$	KF observation matrix	
H	altitude	m
k	minimum lanecount	
k_0	grid scale constant on CM	
L	lane number fraction	
l	vector of observed values	m
N_T	Transit Dopplercount	
N	for ionospheric disturbances corrected Dopplercount	
$\mathbf{P}$	KF state vector error VCV matrix	
$\mathbf{Q}$	KF system noise error VCV matrix	
$\mathbf{R}$	KF measurement noise error VCV matrix	
R	range	m
R_c	pseudo range corrected for ionospheric effects	m
R_p	measured pseudo-range satellite - receiver	m
R_t	true range satellite - receiver	m
v	velocity	m/s
$\mathbf{v}$	vector of LS residuals	m
$\mathbf{v}$	white noise process vector of KF measurements	
v_c	speed of current	m/s
v_l	log speed	m/s
$\mathbf{W}$	LS weight matrix	$1/m^2$
w	w-test observable	
$\mathbf{w}$	white noise process vector of KF system	
$\mathbf{x}$	LS vector of parameters	m

PDOP	Position Dilution of Precision
PLL	PhaseLock Loop
POE	Pool Of Errors
PPS	Precise Positioning Service
PVC	Phase Velocity Correction
RC	Radio beacon
RMS	Root Mean Squared
RPF	Radio Position Fixing
SA	Selective Availability
SAM	Service Area Monitor
SEP	Spherical Error Probable
SID	Sudden Ionospheric Disturbance
SINS	Ships Inertial Navigation System
SIR	Signal-to-Interference Ratio
SMR	Signal-to-Multipath Ratio
SNR	Signal-to-Noise Ratio
SPS	Standard Positioning Service
SV	Space Vehicle
TM	Transverse Mercator
TRANET	TRAcking NETwork
TOT	Time Of Travel
TRP	Time Reference Point
TTFF	Time To First Fix
UEE	User Equipment Error
UERE	User Equivalent Range Error
URE	User Range Error
UTC	Universal Time Coordinate
UT1	Universal Time 1 (= GMT)
VCV	Variance Covariance
VLF	Very Low Frequency
WGS84	World Geodetic System 1984

A1.2 Abbreviations

2D	Two Dimensional
ASF	Additional Secondary Factor
BLUE	Best Linear Unbiased Estimate
C/A - code	Coarse / Acquisition code
CD	Coding Delay
CEP	Circular Error Probable
CM	Central Meridian
CRI	Cross Rate Interference
CW	Continuous Wave
CWI	Carrier Wave Interference
DMA	Defence Mapping Agency
DNS	Decca Navigator System
DR	Dead Reckoning
d_{RMS}	Distance Root Mean Squared
ECEF	Earth Centred Earth Fixed
ED	Emission Delay
EM	Electro Magnetic
EP	Estimated Position
EPF	Electronic Position Fixing
E/S	Echo sounder
FE	Fixed Error
FRP	1990 Federal Radionavigation Plan
GDOP	Geometric Dilution of Precision
GMT	Greenwich Mean Time
GPS	Global Positioning System
GRI	Group Repetition Interval
HDOP	Horizontal Dilution of Precision
H_0	Null hypothesis
H_1	Alternative hypothesis
IHR	International Hydrographic Review
IMO	International Maritime Organization
I/O	Input / Output
KF	Kalman Filter
LEF	Lane Expansion Factor
LF	Low Frequency
LOP	Line Of Position
LSE	Least Squares Estimate
MCS	Master Control Station
MP	Multi Pulse
MPP	Most Probable Position
MSL	Mean Sea Level
NAVSTAR	NAVigation System using Timing And Ranging
NNSS	Navy Navigation Satellite System
OPCT	Omega Propagation Correction Tables
OPNET	OPerational NETwork
PCA	Polar Cap Absorption
P - code	Precise code
pdf	probability density function

Reliability[1] : the probability of performing a specified function without failure under given conditions for a specified period of time

Standard Error[1] : a measure of dispersion of random errors about the mean value

Statistic[2] : any function of a number of random variables, usually identically distributed, that may be used to estimate a population parameter (mean, variance etc.)

Systematic Error : an error that follows some law by which it can be predicted

Unbiased[2] : an estimator $g(X_1,...,X_n)$ for a parameter α is said to be unbiased if

$$E[g(X_1,...,X_n)] = \alpha$$

Unbounded System/Sensor : a system/sensor that needs external signals to give its own output

Variate[2] : a random variable or a numerical value taken by it

1. definitions given in the 1990 Federal Radionavigation Plan [FRP,1990]
2. definitions used in mathematical statistics.

streams, sea state etc.

Error : the difference between a specific value and the correct or standard value

Fix rate : number of independent position fixes that can be obtained per unit of time

GDOP[1] : all geometric factors that degrade the accuracy of position fixes derived from externally referenced navigation systems

Hybrid Position fixing : independent positions of two (or more) systems are compared with each other and combined to obtain an MPP

Integrated Position fixing : an MPP is derived by combination of raw data (ranges, range differences, bearings etc) provided by two or more systems and/or sensors

Integrity[1] : the ability of a system to provide timely warnings to users when the system should not be used for navigation

Most Probable Position : the best position that can be derived using all information available

Optimal Estimator : a computational algorithm that processes measurements to deduce a minimum error estimate of a system by utilising :

i. knowledge of system and measurement dynamics
ii. assumed statistics of the system noise and measurement errors
iii. initial conditions of information

Predictable Accuracy : the accuracy of a statistic derived from measurements, taking into account all predicted errors

Predictable Accuracy[1] : the accuracy of a position with respect to the geographic or geodetic coordinates of the Earth

Precision : the degree of agreement between individual measurements in a set of observations

Random Error : an error unpredictable in magnitude and/or sign, but governed by laws of probability

Reliability : the measure of ease with which a blunder in a measurement can be detected

A1. Definitions, Abbreviations and Symbols

A1.1 Definitions

Accuracy[1] : the degree of conformance between the estimated or measured position at a given time and the true position

Accuracy[2] : the degree of agreement between the true value of a quantity and the most probable value derived from a series of measurements

Ambiguity : the identification of two or more possible positions with the same set of measurements, with no indication of which is the most nearly correct position

Availability[1] :
- navigation system availability is the percentage of time that the service of the system is usable
- signal availability is the percentage of time that navigational signals transmitted from external sources are available for use

Bias[2] : the difference in the mean of a sampling distribution of a statistic and the corresponding population parameter

BLUE : an estimate having the following statistical properties
- the estimate is unbiased
- the VCV matrix of the estimate is the one having minimum trace
- derived quantities have minimum variance

Blunder : a measurement which differs significantly from the expected value making it very likely that certain external circumstances are present than the ones that would make it a random error

Bounded System/Sensor : a system/sensor that does not need external signals to give its own output

Dead Reckoning Position : the position derived by using bounded sensors that measure the vessels heading, speed through the water and attitude, together with the estimated effects of wind, currents and tidal

UKMOD, 1984. The Effects of the Environment on Radio and Radar Wave Propagation. UK Hydrographic Department, Naval Oceanography and Meteorology Memorandum No. 1/84. 116 pp.

UKMOD, 1990. The Assessment of the Precision of Soundings. UK Hydrographic Department, Professional Paper No. 25. 41 pp.

US DoD & DoT, 1990. 1990 Federal Radionavigation Plan (FRP). US Departments of Transportation and Defence, report no. DOT-VNTSC-RSPA-90-3/DOD-4650.4. pp A1 - A38, C1 - C5.

Velberg P.J., 1992. The accuracy of depth information of the nautical chart. Hydrographic Journal, April 1993 (in print).

Willigen D. van, 1991. Radio plaatsbepaling (lecture notes). Delft University of Technology, Department of Electrical Engineering. 130 pp.

Wulder J.H., 1992. The implementation of a ship manoeuvring model in an integrated navigation system. PhD thesis. Delft University of Technology, Department of Mechanical Engineering and Marine Technology. Delft University Press 1992, ISBN 90-370-0063-0. 171 pp.

Lu G., 1992. Statistical Quality Control for Kinematic GPS Positioning. The University of Calgary, Department of Surveying Engineering. pp 1 - 7.

Mikhail E.M., Ackermann F., 1976. Observations and least squares. Harper & Row, New York. 484 pp.

Mikhail E.M., Gracie G., 1981. Analysis and adjustment of survey measurements. Van Nostrand Reinhold Company Inc., London, ISBN 0-442-25369-9. 339 pp.

Nanninga M., 1985. The Accuracy of Echosounding - Description in a Mathematical Model. Master thesis. Delft University of Technology, Department of Geodecy. pp. 115.

NATO, 1991:1. NAVSTAR GPS user equipment, introduction. The NATO team US Air Force Space Systems Division. 192 pp.

NATO, 1991:2. Technical characteristics of the NAVSTAR GPS. The NATO NAVSTAR GPS Technical Support Group. 160 pp.

Nee R.D.J. van, 1992. Multipath Effects on GPS Code Phase Measurements. Navigation : Journal of the Institute of Navigation Vol. 39 No. 2, Summer 1992. pp. 177 - 190.

Nee D.J.R. van, 1992. GPS Multipath and Satellite Interference Delft University of Thechnology, Department of Electrical Engineering. 11 pp.

Pierce J.A., 1965. Omega. Selected Papers Related to Long Range Navigation. Pickard & Burns Electronics, Waltham, U.S.A., Pub. No. 989, June 1966. pp 2-1 - 2-26.

SMS, 1988. The TRANSIT system (lecture notes). School of Military Survey Newbury, report no. WP1521/SIFd/2. 15 pp.

Spaans J.A., 1988:1. Application of the Kalman Filter in Marine Navigation. Delft University of Technology, Ship Hydromechanics Laboratory, Report No.803-P, January 1988. 7 pp.

Spaans J.A., 1988:2. Betrouwbaarheid, nauwkeurigheid en precisie van navigatiemethodieken (lecture notes). Netherlands Naval Academy. 131 pp.

Teunissen P.J.G., 1990. An Integrity and Quality Control Procedure for Use in Multi Sensor Integration. Proceedings of ION GPS-90, Colerado Springs, September 1990. pp 513 - 522.

Teunissen P.J.G., 1991. Differential GPS : Concepts and Quality Control. Invited lecture for the Netherlands Institute of Navigation, Amsterdam, September 1991. 46 pp.

Decca, 1979. The Decca Navigator, Principles and Performance of the System. The Decca Navigator Company Ltd, New Malden (UK). 72 pp.

Draaisma Y., Meester J.J. de, Mulders J.H., Spaans J.A., 1982. Leerboek Navigatie deel 2. De Boer Maritiem, ISBN 90-288-8308-6. 219 pp.

Draaisma Y., Mulders J.H., Spaans J.A., 1986. Leerboek Navigatie deel 3. De Boer Maritiem, ISBN 90-228-3809-4. 206 pp.

Gelb A., 1974. Applied Optimal Estimation. M.I.T. Press, London, ISBN 0-262-70008-5. 374 pp.

Haagmans M.E.E., 1986. De bepaling van korrekties voor propagatieeffekten met behulp van een microcomputer. Master Thesis. Delft University of Technology, Department of Geodecy. 93 pp.

IALA, 1990. IALA Aids to Navigation Guide (NAVGUIDE). International Association of Lighthouse Authorities, Paris, February 1990.

Jong G. de, 1989. Propagatie van radiogolven (lecture notes). Delft University of Technology, Department of Electrical Engineering. 184 pp.

Kielland P., Dagbert M., 1992. The Use of Spatial Statistics in Hydrography. International Hydrographic Review, Monaco, Vol LXIX No.1. pp 71 - 91.

Kielland P., Burrows K., Ward B., Dagbert M., Velberg R., 1992. Towards "IHOstat" : IHO approved software which evaluates the quality of bathymetric data. Proceedings Hydro '92, Copenhagen.

Kranendonk P., 1992. The influence of multipath on stationary GPS-range observations. Master thesis. Delft University of Technology, Department of Electrical Engineering. 57 pp.

Kreyszig E., 1970. Introductory mathematical statistics - Principles and Methods. John Wiley & Sons, New York, ISBN 471-50735-0. 470 pp.

Last D., 1992. The Accuracy and Coverage of Loran-C and of the Decca Navigator System - and the Fallacy of Fixed Errors. Journal of Navigation, Vol. 45 No 1. pp 36 - 51.

Lenart A.S., 1989. ARPA Accuracy Testing. Journal of Navigation, The Institute of Navigation, Vol. 42 No. 1. pp 117 - 123.

Literature

AGARD, 1970. Theory and Applications of Kalman Filtering, NATO Advisory group for aerospace research and development, report no. AGARDograph no.139 (AD 706 306). 537 pp.

Alper S., Bossler D., 1985. Toward improved accuracy standards for hydrographic surveying. International Hydrographic Review, Monaco, Vol. LXII no. 2, July 1985. pp. 39 - 49.

Appleyard S.F., Linford R.S., Yarwood P.J., 1988. Marine Electronic Navigation, second enlarged edition. Routledge & Kegan Paul Ltd, London, ISBN 0-7102-1271-2. 605 pp.

Asche G.P., 1972. The Omega System of Global Navigation. International Hydrographic Review, Monaco, Vol. L No. 1, January 1973. pp 87 - 99.

Braasch M.S., 1990. A Signal Model for GPS. Navigation : Journal of the Institute of Navigation Vol. 37, No. 4, Winter 1990-91. pp 363-377.

Braasch M.S., Graas F van, 1991. Guidance Accuracy Considerations for Realtime GPS Interferometry. Proceedings of ION GPS - 91, Albuquerque, New Mexico, September 1991. pp. 373 - 386.

Breach M.C., 1987. The Global Positioning System or ''Where am I ... now!'' The Royal Engineers Journal, Vol.101, no.3 September 1987, pp 196-207.

Brouwer F.J.J., et al., 1989. Navigatie en Geodetische Puntsbepaling met het Global Positioning System. Delftse Universitaire Pers, Delft, ISBN 90-6275-540-2. 322 pp.

Cross P.A., 1983. Advanced Least Squares to Position Fixing. North East London Polytechnic, Department of Land Surveying, Working Paper No. 6, ISBN 0-907382-06-1. 205 pp.

Cross P.A., Hollwey J.R., Small L.G., 1988. Working Paper No.2 Geodetic Appreciation. North East London Polytechnic, Department of Land Surveying, ISBN 0-907382-01-0. 185 pp.

Defence Mapping Agency (DMA), 1981. Omega Propagation Correction Tables for 10.2 kHz. Defence Mapping Agency Hydrographic/Topographic Center, Washington, Pub. 224 (101-C) B, DMA Stock No. OMPUB224101CB. pp I - IX.

- the program must be able to continue its calculation process while the navigator is using the menu program;

- the software needs to be optimized and it should be considered to rewrite time critical parts of the program in assembler;

- the user interface needs to be optimized by putting relevant information for the navigator together in one or more windows. Maybe 'hot keys' should be defined to be able to skip directly to the most relevant data. Furthermore, a help file should be created to give the user on line information about the data provided on the different output screens as well as information on input parameters to be used so sensible input data is provided;

- a detailed user manual, giving background information on the program and its options has to be provided.

The Kalman filter algorithm

In this paper, sensor dynamics of the gyro compass and SINS have not been considered. The measurement and system noise processes are assumed to be Gaussian white noise processes with standard deviations as given by the error budgets in chapter 4. This means that prediction of the state vector is limited considerably as performance of sensors like the gyro compass and INS depends very much on the movements of the vessel. Sensor dynamic models need to be developed. These models can be used to predict the magnitude of low frequency systematic errors in order to adjust the predicted measurements.

The noise in navigation errors is not white noise but displays time correlation. To allow for this, shaping filters have to be added to the navigation filter mathematical model.

As mentioned above, global detection tests and error adaption algorithms need to be incorporated in the navigation filter.

The error budgets

By investigating literature, error budgets could be formed for systems and sensors available. Sea trials have to prove whether the values given agree with reality.

The computer program

Apart from the improvement needed in the algorithms used to calculate the position, there is much scope for improvement of the first version of the simulation program as presented in this paper. This will not only comprise improvement of the simulation program itself, but also adapting it in order to make it work in a real-time environment. The important updates needed are :

- interfaces have to be written to be able to read the raw data from sensors and systems straight into the program;
- routines have to be developed to incorporate Loran-C, Omega and SINS;
- a tidal prediction model needs to be incorporated in order to be able to correct the measured depth for vertical water movement;
- input routines are needed to check whether the information provided by the user is valid. Furthermore wrong input may not lead to runtime errors;

sensors (eg. logs as mentioned above), improvement can also be achieved by improving the navigation filter. In chapter 6 some drawbacks of the mathematical model as implemented in the simulation model were discussed. One way to overcome them is by refining the dynamic model of the submarine and incorporating sensor dynamics. Another possible way of improving position fixing underwater, is by combining a bathymetric database with echo sounder information using map matching techniques the same way as done in land and air navigation. The problem that arises here is twofold. Firstly, the major part of the seabed is either not surveyed at all or a very long time ago. Secondly, sediment transportation can change the bottom contours rapidly, making the database invalid. Only areas with high shipping density are surveyed regularly, leading to usable data. However, in these areas, the submarine will generally be navigating at the sea surface and thus having sufficient external information.

To conclude it is important to stress again that all that is written in this paper is purely theoretical. The next step is to perform field tests. The results of these tests must be used to validate and improve the dynamic model of the submarine and to validate the error budgets used in the statistical models.

8.2 Suggestions on further development

The limited amount of time available compared to the magnitude of the problem has left much scope for improvement and further research on the items described in this paper. In this section the most important improvements needed will be discussed.

The least squares algorithm

The algorithm as it is implemented in the simulation program was described in chapter 5. It will probably suffice for a real world environment as long as raw system and sensor data is used (eg. GPS pseudo-range measurements instead of the position provided by the receiver).

Although the algorithm checks the integrity of LOPs, resulting in discarding of the unreliable LOPs, it does not tell the navigator what the actual source of the error is (eg. laneslip, signal loss etc.). In further development algorithms that are able to give the navigator the likely nature of the error present should be incorporated in the system.

It is however important to realize that the information displayed is only as accurate as the information provided to the system and underlying algorithms.
In order to obtain unbiased minimum variance (position) estimates, observations must be free of blunders and systematic errors, leaving only random errors present. Furthermore, the algorithms must reflect physical reality. This is not always the case in a dynamic environment. To overcome this problem, tests were given in chapter 5 to check whether or not observables are free from outliers whereas in chapter 6 tests were discussed that can be used to check the validity of the mathematical models. In the computer simulation model only the tests given in chapter 5 and the local tests LOM and LS (chapter 6) are incorporated. This will not suffice completely as errors having a slowly increasing or decreasing tendency (drift) will remain undetected. Therefore global tests (GOM and GS) need to be incorporated as well.

A good user interface, showing the important data in a clear way to the navigator is necessary. The combination of figures and graphs seems to be the best combination. Since position information becomes available and future positions based on present and past information can be predicted, a link of the integrated system with an electronic chart (ECDIS) system seems feasible. This way, the navigator does not have to plot any information but still gets a clear view on the current size of the position confidence region with respect to navigational dangers in the area.

As was mentioned at the end of chapter 2, the configuration of sensors available for position fixing and cross-checking on board submarines is not ideal at the moment. Once the submarine is submerged, cross-checking of systems is limited. SINS can partly provide this task but will still show errors increasing in time, especially when estimating depth over longer periods. Furthermore, SINS is very sensitive to errors made by the operator at start-up of the system. A way to improve velocity and course cross-checking is by using a Doppler log or correlation log. The information provided by these types of log can also be used to make a better estimate of the current speed and direction and the EM log and gyro compass systematic errors. Especially the last mentioned estimate is an important feature : when the submarine is frequently changing course and/or speed, the gyro compass systematic error can become quite large and damps out relatively slowly (damping oscillation period is approximately 84 minutes in which the error is reduced only to 1/3 of the error at the start of the cycle !). If continuous use of these logs is not feasible from an operational point of view, they can be used at intervals to update the estimates. The EM log can be used in between.

When the submarine is submerged, only sparse outside information is available at present. Apart from using other

8. Concluding Remarks and Suggestions for Further Development

8.1 Concluding remarks

Taking the concepts of the POE as starting point has lead to the design of a simulation program of an integrated navigation system that can serve as basis for an integrated system to be used on board submarines. The simulation program clearly shows advantages of integrated navigation with respect to the currently used concept of the POE :

1. it provides an almost instantaneous and continuous calculation of a 3D MPP with associating confidence region based on rigorous algorithms; Furthermore, other parameters such as systematic errors in log and gyro compass can be estimated by the navigation filter, leading to a better performance of the integrated system.

2. observables from different systems and sensors can be combined to obtain an MPP : there is no restriction to the type of the observations that can be incorporated. Furthermore there is no maximum to the number of observables that can be used. The minimum number is given by the number of parameters to be estimated.

3. cross-checking of systems is done by a device which is not subject to fatigue or stress : it provides warnings of incompatibility of receiver information such as for example gyro failure or interruption of signals (**detection**) and gives the possible source (**identification**). Algorithms have been developed that can be used to correct the error (**adaption**). The last-named is not considered in this paper.

4. once the system is started, no input is required from the navigator to keep the system running. He will therefore be able to concentrate more on the quality control side of the position fixes provided instead of having to spend much time on the construction of the position fix with associating error ellipse in the chart.

5. based on the present state of the submarine, it is easy to predict the future position with associating confidence region. No special construction is needed.

6. automatic accurate logging of position information and ship's attitude is provided. This information can be used for post mission analysis (track reconstruction or smoothing) and improvement of the integrated system.

7.3.3 Error messages

In order to make the system as 'fool proof' as possible, most input provided by the operator is checked on consistency. If wrong input is detected, an error message will appear and the operator is requested for new input.

ED50	ED87	OSGB 1936
ED50 [DMA]	NAD 1927	TOKYO
ED50 [UKOOA]	NAD 1983	WGS84

The datum data file contains the ellipsoid and geodetic datum parameters and is given in appendix 8. As this data file is an ASCII file, it is easy to add datum definitions to the file and/or change definitions already given.

Positions

When a position input is needed, this must be a geodetic position (i.e. a position on the spheroid). Each of the following input formats (and combinations of them) can be used :

52 30 00.000 N 003 34 24.735 E[1,2]
52 30 00 N 003 34 24 E
52 30.0 N 003 34.3 E
52 30 N 003 34 E
52.3 N 003.7 E
52 N 003 E

For latitude South, the letter S should be used and for longitude West, the letter W.
The position input format is checked and when a wrong format is detected, the error message **'Wrong position input format !'** will be given after which new input is requested. This is repeated until a correct format is used.

Bearings, distances and speeds

A bearing is given in degrees, counting from zero (North), clockwise to 360°.

Distances are given in nautical miles[3]. For this unit is chosen because most distances used in sea navigation are expressed in nautical miles.

Speeds are given in knots[4]

---

[1] Spaces or commas can be used as field separators.

[2] leading zeroes can be left out

[3] One nautical mile is 1852 metres precise.

[4] One knot equals one nautical mile per hour which can be approximated by 0.5 m/s.

- A-prio :
  this option will show the contents of design matrix, observation vector and a priori VCV matrix used in the least squares calculations.

- Weights :
  this option shows the contents of the weight matrix used in the least squares calculations.

- Kalman :
  this option will show the contents of the matrices and vectors used in the Kalman filter calculations.

- Ellipsoid :
  this option gives the dimensions and orientation of the standard, 95% and 99.9% error ellipse/ellipsoid of the least squares estimate, the predicted Kalman filter estimate and navigation filter estimate of the MPP.

Run

This item is used to control the has three options :

- Continue :
  used to restart the simulation program after it had been suspended.

- Suspend :
  this option is used to halt the program and keep its current state after the last calculation cycle. This way all display pages can be looked at without having the program continuing its calculations. This option can be used to compare data on different pages with each other without having them in the mean time changed by the program.

- Start :
  this item is used to start another simulation run without having to stop quit the program first. All parameters will be set to their default values, bringing the program in the initial startup state. The user gets the main page with the 'Parameter Initialization' cadre.
  In order to avoid accidental activation of this item, the user is asked for confirmation.

7.3.2 Data input formats

Geodetic Datums

The following geodetic datums are supported by the program :

least squares calculations performed on the LOPs, such as : magnitude of residuals, estimated corrected observations, azimuth of the observations and the outcome of the statistical tests.
It will also give the least squares estimate of the MPP with dimensions and orientation of the associating 95% error ellipse.

- KF :
  This item will provide the navigator with the results of the Kalman filter calculation of the state vector and statistical tests performed.
  It will also give the navigation filter estimate of the MPP with dimensions and orientation of the associating 95% error ellipse / ellipsoid.

- Position :
  Under this item, the information about the different positions calculated (LS,KF predicted and MPP) will be given. Also a plot of the current situation will be provided.
  Here also an option 'History' is shown. Its purpose will be to show the effect of filtered positions compared to the positions obtained by the least squares process. This option has not been implemented yet.

- MainPage :
  Gives the user the possibility to return to the main output screen.

Demo

In this menu several parameters having an effect on the program performance and screen layout can be changed : the plot-scale, plot-size, specific weight of seawater, sample interval, corrections to be applied to the observations, scaling of a priori VCV matrix of the least squares algorithm etc.
For each of these parameters a default value is given at startup of the program. The default values are listed in appendix 8.

Data

In this menu the contents of matrices and vectors used in the least squares and Kalman filter process (described in chapters 5 and 6) are shown. Each of the items in this menu has its own output screen. An example of each screen is given in appendix 8. Five options exist:

- Decca :
  Gives the chain data when the Decca navigation system is switched on.

bearing would be used.

Kalman

This menu is used for Kalman filter control (see chapter 6) and consists of three options :

- Filtering :
  Used to switch the navigation filter on and off. The default value is ON. If switched on, the predicted position based on the submarine's dynamic model and bounded sensor input is combined with the least squares position estimate, when available. The filter has to be switched on when the vessel is submerged.

- Predict. :
  Based on the current state vector with associating error VCV matrix, the state vector with associated error VCV matrix at some future time is predicted (see section 6.3.3). The time interval and step size to be used need to be given as input.

- Smoothing :
  Used for post processing of data (not implemented yet).

Statistics

This menu is used to display all calculation results and statistical information provided by the simulation program. The statistical information not only comprises the results of the statistical tests performed by the least squares and Kalman filter algorithm, but also additional information of the systems and sensors used. Each of the items in this menu has its own output screen. An example of each screen is given in appendix 8. The following options are available:

- Systems :
  This option will show some data concerning the systems GPS and Decca. It will show the observed values, corrections applied by the user to the observations, the calculated values by the least squares program, the a priori VCV matrix used for this system etc. Per system, one display page is used.

- Sensors :
  This item will show a display page of the sensor statistics, such as their measured value, systematic error estimate provided by the Kalman filter and standard deviation.

- LSE :
  This item will provide the navigator with the results of the

environment.

- Systems :
This option is used to control the systems to be used for position fixing. In this version of the program only GPS and Decca are implemented as system that can be used. The items are used to switch the systems on and off. Apart from this it is possible to select observables to be used for position fixing.
No systems are switched on as default at program startup.

- Sensors :
This item is used to switch the sensors on and off. In this version of the program the gyro compass, log, inclinometer and pressure sensor can be used. The gyro compass and log are switched on as default when starting up the program.
In order to get sensible results from the navigation filter, it is assumed that log and gyro will not be switched off.

Obs

This menu has three options :

- Bearing :
this item is chosen when a bearing is taken using either a compass repeater, periscope or radar. First the name and coordinates of the station to which the bearing is taken must be given. It is assumed that the coordinates of the station are geodetic coordinates on the datum chosen. Next the bearing is given and the operator has to specify which sensor was used for the observation. This is necessary because the three sensors make use of the same observation equation (see appendix 5) but a different value of standard deviation is used for each of the systems (see chapter 4).

- Distance :
this option is chosen when a radar distance is taken. First the name and coordinates of the station to which the distance is taken must be given. It is assumed that the coordinates of the station are geodetic coordinates on the datum chosen. Next the observed distance is specified.

- Radar :
this option is a combination of the two mentioned above. Instead of having to give the name and position of the station twice when both a bearing and a distance are taken using the radar, this needs to be done only once.
It should be noted that this option must not be used to combine a radar distance with for example a periscope bearing to the same station as the wrong standard error for the

by the program will be given.

7.3.1 Description of menu items

Quit

This item is used to quit the program and to go back to DOS. Before closing down the program, the user is asked for confirmation. If the user decides not to end the program after all, the program is resumed.

Edit

This item has only one option. It allows the user to make a hard copy of the current screen. Although it is possible to make use of the 'Print Screen' key on the key board to make a hard copy of one of the text screens, this is normally not possible for a graphics screen. This menu item allows the user to also make a hard copy of a graphics screen.

Setup

This menu is used to control the sensors and systems to be used for position fixing. It contains three items :

- Params :
  This option allows the user to

  - choose a new datum (Datum) : At startup, a geodetic datum is chosen. All calculations and geodetic positions given will be referred to this datum. With this option it is possible to change the datum to be used.
  - give a new DR position (DR Pos) : When the submarine is at the sea surface, this position will provide the provisional coordinates for the least squares algorithm. If the submarine is submerged, this position will be taken as new starting point for position estimation in the Kalman filter.
  - change current parameters (Current) : at startup, the user is asked for the direction and speed of current. With this option it is possible to change the direction and/or speed when needed.
  - activate the developers menu (Develop) : this option gives access to additional data and output screens, containing information normally of no use to the navigator
  - activate user menu (Reset) : When in the developers menu environment, this option becomes available as an extra option. It is used to switch back to the user menu

If, on the other hand, not enough LOPs are available to calculate the least squares estimate (LSE) based on the observations made, the LOPs are combined with the position based on bounded sensor information and dynamic model in a least squares solution in which the navigation filter provides two position LOPs (one for latitude and one for longitude).

Display of results conform chosen output screen

In order to present results to the navigator, a number of different 'screen pages' have been defined. Each of these pages gives a different kind of information. The pages are chosen via the menu and an example of each of the pages is given in appendix 8.

As was the case with accessibility to certain parameters, again a partition has been made between the output screens accessible to both the navigator and developer and those accessible to the developer only. The pages available to the navigator show the results of calculations (MPP and associated confidence area) and statistical tests, whereas the developer has access to additional pages showing contents of matrices and other parameters used during the calculation process.

7.3 Concise user manual

The program is started from the sub-directory \UVNAV by the command : UVNAV <return>. This will bring up the main page and a 'Parameter Initialization' cadre. Here the user has to specify successively

- the geodetic datum (eg. WGS84, ED50, NAD27 etc.) on which the calculations will be performed;
- the initial DR position;
- the depth below the sea surface;
- the speed and direction of current;
- the input data file (given without extension .RUN) to be used.

Once this is done, the Main Screen is fully shown and the program is ready to be used. The whole simulation process is menu driven and the menus are activated by pressing any key. Once in the main menu, the cursor keys and return key are used to go to the desired menu item and to activate it. A layout of the complete menu is given in appendix 8.

As it is not feasible to explain all the details and options of the program, only a short description of each of the menu items will be given. Also important input formats of data needed

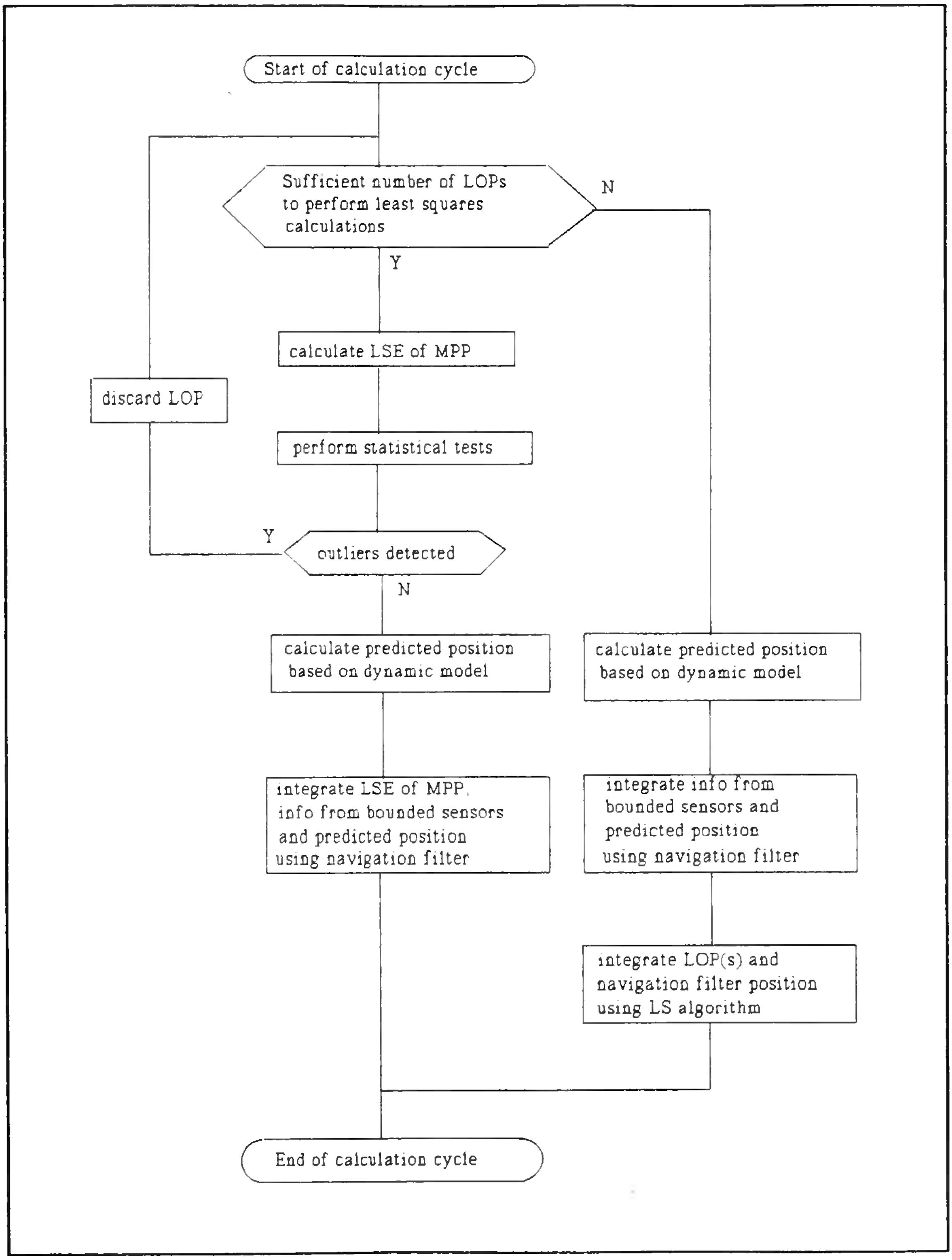


Figure 7.4 Flow diagram of the MPP calculation cycle.

- frequency of pressure sensor crystal is converted to depth in metres

Calculation of position based on dynamic model

When no LOPs are available from external sources, either because the submarine is submerged or simply because the systems have been switched off or not selected by the navigator, an MPP can only be calculated based on the dynamic model of the ship and information from its bounded sensors. It is assumed that gyro and log information is always available as otherwise no MPP could be calculated at all.

The way the position is calculated using the ship's dynamic model is described in chapter 6. In section 5 of this chapter the dynamic measurement and statistical models implemented are described whereas section 4 gives the navigation filter formulae to be used to obtain the position with associating confidence region.

The implementation of the model is performed by using the matrix calculation routines and is as such straightforward. Figure 7.3 shows the flow diagram of this calculation process as implemented.

It should be noted that although statistical tests are performed and results are shown to the user, no attempts are made to correct any outliers detected. The reason for this is that the dynamic model used is very simplified and field tests and further development are needed to improve and validate the model.

MPP calculation cycle

The MPP calculation cycle is used when LOPs are available from observations using unbounded systems and/or sensors. If sufficient LOPs are available (see table 5.1), a least squares estimate of the MPP with associating confidence region can be calculated from these LOPs according to the mathematical model described in section 3 of chapter 5, using equation (5.1).
On the results, statistical tests as described in section 4 of chapter 5 are performed and one or more LOPs are discarded if necessary.
Once the least squares estimate of the MPP has been calculated, it will be combined with the data provided by bounded sensors and the predicted position based on the dynamic model in order to get the filtered MPP with associating confidence region. This is done according to the mathematical and statistical models described in chapter 6 and outlined above.

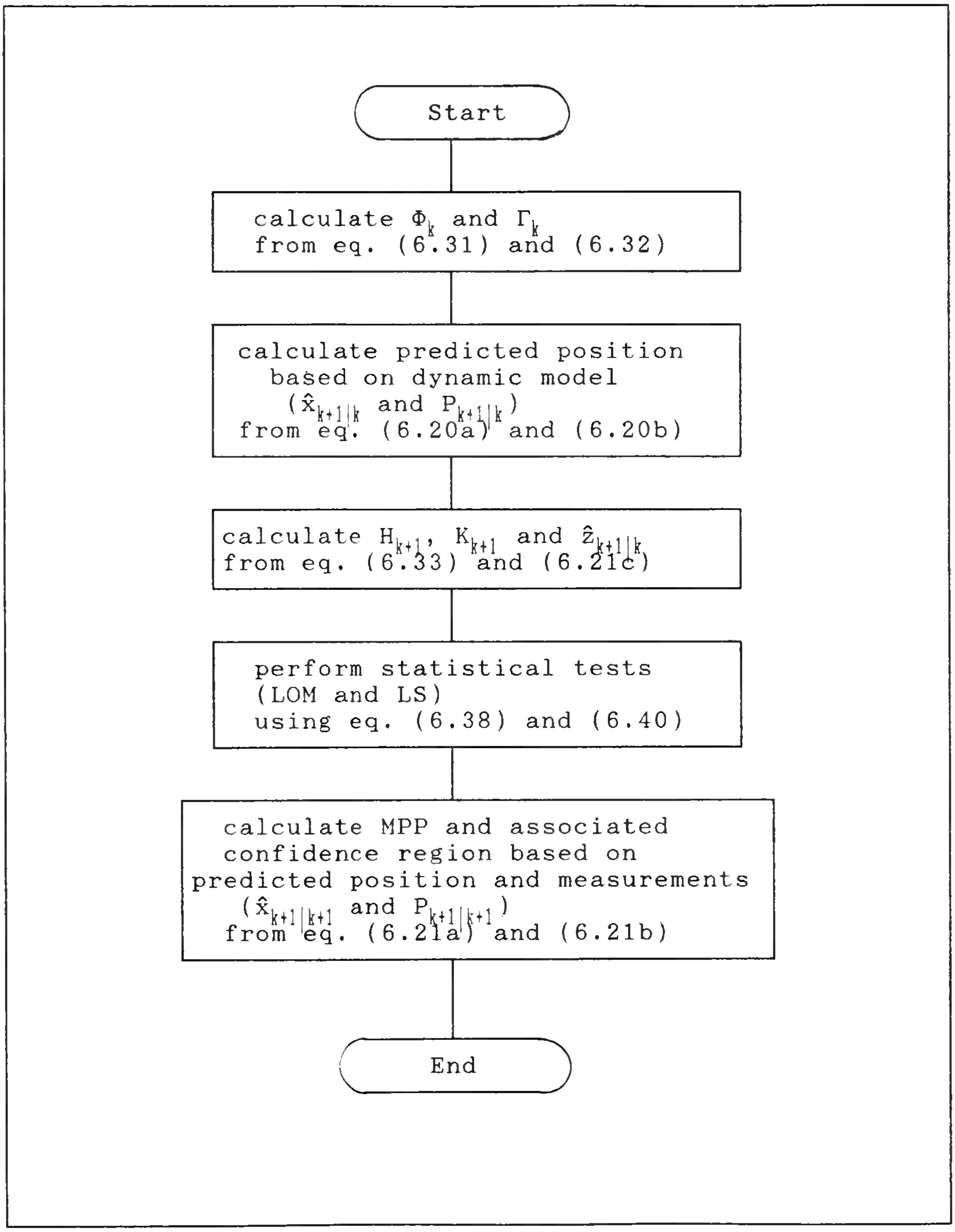


Figure 7.3 Flow diagram of navigation filter position calculation.

- startup position : approx. 52 50 00 N 004 30 00 E
- GPS receiver type : C/A code receiver
- Decca chain : 9B
- Decca pattern Red : B20
- Decca pattern Green : F34
- Decca pattern Purple : A51 (pattern should not be selected because of its large LEF)

- Output Data File

  In order to be able to use system and sensor input data and calculated data for post mission analysis - such as track reconstruction or evaluation of the navigation filter performance -, all input data, system and sensor status (i.e. what was actually selected to be used for position fixing at the time) is written to an output file (.OUT).

7.2 Software description

In this section an outline of the position fixing and quality control part of the program will be given. It will not be an in-depth description. The fundamental algorithms and models implemented in the simulation program are described in detail in chapters 5 (least squares) and 6 (Kalman filter).

Sampling of systems and sensors

In this section of the program, one line of the simulation data file is read at the time. This information is used as input for the system and sensor simulation modules.

The information contained in the data file is signal output of each system and sensor sampled on board HNlMS Buyskes. This information can not be used directly in the computer program. In the system and sensor modules, this information is converted to useable data for the program :

- GPS position string, which is based on WGS84 is converted to information on the datum chosen by the user;

- Decca pattern lane fractions are combined with zone and whole lane number given by the user

- angles given by gyro, rudder and inclinometer are converted to radians;

- Speed given by the log is converted into metres per second

- Geodetic Datum Data File

The Geodetic datum data file (TO_WGS84.DAT) contains the parameters of the most common geodetic datums. This information is used in geodetic calculations on the spheroid (distances, bearings, datum transformations etc.) as well as in grid calculations (conversion of geodetic coordinates to the grid and vice versa, scale factors, grid convergence etc.).

- Decca Chain Data File

The Decca chain data files (DECCA.CHN) contains information about each Decca chain such as the geodetic coordinates of the transmitting stations, the datum on which these coordinates are given, minimum lane counts and assumed signal propagation velocity.

- Simulation Data File

The simulation data file (.RUN) contains system and sensor input data. Each line in this file consists of date, time, GPS position, Decca pattern lane fractions, gyro heading, log speed and rudder angle.

The data file provided contains data collected on board the survey vessel HNlMS Buyskes. It was not possible to collect raw GPS data (pseudo-ranges) on board. In order to save time in developing a simulation package for GPS, the GPS position string given by the receiver is used. This has some drawbacks. Firstly, the data given by the GPS receiver is already filtered in a navigation filter in the receiver. This means that data has been smoothed and incorrect data has been discarded. Secondly, only two LOPs that can be used for position fixing are provided by the GPS receiver. This means that redundancy is decreased and performance of statistical tests limited. Finally, no VCV matrix of the position provided by the GPS receiver is given. To overcome this, a diagonal matrix with a $2d_{rms}$ of 100 m is used for a GPS position provided.

The data collected has been used to check the program for correct working. The amount of data collected was however not sufficient to validate the error budgets given in chapter 4. Furthermore, no special selected tracks could be sailed in order to collect sensor data that could be used to evaluate the filter performance under different circumstances.

In order to be able to use the data in the simulation data file provided, the following initial position and system parameters should be used at startup of the program :

'common block' in which all parameters used by the various routines are placed. This common is used by routines on all levels. The next level down contains the menu control routine, the data input and output routines, the general least squares loop and Kalman filter loop. The next level down contains the routines simulating systems and sensors, the least squares calculation (model and statistical) routines, Kalman filter calculation (model and statistical) routines and the screen layout routines. The bottom level consists of the utility routines such as matrix calculation routines, statistical routines, error routines, geodetic routines etc. Figure 7.2 gives a diagrammatic overview of the different levels.

In order to make a distinction between the information a system developer would need and the information needed by the navigator, a partition has been made in the menu :

- The navigator will have access to all system and sensor setup routines, observation routines, the Kalman filter routines and will be provided with positional and statistical data resulting from calculations. He will also have access to some parameters used for setup and performance of the simulation program.

- The developer will have access to all information provided by the system. This is the information provided to the navigator plus calculation results from several routines, such as for example the contents of the design matrix of the least squares algorithm or the system noise VCV matrix of the Kalman filter. Furthermore, he will be presented with for example chain data of the Decca chain used. The developer is also allowed to change crucial parameters of the program, such as switching scaling of the least squares a priori VCV matrix on and off, introducing offsets in observations etc.

This approach will not be relevant for a simulation program as such as it is the meaning of a program like this to show also for example the effects of ill performance of the sensors and systems on obtained results. It does show however how, in a real-time system on board, information can be shielded off from the navigator.

Data Files used by the program

With the program some data files, necessary for correct working of the simulation program, are provided. Each of these files is an ASCII file, making editing by an ordinary word processor possible. This way, updating of files as well as adding new data to each file is made easy.
The layout of each of these files will be described below. In appendix 8 an example of each of the files will be given.

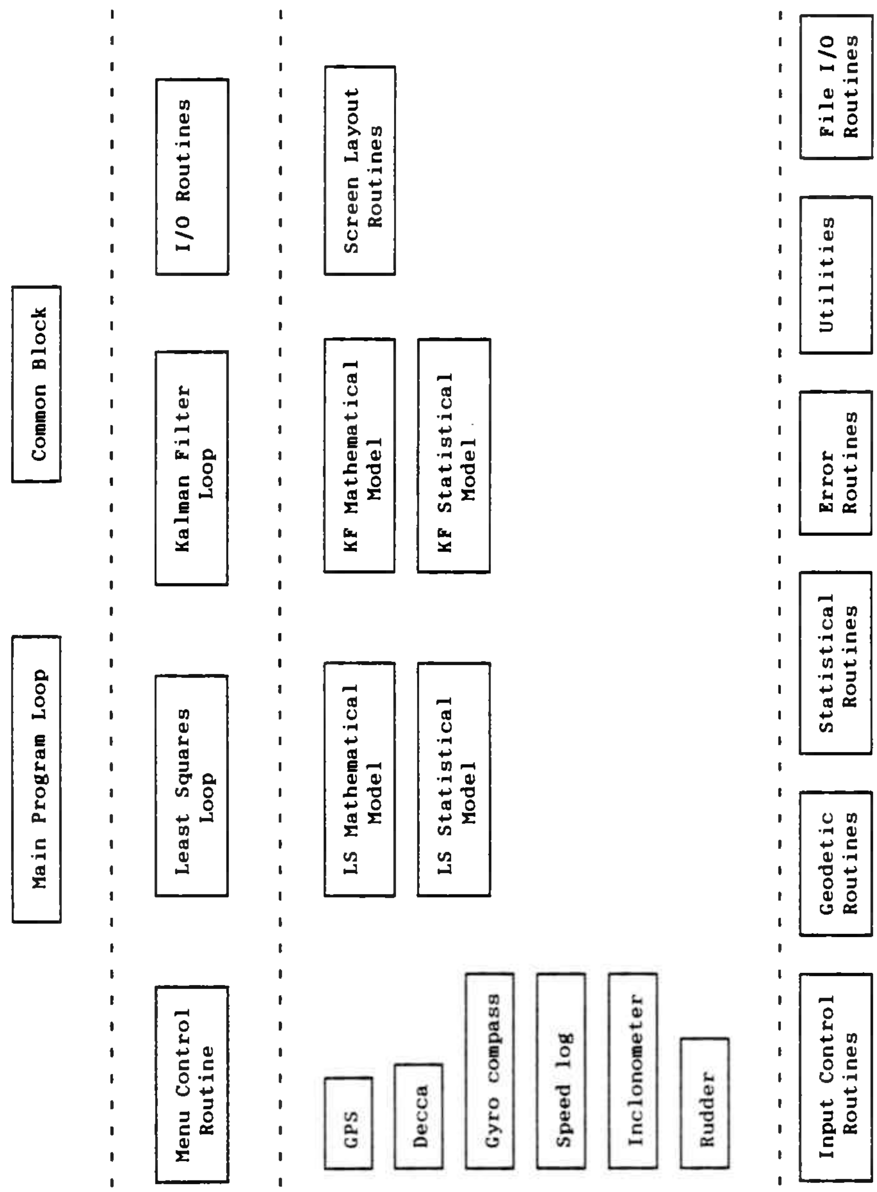


Figure 7.2 Diagrammatic overview of program modules.

operating system (version 3.x and higher). In order to keep the program machine independent no special calls are made to the BIOS.

Using the INSTALL program, a sub-directory '\UVNAV' is created on hard disk (drive C:). This directory contains the program itself (called UVNAV.EXE) plus a sub-directory '\DATA' in which the data files used by the program are placed. The software plus data files take about 400 kB disk space. Running the program from floppy disk is not possible as the program assumes the files to be on hard disk. The program is started by typing UVNAV <return> when in the directory \UVNAV.

General program layout

The simulation program main loop is given in figure 7.1. In the next section each of the blocks will be described in more detail. The program is completely menu-driven, using a pulldown menu system.

As the program is intended to be a simulation program, it was not considered important to put much effort in finding ways to keep the program continuing with the calculation of position fixes while the operator is using the menu. Instead, the menu operation process and calculation process are considered to be two independent processes that do not run simultaneously. This approach leads to the following three states that can be distinguished, once the program is running :

1. **working** : data is read from the simulation data file, the position is calculated and statistical tests are performed. If the cycle is completed, the program starts the next cycle after a time interval given by the operator. The default value of one cycle time is set to one second. This is the minimum time that can be chosen with the data file provided.

2. **interrupted** : the program is stopped and the user has access to the menu. The calculations are resumed by the program itself as soon as possible.

3. **suspended** : the calculation process is halted. This is a special feature of the simulation program. It is now possible to go through all output screens to check the results of the last calculation cycle. The program will resume the calculating process after the user has chosen the menu option 'Continue'.

Apart from the main program loop, several routines are used to perform the calculations and other tasks. Four levels can be distinguished. At the top level is the main program loop with a

Underwater Vehicle Integrated Navigation

author : A.F.Neele

TECHNISCHE UNIVERSITEIT DELFT
Faculteit der Elektrotechniek
Vakgroep Telecommunicatie- en
Verkeersbegeleidingssystemen

Titel : Underwater Vehicle Integrated Navigation

Auteur : A.F. Neele

Aard : Afstudeerverslag

Omvang : 222 pagina's

Datum : 07 april 1993

Hoogleraar : Prof. Dr. Ir. D. van Willigen
Prof. Ir. J.A. Spaans

Mentor(en) : -

Opdrachtnr. : A - 490

Periode : januari 1992 - maart 1993

Most of the research done on navigation nowadays, deals with the optimization of space, air, land and sea-surface navigation. Research into underwater navigation however, is a somewhat neglected area, although there are many submersible vehicles in use, ranging from small unmanned remotely operated vehicles (ROVs) to large nuclear submarines.
Navigators of submarines - these vessels probably representing the largest group of submersibles - are particularly in need of accurate underwater navigation methods. At present, navigators on board submarines of some European countries use a concept called the 'Pool of Errors' (POE) to establish position with associated confidence region. This is a purely graphical method, the accuracy of which is questionable. As this is recognized, a request for an investigation on how to improve underwater navigation on board submarines was forwarded by the Royal Netherlands Navy.

The paper starts with an evaluation of the concept of the POE as currently used, pointing out some important shortcomings in its use. In order to improve underwater navigation and to overcome the shortcomings, a mathematical and a statistical model, based on rigorous formulae, to be used for position fixing and quality control, are proposed. These models are implemented into a computer simulation program which is developed to show the main features of integrated navigation and can serve as a basis for an integrated navigation system to be implemented in a real-time environment on board submarines.

Table of Contents

1. Introduction

Most of the research done on navigation deals with systems that can be used for navigation of vehicles on land, in the air or at the sea surface. There is however a category of users that can only make use of these systems for a limited amount of time. This is the group of navigators on board submersible vehicles, such as submarines. While the vehicle is at the sea surface, use can be made of, for example, satellite and terrestrial navigation systems to obtain an accurate position. Once submerged, information is very much limited to the use of the vessel's bounded sensors or special underwater navigation systems. Acoustic bottom references such as bottom mounted acoustic transponders are impractical because of their limited area of coverage. Perhaps an occasional check-point from a recognizable terrain feature is the best that can be hoped for as a bottom reference. One must therefore fall back on dead reckoning to establish a position.

At present day, navigators on board submarines belonging to navies of some European countries use a concept called 'Pool of Errors' (POE) to establish their position with associating confidence region. This is a purely graphical method which usefulness and accuracy is questionable. Furthermore, it is a time consuming method especially since all factors affecting position accuracy need to be included, leaving no room to identify any errors present in information provided by the systems.

A few reasons may be given to illustrate why a more accurate method of position fixing underwater has to be developed. One of them is the fact that the 'Marineraad'[1] of the Royal Netherlands Navy urged for research on underwater position fixing in one of its verdicts after investigating an accident that had happened to one of the Dutch submarines as a result of misinterpretation of information leading to a wrong position. Another reason is that over the past few years several incidents took place between fishermen and submerged submarines of the Royal Navy (UK) in which the fishermen claimed damage allegedly caused by submarines. In court the evidence of the submarine's position at the time of the accident based on the concept of the 'Pool of Errors', was not admitted as sufficient, resulting in claims by the fishermen being granted. Inaccurate position fixing can also lead to trespassing in territorial waters by any vehicle navigating underwater. This has always been a delicate matter easily escalating into a diplomatic incident.

---

[1] Naval Council

1.1 Aim of the research

Before starting the research, it is important to state what will be investigated and what will be left out of consideration, especially because the variety of sensors and system that can be used for navigation is immense. Furthermore, the group of underwater vehicles comprises the whole range between a small unmanned remotely operated vehicle (ROV) and large nuclear submarines, so a choice has to be made what sort of vehicle and application to focus on.
As the question to investigate how navigation underwater can be improved was posed by the Royal Netherlands Navy, and since the group of submarines probably forms the largest and most important group of submersible vehicles in need for accurate underwater navigation, the research will focus on the way navigation can be improved on board submarines of the Royal Netherlands Navy. This choice automatically leads to the sensors and systems to be considered, namely those present on board the mentioned submarines.

After having chosen platform and range of sensors and systems to be used, the aim of the research has to be described. The aim is threefold :

1. Evaluation of the concept of the POE as it is currently used on board submarines of the Royal Netherlands Navy;

2. Find a suitable mathematical and statistical model that can be used to calculate the Most Probable Position (MPP) by combining information provided by the vessel's bounded and/or unbounded systems and sensors. It must also be able to give figures of position accuracy and perform quality control of information provided.

3. The development of a computer simulation program based on the mathematical and statistical model, that shows the main features of integrated navigation and that can serve as basis for an implementation in a real-time environment (i.e. on board the submarine).

It should be noted that it is not the intention of this research to give the optimum solution for integration of systems, but merely to show how integration of navigation systems using a computer can lead to more reliable position fixing. So each system is regarded as an entity and no special study has been made on integration of a minimum number of systems to reach maximum performance (eg. GPS with an Inertial Navigation System), although when giving suggestions for further research and improvements in the final chapter of this paper, some remarks will be made on this matter.

1.2 Preview

In order to get an impression on the contents of the chapters and their relation with respect to each other, a brief summary of each chapter will be given here.

In chapter 2 the concept of the POE as it is currently used on board submarines will be explained, together with the problems encountered when using it.

After having evaluated the concept of the POE, in chapter 3 the main characteristics of the systems and sensors available at present on board submarines to obtain information for position fixing, are described. The chapter does not deal with the characteristics in great detail since plenty of good textbooks are available. The chapter is merely included to state what is available. Next, Chapter 4 deals with the errors present in the data provided by the systems and sensors described in chapter 3. Although this chapter forms the basis for the statistical models used in calculations, care had to be taken again not to go into too great detail when describing the error sources. Therefore, only of those error sources that contribute significantly to the total error of each system and sensor are briefly described, resulting in an error budget for each observable.

Having dealt with the main characteristics and error sources of the systems and sensors available, the algorithms that can be used to calculate the MPP from observations, need to be looked at. These algorithms, that form the main calculation part of the computer simulation program, are provided in the next two chapters of the paper. In chapter 5 a description is given of the least squares algorithm used to derive an MPP from LOPs obtained from unbounded systems (i.e. external information). The algorithm can only be used when the submarine is at the sea surface. The chapter also contains a section on statistical tests performed to evaluate the quality of the MPP derived by means of least squares. Chapter 6 deals with the Kalman filter algorithm, which is either used to combine the MPP derived using the least squares algorithm with a position based on a ships model and bounded sensor (i.e. internal information) or to predict the probable (future) position of the vessel based on the ships model and bounded sensor information only. The advantage of this algorithm is that it can be used both when the submarine is at the sea surface or submerged.

Chapter 7 will focus on the developed computer simulation program. It will provide an outline of the computer program, a general description of the software, its outputs and a concise user manual. Furthermore some limits of the program will be given.

Finally, chapter 8 contains concluding remarks and suggestions for further research. Here an evaluation of the integrated system - as simulated by the computer program -, will be given. Furthermore some recommendations will be made on the optimum use of already available systems and sensors and possible changes in equipment needed to improve the accuracy of position fixing underwater in particular. Also suggestions for further development of the simulation program and underlying mathematical and statistical models will be given in order to make it ready to be used in a real-time environment.

2. Concept of the Pool of Errors

2.1 Introduction

In this chapter the concept of the Pool of Errors (POE) as it is currently used on board submarines will be discussed. The chapter starts however, with a section containing definitions of important terms that will be used throughout the paper. Since many subject related terms and abbreviations will be used, appendix 1 has been added in which a complete list of definitions, terms and abbreviations used, can be found. In the next section current use of the concept of the POE will be described, using information from existing naval publications on this subject. The third section describes in short how the theory, as discussed in the second section, is actually used on board. In the final section of this chapter some of the most important shortcomings in the use of the concept are listed, giving a good idea why a more rigorous treatment of this subject is needed.

2.2 Definitions

Although appendix 1 will contain a list of definitions, it is important to discuss a few of them here in greater detail before starting to investigate any theories, since a good understanding of their meaning is relevant for a good understanding of the theory and conclusions.

Because a considerable part of the theory of statistics is used in the theory described in this paper, related terms will be explained first. The most important terms to be used are precision, reliability and accuracy. There is generally much confusion on the interpretation of these terms.

- **PRECISION** is the degree of agreement between repeated measurements of the same quantity to each other.

When an instrument is used for observing a quantity, precision gives an indication of the spreading of the measurements when the observation is repeated many times under the same conditions. The need for same conditions is important for the following reasons:

- systematic errors present in the observations will not have particular when the signals travel from East to West the interference is pronounced over a wide area.

B. Irregular errors in assumed propagation model

Forecasting based on long term observations only (on which the OPCT is based), gives the general aspect of phenomena with associating errors that can be expected and is therefore lacking in refinement, since certain sudden irregularities cannot be predicted sufficiently. Consequently, these irregularities lead to degradation of fix accuracy since no corrections can be promulgated in order to reduce their effect.

1. Diurnal ionospheric changes

Propagation conditions during daytime and nighttime are relatively stable, making predictions possible to great extend (< 0.05 lane for daytime and < 0.10 lanes for nighttime). Transition periods (sunrise and sunset) however, are of intermediate stability and present complications in prediction, resulting in considerable departures from predicted corrections, particularly near the end of the sunrise transitions.
It should be noted that even when day or night conditions are present at the receiver station, transition conditions can be present at some part of the propagation path between transmitter and receiver, still influencing the observations.

2. 'Long path' signal reception

Omega signals propagating from the transmitting stations to the magnetic east are attenuated less severely than signals propagating to the magnetic West. Because of this characteristic, an Omega receiver located well west of a station may receive a signal which has come the 'wrong way' round, that is over the long instead of the short path around the Earth. Since Omega lattices on charts and algorithms for position calculation are based on reception of signals using the shortest path from station to receiver, large errors can be introduced.

3. Solar propagation anomalies

Several Omega propagation anomalies result from solar activity. The main sources are :

- Sudden Ionospheric Disturbances (SID) caused by X-ray emitting solar flares;

- Polar Cap Absorption (PCA) caused by proton bombardment of the magnetic polar regions, leading to a concentration of high-energy particles in the region of the magnetic pole with the result that normal VLF transmission is disrupted. The effect of PCA may be to shift an LOP 11 - 15 km (6 - 8 nm) for a period of several days.

C. Receiver errors

For receivers used nowadays, the phase-tracking error is in the order of 0.01λ (depending on C/N_0 and loop bandwidth). The phase-tracking errors can increase due to course and/or speed changes.
If a receiver performs automatic transmitter identification, which is based on pattern matching, wrong identification of the transmitting stations will lead to laneslips.

D. User errors

Errors similar to those existing with the Decca system can be observed when using Omega. An extra error inherent to the Omega system is misidentification of stations. A wrong station identification when the receiver is manually synchronized with the signal transmission format will result in laneslips.

It can be concluded that total error in Omega LOPs obtained is mainly dependent on the validity of corrections promulgated (accuracy depending on propagation models available) and the means to react adequately on sudden disturbances (PCA and SID), and not due to synchronization of transmissions, noise or receiver performance.
Because of the unpredictable nature of errors present, no definite value for the total standard error of LOPs can be given. Under normal conditions (no PCA or SID present) the following approximated values can be used [Pierce,1965;Draaisma,1982] :

	σ_{LOP}
day path	4.9 μsec
transition path	8.7 μsec
night path	8.9 μsec

These values are transformed into standard errors in metres:

$$\sigma_m = (2 f_c \sigma_{LOP}) \frac{1}{2}\lambda_c \operatorname{cosec}(\frac{1}{2}\gamma) \quad \text{[metres]} \qquad (4.13)$$

In the Federal Radio Navigation Plan [1990] it is stated that "*in most cases the predictable accuracy is consistent with the 2 - 4 nm (2d$_{rms}$) system design goal*".
Field observations show that an error of 2 nm (2d$_{rms}$) during daytime and 4 - 5 nm (2d$_{rms}$) at night are good working values. Propagation anomalies can, however, cause large errors up to 10 nm lasting for days at a time.

A way to increase accuracy is by using Differential Omega. This system is based on continuously monitored Omega signals at monitor stations. By comparing at the station the received signal with a predicted signal, corrections can be calculated. These corrections can then be transmitted to users in the area. Using this technique, Omega predictable accuracy (2d$_{rms}$) can be increased to 450 m (0.25 nm) for receivers within 100 km (50 nm) from the monitor station and to 2 km (1 nm) at daytime and 3 km (1.5 nm) at nighttime when the receiver is 550 km (300 nm) from the monitor station.

LORAN-C

This system is based on measuring the difference in time of arrival (TOA) between pulses from the master- and a secondary station. At the receiver the measured time difference is given by :

$$\Delta t = \left[\frac{D_{ms}}{c_a} + CD + \frac{D_s}{c_a}\right] - \frac{D_m}{c_a} = \frac{(D_s - D_m)}{c_a} + \left[\frac{D_{ms}}{c_a} + CD\right] \quad \text{[sec]} \qquad (4.14)$$

where

D_{ms} = distance between master station and secondary
D_m = distance between master station and receiver
D_s = distance between secondary station and receiver
CD = secondary coding delay
c_a = propagation velocity of radio waves

From this equation it can be seen that lines of equal time difference are geodetic hyperbolas with the master- and secondary station at the focal points.

Several factors are responsible for unwanted phase shifts in the pulses received at the receiver, thus causing errors in measured time differences. The main sources are :

A. System errors

1. 'Man made' interference

The Loran-C system suffers from two types of interfering signals [Beckmann,1992] :

- signals generated by the Loran-C system itself : these originate from a different chain operating close to the chain used for positioning. This type of interference is called Cross Rate Interference (CRI);
- all other man-made signals in the spectrum from 50 to 150 kHz. These signals are called Carrier Wave Interference (CWI).

The mentioned groups of interference signals can be subdivided into the following categories :

- synchronous interference signals;
- near-synchronous signals;
- a-synchronous signals.

Each of these three classes generates different problems for Loran-C receivers. The effect of these types of interference on position accuracy is subject of current research. The reader is suggested to Beckmann [1992] for detailed information on this subject.

It is assumed that the effect of CWI and CRI can be equated to an atmospheric noise level of 61 dBµV/m [Last,1992]. This way existing prediction methods, that take only atmospheric noise into account, can be used to calculate the coverage area for which position accuracy is stated below.

2. Sky wave interference

Apart from the above mentioned man-made interference signals, sky wave interference is also present with the Loran-C system. Depending on the distance between transmitting station and receiver, and on effective ionospheric height, sky waves can arrive at the receiver as early as 35 µsec and as late as 1000 µsec after the ground wave. In the first case, sky wave and ground wave

from the same pulse interfere with each other. To overcome this problem, only the first part of the ground wave pulse is used for measurement of time difference. In the second case, the sky wave of one pulse interferes with the ground wave of the succeeding pulse of the eight pulse transmission sequence. By using phase coding on the pulses, the receiver is able to discriminate between ground wave and sky wave of a composite pulse.
The above mentioned provisions make the Loran-C system essentially immune to sky wave signals at most times over most of its coverage area.

3. Inaccuracies in assumed propagation model

The propagation time needed for a signal to travel from a transmitter to receiver is taken as the time needed if the path had been entirely over sea water plus an 'Additional Secondary Factor' (ASF) to allow for land stretches along the path. The ASF values are published and are either entered manually into the receiver or stored in a ROM in the receiver. This way, the effects of land paths on LOP precision are reduced to the residual inaccuracies of ASF mapping and to the effects of seasonal variation in velocity of propagation.

4. Transmitter synchronization

All transmitting stations are equipped with Cs frequency standards. The high stability and accuracy of these standards permit each station to derive its own time of transmission without reference to another station. The objective for control of a Loran-C chain is to maintain constant the observed time difference (Δt) of each master-secondary pair throughout the coverage area. Frequency offsets in the cesium standards and changes in propagation conditions can cause fluctuations in the observed time differences. Therefore, one or more **Service Area Monitor (SAM)** stations with precise receiving equipment are established in the coverage area to monitor continuously the time differences of the master-secondary pairs. When the observed time difference varies from a control time difference (which is established during chain calibration) by one-half of the prescribed tolerance (typically 200 nsec or better) the SAM directs a change in the timing of the secondary station to remove the error.

B. Receiver errors

1. Filtering in the receiver

Because of its burst-type character, the spectrum of the Loran-C signal takes up frequencies roughly between 50 and 150 kHz. In this frequency band a lot of CWI-signals are present. However, the frequency band from 90 -110 kHz only contains Loran-C transmissions.
In order to get rid of interfering signals a band-pass filter is used in the receiver. The effect of this filter on the received Loran-C signal is that the modulation waveform is delayed (40 - 60 µsec !) and distorted : if a band-pass filter with steep slopes is used, most of the CWI signals are suppressed. However, the rising edge of the burst is lost making the sky wave rejection capability of the receiver difficult if not impossible. On the other hand, if the filter has gentle slopes, distortion of the modulation waveform is less severe leading to good sky wave suppression but allowing more CWI signals to distort the phase.
Good filter design is therefore very important and many research is done in developing and optimizing filters.

2. Cycle identification

This process determines the proper cycle of the Loran-C cycle to be used for time difference measurements. An error of one cycle results in a time range error to a station of 10 µsec.

3. Zero crossing phase tracking

Especially synchronous and near-synchronous interference can cause errors in range measurements as these interferences cannot be distinguished from a frequency shift due to a change in receiver position. The error introduced cannot be detected or removed by analysis and filtering of the tracking data.
Asynchronous signals with very large amplitudes can cause the zero-crossing loop to fall out of lock due to an apparent noise increase caused [Beckmann,1992].

C. User errors :

The same errors as mentioned with the DNS can be observed here.

On the accuracy to be expected using Loran-C for position fixing the following, based on Last [1992], can be said :

Improvements in receiver and chain control techniques have reduced resulting timing uncertainties, making them negligible in comparison with uncertainties due to SNR. Traditionally, the noise had been considered to be atmospheric in origin and random in nature. In NW Europe however, the dominant source of noise is CWI. The limit of coverage is assumed to be reached when the SNR (including CWI in NW Europe) at the sampling point falls to -10 dB, (this is equal to C/N_0 = 10 dB for a typical bandwidth of 50 Hz) resulting in a carrier tracking error due to noise of approximately 0.15 μsec (tracking loop bandwidth of 0.1 Hz).

The precision of a Loran-C LOP can be stated as :

	σ_{LOP}
tracking loop (noise)	0.15 μsec
ASF calculation	0.1 μsec
station synchronization	0 - 0.1 μsec
total error (1σ)	0.18 - 0.21 μsec

These errors can be converted to metres using (4.13).

The Federal Radionavigation Plan [1990] states that within the coverage area of a chain the predictable accuracy of a position has to be smaller than 463 m (0.25 nm; $2d_{rms}$).
At the edge of the area cycle matching is more difficult and can result in cycleslips. Each cycleslip gives a time error of 10 μsec to a station.

Field observations give that within the maximum range of the ground wave coverage (approximately 1500 km) predictable fix accuracies of 50 - 450 m ($2d_{rms}$) can be expected.

4.2.3 Radar

A radar can be used to observe both bearing and distance to an object. In both observations random errors will be present. First an overview of errors in bearings will be given , followed by the overview of errors in a distance measurement.

A. Bearings

The main error sources in a relative radar bearing are :

1. misalignment of heading marker with ships heading

   This error can be minimised by carefully comparing the direction given by the heading marker against the ships heading given by a gyro compass. When the heading marker is aligned this way, the systematic error can be reduced to approximately 0.1°, provided only random errors are present in the gyro course.

2. 'own ship' positioning

   If the position of the 'own ship' is not centred, an error will be introduced in observed bearings, depending on the direction and magnitude of displacement, and bearing. This error will not be constant in magnitude but change according to :

$$\delta\alpha = \frac{d\sin\alpha}{D - d\cos\alpha} \qquad (4.15)$$

   where

   α = bearing to object relative to direction of displacement
   D = distance to object
   d = displacement

   With most modern radars it is possible to automatically align the position of the 'own ship' with the centre of the screen.

3. Electronic Bearing Line (EBL) not positioned on 'own ship'

   When taking bearings, this gives a displacement in the same way as with misalignment of the 'own ship' position.

4. size of object on screen

   - the size of an object is increased on the screen as a result of beamwidth. This effect increases with increasing distance
   - if the bearing to a physically large object is measured, it is important to know which part of the object is used. This error becomes less important as the distance to the object becomes larger.

5. quantization

   Normally a graduation circle is engraved around the border of the CRT. If this is used, an error can be introduced due to parallax. With modern radars, a digital bearing indicator is part of the radar system. This indicator is connected to the position of the synthetic radar cursor, giving the bearing of the cursor. Apart from wrong positioning of the cursor due to parallax, the indicator gives the bearing of a target to the nearest 0.1°.

6. user rounding off

   In most cases, when a radar bearing is plotted, the bearing is rounded off to the nearest one tenth of a degree.

The above mentioned errors make up the total error for a relative bearing. To obtain the absolute error, the total gyro compass error (see section 4.3.3) has to be included since the largest part of the total error depends on the error of the gyro compass. Considerable errors (up to a few degrees) can arise from many speed and/or course changes made in a small time interval.
The total error of an absolute radar bearing is given by [Draaisma et al.,1982;Lenart,1989] :

total gyro compass errors (1σ)	0.65°
misalignment heading marker	0.1°
cursor setting	0.5°
beam shape errors	0.05°
quantization	0.14°
total angle error (1σ)	0.85°

This error is converted to meters in a direction perpendicular to the bearing line by

$$\sigma_{E} = \sigma \frac{\pi}{180} d \quad \text{[metres]} \qquad (4.16)$$

where

d = distance to object in metres

B. Distances

The main sources, causing errors in distance measurements using a radar are :

1. time base errors

   At the instant the RF energy starts to leave the antenna, the electron beam in the CRT comes under the deflecting influence of a linearly increasing magnetic force. The deflective force is produced by the action of a sawtooth current waveform produced by the **timebase generator** which is synchronized to start the deflecting process at the onset of each transmitted pulse. Video pulses of received echoes are applied to the CRT so that the target appears. Any errors in synchronization will lead to errors is position of the target displayed on the CRT. When the radar circuitry is well adjusted, errors due to synchronization are negligible with respect to errors described below.

2. errors in Variable Range Marker (VRM)

   - accuracy of VRM : IALA [1990] states that the 95 percent radial error should be better than 1.5 percent of the maximum range of the scale in use or 70 metres whichever is greater.
   - correct setting of VRM on the contact since the size of the contact is given enlarged on the radar screen, due to pulse length

   Although these errors will be the same in magnitude on the radar screen, their absolute magnitude depends on the range scale used.

3. elevation of the object with respect to antenna height

   If the target is elevated with respect to antenna height, a slant range is measured instead of the horizontal range plotted or used in calculations. The error introduced is generally negligible, especially if the target is at great distance.

4. pulse shape

   - the size of an object is increased on the screen as a result of pulse length. This effect decreases on screen with larger range scale used
   - if the distance to a physically large object is measured, it is important that the point of the echo closest to the centre of the screen is used for measurement.

5. quantization

   The digital distance indicator is connected to VRM, giving the distance a target to the nearest 0.01 nm.

6. user round off error of distance to 0.1 nm.

The error budget of a radar distance is given by [Draaisma et al.,1982; Lenart,1989] :

accuracy VRM	10 - 35 m
positioning VRM	10 - 80 m
quantization	5 m
pulse shape errors	20 m
reading rounding error	55 m
total distance error (1σ)	60 - 105 m

4.3 The error budget of bounded and unbounded systems

4.3.1 Ships Inertial Navigation System

The correct working of the SINS is depending on the following factors introducing errors in the measured accelerations and therefore in the calculated velocities and distance travelled :

1. Timekeeping errors

   In order to obtain the distance travelled, acceleration have to be integrated twice with respect to time. The result of these integrations is also used to calculate the rotation angles needed to correct for rotations of the reference platform due to the ships's movement in N-S and/or E-W direction.

2. Dislevelment of the platform

   It is quite a difficult task to keep the platform at right angles to the local vertical at any time. Transient accelerations introduced by course and/or speed changes, roll, pitch and irregularities in the earth gravitational field can cause platform oscillations. If the platform is not completely horizontal, accelerations are introduced due to gravity. Since the accelerometers cannot make a distinction between accelerations resulting from ship movement and components of the Earth's gravity, both are combined in the calculations of the ship's speed and distance travelled. In order to reduce susceptibility to erroneous accelerations, the reference platform is 'Schüler-tuned'. The result is that errors do not increase in time but oscillate round a mean value.

3. Coriolis force

When the ship is under way, the platform is subject to the Coriolis force, introducing extra accelerations, which have to be corrected for. The magnitude and direction of the acceleration vector can be computed and errors introduced in calculation of velocity and distance travelled can be allowed for.

4. Accelerometer errors

The most important errors in modern inertial quality accelerometers are :

- bias : an output when no acceleration is applied;
- non-linearity : deviations from the least squares straight line for input-output relationships;
- threshold : minimum detectable change in accelerometer output;
- misalignment with respect to direction in which the accelerations are measured.

5. Gyro errors

Drift of the gyro's are caused by internal torques caused by mechanical wear, friction and mass imbalance will lead to a wrong 'horizontal' position of the reference frame.

Misalignment of gyros will lead to wrong stabilization signals, resulting dislevelment of the platform.

It can be shown [Draaisma,1986] that for a single channel a constant error in a measurement made by an accelerometer, an incorrect correction applied to allow for Coriolis force or an error in the gravitational force each will lead to a displacement error (δ_π) of

$$\delta_{\blacksquare}(t) = \frac{R}{g}\left[1 - \cos\left(\sqrt{\frac{g}{R}}\, t\right)\right]\delta a \quad [\text{metres}] \tag{4.17}$$

where

R = earth radius
g = acceleration due to gravity
δa = error in acceleration measurement

whereas a constant gyro drift or an incorrect correction applied to allow for earth rotation or ship's speed in E/W direction will each lead to an error of

$$\delta_{\blacksquare}(t) = \left[t - \sqrt{\frac{R}{g}} \sin\left(\sqrt{\frac{g}{R}}\, t\right)\right] \frac{\pi R}{180}\delta_{\omega} \quad [\text{metres}] \tag{4.18}$$

where

δ_{ω} = error in rotation angle correction

4.3.2 Electromagnetic Log

The EM log measures the ship's speed through the water within the hydrodynamic influence of the ship's hull. This will lead to the following errors in speed measurements :

1. boundary layer

   The velocity of the ship is measured with respect to a small volume of water in the direct vicinity of the ship's hull. Due to the viscosity of water a friction boundary layer is carried along with the hull. The speed is measured with respect to this boundary layer, which differs from the actual speed with respect to the surrounding water.

2. shallow water effect

   When the hull of the vessel is close to the bottom, the water flow velocity distribution changes due to restriction of the region in which the water can flow around the hull, causing an increase in the speed reading of a bottom-mounted sensor.

Since the log is rigidly attached to the hull and not in the centre of gravity of the ship, it undergoes displacements (both translations and rotations) proportional to the distance from the centre of gravity. The errors resulting from this can be divided into two categories :

1. errors which result from the dynamic orientation being different from the designed sensor orientation due to

   · trim
   · instantaneous pitch, roll and yaw angles
   · drift component due to wind

   These sources introduce errors since the log measures the actual speed through the water multiplied by the cosine of the angle

2. manoeuvring errors caused by controlled motions of the vessel, such as

   · turns : introduce an angle between the heading and the actual water velocity vector at the sensor.

Finally there are errors introduced when

· a misalignment exists between the ships fore-and-aft axis and the sensor axis;

· calibration of the log has not been performed correctly, resulting in erroneous settings

The total log error can be divided into three parts : a constant part, a part which is proportional with speed and a random part. If the vessel sails with constant speed, the error proportional to speed can be seen as part of the systematic error.
For an EM log that has been calibrated and correctly aligned, the standard deviation of the vessel's random error can be approximated as [Draaisma et al.,1982] :

$$\sigma = 0.02\ v \qquad (4.19)$$

where
v = speed of vessel through the water.

4.3.3 Gyrocompass

The gyrocompass is subject to several errors, some of which can be eliminated in the design of the compass, while others require manual adjustment. The main error sources are :

1. speed error

If the vessel is moving in another direction than due East or West, the compass is, in effect, being carried in a direction perpendicular to the resultant of the sum of the ship's velocity vector and the Earth rotation velocity vector. This error angle introduced is well known and the error angle (δ_v) can be approximated by [Draaisma et al.,1986] :

$$\delta_v \approx 0.1232\ V \cos(C) \sec(\varphi)\ [\text{degrees}] \qquad (4.20)$$

where

V = ship's velocity in m/s
C = ship's heading
φ = latitude

In modern gyrocompasses this error is corrected mechanically. Speed and latitude are set by hand and the cosine of the course is introduced automatically. Wrong setting of the latitude and/or speed will still lead to an incorrect velocity correction which is considered to have a uniform pdf with a standard deviation of $\sigma = 0.2°$

2. latitude error

When the gyro compass is equipped with a vertical damping system, a difference exists between the direction of steady state of the rotation axis of the gyroscope and the direction of the local meridian. The error is called the latitude error and has a magnitude

$$\delta_\varphi \approx 57.3\ f \tan(\varphi)\ \ [\text{degrees}] \qquad (4.21)$$

where

f = ratio between magnitude of horizontal and vertical damping forces

This error has a uniform pdf with a standard deviation of approximately 0.1° (latitude set to the nearest 10° and $f = 0.04$).

Gyro compasses equipped with a horizontal damping system do not experience a latitude error as such. This error is still introduced, however, when an incorrect latitude is set to correct for speed error.

3. error due to ship's motion

   When the North-South component of the ship's speed changes, an accelerating force acts on the gyroscope which results in a precessing force in the horizontal plane, introducing a temporary error in the readings of magnitude :

$$\delta = \alpha \, \Delta(V \cos(C)) \quad \text{[degrees]} \tag{4.22}$$

   where

   α = multiplication factor depending on gyro compass type used ($\alpha \leq 0.1$)

   This error will be systematic in nature and can be as large as a few degrees after manoeuvres.

4. ballistic damping error

   A temporary oscillatory error of a gyrocompass introduced during changes of course and/or speed as a result of the means used to damp the oscillations of the spin axis. Provisions are made to counteract this effect when rates of change of course and/or speed exceed certain limits.

5. mechanical wear

   Mechanical wear of the compass will lead to precessions different from those that are corrected for. This will result in a deviation of the heading.

6. errors in construction

   This group of errors includes

   - shift in centre of gravity of gyroscope, resulting in an additional precession;
   - errors in the compass corrector-circuits
   - misalignment of compass housing with respect to ship's fore-and-aft axis;
   - follow-up error of repeater.

   For a good working compass a maximum error due to mechanical wear and construction of less than 1° can be expected [Draaisma et al.,1986]. This error will be considered to have a uniform pdf with standard deviation of $\sigma = 0.6°$.

It has to be remembered that temporary disturbances like course and/or speed changes can influence the compass deviation over a long period since the oscillation period is approximately 84 minutes and per oscillation the deviation is reduced to 1/3 of its value. When many changes in course and/or speed are done in a short period of time a large deviation (up to a few degrees) can be the result.

The total gyro compass error can be divided into two parts : a systematic error introduced by course and/or speed changes and a random part. If the vessel sails with constant course and speed, the standard deviation of the random part can be approximately given as $\sigma = 0.65^\circ$.

If a compass repeater is used to take bearings to objects, the following additional errors are introduced :

1. error in the repeater
2. round off error due to rounding to the nearest 0.5°

The total error of a bearing line is given by

error in gyro tc	0.65°
follow-up error of repeater	0° - 0.5°
reading rounding error	0.14°
total bearing error (1σ)	0.7° - 0.8°

This error is converted to meters in a direction perpendicular to the bearing line by using equation (4.16).

4.3.4 Inclinometer

During this research, no information was found on accuracy of attitude angles provided by SINS.

4.3.5 Echo sounder

When a depth is measured using an echo sounder, a combination of factors influence the accuracy and precision of the depth obtained. These factors are partly dependent on the echo sounder used but also on the bottom profile. When the depth obtained is used as bathymetric LOP and as such compared with depths given in charts, the whole process of chart compilation has to be considered as well. The compilation of a complete error budget, including all factors is outside the scope of this paper. Here, only the important factors are given. Nanninga [1985] and

Alper et al. [1985] have made an extensive analysis of precision of echo soundings obtained during surveys. The Royal Navy (U.K.) hydrographic department has also published a professional paper in which precision of soundings is assessed [MODUK,1990]. The description of errors as presented in this subsection is mainly based on these papers.

1. assumed speed of sound in water

   The speed of sound in water is dependent on the water temperature, salinity and water pressure. These quantities change with depth and position. In order to be able to measure depth accurately, sound velocity needs to be known over the whole water column. In most cases, only the velocity near the transducers is measured, introducing errors. As a guideline for errors in sound velocity, the following values can be used :

   - temperature : an error of 1°C results in an approximate velocity error of 3.6 m/s
   - salinity : an error of 1 ppt results in an approximate velocity error of 1.5 m/s
   - depth : an error of 100 m results in an approximate velocity error of 1.5 m/s

   The best way to establish sound velocity over the water column is by using a bathy-thermograph (BT) or a sound velocity probe. This way the sound velocity can be measured accurately with a standard deviation of σ_v = 1 m/s, leading to a standard deviation for a depth measured of:

$$\sigma = \frac{d}{v}\sigma_v \quad \text{[metres]} \qquad (4.23)$$

   where
   v = assumed sound velocity
   d = measured depth

2. timing accuracy

   The error in depth measurements is directly proportional to the accuracy with which the time interval between transmitted and received pulse can be measured. The error is normally negligible with respect to errors resulting from using a wrong propagation velocity and bottom profile.

3. bottom profile and composition

   When a depth is measured, this is assumed to be the distance between transducer and top of seabed. It depends however on the frequency used and composition of the top layer of the bottom, whether this is the case or not. Another factor that plays an important role is sediment transportation as this can change the bottom profile considerably.

When the depth is measured, it is assumed that this is the depth directly under the transducer. However, the combination of transducer beamwidth and seabed slope cause errors in assumed position and depth. If a slope of α° exists and the beamwidth of is 10° used the following errors are introduced:

displacement	:	$\delta = d \sin(5^\circ) = 0.09\ d$	[metres](4.24a)
depth measured	:	$\delta = 0.09\ d \tan(\alpha^\circ)$	[metres](4.24b)

4. ship motions due to sea and swell

In general, ship motions like roll and pitch and swell can have an effect on the depth measurement of over a metre in bad weather conditions. Under normal circumstances, the effect will only be in the order of a few decimetres. The effect is reduced when the submarine is submerged and negligible at depths greater then 0.5 times the wavelength of the surface waves.

5. squat and settlement

Squat is the change in trim of a vessel under way with respect to that of the vessel stopped.
Settlement is the lowering of a vessel in the water due to the interaction between the hull and seabed. It occurs only in waters of depth less than approximately six times the vessel's draught.

6. survey accuracy and plotting of depth in charts

This comprises both the accuracy and precision of the depth measurement as well as positional accuracy of the depth measured.

The standard for precision of depths (combining both observation and position accuracy) as set from January 1991 by the International Hydrographic Organization is (1σ) :

$$\sigma = 0.25 + 0.0045d \quad \text{[metres]} \qquad (4.25)$$

This value can also be used as approximate value for precision of a depth measured on board the submarine.

To use a measured depth in combination with a depth contour given in a chart to obtain a 'horizontal' LOP, precision of both observations need to be combined. In order to give a value of precision for this bathymetric LOP in a horizontal direction, bottom topography, especially the bottom slope, and survey accuracy, especially in post processing of data, play an important role. This needs further investigation which falls outside the scope of this paper. Some results from current investigation have been published [Kielland et al.,1992; Kielland et al.,1992; Velberg,1992], although no definite values are given yet. Therefore no value will be presented here.[1]

4.3.6 Pressure sensors

The following error sources can be distinguished when using the pressure sensor to obtain a depth :

1. measurement precision

   The measurement precision is stated by the manufacturer as 0.02 percent of the frequency cycle time measured. Regular calibration of the sensor will guarantee that this is achieved.
   The measurement precision is effective on both the measurement of the reference frequency at the sea surface and the measured frequency at depth.

2. sea and swell

   Since a pressure based on the water column above the sensor is measured, water motions at the sea surface such as sea and swell cause fluctuations in the measurements. In order to reduce this effect, depth measurements have to be filtered resulting in a mean value.

3. reference pressure

   Before the submarine submerges, the reference frequency is determined. This is based on the present atmospheric pressure. Any changes in atmospheric pressure will lead to an error in the depth measurement. This error is directly proportional to the difference in surface air pressure.

---

[1] A horizontal precision of 1 - 10 km (1σ) is currently used as value in the Royal Netherlands Navy. This is mainly based on sparse field observations.

4. specific weight of seawater

When the measured frequency is converted to depth, one of the multiplication factors is to allow for specific weight with respect to water with a density of 1000 kg / m^3. The error is directly proportional to the depth.

The error budget of a depth obtained by using the pressure sensor can be given as

reference frequency	1.00 m
frequency at depth	1.00 m
sea and swell	0.25 m
surface pressure	0.05 m
specific weight seawater	0.10 m
total error in depth (1σ)	1.45 m

4.3.7 Periscope

The information on the accuracy with which bearings can be taken using a periscope when the submarine is at periscope depth is regarded as classified information, so no specific figures can be given.
The errors present in a bearing can, however, be divided into two categories :

A. errors present in the gyrocompass system as described in section 4.3.3

B. errors inherent to the periscope itself, such as :

- position and alignment with submarine's fore-and-aft axis;
- optics;
- follow-up errors of the periscope synchros;
- magnification used.

The total error of a bearing line is given by

total gyrocompass error	0.65°
total periscope error	0.25°*
rounding off error	0.14°
total bearing error (1σ)	0.7°

* Although the magnitude of the errors of the periscope itself are classified, a the theoretical value (specs.) is given.

This error is converted to meters in a direction perpendicular to the bearing line by using equation (4.16).

5. Least Squares

5.1 Introduction

In chapters 3 and 4 an overview was given of the sensors and systems available for collecting data to be used for position fixing, on board a submarine. In this chapter a mathematical model that combines the information provided by the unbounded systems and sensors to obtain an MPP will be developed. The model is based on the theory of least squares.
In the first section a short explanation is given why the method of least squares is the preferred method to combine observations into an MPP. Next the mathematical and statistical models on which the least squares method is based are discussed. The least squares algorithm provides unbiased minimum variance estimates of the parameters assuming the mathematical model reflects physical reality and the observations only contain random errors. In real-life this is not always the case. For this reason, statistical tests are performed on the results to check the validity of these assumptions. These statistical tests are described in section four. This section is followed by a section on position confidence regions. The chapter ends with some concluding remarks on how to expand the least squares to incorporate optimal parameter estimation of bounded sensor observables.

5.2 Justification for least squares

When a redundancy in measurements exists, an adjustment is necessary in order to get a unique solution to the problem at hand. The adjustment according to the principles of least squares provides a general and systematic procedure for applications to all sorts of situations. It states that '*the most probable values of measured quantities are those which make the sum of the weighted squares of the residuals a minimum*'. The method is limited when blunders and/or systematic errors are present. The adjustment does not correct the observations though it may improve them. After the adjustment, the observations should be consistent.

Assume blunders and systematic errors to be removed from the observations, leaving only random errors. From a statistical point of view, the least squares estimates of parameters can be considered as 'best estimates' under the assumption made, which means that the following statistical properties apply [Cross, 1983; Spaans,1988:2] :

1. the estimates are unbiased;

2. the VCV matrix of the estimates is the one having minimum trace;
3. derived quantities have minimum variance.

These properties are independent of the pdf of the observational errors, in particular irrespective of whether or not they are normally distributed. As a result of these properties, the least squares estimates are often referred to as the **Best Linear Unbiased Estimates (BLUE)**.
If the observational errors are normally distributed, then the least squares estimates have the additional property of being **maximum likelihood estimates**.

If, however, systematic and/or random errors are still present in the observations, the estimates will be biased, therefore leading to degradation of predictable accuracy. Statistical tests on the results need to be performed in order to indicate possible presence of non-random errors. If an error is suspected to exist in one or more of the observations, the observation(s) will be removed from the set and estimates will be calculated using the remaining observations. How this is done will be explained in section 4 of this chapter.

Apart from the mentioned statistical arguments, there are also some practical reasons for using the least squares method for calculating an MPP :

1. the method is extremely easy to apply because it yields a linear set of normal equations;
2. different types of observations can be mixed to obtain estimates of the parameters;
3. there is no upper limit to observations that can be incorporated in the calculation of parameters; the lower limit is determined by the dimension of the problem (i.e. the number of parameters to be estimated) and statistical tests to be performed;
4. it is flexible with respect to the number of observations being used for calculations, i.e. adding observations to or deleting them from the set of observations is easy;
5. it gives a unique solution;
6. it is, generally speaking, 'unobjectionable' - it is very difficult to form an argument against least squares and in favour of some other method;
7. the method leads to an easy assessment of quality via the a posteriori VCV matrix $C_{\hat{x}}$.

5.3 The mathematical model

In order to be able to use least squares, a mathematical model has to be constructed. This model is often thought of as being composed of two parts [Mikhail, 1976] :

1. functional model : describes the deterministic properties of the physical situation or event under consideration,

2. stochastic model : designates and describes the non-deterministic properties - i.e. the totality of the assumptions of statistical properties - of the variables involved. It includes all model variables and designates those that are considered fixed (constants) and those that are considered free (parameters).

The usefulness of the least squares technique and results obtained, depends very much on the way the model reflects physical reality.

There are three general methods of deriving most probable values. The functional model, as implemented in the computer program, makes only use of the class of observation equations, leading to the following set of equations [Cross, 1983; Spaans,1988:2] :

1. mathematical model	: $F(x) - l = 0$	(5.1a)
2. normal equations	: $A\,\delta x = b + v$	(5.1b)
3. LSE parameter corrections	: $\delta\hat{x} = (A^TWA)^{-1}A^TWb$	(5.1c)
4. LSE parameters	: $\hat{x} = x_0 + \delta\hat{x}$	(5.1d)
5. VCV matrix of parameters	: $C_{\hat{x}} = (A^TWA)^{-1}$	(5.1e)
6. residuals	: $\hat{v} = A\,\delta\hat{x} - b$	(5.1f)
7. corrected observations	: $\hat{l} = l + \hat{v}$	(5.1g)

5.3.1 Use of meter as calculating unit

Because different systems and sensors are used to obtain the MPP, each having its own unit of measurement (and standard error given accordingly), relative performance of the systems and sensors is difficult to judge from results. Although it is not necessary at all to transform all units to meters in order to obtain the estimate of parameters, since the least squares method is very well capable of dealing with different units, there are some advantages to favour for the meter as unit for calculations:

- the design matrix A is better balanced since all its elements will be in the same order of magnitude. This will result in better accuracy when performing matrix

number of LOPs in observation set	2D	3D
1 LOP	I	I
2 LOPs	II	I
3 LOPs	III	II
4 LOPs	IV	III
≥ 4 LOPs	IV	IV

explanation of categories :

I. no least squares estimates can be calculated from the set of observations. The observations can only be used to update an MPP calculated by the Kalman filter. No statistical tests can be performed;

II. least squares estimates can be calculated from the set of observations. However, the redundancy is zero so no statistical tests can be performed;

III. least squares estimates can be calculated from the set of observations. The redundancy is one, leading to a limited set of statistical tests that can be performed (only detection of outliers is possible);

IV. least squares estimates can be calculated and all statistical tests can be performed. (both detection and identification of outliers is possible).

Table 5.1 Overview showing the possibility to calculate a least squares solution and to perform statistical analysis for observation sets of different size.

operations, especially when inverting matrices;

- the a-priori VCV matrix of the observations (C_l) gives a better overview of the relative weights of the observations with respect to each other, leading to a better and easier evaluation of the results to be expected;
- interpretation of the residuals is easier when given in meters, especially when hyperbolic RPF systems which are sustaining from an LEF, or bearings are used;
- interpretation is easier on the whole since the unit of meter is well known as a unit.

5.3.2 Number of observations

In order to be able to calculate a least squares solution of a position from observations and to perform statistical tests, a minimum number of observations is needed. Table 5.1 gives for observation sets of different size (number of LOPs available) whether or not a least squares solution can be found and whether or not statistical analysis can be performed. A distinction is made between 2D and 3D position fixing.
If the number of observations available is not enough to calculate a least squares estimate of the MPP, the observations will be combined directly with the position based on the dynamic model and bounded sensors provided by the navigation filter (described in chapter 6). This is also done by using the least squares algorithm, where the navigation filter provides two LOPs (latitude and longitude).

5.3.3 The functional model

<u>The design matrix and observation vector</u>

The design matrix **A** contains for each observation the partial derivatives of $F(\mathbf{x},\mathbf{l})$ in x and y (and also z and $c\Delta t$ when GPS pseudo-ranges are observed). Each row of **A** represents one observation. The associated observation vector **b** contains for each observation the result of the observed minus calculated (O - C) value, calculated using provisional parameter values. For the different types of observation available in the simulation program, the functional model and its derivatives are given in appendix 5.

A fixed order in which the observations are placed in the matrix is used. This means that, in the event observations need to be discarded - based on eg. the outcome of statistical tests

or user choice -, no row-shifts need to be performed. This again leads to an easier interpretation of the matrices. LOPs are in- or excluded from calculations by adjusting the a-priori VCV matrix of the observations and the weight matrix (see section 5.3.4).
The top part of the design matrix contains the observations from GPS (measured pseudo-ranges) and Decca (measured lane fraction in combination with the whole lane number and zone given by the navigator) - the only systems implemented so far.
The bottom part of the design matrix contains one or more bearings, distances and/or radar (BRR) observations. At present the computer program allows up to a maximum of 5 observations to be processed at one time. This suffices under normal conditions since the process of taking more than one or two BRR observations is normally too slow compared with calculation time needed for a complete least squares calculation cycle. This means that the elements in this part of the matrix will only be calculated occasionally when observations are made. The order in which the observations were fed into the computer is kept throughout calculations. Once the MPP is calculated and a new cycle starts, all previous BRR observations are discarded, i.e. the elements in this part of the matrix are set to zero.
The design matrix has therefore the following general form :

$$A = \begin{bmatrix} \text{GPS observation(s)} \\ \text{-------------} \\ \text{Decca observation(s)} \\ \text{-------------} \\ \text{BRR observation(s)} \end{bmatrix} \qquad (5.2)$$

The observation vector **b** is adjusted according to the design matrix.

The depths observed by the pressure sensor and echo sounder are not incorporated in the least squares algorithm. They are taken directly as measurements in the measurement model of the Kalman filter, which will be described in the next chapter. The reason for doing this is based on the ease with which the navigation filter can be adapted to the different situations distinguished in submarine navigation. These situations will be given below. This approach does not affect the final accuracy of the MPP.

In the mathematical model both the design matrix and observation vector can be expanded easily by adding an appropriate number of rows via adjusting dimensions of matrix and vector. This way implementation of new system observations is an easy task.

Calculating surface

A choice had to be made whether calculations should be performed on the geodetic datum or on a plane projection. Since Omega and Loran-C have not been implemented into the model yet, all calculations can be performed without problems on a plane surface. The reason to use a plane in favour of the geodetic datum is merely based on time available and ease of graphics implementation. This choice does not influence the final results. It is, however, assumed that observations have already been reduced to the geodetic datum when given as input into the computer program. So far, no provisions have been made in the program in order to be able to reduce the observations from the Earth surface to the datum (spheroid). No systematic errors are introduced in the transformations of observations from spheroid to the grid since all observations are reduced to the grid using rigorous formulae.
Use is made of the Transverse Mercator (TM) projection in conjunction with a grid having the following parameters :

1. Central Meridian (CM) : longitude of calculated MPP rounded to the nearest degree;

2. grid origin : latitude = 0°
   longitude = CM

   The geodetic position (0°,CM) has the grid values (0,0);

3. grid scale constant on CM : k_0 = 1.0000;

At startup is the user will choose a geodetic datum[1]. Once the program is running, this datum can be changed if desired. The coordinates of fixed stations such as the transmitting stations of RPF systems will automatically be transformed to the new datum by the program. If coordinates of stations are requested as input by the program, these are assumed to be geodetic coordinates on the datum currently in use. All grid positional results are transformed back to geodetic coordinates on the datum currently

[1] The combination of a spheroid (axis and flattening) and origin (a position where the geoid-spheroid separation and deflection of the vertical are defined) constitutes a datum. For example, the International spheroid (a and f) located by the Potsdam 1950 origin (ξ_0, η_0, N_0) constitutes the European 1950 Datum (ED50).

used by means of rigorous formulae, and as such presented to the user.

Parameters

The parameters to be optimized are : latitude, longitude and depth[1]. These quantities are optimized indirectly since all calculations are performed on a plane projection, leading to the parameters x,y and z to be optimized. Once the latter are optimized, the transformation to geodetic coordinates results in optimized values for the first parameters. Three situations can be distinguished :

a. submarine at sea surface

In this case the model is automatically reduced to a 2D problem, taking depth as being equal to zero and having no error (i.e. a standard deviation equal to zero). The z coordinate obtained from GPS pseudo-ranges is taken as a value for the geoid-spheroid separation.
In this situation use can be made of all EPF systems and all sensors the submarine has available on board.

b. submarine submerged

In this case the submarine is below the sea-surface so no use can be made of the RPF systems, radar, periscope etc. In this situation no least squares estimate of the MPP will be obtained. The information from the depth sensor and echo sounder are directly incorporated into the measurement model of the Kalman filter.

c. submarine at periscope depth

In this situation it is possible to obtain one or more bearings with the periscope which can be used to update the horizontal position parameters. If sufficient observations are available, an estimate of the MPP is calculated. This situation is also considered to be two dimensional (only x and y are estimated).

In the situation described under a., the z coordinate is the sum of geoid-spheroid separation and tidal height. In the simulation program developed, tide has been assumed non-existent, and thus not corrected for. In a real-time situation, tidal height can be incorporated using a tidal prediction program.

[1] In the position calculation, mean sea level (MSL) is taken to be zero height. MSL is considered to coincide with the geoid.

In the computer program optimization of assumed propagation velocity of radio waves by introducing for example scale factors is not considered. The signal propagation velocity of terrestrial EPF systems is assumed to be equal to those velocities used to draw lattices on charts. However, provisions are made for input of 'fixed error corrections' (DNS) or ASF corrections (Loran-C) in order to be able to correct for (local) propagation velocity anomalies due to land path, when known. Field test should show whether this approach is sufficient for surface navigation or not.

Provisional coordinates

At startup of the program, the user has to give a DR position in combination with depth, which is taken as initial MPP. This position provides the initial provisional coordinates for the least squares calculation cycle. When the program is running, the coordinates of the last MPP derived are used to provide the provisional values of parameters for the following cycle. The user has, however, the possibility to input a new DR position and to use this position as new starting point.

If the approximate values used to obtain the numerical values of the elements of the **A** matrix are not close to the final least squares estimates, then the normal equations will not be a true linearization of the functional model. In such case it is necessary to iterate using the least squares estimates from the i-th computation as approximate values for the (i+1)-th computation. The iteration is stopped when the vectors of the parameter corrections and residuals change by insignificant amounts. In the computer program the change in value of the provisional coordinates between iterations is checked : if the correction ($\delta\hat{x}$) to each element of the parameter vector ($\hat{x}$) is less than the default value of one metre - which is considered to be sufficient for navigational purposes -, iteration is stopped. It is possible to change the value to any desired value.

5.3.4 The stochastic model

The a-priori VCV matrix

The a-priori VCV matrix of the observations ($\mathbf{C}_l$) contains the variances and covariances of the observations based on prior knowledge. The ways to obtain values for these elements are many. The a priori VCV matrix used in the computer model is based on the error budgets given in chapter 4. These values are a combination of values given by manufacturers and found in literature. The values are either theoretical or derived from many observations made in the field under various conditions. It

is therefore assumed that these values reflect the actual standard errors of the observations quite well.
The layout of the VCV matrix is associated with the design matrix **A** and has the following form :

$$C_l = \begin{vmatrix} C_{GPS} & 0 & 0 \\ 0 & C_{Decca} & 0 \\ 0 & 0 & C_{obs} \end{vmatrix} \qquad (5.3)$$

where

- C_{GPS} an n x n VCV matrix belonging to the GPS pseudo-ranges
- C_{Decca} a 3 x 3 matrix formed by using the formula (A4.2)
- C_{Obs} a 5 x 5 diagonal matrix containing the variances for bearings, distances or radar observations as given by the error budgets and equation (4.16)

It is assumed that there does not exist any mutual correlation between the observations presented by different systems, bearings, distances or radar observations.

It is very important to have the variances of observations from different sources scaled correctly with respect to each other, i.e. given the correct **relative weight**. By having analyzed the error sources for data presented by the systems and other observations, resulting in error budgets as given in chapter 4 and converted to metres where necessary, this relative scaling of variances between systems and sensors should be correct. Future field tests have to prove whether these are chosen correctly, since relative scaling may well depend on the actual physical conditions affecting the performance of some systems, but not all. No provisions have been made yet for the user to adjust any variance or covariance values for the systems or other observations since estimating these has to be done as part of the system and model calibration using statistical analysis. It is very likely that input of the wrong values will lead to unwanted results. **Absolute** scaling of variances will be looked at in section 4.3 of this chapter.

In order to be able to leave observations out while retaining the correct order of observations in the various matrices, but still being able to calculate the inverse of the matrix, values of the a-priori VCV matrix are manipulated by the computer program as follows :

- if an observation has to be left out from the least squares calculations, its variance is set to one[1] while all its associated covariances are set to zero;
- the weight matrix (see below) is adjusted accordingly.

Expanding the a-priori VCV matrix

As it is possible to expand the **A** matrix, allowing new systems and/or more observations to be incorporated, the a-priori VCV matrix needs to be expanded accordingly. This is done simply by adding the VCV matrix of the new system as a sub-matrix at the appropriate position in the matrix or by increasing the sub-matrix $\mathbf{C}_{Obs}$.

The Weight Matrix

The weight matrix (W) of the observations is defined as :

$$W = C_l^{-1} \qquad (5.4)$$

If the a-priori VCV matrix has been modified in order to leave one or more observations out of the calculations, the weight matrix needs to be modified as well. This is done simply by setting the appropriate diagonal element of the weight matrix to zero (in order to avoid rounding errors in the computer), thus effectively introducing a zero weight for the observation under consideration.

A - posteriori VCV matrices

i. parameters : $\mathbf{C}_{\hat{x}} = (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}$ (5.5a)

This matrix is used for evaluation of the precision of the parameters and accuracy of the position derived using a given set of observations;

---

[1] The variance of an observation not used in calculations is technically equal to infinity. Since the observation is made uncorrelated with other observations, its calculated weight will be 1 over infinity. In order to avoid rounding errors, the weight is set to zero manually, which means that any value (except zero) can be chosen in the a priori VCV matrix.

ii. residuals : $C_{\hat{v}} = C_l - AC_{\hat{x}}A^T$ (5.5b)

This matrix is used in calculations of reliability figures for observations (τ) and in the statistical test for detection of outliers.

iii. observations : $C_{\hat{l}} = C_l - C_{\hat{v}} = AC_{\hat{x}}A^T$ (5.5c)

This matrix is normally not used. It shows however a theoretical VCV matrix of the adjusted observations and should ideally be the same as the a-priori VCV matrix of the observations if the latter one was given correctly.

It should be noted that these matrices can be calculated based on the mathematical model only, i.e. no actual observation values are needed. Therefore the matrices could be used in advance to determine the minimum number and geometry - and therefore optimum set - of observations needed to provide results meeting specified precision criteria. In the computer program this analysis part has not been implemented for the simple reasons that generally no time is available for pre-analysis and that the navigator will always use as many observations as possible to calculate the MPP at any given moment.

The a-posteriori VCV matrices are only a measure of the precision of the position fixes. This is not sufficient to describe fully the quality of a fix; it is essential to quote also some measures of reliability and to have some indication of whether or not systematic errors and/or blunders are present in the observations. The way this can be done is discussed in the following section.

5.4 Statistical tests

The method of least squares is used to calculate estimates of random variables (parameters) from samples (set of observations). Related with this estimation is the task of determining accuracy and reliability with which the estimates are obtained (confidence measures) and whether the results of the estimations are in agreement with the initial assumptions (hypothesis). Therefore, once the MPP is calculated using the observations made, statistical tests are performed to validate the observations and final results derived. The general procedure of statistical testing always refers to a **null hypothesis (H_0)** - that is, the set of population parameters (mean, variance etc.) to which the statistics are compared. The result of the test is a statement whether, according to the available evidence, H_0 can be considered acceptable or not. In the developed computer model the following tests are performed:

- test for presence of outliers (blunders and large systematic errors);
- test for sufficient reliability;
- check on correct assumed value of unit variance.

The following type of test for hypotheses is performed in these situations :

$$H_0 \quad : p = p_0 \qquad (5.6a)$$
$$H_1 \quad : p \neq p_0 \qquad (5.6b)$$

where

p_0 represents a given standard value of the parameter
H_1 is the alternative hypothesis.

The significance level of the tests is set to $\alpha = 0.05$ (5%) and $\alpha = 0.01$ (1%).

Based on the outcome of the tests, a decision is made whether a LOP needs to be discarded from the observation set and the MPP to be recalculated. Additionally an advice is given whether or not the a priori VCV matrix should be scaled.

It is not important that the errors in observations are distributed according to a normal distribution to apply the least squares algorithm to calculate an MPP from LOPs. However, statistical tests as described in this section are based on a multivariate normal distribution. In chapter 2 it was reasoned that, due to the central limit theory, the distributions of the total error of observations (LOPs) are approaching the normal distribution, leading to a multivariate normal distribution of the calculated parameters.

5.4.1 Test for identification of outliers

Outliers in observations can either be blunders (eg. laneslips, multipath, malfunction of a system, sky wave interference, user input errors when for example bearings are included etc.) or large systematic errors (eg. in time changing signal interference, propagation velocity model errors etc.). In order to obtain an accurate position fix, observations sustaining errors like this should not be used. The test described here is used to detect and identify any outliers present.

When testing the observations for outliers, the following statistical test, known as B-method of testing after Baarda who introduced it, is performed [Spaans,1988:2] :

$$H_0 : \hat{l}_i = l_i - \hat{v}_i \tag{5.7a}$$

$$H_1 : \hat{l}_i = l_i + \Delta - \hat{v}_i \tag{5.7b}$$

were

- l = true observed value
- $\hat{l}$ = least squares estimate of observation
- $\hat{v}$ = residual
- Δ = blunder present in the observation.

The test is based on the following reasoning. If the least squares process is repeated without observation l_i, a new solution $\hat{x}'$ with associating $\hat{v}'$ is found. Using this information, a new value $\hat{l}_i'$ is calculated. From this the following quantities can be defined :

$$d_i = \hat{l}_i - \hat{l}_i' \tag{5.8a}$$

$$\sigma_{d_i} = \left(e_i^T W C_{\hat{v}} W e_i \right)^{-1/2} \tag{5.8b}$$

where

e_i = a null vector except for unity at the corresponding observation (i) being tested

The test statistic is consequently defined as

$$w_i = \frac{d_i}{\sigma_{d_i}} = - \left(e_i^T W \hat{v} \right) \sigma_{d_i} \tag{5.8c}$$

If the observation l_i does not contain a blunder, the test statistic will be a normalized statistic, having a normal distribution with zero mean and standard deviation one. If the observation does contain a blunder, w_i will have a normal distribution with a mean (actually a bias) equal to

$$m_w = \frac{\Delta_i}{\sigma_{d_i}} \tag{5.8d}$$

With this information, the hypotheses to be tested as given in (5.7) will be restated as :

$$H_0 \quad : E[w_i] = 0 \qquad (5.9a)$$
$$H_1 \quad : E[w_i] \neq 0 \qquad (5.9b)$$

which is a standard statistical test for sample mean with known standard deviation (σ = 1). From this it follows that, with significance level α given, H_0 will be accepted if :

$$P(-x < w_i < x) = 1 - \alpha \qquad (5.10a)$$

or when

$$P(abs(w_i) < x) = 1 - \alpha/2 \qquad (5.10b)$$

In the computer program each observation will be tested for given confidence level in the following way :

1. $abs(w_i) \leq 1.960$: H_0 will be accepted for both $\alpha = 0.05$ and $\alpha = 0.01$. The observation is not considered to be an outlier and will therefore always be accepted;

2. $1.960 < abs(w_i) \leq 2.576$: H_0 will be accepted when $\alpha = 0.01$ but rejected when $\alpha = 0.05$. In this case the observation will be accepted, but monitored in order to get additional information;

3. $2.576 < abs(w_i)$: H_0 will be rejected for both $\alpha = 0.05$ and $\alpha = 0.01$. The observation is assumed to be an outlier and will therefore always be rejected.

The values for confidence levels as stated above have been chosen to avoid too many unjust rejections of LOPs ($\alpha = 0.05$ means a chance of 1 in 20 of a LOP being rejected while H_0 is true, whereas $\alpha = 0.01$ gives a chance of 1 in 100), but to give the navigator on the other hand ample warning that limits are being reached due to for example low frequency systematic errors present in one or more LOPs. This makes it possible for the navigator to react adequately by for example de-selecting a LOP from the set of LOPs used for MPP calculation and/or choosing a more optimal set of observations to be used.
For the criteria 2 and 3 as described above, further investigation of the source of the error should be performed. It may well be caused by a constant systematic error not corrected for during system calibration (eg. wrong assumed propagation velocity), by low frequency disturbances resulting in a temporary large value of w_i (eg. due to noise, interference etc.) or by sudden unfavourable situations (eg. loss of receiver phase lock resulting in a laneslip, interference, multipath etc.).

If for one or more observations the value of w_i is larger than 2.576 (criterion 3), the observation having the largest absolute value will be considered to contain a blunder and will therefore be discarded from the set before recalculation of the position is performed. This process is repeated until all observations in the remaining set suffice either criterion 1 or 2 as given above, or when the remaining set of observations becomes too small to perform identification of outliers (see table 5.1).

5.4.2 Check on sufficient reliability

In chapter 2 reliability has been defined as 'the ease with which a blunder in a measurement can be detected'.

To give a measure for reliability, the reliability factor (τ) can be used, which is defined as [Cross,1983; Spaans,1988:2] :

$$\tau_i = (e_i^T C_l e_i e_i^T W C_{\hat{v}} W e_i)^{-0.5} = \left(\frac{\sigma_i^2}{\sigma_{d_i}^2} \right)^{-0.5} = \frac{\sigma_{d_i}}{\sigma_i} \qquad (5.11)$$

The larger the value of τ ($\tau > 1$), the less is the reliability of the LOP under consideration. The reliability factor can be used to calculate the maximum undetectable blunder under H_0. Associated with the reliability factor is the variance factor (VF), which is defined as :

$$VF_i = \frac{e_i^T (C_l - C_{\hat{v}}) e_i}{e_i^T C_l e_i} = \frac{\hat{\sigma}_i^2}{\sigma_i^2} \qquad (5.12)$$

where

σ_i = a priori standard deviation of observation (from C_l)
$\hat{\sigma}_i$ = a posteriori standard deviation (from equation 5.5c)

If VF > 0.9, the observation is considered to be unreliable, i.e. not independently checked which means that a blunder in the observation would remain undetected.

Both quantities are given in the computer simulation program and can help the navigator to decide whether a LOP should be used for position fixing or not.

5.4.3 Unit variance

A statistic known as the **standard error of unit weight** ($\hat{\sigma}_0$) is normally computed after the least squares estimates of the parameters and associating residuals have been obtained. It is a test statistic used to assess the a priori variances and covariance, and is defined as :

$$\hat{\sigma}_0 = \sqrt{\frac{\hat{v}^T W \hat{v}}{n - m}} \quad (5.13)$$

where

n = number of observations
m = number of parameters estimated

It can be shown [Cross,1983;Spaans,1988:2] that if the correct a priori VCV matrix is used :

$$E[\sigma_0^2] = 1 \quad (5.14)$$

If $\hat{\sigma}_0^2$ differs significantly from unity and statistical testing has shown that no blunders or large systematic errors are present in the observations, it is normally assumed that the a priori error VCV matrix has on average been underestimated by a factor $1/\hat{\sigma}_0^2$, i.e. **absolute scaling** of the variances and covariances had been wrong given the current physical conditions. One has to be careful with this assumption as errors in the mathematical model can also lead to a wrong value of $\hat{\sigma}_0^2$.

Using $\hat{\sigma}_0^2$ as scaling factor for the a priori VCV matrix will not have any effect on the least squares estimates of the parameters. It will however affect the a posteriori VCV matrices (equations 5.5a - 5.5d) and therefore the size of the error ellipse / ellipsoid, making assessment of fix quality unreliable.
In order to decide whether $\hat{\sigma}_0^2$ differs significantly from unity, the following hypotheses are tested :

$$H_0 \quad : \sigma_0^2 = 1 \quad (5.15a)$$

$$H_1 \quad : \sigma_0^2 \neq 1 \quad (5.15b)$$

The test statistic to be used is given by Mikhail [1976] as :

$$T = \frac{[(n - m) - 1] S^2}{\sigma_0^2} \tag{5.16}$$

where
- S^2 = sample unit variance, calculated using equation (5.13)
- σ_0^2 = a priori unit variance
- n = number of observations
- m = number of parameters

It can be shown that this test statistic has a chi-square distribution with (n-m)-1 degrees of freedom. The test consists of checking if the value of the sample is within the confidence region given by

$$P\left[\frac{\sigma_0^2}{(n-m)-1}\chi^2_{\frac{\alpha}{2},(n-m)-1} < S^2 < \frac{\sigma_0^2}{(n-m)-1}\chi^2_{1-\frac{\alpha}{2},(n-m)-1}\right] = 1 - \alpha \tag{5.17}$$

If H_0 is rejected and no blunders are identified, it can be concluded that the a priori VCV matrix had been scaled incorrectly and should be scaled by S^2.

In the computer program, scaling is not automatically performed as this is not desirable because of the small number of redundancy. As long as a least squares MPP is calculated, not only the variance of unit weight based on the sample is calculated using equation (5.13), but also a 'cumulative' unit variance, using the following recursive formula

$$\hat{\sigma}^2_{i+1} = \hat{\sigma}^2_i + \frac{r_{i+1}}{R_{i+1}}\left(S^2_{i+1} - \hat{\sigma}^2_i\right) \qquad i = 0,1,2,\ldots \tag{5.18}$$

where
- $\hat{\sigma}^2$ = cumulative unit variance with initial value $\hat{\sigma}_0^2 = 0$
- S^2 = sample unit variance calculated using equation (5.13)
- r = redundancy of sample
- R = cumulative redundancy with initial value $R_0 = 0$

This way a more reliable test can be performed as the cumulative redundancy R increases. As an indication of the number of position fixes needed before a reliable test can be performed, Spaans [1988:2] gives a total redundancy of **R** = 200, which means that for a 4 LOP - 2D position fix with sample redundancy r = 2, at least 100 consecutive position fixes have to be made. On the other hand, Cross [1983] gives an example of the danger of using S^2 as scaling factor. This means that special techniques have to be developed to decide whether or not the a priori VCV matrix should be scaled. This has not been performed as part of the research. Therefore, automatic scaling of the a priori VCV matrix is not performed and it is left to the navigator to decide whether this should be done or not. In the computer simulation program a provision is made to switch automatic scaling on and off.

5.5 Ellipsoids of constant probability

5.5.1 Multidimensional distribution

Although the least squares theory of adjustment does not require a specified distribution, the statistical testing following the adjustment is based on random vectors with elements having a normal distribution, leading to the multivariate density function :

$$f(x_1 \ldots x_n) = f(x) = \frac{1}{(2\pi)^{n/2}\sqrt{\det C_x}} e^{-\frac{1}{2}(x - \mu_x)^t C_x^{-1} (x - \mu_x)} \qquad (5.19)$$

The function

$$h(x) = (x - \mu_x)^T C_x^{-1} (x - \mu_x) = k^2 \qquad (5.20)$$

represents a family of hyper-ellipsoids of constant probability when the quadratic form is positive definite [Mikhail,1976].

In this paper only the situations for n = 2 (the ellipse) and n = 3 (the ellipsoid) will be considered since these are used for 2D and 3D position fixing respectively.

It can be shown [Mikhail,1981] that for every symmetric VCV matrix C_x^{-1} there exists an orthogonal rotation matrix **R** such that $\mathbf{R}^T C_x^{-1} \mathbf{R}$ is a diagonal matrix. The columns of **R** are the normalized eigenvectors of C_X^{-1} and the elements of the diagonal matrix are the corresponding eigenvalues. From this it can also be shown that the eigenvalues of the VCV matrix C_X represent the squares of the lengths of the primary axes of the ellipsoids.

5.5.2 The error ellipse

In the 2D situation we find for (5.20) :

$$h(x,y) = \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix}^T C_X^{-1} \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \tag{5.21}$$

$$= \frac{1}{1-\rho^2}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho \frac{(x-\mu_x)(y-\mu_y)}{\sigma_x \sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right] = k^2$$

The equation $h(x,y) = k^2$ for a specific value of k is an ellipse which is known as an **ellipse of constant probability.** The value of the probability depends on the value of k. If k = 1, the ellipse is called the **standard ellipse.** The shape of the ellipse is fully determined by σ_x, σ_y and ρ.

There are several ways of calculating the lengths of the semi-major and -minor axes of the ellipse and the angle between its major axis and the y-axis (direction of North) of the local reference system. In the computer program these values are obtained by calculating the eigenvalues and eigenvectors of C_X using the following characteristic equation :

$$\lambda^2 - Tr(C_X)\lambda + \det(C_X) = 0 \tag{5.22}$$

The parameters for the standard ellipse are then given by

$$a = \left[\frac{1}{2}(\sigma_1^2 + \sigma_2^2) + \frac{1}{2}\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + 4\sigma_{12}^2} \right]^{1/2} \quad \text{[metres]} \tag{5.23a}$$

$$b = \left[\frac{1}{2}(\sigma_1^2 + \sigma_2^2) - \frac{1}{2}\sqrt{(\sigma_1^2 - \sigma_2^2)^2 + 4\sigma_{12}^2} \right]^{1/2} \quad \text{[metres]} \tag{5.23b}$$

$$\alpha = \arctan\left[\frac{\sigma_{12}}{\lambda_1 - \sigma_1^2} \right] \quad \text{[degrees]} \tag{5.23c}$$

Having defined the standard ellipse this way, it is now possible to calculate the probability that the random vector x takes values within or on an ellipse with semi principal axes ka and kb. The general expression is :

$$P\left\{ \chi_2^2 < k^2 \right\} = 1 - e^{-\frac{k^2}{2}} \tag{5.24}$$

where

χ_2^2 = chi-square distribution with two degrees of freedom

In order to establish confidence regions, the confidence level (γ) is selected and the multiplier k is calculated from :

$$P\left\{ \chi_2^2 < k^2 \right\} = 1 - e^{-\frac{k^2}{2}} = \frac{\gamma}{100} \tag{5.25}$$

Table 5.2 gives typical values of k for given P.
The computer model will provide the navigator with the dimensions of both the 95% ellipse (IMO recommended) and the 99.9% ellipse since submarine navigation is more hazardous than normal surface navigation.

	k	P
standard ellipse	1.000	0.3935
50% ellipse	1.177	0.5000
95% ellipse	2.447	0.9500
99.9% ellipse	3.717	0.9990

Table 5.2 Typical values of multiplication factor k (2D)

	k	P
standard ellipsoid	1.0000	0.1988
50% ellipsoid	1.5382	0.5000
95% ellipsoid	2.7955	0.9500
99.9% ellipsoid	4.0332	0.9990

Table 5.3 Typical values of multiplication factor k (3D)

5.5.3 The error ellipsoid

In the 3D situation we find for (5.20), defining the ellipsoid of constant probability :

$$h(x,y,z) = \begin{pmatrix} x - \mu_x \\ y - \mu_y \\ z - \mu_z \end{pmatrix}^T C_x^{-1} \begin{pmatrix} x - \mu_x \\ y - \mu_y \\ z - \mu_z \end{pmatrix} = k^2 \tag{5.26}$$

Again, the value of probability depends on k. If k = 1, the ellipsoid is called the **standard ellipsoid**. As was the case with the error ellipse, several ways are possible to calculate the lengths of the three main axes of the ellipsoid and its orientation in the local reference system. In the computer program the eigenvalues of $C_{\hat{x}}$ are computed from the characteristic equation :

$$\det\left(C_x - \lambda\, I\right) = 0 \tag{5.27}$$

Since this is a third order polynomial in λ, its roots can be found using rigorous formulae. Once the eigenvalues are calculated, the normalized eigenvectors of $C_{\hat{x}}$ are calculated. Having found these vectors, the orientation of the ellipsoid is found by solving the equation :

$$E = R\,(\rho\,,\theta,\varphi\,) = R_z(\varphi\,)\,R_y(\theta)\,R_x(\rho\,) \tag{5.28}$$

where

- E = matrix containing normalized eigenvectors of $C_{\hat{x}}$
- R = rotation matrix
- ρ = rotation angle around x-axis
- θ = rotation angle around y-axis
- φ = rotation angle around z-axis

Finally, the general expression giving the probability that the random vector x takes values within or on the surface of the ellipsoid with principle axes of length ka, kb and kc is :

$$P(\chi_3^2 < k^2) = \int_0^{k^2} \frac{1}{2\sqrt{2}\,\Gamma\left(\frac{3}{2}\right)} \sqrt{t}\, e^{-t/2} dt = \frac{2}{\sqrt{\pi}} \int_0^{\frac{k^2}{2}} \sqrt{p}\, e^{-p} dp \tag{5.29}$$

$$\Gamma\left(\frac{3}{2}\right) = \int_0^{\infty} \sqrt{t}\, e^{-t} dt = \frac{1}{2}\sqrt{\pi}$$

Table 5.3 gives typical values of k for given P.
The computer model will provide the navigator with the dimensions of both the 95% and the 99.9% ellipsoid.

5.5.4 Radial standard deviation

The result of the least squares calculation is an MPP with associating error ellipse / ellipsoid based on LOPs available. If the navigator wants to plot the position with associating confidence region in the chart, it is easier to plot a circle than an ellipse. Since the relationship between the standard ellipse / ellipsoid and radial standard deviation is normally not straight forward, except in the case when the standard ellipse or ellipsoid becomes a circle or sphere respectively, curves and tables have been made available to convert from one to the other.

The only radial standard deviation provided by the computer program is the distance root mean squared (d_{rms}), which is defined as :

$$d_{rms} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \tag{5.30}$$

where

- σ_x = length of semi major axis of standard error ellipse / ellipsoid
- σ_y = length of semi minor axis of standard ellipse or semi 'medium' axis of standard ellipsoid
- σ_z = length of semi minor axis of standard ellipsoid and zero for the ellipse

Care has to be taken when interpreting the d_{rms} value : although the standard ellipse / ellipsoid represents an area of constant probability (39.4% and 19.9% respectively), the probability associated with a given value d_{rms} varies as function of ellipse / ellipsoid eccentricity.

5.6 Concluding remarks

In this chapter, the parameter vector only contained position parameters (and $c\Delta$ if GPS pseudo-ranges are observed). The least squares is however not restricted to the estimation of only these parameters. The parameter vector can easily be expanded to incorporate other parameters such as velocity, heading etc. It is possible to use the least squares to estimate all elements of the submarine's state-vector which will be given in the next chapter. This is done by incorporating the relevant observation equations in the design matrix and observation vector and by including the error VCV matrix of the sensor or system to be added as sub-matrix in $C_{\hat{x}}$. Normally this method is not used. Instead a filtering algorithm is used to estimate these parameters. This algorithm is discussed in the next chapter. The least squares algorithm only provides position parameter estimates to this filtering algorithm.

6. Kalman Filter

6.1 Introduction

The estimation of the position of the submarine (MPP) is derived from information provided by systems and sensors. This information is generally corrupted by additive noise. Using external measurements together with information from bounded sensors in an optimal way, can result in a positional accuracy which is better than obtainable by either external measurements or sensor information alone.
In the previous chapter a description was given how the least squares algorithm is used to derive an MPP based on measurements from EPF systems. But, this way not all available information is used because sensor output (from log, gyro compass, SINS, depth sensor etc.) and past information is not considered. The Kalman filter, which is the subject of this chapter, provides a recursive computational algorithm which 'remembers' past data and uses this in combination with present measurements and sensor information to calculate the best estimate of the present and future state of the submarine. The state vector considered in this chapter only contains the vessel's 3D position, speed, heading and inclination along with estimates of current velocity and biases in log, gyro compass and inclinometer.
Before giving the Kalman filter equations, the first section looks to the justification of such a filter. The following two sections describe the filter algorithms for linear and non-linear systems. This is followed by a section giving the model as it is implemented in the computer program. In the next section some problems encountered when implementing the algorithm into a computer will be discussed, including an indication how to reduce their effect on performance. If the assumed models used were correct, the navigation filter would provide minimum variance unbiased estimates. In a dynamic environment deviations from the assumed models are encountered. Therefore, statistical tests are needed to monitor correct functioning of the filter in order to be able to react adequately to malfunctioning. These test will be discussed in the penultimate section of this chapter. The final section of this chapter will provide a short evaluation of the expected performance of the navigation filter, when used on board the submarine.

6.2 Why a Kalman filter ?

When in a dynamic environment, such as a submarine at sea, the parameters to be estimated change with time because the submarine is moving and because the system and sensor observables show temporal variations. At any moment in time it is possible to use the least squares algorithm as described in the previous chapter, to calculate single MPPs based on the present observations made. It is possible to incorporate past measurements in the least squares algorithm, but this would lead to an ever increasing number of observation equations as a number of observations is added every time new observations become available. Soon, the calculation time to obtain a new MPP would become unacceptable. The Kalman filter improves the way parameter estimation is performed by the least squares algorithm by relating parameters calculated at previous moments in time to the parameters calculated at present time in a recursive way. The Kalman filter is considered the most common optimal filtering technique for estimating the state of linear systems. Before starting the actual discussion of the Kalman filter, a definition of an optimal estimator, should be given because of its fundamental importance for the theory :

> 'An optimal estimator is a computational algorithm that processes measurements to deduce a minimum error estimate of the state of a system by utilising :
>
> i. knowledge of the system dynamics and measurements;
> ii. assumed statistics of the system noise and measurement errors; and
> iii. initial conditions of information.'

This leads to three main types of estimation problem :

1. **filtering** : the time at which an estimate is desired coincides with the last measurement point;

2. **smoothing** : the time of interest falls within the span of available measurement data;

3. **prediction** : the time of interest occurs after the last available measurement.

Since the computer model developed is designed to be used to give the navigator on line information on the submarine's current and future position, only the filtering and prediction estimation problems are of interest. Smoothing will therefore not be considered in this paper.

When obtaining a position from measurements, the following has to be kept in mind :

- external measurements contain random errors that may be significant with respect to the errors rising from the use of a bounded-sensor navigation system;
- bounded-sensor navigation system errors are primarily caused by the random, time-varying errors of the sensors

used.

Considering this, the Kalman filter algorithm shows the advantage that it uses all measurement data available, regardless of their errors, plus the prior knowledge about the system and its environment. This leads to the following characteristics [Gelb,1974] :

- it provides useful estimates of all sensor error sources with significant correlation time;
- it can accommodate non-stationary error sources when their statistical behaviour is known;
- configuration changes in the availability of navigation systems and sensors can easily be accounted for;
- it provides for optimal use of any number, combination and sequence of external measurements.

In order to avoid a growing memory filter resulting from storing all past measurement data, the estimate is sought in a linear recursive form so that there is no need to store past measurements for the purpose of computing present estimates. Again, the Kalman filter provides such a recursive algorithm.

Apart from the advantages mentioned above, there are also some important disadvantages to the filtering technique :

- the filter is based on the assumptions that the system is linear and the random errors are described by Gaussian white noise processes;
- it is sensitive to erroneous a-priori models and statistics;
- the computational burden when exact system and error models are to be used.

The first disadvantage can be overcome by linearization of the system equations (described in section 6.4) and use of shaping filters (described in section 6.3.2), whereas the other two disadvantages can be overcome to a fair extent by carefully selecting model- and error descriptions.

6.3 Linear dynamic systems

In this section the algorithm for optimal filtering will be given. It is based on the equations describing the dynamics of time continuous systems which are transformed into discrete-time equations in order to be able to implement the algorithms on a computer. Gaussian white noise is assumed but some consideration will be given on implementing the filter algorithm in the presence of coloured noise. The final subsection deals with prediction.

6.3.1 The discrete-time model filter equations

The expected future state of a linear dynamic system under the influence of (external) disturbances, may be predicted by the equations of motion, represented by the following linear first order vector differential equation :

$$\dot{x}(t) = F(t)x(t) + G(t)w(t) \tag{6.1}$$

where

$x(t)$	n dimensional continuous time state vector
$F(t)$	n x n system dynamics matrix
$G(t)$	n x m matrix representing the effect of noise on the state
$w(t)$	m dimensional Gaussian white noise process vector having the following statistics

$$E[w(t)] = 0 \quad \forall t$$
$$E[w(t)\, w^T(\tau)] = Q(t)\delta(t - \tau)$$

$Q(t)$	m x m symmetric positive semi definite matrix
$\delta(t)$	Dirac delta function

The matrix Q(t) is the spectral density matrix of the system noise. It is used when calculating the error VCV matrix of the predicted state vector. Its contents will be defined in section 5.4 of this chapter.

Given the state vector at a particular point in time, and a description of the system forcing functions from that point in time forward, the state vector can be computed at any other time from the solution of equation (6.1), which is given by :

$$x(t) = \Phi(t,t_0)\,x(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\,G(\tau)w(\tau)\,d\tau \tag{6.2}$$

where

$\Phi(t,t_0)$ n x n transition matrix which is the solution of the matrix differential equation

$$\dot{\Phi}(t,\tau) = F(t)\Phi(t,\tau)$$
$$\Phi(t,t) = I \quad \forall t$$

Since a computer is used to obtain the most probable position of the submarine, its state vector can only be computed at discrete times using numerical solutions to integration and differentiation. Therefore discrete-time dynamic system equations need to be used to describe the motion of the submarine. These discrete-time equations arise from sampling the continuous system described by (6.1) and (6.2). The system vector difference equations allowing the state x_{k+1} of the submarine at time t_{k+1} to be calculated from the state x_k at time t_k is given by :

$$x_{k+1} = \Phi_k x_k + \Gamma_k w_k \tag{6.3a}$$

where

$$\Phi_k = \Phi(t_{k+1}, t_k) \tag{6.3b}$$

$$\Gamma_k w_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) G(\tau) w(\tau) d\tau \tag{6.3c}$$

The stepsize $\Delta t = t_{k+1} - t_k$ will be chosen to ensure that the transition matrix Φ and the variance of model and measurement noise together with their effects on the system can be considered constant.

Observation data is obtained at discrete moments in time. This data is considered to be linearly related to the state of the system according to

$$z_{k+1} = H_{k+1} x_{k+1} + v_{k+1} \tag{6.4}$$

where

- z p dimensional measurement vector
- H p x n observation matrix
- v p dimensional Gaussian white noise process vector having the following statistics

$$E[v_k] = 0 \qquad \forall\ k$$

$$E[w_k w_l^T] = R_k \delta_{kl}$$

- R_k p x p symmetric positive definite VCV matrix

The matrix R_k is the error VCV matrix of the measurement noise. It is used when calculating of the Kalman Gain, updated state vector and error VCV matrix of the updated state vector. Its contents will be defined in section 5.4 of this chapter.

The noise sequences $\{w_k\}$ and $\{v_k\}$ are assumed to be uncorrelated:

$$E[v_k w_l^T] = 0 \qquad \forall\ j,k \tag{6.5}$$

The initial state of the discrete-time system described by (6.3) and (6.4) is defined as :

$$E[x_0] = \hat{x}_0 \qquad (6.6a)$$

$$E[\tilde{x}_0 \tilde{x}_0^T] = P_0 \qquad (6.6b)$$

where

P n x n error VCV matrix of the state estimate

$\tilde{x}$ error in the state vector estimate $\tilde{x} = x - \hat{x}$

Given the linear system described by (6.3) and (6.4) with initial conditions (6.6), the Kalman filter algorithm provides the estimate $\hat{x}_{k|k}$ of the state x_k that is a linear function of all measurement data $(z_1, \ldots\ldots, z_k)$ which satisfies the following conditions :

- $\hat{x}_{k|k}$ is unbiased;
- $\hat{x}_{k|k}$ is the minimum variance estimate;
- $\hat{x}_{k|k}$ is consistent.

The estimate of the state vector is obtained in two steps :

Step 1 : State vector propagation.

This is a 'forecasting' process describing the discontinuous state estimate and error VCV matrix behaviour between measurement times t_k and t_{k+1}. Since system noise has zero mean, an estimate of the state vector and its VCV matrix at time $t_{k+1} > t_k$ is given by [AGARD,1970] :

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k} \qquad (6.7a)$$

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \qquad (6.7b)$$

Step 2 : State vector update.

This is the discontinuous state vector estimate and error VCV matrix behaviour across a measurement. The difference between the actual and predicted measurements is used as a basis for calculation of corrections to the estimate of the state vector. The algorithm combines the measurement vector z_{k+1} with the prediction of the state vector as derived in step 1 and is given by the following set of equations [AGARD,1970] :

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\left[z_{k+1} - \hat{z}_{k+1|k}\right] \tag{6.8a}$$

$$= \hat{x}_{k+1|k} + K_{k+1}\left[z_{k+1} - H_{k+1}\hat{x}_{k+1|k}\right]$$

$$P_{k+1|k+1} = \left[I - K_{k+1}H_{k+1}\right]P_{k+1|k} \tag{6.8b}$$

$$K_{k+1} = P_{k+1|k}H_{k+1}^T\left[H_{k+1}P_{k+1|k}H_{k+1}^T + R_{k+1}\right]^{-1} \tag{6.8c}$$

where

K	n x p optimal gain matrix for the unbiased minimum variance filter
$\hat{z}_{k+1\|k}$	predicted measurement vector

Equations (6.7) and (6.8) describe the complete discrete-time Kalman filter algorithm for linear systems.

6.3.2 Coloured noise

In the derivation of the Kalman filter equations, system and measurement noise are considered to be Gaussian white noise processes. However, white noise is physically not realizable. Instead, most system and measurement noise processes exhibit time correlation (coloured noise). In this subsection a method will be given by which coloured noise can be allowed for in the description of a dynamic time continuous system.

The continuous-time equations for a dynamic system and measurements having coloured noise are given by [AGARD,1970] :

$$\dot{x}(t) = F(t)x(t) + Gn_s(t) \tag{6.9a}$$

$$z(t) = H(t)x(t) + n_m(t) \tag{6.9b}$$

where $n_s(t)$ and $n_m(t)$ are coloured noise processes having the following statistical properties

$$E[\mathbf{n}_s(t)] = 0$$
$$E\left[\mathbf{n}_s(t)\mathbf{n}_s^T(\tau)\right] = D_s(t,\tau) \qquad (6.10a)$$

$$E[\mathbf{n}_m(t)] = 0$$
$$E\left[\mathbf{n}_m(t)\mathbf{n}_m^T(\tau)\right] = D_m(t,\tau) \qquad (6.10b)$$

When the coloured noise can be modelled by a shaping filter, having Gaussian white noise as forcing function (see Appendix 6), state vector augmentation can be used to transform equations (6.9a) and (6.9b) as follows :

$$\dot{\mathbf{x}} = F(t)\mathbf{x} + G\mathbf{n}_s \qquad (6.11a)$$
$$\dot{\mathbf{n}}_s = A_s\mathbf{n}_s + \mathbf{w} \qquad (6.11b)$$
$$\dot{\mathbf{n}}_m = A_m\mathbf{n}_m + \mathbf{v} \qquad (6.11c)$$

$$\mathbf{z} = H\mathbf{x} + \mathbf{n}_m \qquad (6.11d)$$

where

A_m, A_s system transfer function of shaping filters
$\mathbf{w}, \mathbf{v}$ Gaussian white noise processes

or in matrix form :

$$\begin{pmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{n}}_s \\ \dot{\mathbf{n}}_m \end{pmatrix} = \begin{pmatrix} F & G & 0 \\ 0 & A_s & 0 \\ 0 & 0 & A_m \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{n}_s \\ \mathbf{n}_m \end{pmatrix} + \begin{pmatrix} 0 \\ \mathbf{w} \\ \mathbf{v} \end{pmatrix} \qquad (6.12a)$$

$$\mathbf{z} = \begin{pmatrix} H & 0 & I \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{n}_s \\ \mathbf{n}_m \end{pmatrix} \qquad (6.12b)$$

Equation (6.12a) gives the description of an equivalent system having a 'white noise' driving force and equation (6.12b) describes 'error free' measurements. Before the Kalman filter algorithms (6.7) and (6.8) can be used with these equations, an additional transformation is needed to allow for the 'error free' measurements since the equivalent **R** matrix is singular thus preventing the Kalman gain matrix **K**, required for obtaining the optimal state estimator, to be calculated. Next the equations need to be written in discrete form in order to be able to implement them in the computer model. This is done in analogy with the system described in subsection 6.3.1.

In this paper the system and measurement model described by equation (6.9) is not considered. In order to get a good working filter, extensive analysis of system and sensor information is needed to obtain correct correlation times for the shaping filters. This analysis has not been performed as part of the research.

6.3.3 Prediction

Optimal prediction can be thought of in terms of optimal filtering in the absence of measurements. This is equivalent to optimal filtering with arbitrarily large measurement errors ($R^{-1} \to 0$ so $K \to 0$). Therefore, if measurements are unavailable beyond some time t_k, the optimal prediction of x_{k+1} for $t_{k+1} > t_k$, given $\hat{x}_{k|k}$ must be

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k} \tag{6.13}$$

having an error VCV matrix given by :

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \tag{6.14}$$

6.4 Discrete-time non-linear estimation

The general Kalman filter algorithm is based on linear differential equations. However, the equations used to describe the motions of the submarine in 3D space are non-linear time continuous differential equations. This means that the equations need to be linearized. Furthermore, since the Kalman filter is to be implemented on a computer, the differential equations need to be converted to discrete-time difference equations.
In this section the linearized, discrete-time Kalman filter equations, based on a time continuous system model and discrete-time measurements are given. They form the basis of the filter to be implemented in the computer model, which will be discussed

in section 6.5.

Non-linear systems having continuous dynamics are described by [Gelb,1974] :

$$\dot{x}(t) = f(x(t),t) + Gw(t) \tag{6.15a}$$

$$z_{k+1} = h_{k+1}(x(t_{k+1})) + v_{k+1} \tag{6.15b}$$

Given the minimum variance estimator $\hat{x}(t_k)$ at time t_k containing all measurement data up to t_k, the best estimate at t_{k+1} is found in two steps :

Step 1 : State vector propagation.

Between measurement times t_k and t_{k+1} no measurements are taken and the state propagates according to (6.15a). On the interval $t_k \le t \le t_{k+1}$, the conditional mean of x(t) is the solution of the equation

$$\dot{\hat{x}} = \hat{f}(x(t),t) \qquad t_k \le t \le t_{k+1} \tag{6.16}$$

In order to be able to compute $\hat{f}(x,t)$, the pdf p(x,t) needs to be known. To obtain practical estimation algorithms, methods of computing the mean and VCV matrix which do not depend on knowing the pdf are needed. One way to do this is by expanding f(x,t) in a Taylor series about the current estimate of the state vector:

$$f(x,t) = f(\hat{x},t) + \left.\frac{\partial f}{\partial x}\right|_{x=\hat{x}} (x - \hat{x}) + \ldots . \tag{6.17}$$

which leads to

$$\hat{f}(x,t) = E\left[f(x,t)\right] = f(\hat{x},t) + o + \ldots . \tag{6.18}$$

Using (6.18) in (6.16), the first order approximation of the state propagated state vector and its error VCV matrix can be written as [Gelb,1974]

$$\dot{\hat{x}}(t) = f(\hat{x},t) \tag{6.19a}$$

$$\dot{P}(t) = F(\hat{x}(t),t)P + PF^T(\hat{x}(t),t) + G(t)Q(t)G^T(t) \tag{6.19b}$$

where $F(\hat{x}(t),t)$ is the matrix whose elements are given by

$$F_{ij}(\hat{x}(t),t) = \left.\frac{\partial f_i(x(t),t)}{\partial x_j(t)}\right|_{x(t) = \hat{x}(t)} \tag{6.19c}$$

In order to be able to use this propagation of the state vector in the computer model, equations (6.19a) and (6.19b) need to be written in their discrete-time equation equivalent, which are given by :

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k},t_k) \tag{6.20a}$$

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \tag{6.20b}$$

where

$$\Gamma_k w_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1},\tau)G(\tau)w(\tau)d\tau \tag{6.20c}$$

$$f(x_k,t_k) = x(t_k) + \int_{t_k}^{t_{k+1}} f(x(t_k),\tau)\,d\tau \tag{6.20d}$$

$$\Phi_k = \Phi(t_{k+1},t_k) = I + \int_{t_k}^{t_{k+1}} \nabla_x\,d\tau \tag{6.20e}$$

$$\nabla_x = \left.\frac{\partial f(x,\tau)}{\partial x}\right|_{x = \hat{x}_{k|k}} \tag{6.20f}$$

Step 2 : State vector update equations.

After the measurement z_{k+1} is taken, the state vector and its error VCV matrix are improved by

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}\left[z_{k+1} - h_{k+1}(\hat{x}_{k+1|k})\right] \tag{6.21a}$$

$$P_{k+1|k+1} = \left[I - K_{k+1}H_{k+1}\right] P_{k+1|k} \tag{6.21b}$$

$$K_{k+1} = P_{k+1|k}H^{T}_{k+1}\left[H_{k+1}P_{k+1|k}H^{T}_{k+1} + R_{k+1}\right]^{-1} \tag{6.21c}$$

where H_{k+1} is the matrix whose elements are given by

$$H_{ij} = \left.\frac{\partial h_i(x(t_{k+1}))}{\partial x_j}\right|_{x(t_{k+1}) = \hat{x}_{k+1|k}} \tag{6.21d}$$

The formulae given in this section were based on truncated Taylor series approximations for computing the estimates. By being linearized about $\hat{x}(t)$, the equations show similarity with the Kalman filter equations for linear systems. They are therefore often referred to as the Extended **Kalman filter** propagation equations.
The Extended Kalman filter has been found to yield accurate estimates in a number of practical applications. However, when measurement and dynamic non-linearities become stronger, higher order methods such as the ***Iterated extended Kalman filter*** or a ***second-order filter*** yield better estimates at the expense of greater computational burden.

6.5 Implementation of the Kalman filter

6.5.1 Dynamic model of the submarine

The motions of the submarine are described by the theory of hydrodynamics. To model the ship dynamics a mathematical model, describing the relation between the input (propeller revolutions and rudder angle) and the state of the ship, is used. The

manoeuvring model described by Inoue, which is based on the Newtonian force equations, can be used best as mathematical model for the Kalman filter [Wulder, 1992].
The model implemented here is a very simplified version of the 2D model described by Wulder [1992]; no external forces or moments are considered. As the submarine can also navigate underwater, making its manoeuvring space 3D, the equations are rewritten to allow for this third dimension.

The dynamic manoeuvring model of the submarine will be described by the following set of differential equations :

$$\dot{x}(t) = v(t)\cos(\epsilon(t))\sin(\psi(t)) \quad (6.22a)$$
$$\dot{y}(t) = v(t)\cos(\epsilon(t))\cos(\psi(t)) \quad (6.22b)$$
$$\dot{z}(t) = v(t)\sin(\epsilon(t)) \quad (6.22c)$$
$$\dot{v}(t) = 0 \quad (6.22d)$$
$$\dot{\psi}(t) = 0 \quad (6.22e)$$
$$\dot{\epsilon}(t) = 0 \quad (6.22f)$$

where

- x,y,z = position of the submarine
- v = longitudinal ship velocity
- ψ = submarine's heading
- ϵ = inclination angle

According to this set of differential equations, the motion of the submarine is considered to be uniform along a straight line, taking accelerations as disturbances. This initial assumption can be made as long as the submarine can be considered to be a low dynamic system and a relatively high sample rate (approximately once every second) of gyro compass, log and inclinometer can be achieved.

Three main external disturbances, which influence the ship dynamics, act on the submarine :

i. waves : described by their height and direction with respect to the course steered. It acts on all motions of the submarine;

ii. wind : causes a force and moment on the submarine depending on the relative speed and direction of the wind (with respect to the submarine's heading and speed) and on the shape of the submarine;

iii. current : causes the submarines heading and speed to be different from its ground heading and speed.

From these, only the current will be considered, as this is the only disturbance that will influence the submarine's motion both when the submarine is at the sea surface and when submerged. It can be modelled as a velocity vector

$$v_c(t) = \begin{pmatrix} v_{cx}(t) \\ v_{cy}(t) \\ v_{cz}(t) \end{pmatrix} = \begin{pmatrix} v_c \sin(\psi_c(t)) \\ v_c \cos(\psi_c(t)) \\ 0 \end{pmatrix} \qquad (6.23)$$

where

v_c = current velocity
ψ_c = current direction

that can be added to the submarine's velocity vector in order to obtain the vessel's groundspeed and course.

Since current velocity and direction are not measured, they have to be estimated by the Kalman filter. The initial values must be given by the navigator, based on local knowledge. The current speed vector is assumed to be constant. However, current direction and speed change over a period of approximately 12^h30^m when the tidal motion is semi-diurnal, or 25^h when the tidal motion is diurnal. As an approximation of the uncertainty in the current model, the following can be used [Wulder, 1992]

$$\dot{v}_{cx}(t) = V_{cm}\omega_{cv} + V_{cm}\omega_{c\psi} \qquad (6.24a)$$
$$\dot{v}_{cy}(t) = V_{cm}\omega_{cv} + V_{cm}\omega_{c\psi} \qquad (6.24b)$$

where

V_{cm} = maximum current velocity
ω_{cv} = frequency of the current velocity
$\omega_{c\psi}$ = frequency of the current direction

Combination of the manoeuvring model (6.22) and current disturbance (6.23) makes up the complete system model to be implemented, which is given by the following set of equations

$$\dot{x}(t) = v(t)\cos(\epsilon(t))\sin(\psi(t)) + v_{cx}(t) \qquad (6.25a)$$
$$\dot{y}(t) = v(t)\cos(\epsilon(t))\cos(\psi(t)) + v_{cy}(t) \qquad (6.25b)$$
$$\dot{z}(t) = v(t)\sin(\epsilon(t)) \qquad (6.25c)$$
$$\dot{v}(t) = 0 + w_v(t) \qquad (6.25d)$$
$$\dot{\psi}(t) = 0 + w_\psi(t) \qquad (6.25e)$$
$$\dot{\epsilon}(t) = 0 + w_\epsilon(t) \qquad (6.25f)$$
$$\dot{v}_{cx} = 0 + w_{cx} \qquad (6.25g)$$
$$\dot{v}_{cy} = 0 + w_{cy} \qquad (6.25h)$$

where

w_v = noise signal acting on the longitudinal velocity
w_ψ = noise signal acting on the heading
w_ϵ = noise signal acting on the inclination angle
w_{cx} = noise signal acting on the x-component of the ship's ground velocity caused by uncertainties in the current model
w_{cy} = noise signal acting on the y-component of the ship's ground velocity caused by uncertainties in the current model

The noise signals describe the uncertainty of the assumption that the derivatives (6.25d) - (6.25h) are zero.

6.5.2 Observation equations

Now the system model has been defined, the observables need to be discussed. Measurements are taken at discrete moments in time and are related to the elements of the state vector by observation equations as given by equation (6.15b).

- submarine's horizontal position observable

  The horizontal position of the submarine is determined by using observations from EPF systems, bearings and distances. These observations are combined using the least squares algorithm as discussed in chapter 5. Therefore, the submarine's position observable can be written as

$$\begin{pmatrix} x_{LSE} \\ y_{LSE} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + v \qquad (6.26a)$$

  where
  v is a noise vector having zero mean and error VCV matrix given by $C_{\hat{x}}$.

- submarine's vertical position observable

  The depth of the submarine with respect to the sea surface is observed by the pressure sensor. The relation between the measured depth and true depth is given by

$$z_m = z + v_p \qquad (6.26b)$$

  where

z_m depth given by the pressure sensor
v_p a random noise signal with zero mean and standard deviation σ_p acting on the measurement.

In order to agree with the definition given for zero depth, the depth given by the pressure sensor should also be corrected for tidal height. This can be done using a tide prediction program giving the height of tide above or below MSL. A prediction program has not been incorporated in the simulation program and is a matter for further investigation since tidal ranges of a few metres are quite common, therefore introducing a low frequency systematic error which may be significant.

- submarine's velocity observable

The speed of the submarine through the water is observed by an EM log. The relation between the measured speed and true speed will be given by (see section 3.2 of chapter 4)

$$v_m = v + \Delta v + v_v \qquad (6.27)$$

where

v speed given by EM log
Δv systematic deviation of the log
v_v a random noise signal with zero mean and standard deviation σ_v acting on the speed measurement

- submarine's heading observable

The submarine's heading is measured by using a gyro compass, giving the course steered through the water. The relation between the true heading and measured heading will be given by (see section 3.3 of chapter 4)

$$\psi_m = \psi + \Delta\psi + v_\psi \qquad (6.28)$$

where

ψ course given by gyro compass
$\Delta\psi$ deviation of the compass caused by unknown part of the speed error, instrument correction etc.
v_ψ a random noise signal with zero mean and standard deviation σ_ψ acting on the heading measurement

· submarine's inclination angle

The inclination of the submarine is measured by an inclinometer. The relation between the measured inclination angle and true angle will be given by

$$\epsilon_m = \epsilon + \Delta\epsilon + v_\epsilon \tag{6.29}$$

where

ϵ	inclination angle given by inclinometer
$\Delta\epsilon$	systematic deviation of the inclination angle caused by trim, ship motion etc.
v_ϵ	a random noise signal with zero mean and standard deviation σ_ϵ acting on the inclination angle measurement

Although the measurement noise of the sensors described above shows time-correlation in real life, it is considered to be zero mean Gaussian white noise for simplicity, since measuring of correlation times in order to be able to use shaping filters falls outside the scope of this thesis.

6.5.3 The State vector and measurement-vector

The state vector contains the variables describing the dynamics of the submarine, which are

- position : x,y,z
- speed : v
- heading : ψ
- inclination angle : ϵ
- current velocities : v_{cx}, v_{cy}

This state vector is extended with parameters used to eliminate the systematic errors in the (measurement) sensors, which are

- systematic error in log : Δv
- systematic error in gyro compass : $\Delta\psi$
- systematic error in inclinometer : $\Delta\epsilon$

The complete state vector to be used now reads :

$$x = (x,y,z,v,\psi,\epsilon,v_{cx},v_{cy},\Delta v,\Delta\psi,\Delta\epsilon)^T$$

The measurement vector contains the data given by the sensors available plus the positional data derived from observations using least squares (in the event enough observations were available to do this), which gives

- horizontal position : x_{LSE}, y_{LSE}
- vertical position : z_m
- log speed : v_m
- heading of gyro compass : ψ_m
- inclination angle : ϵ_m

The complete measurement vector therefore reads :

$$z = (x,y,z,v,\psi,\epsilon)^T$$

6.5.4 Discrete-time Kalman filter equations

The discrete-time Kalman filter algorithm is given by equations (6.20) and (6.21). Using the information given in subsections 6.5.1 to 6.5.3, the linearized model can be derived.

1. propagation of the state vector

The propagation of the state vector is derived from the dynamic system model equations (6.25) using equations (6.20a) and (6.20d), resulting in

$$\begin{bmatrix} x \\ y \\ z \\ v \\ \psi \\ \epsilon \\ v_{cx} \\ v_{cy} \\ \Delta v \\ \Delta \psi \\ \Delta \epsilon \end{bmatrix}_{k+1|k} = \begin{bmatrix} x + v\cos(\epsilon)\sin(\psi)\Delta t + v_{cx}\Delta t \\ y + v\cos(\epsilon)\cos(\psi)\Delta t + v_{cy}\Delta t \\ z + v\sin(\epsilon)\Delta t \\ v \\ \psi \\ \epsilon \\ v_{cx} \\ v_{cy} \\ \Delta v \\ \Delta \psi \\ \Delta \epsilon \end{bmatrix}_{k|k} \qquad (6.30)$$

2. transition matrix (Φ) and system disturbance matrix (Γ)

In order to be able to calculate the error VCV matrix of the propagated state, the transition matrix Φ_k and disturbance matrix Γ_k are needed. They are derived from the system model by linearizing $f(\hat{x}_{k|k},t_k)$ using equations (6.20e) and (6.20c) respectively

$$\Phi_k = \begin{bmatrix} 1 & 0 & 0 & F_2\Delta t & vF_1\Delta t & -vF_3\Delta t & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & F_1\Delta t & -vF_2\Delta t & -vF_4\Delta t & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 1 & \sin(\epsilon)\Delta t & 0 & v\cos(\epsilon)\Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}_{k|k} \quad (6.31)$$

$$\Gamma_k = \begin{bmatrix} \frac{1}{2}F_2\Delta t^2 & \frac{1}{2}vF_1\Delta t^2 & -\frac{1}{2}vF_3\Delta t^2 & \frac{1}{2}\Delta t^2 & 0 & 0 & 0 & 0 \\ \frac{1}{2}F_1\Delta t^2 & -\frac{1}{2}vF_2\Delta t^2 & -\frac{1}{2}vF_4\Delta t^2 & 0 & \frac{1}{2}\Delta t^2 & 0 & 0 & 0 \\ \frac{1}{2}\sin(\epsilon)\Delta t^2 & 0 & \frac{1}{2}v\cos(\epsilon)\Delta t^2 & 0 & 0 & 0 & 0 & 0 \\ \Delta t & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Delta t & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Delta t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Delta t \end{bmatrix}_{k|k}$$

(6.32)

where

$$F_1 = \cos(\epsilon)\cos(\psi)$$
$$F_2 = \cos(\epsilon)\sin(\psi)$$
$$F_3 = \sin(\epsilon)\sin(\psi)$$
$$F_4 = \sin(\epsilon)\cos(\psi)$$

3. observation matrix

Since the observation equations as given in subsection 6.5.2 are all linear with respect to the state variables, no linearization is needed and the observation matrix is thus given by

$$H_{k+1} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{vmatrix} \qquad (6.33)$$

4. Error VCV matrices Q_k and R_k

In order to be able to use the filter, statistical models of the system and measurement noise processes (w and v respectively) have to be estimated. This is normally done by using model identification techniques in which output of the model based on measurements up to time t_k is compared with observations at time t_k. Measurement data to be used has to be collected during sea trials.
The way model identification is performed falls outside the scope of this paper, but the reader is referred to Gelb [1974] for a description.

As no sea trials were performed as part of the research, the variances of the different noise processes are estimated to the writer's best knowledge.

The statistics of the system disturbances are given by the error VCV matrix Q_k. As zero mean gaussian white noise processes are assumed, the VCV matrix is diagonal and defined as

$$E[w_k w_l^T] = Q_k \delta_{kl} \qquad (6.34)$$

The random process vector $\mathbf{w}_k$ contains all the statistical information of the disturbances acting on the system described by equation (6.25)

$$w_k = (w_v, w_\psi, w_\epsilon, w_{vcx}, w_{vcy}, w_{\Delta v}, w_{\Delta\psi}, w_{\Delta\epsilon})^T$$

The terms $w_{\Delta v}, w_{\Delta\psi}$ and $w_{\Delta\epsilon}$ have been included to describe the uncertainty of the assumption that the derivatives of Δv, $\Delta\psi$ and $\Delta\epsilon$ are zero.

The diagonal elements (variances) of Q_k are estimated as follows:

- acceleration (w_v)

  The pdf for accelerations will be seen as a combination of a discrete and continuous function. If the probability of having no acceleration is p_0 and the probability of acceleration at maximum rate is p_{max}, while the acceleration probability for all other accelerations is described by a uniform pdf, the variance of acceleration is given by [Gelb,1974] :

$$E[w_v^2] = \sigma_a^2 = \frac{a_{max}^2}{3}\left[1 + 4p_{max} - p_0\right] \tag{6.35}$$

  If a maximum acceleration of $a_{max} = 0.25\ m / s^2$ is assumed and the probability for maximum acceleration and no acceleration are estimated to be 0.01 and 0.75 respectively, the variance is equal to $\sigma_a^2 = 0.006\ m^2 / s^4$.

- rate of turn (w_ψ, w_ϵ)

  The standard deviation of rate of turn in horizontal and vertical direction is taken to be half the maximum rate of turn.

- rate of change of current vector (w_{vcx}, w_{vcy})

  Uncertainties in the current model are caused by the assumption that the current vector remains unchanged. As this is not the case, the rate of change has to be estimated. As this is difficult to do, only the maximum rate of change is given (eq 6.24).
  The standard deviation is taken to be half of the maximum rate of change.

In the computer simulation model, a semi diurnal tide with maximum velocity of 1 m / s (2 knots) is assumed, leading to the following variances :

$$\text{x direction} \quad : \sigma^2_{vcx} = 2\ 10^{-8}\ m^2 / s^4$$

$$\text{y direction} \quad : \sigma^2_{vcy} = 2\ 10^{-8}\ m^2 / s^4$$

- rate of change of systematic errors ($w_{\Delta v}, w_{\Delta\phi}, w_{\Delta\epsilon}$)

  It is assumed that the rate change of systematic errors of gyro, log and inclinometer are negligible compared to the errors caused by accelerations and drift rate.
  In the computer program it is therefore assumed that the variances are zero.

The statistics of the measurement disturbances are given by the error VCV matrix R_k, which is defined as

$$E\left[v_k v_l^T\right] = R_k = \left|\begin{array}{c|cccc} C_{\hat{x}} & & 0 & & \\ \hline & \sigma_p^2 & 0 & 0 & 0 \\ & 0 & \sigma_v^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_\phi^2 & 0 \\ & 0 & 0 & 0 & \sigma_\epsilon^2 \end{array}\right| \qquad (6.36)$$

where $C_{\hat{x}}$ is the error VCV matrix of the position obtained via least squares as given in chapter 5, and as values for σ_p, σ_v, σ_ϕ and σ_ϵ the standard errors of the sensors as given in chapter 4 are used.

5. Initial state

In order to be able to use the Kalman filter, the values of the state vector at k = 0 and its error VCV matrix must be known. They are defined as follows :

· state vector

$$\hat{x}_{0|0} = (x(0),y(0),z(0),v(0),\psi(0),\epsilon(0),v_{cx}(0),v_{cy}(0),0,0,0)^T$$

with

$x(0),y(0),z(0)$ a start position obtained from external information resulting in an initial estimate of the MPP by means of the least squares algorithm. If the least squares estimate of the MPP cannot be calculated, either because not enough LOPs are available or when the submarine is submerged, the initial position given by the navigator is taken as starting position

$$v_{cx}(0) = v_c(0)\sin(\psi_c(0))$$
$$v_{cy}(0) = v_c(0)\cos(\psi_c(0))$$

$v_c(0)$ current speed given by the navigator
$\psi_c(0)$ current direction given by the navigator

$$\psi(0) = \psi_{GK}(0)$$

· error VCV matrix

$$P_{0|0} = I$$

where

I is the identity matrix.

6. Confidence areas

In section 5 of chapter 5, expressions were given to obtain the error ellipse (2D) and error ellipsoid (3D). These equations can also be used to calculate the confidence regions of the MPP calculated by the navigation filter. In the equations (5.21) and (5.26), $C_{\hat{x}}$ should then be replaced by the top left-hand submatrix of $P_{k+1|k+1}$.

6.6 Implementation considerations

When the Kalman filter is implemented on a computer, some problems affecting the correct working are encountered. These problems can roughly be divided into two groups :

A. problems related to the model

In order to obtain a truly optimal filter, an exact description of the system dynamics, error statistics and measurement process are assumed to be used. This is not always the case either because an exact model would lead to a computational burden too great for the computer used or because exact statistical characteristics are simply not known.
Therefore, the system is normally simplified leading to a sub-optimal filter. As a result, a discrepancy between the filter- and theoretical state, referred to as **divergence**, can start to exist. To minimize this discrepancy, it is important to carefully evaluate the dynamic model used. This is normally done by sensitivity analysis and error budget calculations for which special computer test programs are developed. This falls outside the scope of this thesis.

B. problems related to the computer

Firstly, there are the constraints imposed by the computer. These are mainly of physical nature such as size, weight, peripherals, memory capacity and processor speed. This will lead to reduction of the complexity of the Kalman filter equations in order to reduce computational burden. Deleting states can be used as a method of simplifying the model - at the cost of introducing problems discussed in point A -, whereas prefiltering can be used when measurements are available more frequently than possible to process them.
Secondly, the finite nature of the computer will lead to truncation errors when approximating integration and differentiation using numerical algorithms - as is the case when using difference equations instead of differential equation -, whereas finite word length of the computer memory will lead to rounding errors. The latter could lead to unsensible results. One should therefore very carefully select algorithms used.

In order to obtain the 'best' Kalman filter for a certain application, the following steps should be part of the development process :

1. obtain a description of the initial model and error statistics;
2. implement the model on the computer;
3. perform sensitivity analysis and error budget calculations;
4. test the filter in real-time environment;
5. evaluate model and use results to modify the model.

Steps 2 through 5 need to be repeated until the chosen model satisfies the real-time situation best. In this thesis only steps 1 and 2 are performed. Steps 3 to 5 are suggested to be part of further investigation to improve the model as given in this chapter.

6.7 Statistical testing

As is the case with the least squares, statistical tests have to be performed to validate the reliability of the Kalman filter. The tests described here are based on Teunissen [1990] and Lu [1992], but parameters used in the reference have been adjusted to agree with those used in chapter 5 of this paper. This way the comparison with the test performed to detect outliers in observations becomes apparent.

The tests described are used to detect outliers originating from model misspecification. Two tests can be distinguished : the Local Overall Model (LOM) test and the Local Slippage (LS) test. Both tests look at only one epoch at a time and are used for detection and identification of outliers. The tests can be expanded to Global tests (GOM,GS), when more epochs are taken into account. The global tests are used to find drifting errors. In the computer simulation program, only the LOM and LS test are implemented.

The LOM and LS test statistics are based on the predicted residual which is given as

$$\hat{v}_{k+1} = z_{k+1} - \hat{z}_{k+1|k} = z_{k+1} - H_{k+1}\hat{x}_{k+1|k} \qquad (6.37a)$$

having as error VCV matrix

$$C_{\hat{v}_{k+1}} = H_{k+1}P_{k+1|k}H_{k+1}^{T} + R_{k+1} \qquad (6.37b)$$

Local Overall Model test

The Overall Model test is performed to detect whether an outlier is present or not. It is a test on the overall validity of the mathematical model.

If the dynamic model is valid, the predicted residuals should represent a Gaussian zero mean white noise process. If the model is not correct due to misspecification, one or more of the predicted residuals will not have zero mean any more but will show a bias. Therefore the following hypotheses are tested :

$$H_0^{k+1} : \hat{v}_{k+1} \sim N\left(0, C_{\hat{v}_{k+1}}\right) \qquad (6.38a)$$

$$H_1^{k+1} : \hat{v}_{k+1} \sim N\left(m_{\hat{v}_{k+1}}, C_{\hat{v}_{k+1}}\right) \qquad (6.38b)$$

where

$m_{\hat{v}}$ = mean of predicted residuals in the event of an outlier

The test statistic for local detection is given as

$$T_{LOM}^{k+1} = \frac{\hat{v}_{k+1}^T C_{\hat{v}_{k+1}}^{-1} \hat{v}_{k+1}}{n_{k+1}} \qquad (6.39)$$

where

n number of observations in measurement vector z

It can be shown that this test statistic has a chi-square distribution with n_{k+1} degrees of freedom. H_0 is accepted if the test statistic is within the confidence region, i.e. :

$$P\left|\chi^2_{\frac{\alpha}{2}, n_{k+1}} < T_{LOM}^{k+1} < \chi^2_{1-\frac{\alpha}{2}, n_{k+1}}\right| = 1 - \alpha \qquad (6.40)$$

If the test statistic falls outside the region, an unspecified local model error is considered to be present.

Local Slippage Test

Once an outlier is suspected to exist as result of the LOM test, the potential source has to be identified. This identification is performed by the Slippage test, which is the next step in the identification process.

To identify the source of the outlier, a new test statistic is defined

$$t_{LS_i}^{k+1} = \frac{e_i^T C_{\hat{v}_{k+1}}^{-1} \hat{v}_{k+1}}{\sqrt{e_i^T C_{\hat{v}_{k+1}}^{-1} e_i}} \tag{6.41}$$

If the observation i does not contain a blunder (H_0), the test statistic will be a normalized statistic, having a normal distribution with zero mean and standard deviation one. Under H_1 the test statistic will have a normal distribution with a non-zero mean and standard deviation one.
The test statistic is calculated for each observation at t_{k+1} and checked if it is within the confidence region as described in section 5.4.1 where w_i is replaced by t_{LS}. If for one or more observations H_0 is rejected, then the observation having the largest absolute value of l_{LS} is considered to be the observation containing the outlier.

Once an outlier has been detected and identified, the state vector has to be corrected. This will be done by estimation of the magnitude of the error and its associated variance. This can be achieved by using a two-stage Kalman filter [Lu,1992]. This will not be considered further in this paper.

6.8 Concluding remarks

The navigation filter is derived by means of the ship dynamic model and Kalman filter theory. It estimates variables concerning the 3D motions of the submarine, i.e. position, heading and inclination, and some variables concerning the sensor dynamics, i.e. deviations of log, gyro and inclinometer.
The ship's manoeuvring model used is a very simple model in which the submarine is assumed to sail with constant velocity along a straight line. The result of this is

- when the submarine accelerates or decelerates, the predicted

positions based on the dynamic model will lag the submarine's true position;
- when the submarine changes course and/or depth, again accelerations are introduced, this time with the effect that the predicted positions will show overshoot.

When the submarine is at the sea surface, additional data provided by the unbounded sensors is available, which will make the predicted position to agree better with the submarine's true position after some time. When the submarine is submerged, only bounded sensor can be used with sparse additional positional information provided by depths measurements (bathy LOP) obtained using the echo sounder. This means that no corrections can be applied to counteract the above mentioned effects due to accelerations. Errors introduced due to the accelerations will not be detected. This results in accumulative errors, making position accuracy provided by the navigation filter less reliable.
To improve performance, the filter gain (K) could be decreased when changing speed and/or course, putting more confidence in the position obtained from the external sources. This is again only possible when the submarine is at the sea surface. It is therefore important to assess how position accuracy degrades in time due to model disturbances. This depends very much on the system model used and sensor data sampling interval.

It is also important to consider the effect of errors in the initial state vector as this will give an indication of the filter capability to react to sudden changes in the environment.

For submarine navigation, currents are considered to be the most important disturbances. This is because their direction and/or velocity may change significantly with depth in general, or with position when the submarine is operating near coastlines in particular. These sudden disturbances can influence filter performance considerably when not taken into account.
The response of the filter may be improved by choosing large standard deviations of system noise at the beginning of the run and to decrease this later on. This way, the influence of current on position accuracy will be small at an early stage and becomes more important once the values are estimated. The simulation program has a provision to scale the error VCV matrix Q_k. Again trials at sea are necessary to find the correct error VCV matrix and 'filter setting' to be used.

7. The Computer Simulation Program

The third aim of the research was to develop a computer simulation program that shows the main features of integrated navigation and that can serve as basis for an integrated system to be implemented in a real-time environment on board submarines. In the previous chapters an inventory was made of the systems and sensors available for position fixing (chapters 3 and 4) and the mathematical and statistical models to be used for position calculation and quality control (chapters 5 and 6). In this chapter a concise description will be given of the computer simulation program, which is based on the information provided by the previous chapters.
The chapter starts with a section giving technical details about the program itself such as hardware and software requirements plus an overview showing the relationship between the main program and routines used. In the next section the routines used for position fixing and quality control will be described in more detail. This section will however not provide an in-depth description of the software. The final section of this chapter provides the user of the simulation program with a concise manual.

7.1 Technical description

Hardware requirements

The program development had started on an Intel 8086 based PC. Soon it became obvious that the calculations performed to process input data to obtain a position with associating confidence region and to perform the statistical tests, required a more advanced PC. By now, the program has been tested on several PCs, using several types of screens, for correct working. In order to get a realistic picture of performance (i.e. short calculation times in order to get the high sampling rate required), it is advised to use at least a 80286 based computer or equivalent with maths co-processor. The program does not use colours so a monochrome screen will suffice. The program presumes a hard disk to be present.

Software requirements

The computer simulation program is provided on one 3.5 inch floppy disk. This disk contains an installation program, the simulation program, a sensor input data file plus additional data files used by the program. The program runs under the DOS

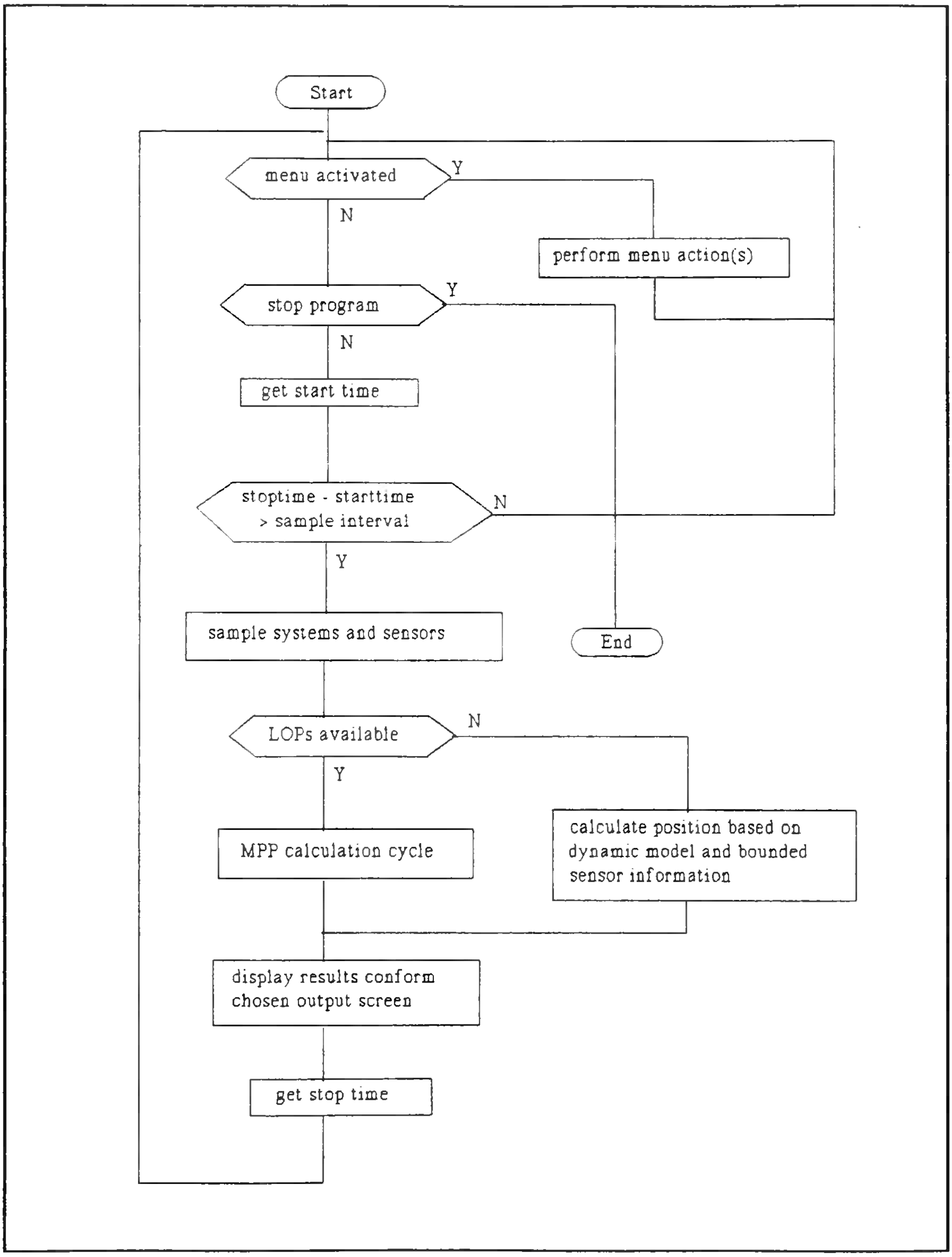


Figure 7.1 Main loop of the simulation program

operating system (version 3.x and higher). In order to keep the program machine independent no special calls are made to the BIOS.

Using the INSTALL program, a sub-directory '\UVNAV' is created on hard disk (drive C:). This directory contains the program itself (called UVNAV.EXE) plus a sub-directory '\DATA' in which the data files used by the program are placed. The software plus data files take about 400 kB disk space. Running the program from floppy disk is not possible as the program assumes the files to be on hard disk. The program is started by typing UVNAV <return> when in the directory \UVNAV.

General program layout

The simulation program main loop is given in figure 7.1. In the next section each of the blocks will be described in more detail. The program is completely menu-driven, using a pulldown menu system.

As the program is intended to be a simulation program, it was not considered important to put much effort in finding ways to keep the program continuing with the calculation of position fixes while the operator is using the menu. Instead, the menu operation process and calculation process are considered to be two independent processes that do not run simultaneously. This approach leads to the following three states that can be distinguished, once the program is running :

1. **working** : data is read from the simulation data file, the position is calculated and statistical tests are performed. If the cycle is completed, the program starts the next cycle after a time interval given by the operator. The default value of one cycle time is set to one second. This is the minimum time that can be chosen with the data file provided.

2. **interrupted** : the program is stopped and the user has access to the menu. The calculations are resumed by the program itself as soon as possible.

3. suspended : the calculation process is halted. This is a special feature of the simulation program. It is now possible to go through all output screens to check the results of the last calculation cycle. The program will resume the calculating process after the user has chosen the menu option 'Continue'.

Apart from the main program loop, several routines are used to perform the calculations and other tasks. Four levels can be distinguished. At the top level is the main program loop with a

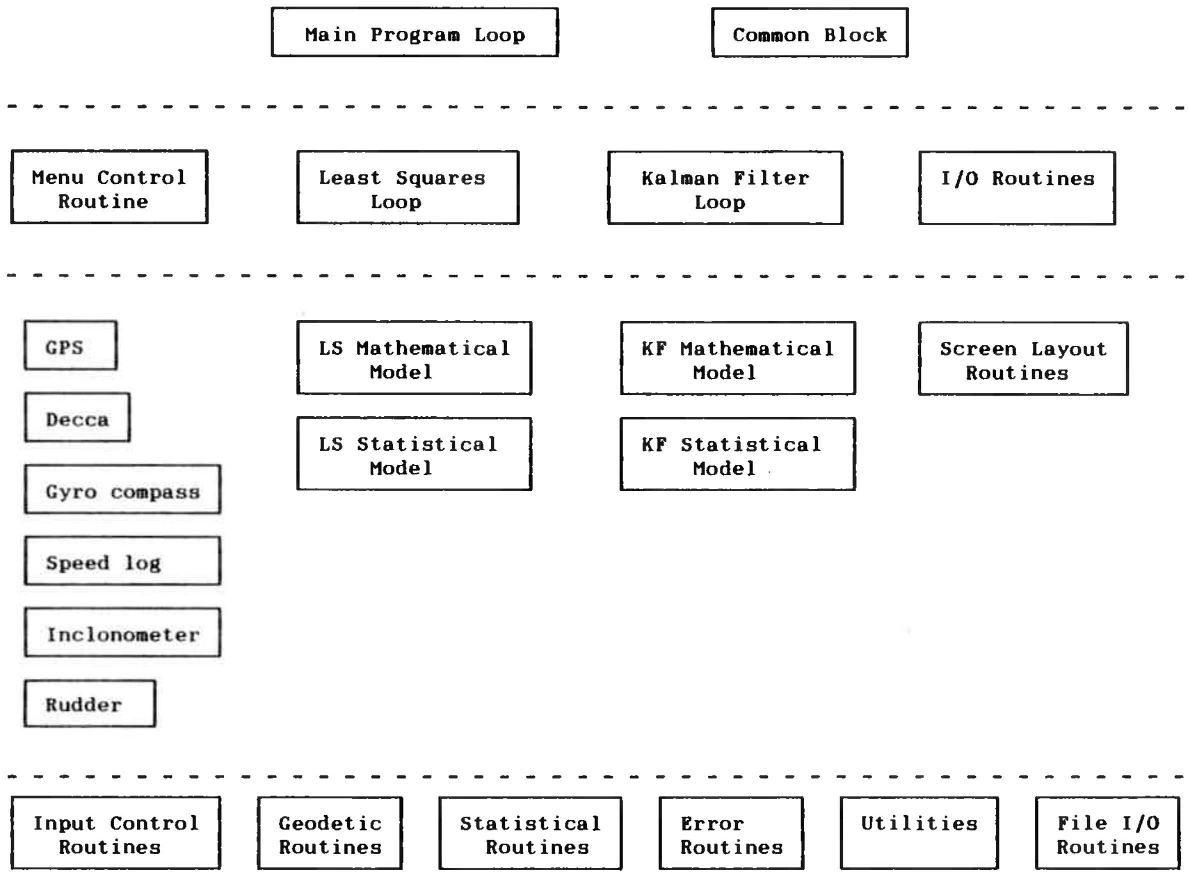


Figure 7.2 Diagrammatic overview of program modules.

'common block' in which all parameters used by the various routines are placed. This common is used by routines on all levels. The next level down contains the menu control routine, the data input and output routines, the general least squares loop and Kalman filter loop. The next level down contains the routines simulating systems and sensors, the least squares calculation (model and statistical) routines, Kalman filter calculation (model and statistical) routines and the screen layout routines. The bottom level consists of the utility routines such as matrix calculation routines, statistical routines, error routines, geodetic routines etc. Figure 7.2 gives a diagrammatic overview of the different levels.

In order to make a distinction between the information a system developer would need and the information needed by the navigator, a partition has been made in the menu :

- The navigator will have access to all system and sensor setup routines, observation routines, the Kalman filter routines and will be provided with positional and statistical data resulting from calculations. He will also have access to some parameters used for setup and performance of the simulation program.

- The developer will have access to **all** information provided by the system. This is the information provided to the navigator plus calculation results from several routines, such as for example the contents of the design matrix of the least squares algorithm or the system noise VCV matrix of the Kalman filter. Furthermore, he will be presented with for example chain data of the Decca chain used. The developer is also allowed to change crucial parameters of the program, such as switching scaling of the least squares a priori VCV matrix on and off, introducing offsets in observations etc.

This approach will not be relevant for a simulation program as such as it is the meaning of a program like this to show also for example the effects of ill performance of the sensors and systems on obtained results. It does show however how, in a real-time system on board, information can be shielded off from the navigator.

Data Files used by the program

With the program some data files, necessary for correct working of the simulation program, are provided. Each of these files is an ASCII file, making editing by an ordinary word processor possible. This way, updating of files as well as adding new data to each file is made easy.
The layout of each of these files will be described below. In appendix 8 an example of each of the files will be given.

bearing would be used.

Kalman

This menu is used for Kalman filter control (see chapter 6) and consists of three options :

- Filtering :
  Used to switch the navigation filter on and off. The default value is ON. If switched on, the predicted position based on the submarine's dynamic model and bounded sensor input is combined with the least squares position estimate, when available. The filter has to be switched on when the vessel is submerged.

- Predict. :
  Based on the current state vector with associating error VCV matrix, the state vector with associated error VCV matrix at some future time is predicted (see section 6.3.3). The time interval and step size to be used need to be given as input.

- Smoothing :
  Used for post processing of data (not implemented yet).

Statistics

This menu is used to display all calculation results and statistical information provided by the simulation program. The statistical information not only comprises the results of the statistical tests performed by the least squares and Kalman filter algorithm, but also additional information of the systems and sensors used. Each of the items in this menu has its own output screen. An example of each screen is given in appendix 8. The following options are available:

- Systems :
  This option will show some data concerning the systems GPS and Decca. It will show the observed values, corrections applied by the user to the observations, the calculated values by the least squares program, the a priori VCV matrix used for this system etc. Per system, one display page is used.

- Sensors :
  This item will show a display page of the sensor statistics, such as their measured value, systematic error estimate provided by the Kalman filter and standard deviation.

- LSE :
  This item will provide the navigator with the results of the

least squares calculations performed on the LOPs, such as : magnitude of residuals, estimated corrected observations, azimuth of the observations and the outcome of the statistical tests.
It will also give the least squares estimate of the MPP with dimensions and orientation of the associating 95% error ellipse.

- KF :
  This item will provide the navigator with the results of the Kalman filter calculation of the state vector and statistical tests performed.
  It will also give the navigation filter estimate of the MPP with dimensions and orientation of the associating 95% error ellipse / ellipsoid.

- Position :
  Under this item, the information about the different positions calculated (LS,KF predicted and MPP) will be given. Also a plot of the current situation will be provided.
  Here also an option 'History' is shown. Its purpose will be to show the effect of filtered positions compared to the positions obtained by the least squares process. This option has not been implemented yet.

- MainPage :
  Gives the user the possibility to return to the main output screen.

Demo

In this menu several parameters having an effect on the program performance and screen layout can be changed : the plot-scale, plot-size, specific weight of seawater, sample interval, corrections to be applied to the observations, scaling of a priori VCV matrix of the least squares algorithm etc.
For each of these parameters a default value is given at startup of the program. The default values are listed in appendix 8.

Data

In this menu the contents of matrices and vectors used in the least squares and Kalman filter process (described in chapters 5 and 6) are shown. Each of the items in this menu has its own output screen. An example of each screen is given in appendix 8.
Five options exist:

- Decca :
  Gives the chain data when the Decca navigation system is switched on.

- A-prio :
  this option will show the contents of design matrix, observation vector and a priori VCV matrix used in the least squares calculations.

- Weights :
  this option shows the contents of the weight matrix used in the least squares calculations.

- Kalman :
  this option will show the contents of the matrices and vectors used in the Kalman filter calculations.

- Ellipsoid :
  this option gives the dimensions and orientation of the standard, 95% and 99.9% error ellipse/ellipsoid of the least squares estimate, the predicted Kalman filter estimate and navigation filter estimate of the MPP.

Run

This item is used to control the has three options :

- Continue :
  used to restart the simulation program after it had been suspended.

- Suspend :
  this option is used to halt the program and keep its current state after the last calculation cycle. This way all display pages can be looked at without having the program continuing its calculations. This option can be used to compare data on different pages with each other without having them in the mean time changed by the program.

- Start :
  this item is used to start another simulation run without having to stop quit the program first. All parameters will be set to their default values, bringing the program in the initial startup state. The user gets the main page with the 'Parameter Initialization' cadre.
  In order to avoid accidental activation of this item, the user is asked for confirmation.

7.3.2 Data input formats

Geodetic Datums

The following geodetic datums are supported by the program :

ED50	ED87	OSGB 1936
ED50 [DMA]	NAD 1927	TOKYO
ED50 [UKOOA]	NAD 1983	WGS84

The datum data file contains the ellipsoid and geodetic datum parameters and is given in appendix 8. As this data file is an ASCII file, it is easy to add datum definitions to the file and/or change definitions already given.

Positions

When a position input is needed, this must be a geodetic position (i.e. a position on the spheroid). Each of the following input formats (and combinations of them) can be used :

```
52 30 00.000 N 003 34 24.735 E[1,2]
52 30 00 N 003 34 24 E
52 30.0 N 003 34.3 E
52 30 N 003 34 E
52.3 N 003.7 E
52 N 003 E
```

For latitude South, the letter S should be used and for longitude West, the letter W.
The position input format is checked and when a wrong format is detected, the error message **'Wrong position input format !'** will be given after which new input is requested. This is repeated until a correct format is used.

Bearings, distances and speeds

A bearing is given in degrees, counting from zero (North), clockwise to 360°.

Distances are given in nautical miles[3]. For this unit is chosen because most distances used in sea navigation are expressed in nautical miles.

Speeds are given in knots[4]

---

[1] Spaces or commas can be used as field separators.

[2] leading zeroes can be left out

[3] One nautical mile is 1852 metres precise.

[4] One knot equals one nautical mile per hour which can be approximated by 0.5 m/s.

7.3.3 Error messages

In order to make the system as 'fool proof' as possible, most input provided by the operator is checked on consistency. If wrong input is detected, an error message will appear and the operator is requested for new input.

8. Concluding Remarks and Suggestions for Further Development

8.1 Concluding remarks

Taking the concepts of the POE as starting point has lead to the design of a simulation program of an integrated navigation system that can serve as basis for an integrated system to be used on board submarines. The simulation program clearly shows advantages of integrated navigation with respect to the currently used concept of the POE :

1. it provides an almost instantaneous and continuous calculation of a 3D MPP with associating confidence region based on rigorous algorithms; Furthermore, other parameters such as systematic errors in log and gyro compass can be estimated by the navigation filter, leading to a better performance of the integrated system.

2. observables from different systems and sensors can be combined to obtain an MPP : there is no restriction to the type of the observations that can be incorporated. Furthermore there is no maximum to the number of observables that can be used. The minimum number is given by the number of parameters to be estimated.

3. cross-checking of systems is done by a device which is not subject to fatigue or stress : it provides warnings of incompatibility of receiver information such as for example gyro failure or interruption of signals (**detection**) and gives the possible source (**identification**). Algorithms have been developed that can be used to correct the error (**adaption**). The last-named is not considered in this paper.

4. once the system is started, no input is required from the navigator to keep the system running. He will therefore be able to concentrate more on the quality control side of the position fixes provided instead of having to spend much time on the construction of the position fix with associating error ellipse in the chart.

5. based on the present state of the submarine, it is easy to predict the future position with associating confidence region. No special construction is needed.

6. automatic accurate logging of position information and ship's attitude is provided. This information can be used for post mission analysis (track reconstruction or smoothing) and improvement of the integrated system.

It is however important to realize that the information displayed is only as accurate as the information provided to the system and underlying algorithms.
In order to obtain unbiased minimum variance (position) estimates, observations must be free of blunders and systematic errors, leaving only random errors present. Furthermore, the algorithms must reflect physical reality. This is not always the case in a dynamic environment. To overcome this problem, tests were given in chapter 5 to check whether or not observables are free from outliers whereas in chapter 6 tests were discussed that can be used to check the validity of the mathematical models. In the computer simulation model only the tests given in chapter 5 and the local tests LOM and LS (chapter 6) are incorporated. This will not suffice completely as errors having a slowly increasing or decreasing tendency (drift) will remain undetected. Therefore global tests (GOM and GS) need to be incorporated as well.

A good user interface, showing the important data in a clear way to the navigator is necessary. The combination of figures and graphs seems to be the best combination. Since position information becomes available and future positions based on present and past information can be predicted, a link of the integrated system with an electronic chart (ECDIS) system seems feasible. This way, the navigator does not have to plot any information but still gets a clear view on the current size of the position confidence region with respect to navigational dangers in the area.

As was mentioned at the end of chapter 2, the configuration of sensors available for position fixing and cross-checking on board submarines is not ideal at the moment. Once the submarine is submerged, cross-checking of systems is limited. SINS can partly provide this task but will still show errors increasing in time, especially when estimating depth over longer periods. Furthermore, SINS is very sensitive to errors made by the operator at start-up of the system. A way to improve velocity and course cross-checking is by using a Doppler log or correlation log. The information provided by these types of log can also be used to make a better estimate of the current speed and direction and the EM log and gyro compass systematic errors. Especially the last mentioned estimate is an important feature : when the submarine is frequently changing course and/or speed, the gyro compass systematic error can become quite large and damps out relatively slowly (damping oscillation period is approximately 84 minutes in which the error is reduced only to 1/3 of the error at the start of the cycle !). If continuous use of these logs is not feasible from an operational point of view, they can be used at intervals to update the estimates. The EM log can be used in between.

When the submarine is submerged, only sparse outside information is available at present. Apart from using other sensors (eg. logs as mentioned above), improvement can also be achieved by improving the navigation filter. In chapter 6 some drawbacks of the mathematical model as implemented in the simulation model were discussed. One way to overcome them is by refining the dynamic model of the submarine and incorporating sensor dynamics. Another possible way of improving position fixing underwater, is by combining a bathymetric database with echo sounder information using map matching techniques the same way as done in land and air navigation. The problem that arises here is twofold. Firstly, the major part of the seabed is either not surveyed at all or a very long time ago. Secondly, sediment transportation can change the bottom contours rapidly, making the database invalid. Only areas with high shipping density are surveyed regularly, leading to usable data. However, in these areas, the submarine will generally be navigating at the sea surface and thus having sufficient external information.

To conclude it is important to stress again that all that is written in this paper is purely theoretical. The next step is to perform field tests. The results of these tests must be used to validate and improve the dynamic model of the submarine and to validate the error budgets used in the statistical models.

8.2 Suggestions on further development

The limited amount of time available compared to the magnitude of the problem has left much scope for improvement and further research on the items described in this paper. In this section the most important improvements needed will be discussed.

The least squares algorithm

The algorithm as it is implemented in the simulation program was described in chapter 5. It will probably suffice for a real world environment as long as raw system and sensor data is used (eg. GPS pseudo-range measurements instead of the position provided by the receiver).

Although the algorithm checks the integrity of LOPs, resulting in discarding of the unreliable LOPs, it does not tell the navigator what the actual source of the error is (eg. laneslip, signal loss etc.). In further development algorithms that are able to give the navigator the likely nature of the error present should be incorporated in the system.

The Kalman filter algorithm

In this paper, sensor dynamics of the gyro compass and SINS have not been considered. The measurement and system noise processes are assumed to be Gaussian white noise processes with standard deviations as given by the error budgets in chapter 4. This means that prediction of the state vector is limited considerably as performance of sensors like the gyro compass and INS depends very much on the movements of the vessel. Sensor dynamic models need to be developed. These models can be used to predict the magnitude of low frequency systematic errors in order to adjust the predicted measurements.

The noise in navigation errors is not white noise but displays time correlation. To allow for this, shaping filters have to be added to the navigation filter mathematical model.

As mentioned above, global detection tests and error adaption algorithms need to be incorporated in the navigation filter.

The error budgets

By investigating literature, error budgets could be formed for systems and sensors available. Sea trials have to prove whether the values given agree with reality.

The computer program

Apart from the improvement needed in the algorithms used to calculate the position, there is much scope for improvement of the first version of the simulation program as presented in this paper. This will not only comprise improvement of the simulation program itself, but also adapting it in order to make it work in a real-time environment. The important updates needed are :

- interfaces have to be written to be able to read the raw data from sensors and systems straight into the program;

- routines have to be developed to incorporate Loran-C, Omega and SINS;

- a tidal prediction model needs to be incorporated in order to be able to correct the measured depth for vertical water movement;

- input routines are needed to check whether the information provided by the user is valid. Furthermore wrong input may not lead to runtime errors;

- the program must be able to continue its calculation process while the navigator is using the menu program;

- the software needs to be optimized and it should be considered to rewrite time critical parts of the program in assembler;

- the user interface needs to be optimized by putting relevant information for the navigator together in one or more windows. Maybe 'hot keys' should be defined to be able to skip directly to the most relevant data. Furthermore, a help file should be created to give the user on line information about the data provided on the different output screens as well as information on input parameters to be used so sensible input data is provided;

- a detailed user manual, giving background information on the program and its options has to be provided.

Literature

AGARD, 1970. Theory and Applications of Kalman Filtering, NATO Advisory group for aerospace research and development, report no. AGARDograph no.139 (AD 706 306). 537 pp.

Alper S., Bossler D., 1985. Toward improved accuracy standards for hydrographic surveying. International Hydrographic Review, Monaco, Vol. LXII no. 2, July 1985. pp. 39 - 49.

Appleyard S.F., Linford R.S., Yarwood P.J., 1988. Marine Electronic Navigation, second enlarged edition. Routledge & Kegan Paul Ltd, London, ISBN 0-7102-1271-2. 605 pp.

Asche G.P., 1972. The Omega System of Global Navigation. International Hydrographic Review, Monaco, Vol. L No. 1, January 1973. pp 87 - 99.

Braasch M.S., 1990. A Signal Model for GPS. Navigation : Journal of the Institute of Navigation Vol. 37, No. 4, Winter 1990-91. pp 363-377.

Braasch M.S., Graas F van, 1991. Guidance Accuracy Considerations for Realtime GPS Interferometry. Proceedings of ION GPS - 91, Albuquerque, New Mexico, September 1991. pp. 373 - 386.

Breach M.C., 1987. The Global Positioning System or ''Where am I ... now!'' The Royal Engineers Journal, Vol.101, no.3 September 1987, pp 196-207.

Brouwer F.J.J., et al., 1989. Navigatie en Geodetische Puntsbepaling met het Global Positioning System. Delftse Universitaire Pers, Delft, ISBN 90-6275-540-2. 322 pp.

Cross P.A., 1983. Advanced Least Squares to Position Fixing. North East London Polytechnic, Department of Land Surveying, Working Paper No. 6, ISBN 0-907382-06-1. 205 pp.

Cross P.A., Hollwey J.R., Small L.G., 1988. Working Paper No.2 Geodetic Appreciation. North East London Polytechnic, Department of Land Surveying, ISBN 0-907382-01-0. 185 pp.

Defence Mapping Agency (DMA), 1981. Omega Propagation Correction Tables for 10.2 kHz. Defence Mapping Agency Hydrographic/Topographic Center, Washington, Pub. 224 (101-C) B, DMA Stock No. OMPUB224101CB. pp I - IX.

Decca, 1979. The Decca Navigator, Principles and Performance of the System. The Decca Navigator Company Ltd, New Malden (UK). 72 pp.

Draaisma Y., Meester J.J. de, Mulders J.H., Spaans J.A., 1982. Leerboek Navigatie deel 2. De Boer Maritiem, ISBN 90-288-8308-6. 219 pp.

Draaisma Y., Mulders J.H., Spaans J.A., 1986. Leerboek Navigatie deel 3. De Boer Maritiem, ISBN 90-228-3809-4. 206 pp.

Gelb A., 1974. Applied Optimal Estimation. M.I.T. Press, London, ISBN 0-262-70008-5. 374 pp.

Haagmans M.E.E., 1986. De bepaling van korrekties voor propagatieeffekten met behulp van een microcomputer. Master Thesis. Delft University of Technology, Department of Geodecy. 93 pp.

IALA, 1990. IALA Aids to Navigation Guide (NAVGUIDE). International Association of Lighthouse Authorities, Paris, February 1990.

Jong G. de, 1989. Propagatie van radiogolven (lecture notes). Delft University of Technology, Department of Electrical Engineering. 184 pp.

Kielland P., Dagbert M., 1992. The Use of Spatial Statistics in Hydrography. International Hydrographic Review, Monaco, Vol LXIX No.1. pp 71 - 91.

Kielland P., Burrows K., Ward B., Dagbert M., Velberg R., 1992. Towards "IHOstat" : IHO approved software which evaluates the quality of bathymetric data. Proceedings Hydro '92, Copenhagen.

Kranendonk P., 1992. The influence of multipath on stationary GPS-range observations. Master thesis. Delft University of Technology, Department of Electrical Engineering. 57 pp.

Kreyszig E., 1970. Introductory mathematical statistics - Principles and Methods. John Wiley & Sons, New York, ISBN 471-50735-0. 470 pp.

Last D., 1992. The Accuracy and Coverage of Loran-C and of the Decca Navigator System - and the Fallacy of Fixed Errors. Journal of Navigation, Vol. 45 No 1. pp 36 - 51.

Lenart A.S., 1989. ARPA Accuracy Testing. Journal of Navigation, The Institute of Navigation, Vol. 42 No. 1. pp 117 - 123.

Lu G., 1992. Statistical Quality Control for Kinematic GPS Positioning. The University of Calgary, Department of Surveying Engineering. pp 1 - 7.

Mikhail E.M., Ackermann F., 1976. Observations and least squares. Harper & Row, New York. 484 pp.

Mikhail E.M., Gracie G., 1981. Analysis and adjustment of survey measurements. Van Nostrand Reinhold Company Inc., London, ISBN 0-442-25369-9. 339 pp.

Nanninga M., 1985. The Accuracy of Echosounding - Description in a Mathematical Model. Master thesis. Delft University of Technology, Department of Geodecy. pp. 115.

NATO, 1991:1. NAVSTAR GPS user equipment, introduction. The NATO team US Air Force Space Systems Division. 192 pp.

NATO, 1991:2. Technical characteristics of the NAVSTAR GPS. The NATO NAVSTAR GPS Technical Support Group. 160 pp.

Nee R.D.J. van, 1992. Multipath Effects on GPS Code Phase Measurements. Navigation : Journal of the Institute of Navigation Vol. 39 No. 2, Summer 1992. pp. 177 - 190.

Nee D.J.R. van, 1992. GPS Multipath and Satellite Interference Delft University of Thechnology, Department of Electrical Engineering. 11 pp.

Pierce J.A., 1965. Omega. Selected Papers Related to Long Range Navigation. Pickard & Burns Electronics, Waltham, U.S.A., Pub. No. 989, June 1966. pp 2-1 - 2-26.

SMS, 1988. The TRANSIT system (lecture notes). School of Military Survey Newbury, report no. WP1521/SIFd/2. 15 pp.

Spaans J.A., 1988:1. Application of the Kalman Filter in Marine Navigation. Delft University of Technology, Ship Hydromechanics Laboratory, Report No.803-P, January 1988. 7 pp.

Spaans J.A., 1988:2. Betrouwbaarheid, nauwkeurigheid en precisie van navigatiemethodieken (lecture notes). Netherlands Naval Academy. 131 pp.

Teunissen P.J.G., 1990. An Integrity and Quality Control Procedure for Use in Multi Sensor Integration. Proceedings of ION GPS-90, Colerado Springs, September 1990. pp 513 - 522.

Teunissen P.J.G., 1991. Differential GPS : Concepts and Quality Control. Invited lecture for the Netherlands Institute of Navigation, Amsterdam, September 1991. 46 pp.

UKMOD, 1984. The Effects of the Environment on Radio and Radar Wave Propagation. UK Hydrographic Department, Naval Oceanography and Meteorology Memorandum No. 1/84. 116 pp.

UKMOD, 1990. The Assessment of the Precision of Soundings. UK Hydrographic Department, Professional Paper No. 25. 41 pp.

US DoD & DoT, 1990. 1990 Federal Radionavigation Plan (FRP). US Departments of Transportation and Defence, report no. DOT-VNTSC-RSPA-90-3/DOD-4650.4. pp A1 - A38, C1 - C5.

Velberg P.J., 1992. The accuracy of depth information of the nautical chart. Hydrographic Journal, April 1993 (in print).

Willigen D. van, 1991. Radio plaatsbepaling (lecture notes). Delft University of Technology, Department of Electrical Engineering. 130 pp.

Wulder J.H., 1992. The implementation of a ship manoeuvring model in an integrated navigation system. PhD thesis. Delft University of Technology, Department of Mechanical Engineering and Marine Technology. Delft University Press 1992, ISBN 90-370-0063-0. 171 pp.

A1. Definitions, Abbreviations and Symbols

A1.1 Definitions

Accuracy[1] : the degree of conformance between the estimated or measured position at a given time and the true position

Accuracy[2] : the degree of agreement between the true value of a quantity and the most probable value derived from a series of measurements

Ambiguity : the identification of two or more possible positions with the same set of measurements, with no indication of which is the most nearly correct position

Availability[1] :
- navigation system availability is the percentage of time that the service of the system is usable
- signal availability is the percentage of time that navigational signals transmitted from external sources are available for use

Bias[2] : the difference in the mean of a sampling distribution of a statistic and the corresponding population parameter

BLUE : an estimate having the following statistical properties
- the estimate is unbiased
- the VCV matrix of the estimate is the one having minimum trace
- derived quantities have minimum variance

Blunder : a measurement which differs significantly from the expected value making it very likely that certain external circumstances are present than the ones that would make it a random error

Bounded System/Sensor : a system/sensor that does not need external signals to give its own output

Dead Reckoning Position : the position derived by using bounded sensors that measure the vessels heading, speed through the water and attitude, together with the estimated effects of wind, currents and tidal streams, sea state etc.

Error : the difference between a specific value and the correct or standard value

Fix rate : number of independent position fixes that can be obtained per unit of time

GDOP[1] : all geometric factors that degrade the accuracy of position fixes derived from externally referenced navigation systems

Hybrid Position fixing : independent positions of two (or more) systems are compared with each other and combined to obtain an MPP

Integrated Position fixing : an MPP is derived by combination of raw data (ranges, range differences, bearings etc) provided by two or more systems and/or sensors

Integrity[1] : the ability of a system to provide timely warnings to users when the system should not be used for navigation

Most Probable Position : the best position that can be derived using all information available

Optimal Estimator : a computational algorithm that processes measurements to deduce a minimum error estimate of a system by utilising :

i. knowledge of system and measurement dynamics
ii. assumed statistics of the system noise and measurement errors
iii. initial conditions of information

Predictable Accuracy : the accuracy of a statistic derived from measurements, taking into account all predicted errors

Predictable Accuracy[1] : the accuracy of a position with respect to the geographic or geodetic coordinates of the Earth

Precision : the degree of agreement between individual measurements in a set of observations

Random Error : an error unpredictable in magnitude and/or sign, but governed by laws of probability

Reliability : the measure of ease with which a blunder in a measurement can be detected

Reliability[1] : the probability of performing a specified function without failure under given conditions for a specified period of time

Standard Error[1] : a measure of dispersion of random errors about the mean value

Statistic[2] : any function of a number of random variables, usually identically distributed, that may be used to estimate a population parameter (mean, variance etc.)

Systematic Error : an error that follows some law by which it can be predicted

Unbiased[2] : an estimator $g(X_1,\ldots,X_n)$ for a parameter α is said to be unbiased if

$$E[g(X_1,\ldots,X_n)] = \alpha$$

Unbounded System/Sensor : a system/sensor that needs external signals to give its own output

Variate[2] : a random variable or a numerical value taken by it

1. definitions given in the 1990 Federal Radionavigation Plan [FRP,1990]
2. definitions used in mathematical statistics.

A1.2 Abbreviations

2D	Two Dimensional
ASF	Additional Secondary Factor
BLUE	Best Linear Unbiased Estimate
C/A - code	Coarse / Acquisition code
CD	Coding Delay
CEP	Circular Error Probable
CM	Central Meridian
CRI	Cross Rate Interference
CW	Continuous Wave
CWI	Carrier Wave Interference
DMA	Defence Mapping Agency
DNS	Decca Navigator System
DR	Dead Reckoning
d_{RMS}	Distance Root Mean Squared
ECEF	Earth Centred Earth Fixed
ED	Emission Delay
EM	Electro Magnetic
EP	Estimated Position
EPF	Electronic Position Fixing
E/S	Echo sounder
FE	Fixed Error
FRP	1990 Federal Radionavigation Plan
GDOP	Geometric Dilution of Precision
GMT	Greenwich Mean Time
GPS	Global Positioning System
GRI	Group Repetition Interval
HDOP	Horizontal Dilution of Precision
H_0	Null hypothesis
H_1	Alternative hypothesis
IHR	International Hydrographic Review
IMO	International Maritime Organization
I/O	Input / Output
KF	Kalman Filter
LEF	Lane Expansion Factor
LF	Low Frequency
LOP	Line Of Position
LSE	Least Squares Estimate
MCS	Master Control Station
MP	Multi Pulse
MPP	Most Probable Position
MSL	Mean Sea Level
NAVSTAR	NAVigation System using Timing And Ranging
NNSS	Navy Navigation Satellite System
OPCT	Omega Propagation Correction Tables
OPNET	OPerational NETwork
PCA	Polar Cap Absorption
P - code	Precise code
pdf	probability density function

PDOP	Position Dilution of Precision
PLL	PhaseLock Loop
POE	Pool Of Errors
PPS	Precise Positioning Service
PVC	Phase Velocity Correction
RC	Radio beacon
RMS	Root Mean Squared
RPF	Radio Position Fixing
SA	Selective Availability
SAM	Service Area Monitor
SEP	Spherical Error Probable
SID	Sudden Ionospheric Disturbance
SINS	Ships Inertial Navigation System
SIR	Signal-to-Interference Ratio
SMR	Signal-to-Multipath Ratio
SNR	Signal-to-Noise Ratio
SPS	Standard Positioning Service
SV	Space Vehicle
TM	Transverse Mercator
TRANET	TRAcking NETwork
TOT	Time Of Travel
TRP	Time Reference Point
TTFF	Time To First Fix
UEE	User Equipment Error
UERE	User Equivalent Range Error
URE	User Range Error
UTC	Universal Time Coordinate
UT1	Universal Time 1 (= GMT)
VCV	Variance Covariance
VLF	Very Low Frequency
WGS84	World Geodetic System 1984

A1.3 Symbols

A	least squares design matrix	m
Az	azimuth	°
B	magnetic field strength	T
B	spheroidal length of baseline	m
b	vector of absolute terms in linearized LS model	m
$C_{\hat{x}}$	error VCV matrix of Least Squares estimates	m²
C_l	a-priori error VCV matrix of observations	m²
c	propagation velocity of EM waves in vacuum	m/s
c_a	propagation velocity of EM waves in air	m/s
D	spheroidal distance	m
d	displacement	m
df	correction on received frequency	Hz
dX,dY,dZ	corrections to Cartesian coordinates in X-, Y- and Z-direction respectively	m
E	electric field strength	N/C
E	elevation	°
f	frequency	Hz
f_g	reference frequency of Transit receiver	Hz
f_s	satellite transmitting frequency	Hz
f_r	frequency of received signal	Hz
F	system dynamics matrix	
G	matrix representing the effect of noise on the state	
H	KF observation matrix	
H	altitude	m
k	minimum lanecount	
k_0	grid scale constant on CM	
L	lane number fraction	
l	vector of observed values	m
N_T	Transit Dopplercount	
N_I	for ionospheric disturbances corrected Dopplercount	
P	KF state vector error VCV matrix	
Q	KF system noise error VCV matrix	
R	KF measurement noise error VCV matrix	
R	range	m
R_c	pseudo range corrected for ionospheric effects	m
R_p	measured pseudo-range satellite - receiver	m
R_t	true range satellite - receiver	m
v	velocity	m/s
v	vector of LS residuals	m
v	white noise process vector of KF measurements	
v_c	speed of current	m/s
v_l	log speed	m/s
W	LS weight matrix	1/m²
w	w-test observable	
w	white noise process vector of KF system	
x	LS vector of parameters	m

x	KF state vector	
z	KF observations vector	
α	significance level	%
α	bearing	°
Γ	time discrete matrix representing the effect of noise on the state	
γ	confidence level	%
γ	angle subtended at receiver by master station and secondary station	°
δN_{400}^{ion}	ionospheric correction for 400 MHz signal Dopplercount of Transit system	
δR_e	range error due to error in ephemeris data	m
δR_m	range error due to multipath	m
δR_n	range error due to noise	m
δR_{tropo}	correction on range for tropospheric effects	m
δt	timing error due to clock offset from GPS time	s
δt_p	timing error due to propagation delay	s
$\delta \varphi_{sync}$	phase angle error due to synchronization	rad
Δt	time difference or time interval	s
Δv	systematic deviation of log	m/s
$\Delta \varphi$	phase difference	rad
$\Delta \psi$	deviation of gyro compass	°
ϵ	inclination angle	°
λ	wavelength	m
λ	longitude	°
σ	standard error	
σ_0^2	unit variance	
τ	reliability factor	
Φ	transition matrix	
φ	phase angle of signal	rad
ψ	submarine's heading	°
ψ_{GK}	gyro course	
ψ_c	direction of current	

subscripts

$k+1 \mid k$	estimate at time t_{k+1} based on observations $(z_1, \ldots, z_k)$
$k+1 \mid k+1$	estimate at time t_{k+1} based on observations $(z_1, \ldots, z_{k+1})$
m	master station
o	initial value
r	receiver
s	secondary station
s	satellite

superscripts

o	observed value
T	transposed of matrix

other

E[.]	mathematical expectation operator
^	estimated value
~	error in estimated value ($\tilde{x} = x - \hat{x}$)
∇	gradient operator
δ	error in measured quantity

A2. Propagation of Radio waves in the Atmosphere

In this appendix some topics related to propagation of radio waves in the atmosphere will be highlighted briefly. For a more detailed description on this subject, the reader is referred to UKMOD [1984], Haagmans [1986] and De Jong [1989].

A2.1 Propagation in the ionosphere

Because the ionosphere is a frequency dispersive medium, group velocity of a modulated wave will be different from the phase velocity of a monochromatic wave. The group velocity in the ionosphere is smaller than in free space and is referred to as group delay, whereas the phase velocity in the ionosphere is larger than in free space and is referred to as **phase advance**. Group delay and phase advance are equal in magnitude but opposite in sign. In [Brouwer et al.,1989] it is suggested that this property could be helpful to correct for ionospheric effect on GPS signals when using a C/A code receiver.
In the next part of this section expressions for the group delay and phase advance due to ionosphere will be given.

The propagation time of a monochromatic signal, travelling through the ionosphere can be derived from its equation of propagation velocity, which is given by

$$v_{ph} = \frac{c_0}{n} \tag{A2.1}$$

where

v_{ph} = phase velocity
n = refractive index
c_0 = propagation velocity in vacuum

Direct integration of (A2.1) in order to obtain **phase advance** is not possible since the refractive index of the ionosphere is dependent on height

$$n(h) = \sqrt{1 - \frac{f_p^2(h)}{f^2}} \approx 1 - \frac{1}{2}\frac{f_p^2(h)}{f^2} \qquad (A2.2)$$

with f_p the plasma frequency which is given by [De Jong,1989] :

$$f_p^2(h) = \frac{N(h)\,e^2}{4\pi^2\epsilon_0 m} \approx 80.5\,N(h) \qquad (A2.3)$$

where

h	= height above the Earth's surface
N	= electron density
m	= electron mass
e	= electron charge
ϵ_0	= permittivity of vacuum
f	= signal frequency

Substituting equations (A2.2) and (A2.3) in (A2.1) and defining the Total Electron Content (TEC) along the ray path between transmitting station and receiver as

$$TEC \triangleq \int_{s_0}^{s} N(h)\,dh \qquad (A2.4)$$

results in a phase advance in time of

$$\Delta t' - \Delta t = t' - t = -\frac{40.25}{c_0 f^2}\,TEC \quad [sec] \qquad (A2.5)$$

where

R		= total distance travelled
$\Delta t'$	$= t' - t_0$	= total propagation time due to ionosphere
Δt	$= t - t_0$	= total propagation time in free space
s		= upper height of propagation path
s_0		= lower height of propagation path

For modulated signals propagating through the ionosphere an equation similar to equation (A2.5) can be derived using the group velocity (v_{gr}) which is given by

$$v_{gr} = n\, c_0 \qquad \text{(A2.6)}$$

Integration of (A2.6), using equations (A2.2) and (A2.3), leads to the expression for group delay :

$$\Delta t' - \Delta t = t' - t = + \frac{40.25}{c_0 f^2}\,\text{TEC} \quad [\text{sec}] \qquad \text{(A2.7)}$$

A2.2 Ionospheric correction for GPS signals

The pseudo-range, containing ionospheric disturbances, can be written using equation (A2.7) as

$$c\Delta t = R_p = R_c + \frac{40.25}{f^2}\,\text{TEC} \quad [\text{metres}] \qquad \text{(A2.8)}$$

If pseudo-ranges are measured using two frequencies, both containing a ionospheric disturbance, then for each frequency equation (A2.8) can be formed

$$R_{p1} = R_c + \frac{40.25}{f_1^2}\,\text{TEC} \quad [\text{metres}] \qquad \text{(A2.9a)}$$

$$R_{p2} = R_c + \frac{40.25}{f_2^2}\,\text{TEC} \quad [\text{metres}] \qquad \text{(A2.9b)}$$

From these two equations, the range corrected for ionospheric disturbance can be calculated as

$$R_c = \frac{f_1^2 R_{p1} - f_2^2 R_{p2}}{f_1^2 - f_2^2} = \frac{R_{p1} - \left(\frac{f_2^2}{f_1^2}\right) R_{p2}}{1 - \left(\frac{f_2^2}{f_1^2}\right)} \quad \text{[metres]} \qquad (A2.10)$$

A2.3 Ionospheric correction for TRANSIT signals

The Dopplercount, given by equation (4.6) can be rewritten to allow for ionospheric disturbances using equation (A2.5)

$$N_i = (f_g - f_s) \Delta t + \frac{f_s}{c} \Delta R - \frac{40.25}{c f_s} \Delta TEC \qquad (A2.11)$$

If the Dopplercount is measured on two frequencies, both containing a ionospheric disturbance, then for each frequency equation (A2.11) can be formed

$$N_{400}^{o} = (f_{g1} - f_{s1}) \Delta t + \frac{f_{s1}}{c} \Delta R - \frac{40.25}{c f_{s1}} \Delta TEC \qquad (A2.12a)$$

$$N_{150}^{o} = (f_{g2} - f_{s2}) \Delta t + \frac{f_{s1}}{c} \Delta R - \frac{40.25}{c f_{s2}} \Delta TEC \qquad (A2.12b)$$

For Transit, a relation exists between the frequencies of the two transmitted signals, given by $k = f_1 / f_2 = 8/3$. Combining (A2.12a) and (A2.12b) results in

$$N_{400}^{o} - k N_{150}^{o} = - \frac{40.25}{c f_{s1}} \left[1 - k^2 \right] \Delta TEC \qquad (A2.13a)$$

or

$$\delta N_{400}^{ion} = - \frac{40.25}{c f_{s1}} \Delta TEC = \frac{N_{400}^{o} - k N_{150}^{o}}{1 - k^2} \qquad (A2.13b)$$

A2.4 Propagation in the troposphere

Propagation of radio waves in the troposphere depends on the refractive index. The refractive index of air is a function of height above the earth's surface and is determined by its pressure, temperature and humidity. Variations in refractive index result in two effects : refraction (the angular bending of radio waves) and change in propagation velocity.
Although all signals in the troposphere suffer from tropospheric effects, it becomes most noticeable for frequencies higher then 100 MHz [Van den Berg,1989]. The reason for this is that for higher frequencies the ground wave dampens more quickly, scattering due to objects in the atmosphere is more pronounced and the curvature of the earth plays a more important role.
The navigation systems discussed in this paper can be grouped into three classes, each having its own mechanism describing signal propagation :

1. for VLF signals (Omega), the propagation can best be described by TM modi in a waveguide bounded by the earth's surface and the ionosphere. The propagation velocity depends on the conductivity of the earth's surface and the ionosphere. For Omega, the TM_1 and TM_2 modi play an important role.

2. for the LF signals (Decca and Loran-C), signal propagation is best described by ground wave in combination with a series of signals reflected at the ionosphere, the so-called sky waves.
   The propagation velocity of the ground wave is a function of ground conductivity along the signal path. Normally, the transmission time of a signal from a transmitter to the receiver is considered to consist of two parts : the time needed if the path is an 'all sea' path plus a correction to allow for land path. This correction is called the Additional Secondary Factor (ASF). The ASF for a given position can be obtained in two ways :

   - measure the ASF value in a number of grid points in the area where operating and apply the corrections by interpolation;
   - calculate the ASF values using the method described by Millington. The way this method can be used to correct observations is described by Haagmans [1986].

   When the signals cross coastlines, transient effects occur, resulting in a sudden changes of phase. This can lead to large errors in position fixing when not corrected for.

3. for UHF signals (GPS and Transit), signal propagation can be approximated by the laws of optics.
Vertical profiles of refractivity are estimated, based on measurements at the surface and standard models of the atmosphere. This way corrections for measured times of propagation can be calculated and allowed for. Normally, the corrections are split into two components : a wet and dry component. The dry component is only dependent on vertical pressure and temperature profiles and can be modelled quite well. The vertical water vapour pressure profile is more difficult to model.
The model most often used for computing the correction for tropospheric refraction is the simplified **Hopfield** model.

A3. Dilution of Precision

The GPS exhibits statistical accuracy distributions because of two important parameters :

- User Equivalent Range Error (UERE) :

  A measure of the error in the range measurement to each satellite as seen by the receiver. It tends to be different for each satellite and tends to be at minimum following an upload.

- Geometric Dilution of Precision (GDOP).

  A measure of the error contributed by the geometric relationship of the satellites as seen by the receiver. It varies because the satellites are in constant motion and their geometric relationships are constantly changing.

A3.1 Geometric Dilution of Precision

The radial range measurement equation for one satellite is given by

$$\sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} + c\Delta t - R_i = 0 \qquad \text{(A3.1)}$$

where

(x,y,z)	coordinates of receiver
(x_i,y_i,z_i)	coordinates if i-th satellite
Δt	receiver clock bias
c	propagation velocity of radio waves
R_i	pseudo-range measurement to i-th satellite

A minimum of four equations is needed to solve for the four unknowns. A best estimate of the parameters is calculated using the least squares algorithm, which also gives an error VCV matrix of the estimates, based on the a priori VCV matrix of the observations

$$C_{\hat{x}} = \left[A^T C_l^{-1} A\right]^{-1} \qquad \text{(A3.2)}$$

where

C_l a priori error VCV matrix of observations

A	design matrix
$\mathbf{C}_{\hat{x}}$	error VCV matrix of parameter estimates

The relationship between the observation errors and parameter estimate errors is a function of satellite geometry only. Therefore, an important consideration in the proper use of GPS is that the four satellites used for the position solution should be arranged in such a geometric relationship that small errors in the pseudo-range measurements (UERE) result in small user position and clock bias errors.
In order to be able to assess the effect of satellite geometry on the position and clock bias errors, it is assumed that each individual pseudo-range measurement has a zero mean error with unit variance, and no correlation exits between the measurements. Under these assumptions, $\mathbf{C}_l$ will be the identity matrix, and the error VCV matrix of the parameters will be given by

$$C_{\hat{x}} = \left[A^T A\right]^{-1} \qquad (A3.3)$$

The Geometric Dilution of Precision (GDOP) in now defined as the square root of the trace of $\mathbf{C}_{\hat{x}}$ when $\mathbf{C}_l$ is the identity matrix. Therefore :

$$GDOP = \sqrt{\mathrm{Trace}\left[\left(A^T A\right)^{-1}\right]} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_T^2} \qquad (A3.4)$$

where

$\sigma_x, \sigma_y, \sigma_z$	standard errors of user position
σ_T	standard error of receiver clock bias.

Depending on the user's application, there are five interrelated DOP statistics. In each case they are the amplification factor of pseudo-range measurement errors into user errors due to the effect of satellite geometry with respect to the receiver. Each DOP statistic is a dimensionless number and independent of the coordinate system used.
In a similar way to GDOP, the four other DOPs can be defined as:

- Position Dilution of Precision (PDOP) - the square root of the sum of the squares of the three components of position error $(\sigma_x, \sigma_y, \sigma_z)$;

- Horizontal Dilution of Precision error (HDOP) - the square root of the sum of the squares of the horizontal components of the position error (σ_x,σ_y). HDOP is sometimes subdivided in its components in the directions East (σ_x) and North (σ_y) which are named EDOP and NDOP respectively.

- Vertical Dilution of Precision (VDOP) - the altitude error (σ_z)

- Time Dilution of Precision (TDOP) - the error in the receiver clock bias, multiplied by the speed of light ($c\sigma_T$).

For submarine navigation, HDOP is most important since GPS can only be used when the submarine is at sea-surface and the vertical component is well known.

A4. Decca and Omega Error VCV Matrix

A4.1 Decca error VCV matrix

For each Decca pattern, the phase deviation in the different signals under sky wave conditions, causing errors in the phase measurements at the receiver, can be combined to give an expression for the total error in phase difference of that pattern at the receiver [Decca,1979] :

- Red : $\delta\Delta\varphi_R = 4\delta\varphi_m - 4\delta\varphi_{mR} - 3\delta\varphi_{syncR} - 3\delta\varphi_R$ (A4.1a)
- Green : $\delta\Delta\varphi_G = 3\delta\varphi_m - 3\delta\varphi_{mG} - 2\delta\varphi_{syncG} - 2\delta\varphi_G$ (A4.1b)
- Purple : $\delta\Delta\varphi_P = 5\delta\varphi_m - 5\delta\varphi_{mP} - 6\delta\varphi_{syncP} - 6\delta\varphi_P$ (A4.1c)

Correlation will exist between the phase difference errors of the LOPs because the same master station is used to determine the phase difference for each pattern LOP. This leads to a 'full' a priori VCV matrix, i.e. having both variances and covariances which can be calculated using the following formulae :

variances :

$$E[(\delta\Delta\varphi_R)^2] = \sigma^2_{\Delta R} = 16\sigma^2_m + 16\sigma^2_{mR} + 9\sigma^2_R + 9\sigma^2_{syncR} \quad [\text{sq.lanes}] \quad (A4.2a)$$

$$E[(\delta\Delta\varphi_G)^2] = \sigma^2_{\Delta G} = 9\sigma^2_m + 9\sigma^2_{mG} + 4\sigma^2_G + 4\sigma^2_{syncG} \quad [\text{sq.lanes}] \quad (A4.2b)$$

$$E[(\delta\Delta\varphi_P)^2] = \sigma^2_{\Delta P} = 25\sigma^2_R + 25\sigma^2_{mP} + 36\sigma^2_P + 36\sigma^2_{syncP} \quad [\text{sq.lanes}] \quad (A4.2c)$$

covariances :

$$E[\delta\varphi_R\delta\varphi_G] = \sigma_{\Delta R\Delta G} = 12\sigma^2_m \quad [\text{sq.lanes}] \quad (A4.2d)$$

$$E[\delta\varphi_R\delta\varphi_P] = \sigma_{\Delta R\Delta P} = 20\sigma^2_m \quad [\text{sq.lanes}] \quad (A4.2e)$$

$$E[\delta\varphi_G\delta\varphi_P] = \sigma_{\Delta G\Delta P} = 15\sigma^2_m \quad [\text{sq.lanes}] \quad (A4.2f)$$

The values for standard deviation of the synchronization errors to be used in equation A4.2 are given by Decca [1979] as

$$3\sigma_{syncR} = 0.012 \quad [lanes]$$
$$2\sigma_{syncG} = 0.009 \quad [lanes]$$
$$6\sigma_{syncP} = 0.015 \quad [lanes]$$

The errors due to sky wave interference have been found to be increasing with covered distance. Table 4.3 given by Decca [1979] shows how standard deviations change with distance. In order to be able to calculate values at intermediate distances, second order polynomials are used to replace the table. These polynomials, calculated using least squares curve fitting are :

$$\sigma = (2.407d^2 - 2.273d + 4.240) * 10^{-3} \quad [lanes] \qquad (A4.3a)$$

$$\sigma = (2.843d^2 - 2.597d + 5.000) * 10^{-3} \quad [lanes] \qquad (A4.3b)$$

where

d = covered distance from station in 100 km

Formula (A4.3b) should be used to calculate a value for σ_{sP}, whereas formula (A4.3a) has to be used for calculation of all the other variances as given in (A4.2a) and (A4.2b).
The calculated values for standard error using above formulae are valid for summer night conditions over seawater. For a Summer's day these values have to be divided by 8, for a winter's day by 2 and for a winter's night to be multiplied by 1.5.

distance to station	σ_{mR} / σ_{sR} / σ_{sG} σ_{msR} / σ_{msG} / σ_{msP}	σ_{msP}
100 km	0.0041	0.0050
200 km	0.0100	0.0118
300 km	0.0187	0.0224
400 km	0.0335	0.0400
500 km	0.0532	0.0632

Table A4.1 Standard deviations in transmitted cycles (lanes) for good soil and seawater during Summer nights [Decca,1979].

When (A4.3a) and (A4.3b) are used, the vessel is assumed to be positioned well inside the area in which ground wave coverage is maximum and interference by sky waves can be neglected. If this is not the case, unpredictable phase shifts might occur by interference, degrading the accuracy of the Decca LOP.

In the above equations, the variances and covariances are given in lanes squared. To get these values in metres squared, the following transformation formula has to be used :

$$\sigma'_{ij} = \frac{1}{4}\sigma_{ij}\lambda_{c_i}\lambda_{c_j}\mathrm{cosec}(\frac{1}{2}\gamma_i)\mathrm{cosec}(\frac{1}{2}\gamma_j) \quad \text{[sq.metres]} \qquad \text{(A4.4)}$$

where σ_{ij} = VCV matrix element in sq.lanes

λ_{c_i} = wavelength of comparison frequency for master and slave i

λ_{c_j} = wavelength of comparison frequency for master and slave j

γ_i = at receiver angle subtended between master and slave i

γ_j = at receiver angle subtended between master and slave j

A4.2 Omega error VCV matrix

Consider two Omega LOPs to be used for a position fix. These LOPs are obtained using stations A,B and C. At the receiver phase differences are measured between the signals from A and B (LOP1) and A and C (LOP2). The error in the obtained phase differences are given by :

$$\delta\Delta\varphi_{AB} = \delta\varphi_A - (\delta\varphi_B + \delta\varphi_{syncB}) \qquad \text{(A4.5a)}$$

$$\delta\Delta\varphi_{AC} = \delta\varphi_A - (\delta\varphi_C + \delta\varphi_{syncC}) \qquad \text{(A4.5b)}$$

where

$\delta\varphi$ = error in single phase measurement at receiver

$\delta\varphi_{sync}$ = synchronization error of station with respect to station A

$\delta\Delta\varphi$ = error in phase difference measurement

Because transmitting station A is used to obtain both LOPs, correlation will exist in errors present, resulting in a 'full' a priori VCV matrix. The variances and covariances can be calculated as follows :

variance :

$$E[(\delta\Delta\phi_{AB})^2] = \sigma^2_{\Delta AB} = \sigma^2_A + \sigma^2_B + \sigma^2_{syncB} \quad \text{[sq.lanes]} \qquad (4.6a)$$

$$E[(\delta\Delta\phi_{AC})^2] = \sigma^2_{\Delta AC} = \sigma^2_A + \sigma^2_C + \sigma^2_{syncC} \quad \text{[sq.lanes]} \qquad (4.6b)$$

covariance :

$$E[(\delta\Delta\phi_{AB})(\delta\Delta\phi_{AC})] = \sigma_{\Delta AB\Delta AC} = \sigma^2_A \quad \text{[sq.lanes]} \qquad (4.6c)$$

In the above equations, the variances and covariances are given in lanes squared. To get these values in metres squared, the following transformation formula has to be used :

$$\sigma'_{ij} = \frac{1}{4}\sigma_{ij}\lambda^2_c \operatorname{cosec}(\frac{1}{2}\gamma_i)\operatorname{cosec}(\frac{1}{2}\gamma_j) \quad \text{[sq.metres]} \qquad (A4.7)$$

In the situation where the two LOPs are obtained using 4 different stations, no correlation exists between the LOPs resulting in a diagonal error VCV matrix (the covariances being equal to zero).

A5. Least Squares Observation Equations

In the least squares calculations, use is made of mathematical models of the observations. In this appendix, the functional model with associating linearized model for the systems as used in the least squares calculations of the simulation program will be given.
In this appendix, the Cartesian coordinates used are coordinates in a TM grid as defined in section 3.3 of chapter 5.

GPS observable

The functional model for a measured pseudo-range is written as

$$F(x,l) = \sqrt{(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2} - c\Delta T - R \qquad (A5.1)$$

The linearized model $Ax = b + v$:

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z + \frac{\partial F}{\partial c\Delta T}\delta c\Delta T = R^0 - (d - c\Delta T)_0 + v \qquad (A5.2)$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{(x_0 - x_s)}{d_0} \qquad (A5.3a)$$

$$\frac{\partial F}{\partial y} = \frac{(y_0 - y_s)}{d_0} \qquad (A5.3b)$$

$$\frac{\partial F}{\partial z} = \frac{(z_0 - z_s)}{d_0} \qquad (A5.3c)$$

$$\frac{\partial F}{\partial c\Delta T} = -1 \qquad (A5.3d)$$

where

$$d_0 = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2 + (z_0 - z_s)^2} \qquad (A5.4)$$

<u>Transit observable</u>

The functional model for a range difference based on Dopplercount is written as

$$F(x,l) = (R_2 - R_1) - \lambda_s[N - (f_g - f_s)\Delta t'] \qquad (A5.5)$$

The linearized model $Ax + Cv - b = 0$:

$$\left|\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z + \frac{\partial F}{\partial f_s}\delta f_s\right| + \left|\frac{\partial F}{\partial N}v_N + \frac{\partial F}{\partial \Delta t'}v_{\Delta t'}\right| + F(x_0, l^0) = 0 \qquad (A5.6)$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{(x_0 - x_2)}{R_2} - \frac{(x_0 - x_1)}{R_1} \qquad (A5.7a)$$

$$\frac{\partial F}{\partial y} = \frac{(y_0 - y_2)}{R_2} - \frac{(y_0 - y_1)}{R_1} \qquad (A5.7b)$$

$$\frac{\partial F}{\partial z} = \frac{(z_0 - z_2)}{R_2} - \frac{(z_0 - z_1)}{R_1} \qquad (A5.7c)$$

$$\frac{\partial F}{\partial f_s} = \frac{c[N - f_g\Delta t']}{f_s^2} \qquad (A5.7d)$$

$$\frac{\partial F}{\partial N} = -\lambda_s \tag{A5.7e}$$

$$\frac{\partial F}{\partial \Delta t'} = \lambda_s (f_g - f_s) \tag{A5.7f}$$

where

$$R_1 = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2} \tag{A5.8a}$$

$$R_2 = \sqrt{(x_0 - x_2)^2 + (y_0 - y_2)^2 + (z_0 - z_2)^2} \tag{A5.8b}$$

<u>Decca observable</u>

The functional model for a Decca lane number is written as

$$F(x,l) = \frac{1}{\lambda_c}\left[\frac{\sqrt{(x - x_m)^2 + (y - y_m)^2}}{L_m} - \frac{\sqrt{(x - x_s)^2 + (y - y_s)^2}}{L_s} + B\right] - N \tag{A5.9}$$

with L the line scale factor used to convert grid distances to spheroidal distances, given by

$$L_i = 1 + \frac{x^2 + x x_i + x_i^2}{6 R^2} \tag{A5.10}$$

The linearized model $Ax = b + v$:

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y = N^0 - \frac{1}{\lambda_c}\left[\frac{d_{m0}}{L_m} - \frac{d_{s0}}{L_s} + B\right] + v \tag{A5.11}$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{1}{\lambda_c}\left[\frac{(x_0 - x_m)}{L_m d_{m0}} - \frac{d_{m0}(2x + x_m)}{L_m^2 6R^2} - \frac{(x_0 - x_s)}{L_s d_{s0}} + \frac{d_{s0}(2x + x_s)}{L_s^2 6R^2}\right]$$

(A5.12a)

$$\frac{\partial F}{\partial y} = \frac{1}{\lambda_c}\left[\frac{(y_0 - y_m)}{L_m d_{m0}} - \frac{(y_0 - y_s)}{L_s d_{s0}}\right] \qquad \text{(A5.12b)}$$

where

$$d_{i0} = \sqrt{(x - x_i)^2 + (y - y_i)^2} \qquad \text{(5.13)}$$

<u>Position observable</u>

An observed position in given in grid coordinates leads to two equations :

· <u>x coordinate</u> :

functional model

$$F(x,l) = x - x^0 \qquad \text{(A5.14)}$$

the linearized model $Ax = b + v$:

$$\frac{\partial F}{\partial x}\delta x = x^0 - x_0 + v \qquad \text{(A5.15)}$$

The differential evaluated at $x = x_0$ is

$$\frac{\partial F}{\partial x} = 1 \qquad \text{(A5.16)}$$

· <u>y coordinate</u> :

functional model

$$F(x,l) = y - y^0 \tag{A5.17}$$

the linearized model Ax = b + v :

$$\frac{\partial F}{\partial y}\delta y = y^0 - y_0 + v \tag{A5.18}$$

The differential evaluated at $x = x_0$ is

$$\frac{\partial F}{\partial y} = 1 \tag{A5.19}$$

<u>distance observable</u>

The functional model for a measured distance is written as

$$F(x,l) = \sqrt{(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2} - D \tag{A5.20}$$

The linearized model Ax = b + v :

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z = D^0 - D_0 + v \tag{A5.21}$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{(x_0 - x_s)}{d_0} \tag{A5.22a}$$

$$\frac{\partial F}{\partial y} = \frac{(y_0 - y_s)}{d_0} \tag{A5.22b}$$

$$\frac{\partial F}{\partial z} = \frac{(z_0 - z_s)}{d_0} \tag{A5.22c}$$

where

$$d_0 = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2 + (z_0 - z_s)^2} \tag{A5.23}$$

<u>azimuth observable</u>

The functional model for an azimuth is written as

$$F(x,l) = \arctan\left(\frac{(x - x_s)}{(y - y_s)}\right) - Az \tag{A5.24}$$

The linearized model Ax = b + v :

$$\frac{\partial F}{\partial x}\delta x + \frac{\partial F}{\partial y}\delta y = Az^0 - \arctan\left(\frac{(x_0 - x_s)}{(y_0 - y_s)}\right) + v \tag{A5.25}$$

Differentials evaluated at $x = x_0$ are

$$\frac{\partial F}{\partial x} = \frac{(y_0 - y_s)}{d_0^2} \tag{A5.26a}$$

$$\frac{\partial F}{\partial y} = -\frac{(x_0 - x_s)}{d_0^2} \tag{A5.26b}$$

where

$$d_0 = \sqrt{(x_0 - x_s)^2 + (y_0 - y_s)^2} \qquad (A5.27)$$

A6. First-Order Markov Process & Shaping Filters

A6.1 First-Order Markov Process

A continuous process x(t) is a first-order Markov process if for every k and $t_1 \le t_2 \le \ldots \le t_k$ it is true that the distribution function for process $x(t_k)$ is dependent only on the value at one point immediately in the past $x(t_{k-1})$

$$F[x(t_k)|x(t_{k-1}),\ldots,x(t_1)] = F[x(t_k)|x(t_{k-1})] \tag{A6.1}$$

A continuous first-order Markov process can be represented by the following differential equation

$$\dot{x} + \beta(t)x = w(t) \tag{A6.2}$$

The autocorrelation function of input and output are given by

$$R_w(\tau) = E[w(t)w(t+\tau)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_w(\omega)e^{-j\omega\tau}d\omega \tag{A6.3a}$$

$$R_n(\tau) = E[n(t)n(t+\tau)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_n(\omega)e^{-j\omega\tau}d\omega \tag{A6.3b}$$

where

$\Phi(\omega)$ noise spectral density function
$R_w(\tau)$ autocorrelation function of input
$R_n(\tau)$ autocorrelation function of output

They are related to each other by

$$R_n(\tau) = h(-\tau) * h(\tau) * R_w(\tau) \tag{A6.4}$$

where

$h(\tau)$ time domain impulse response
$*$ represents the convolution operator.

In the frequency domain the spectral density functions is defined as the Fourier transform of the autocorrelation function

$$\Phi(\omega) = \int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau \qquad (A6.5)$$

The spectral density functions of the input and output signals are related to each other by

$$\Phi_n(\omega) - |H(j\omega)|^2 \Phi_w(\omega) \qquad (A6.6)$$

If the p.d.f. of x(t) and w(t) are restricted to be Gaussian, the process is called a Gauss-Markov process.
The statistics of a stationary first-order Gauss-Markov process are described completely by the following autocorrelation function

$$R_n(\tau) = o^2 e^{-\beta|\tau|} \qquad (A6.7)$$

A6.2 Shaping filters

A shaping filter is a linear dynamical system driven by a Gaussian white noise process whose output has the same statistical characteristics as the Gaussian random process $\underline{n}(t)$ that has zero mean and whose correlation is given by

$$E[\underline{n}(t)\underline{n}^T(\tau)] = D(t,\tau) \qquad (A6.8)$$

The autocorrelation of many physical phenomena is well-approximated by

$$R_n(\tau) = \sigma_n^2 e^{-\beta|\tau|} \qquad (A6.9)$$

where

β correlation time
σ_n standard error of the noise process.

A first-order Gauss-Markov process, generated by passing an uncorrelated signal w(t) through a linear first-order feedback system - the shaping filter -, is often used to provide an approximation for this band limited signal whose spectral density is flat over a finite bandwidth. The feedback system can be described by the following differential equation

$$\dot{n} = -\beta n + w \qquad \text{(A6.10)}$$

The autocorrelation and spectral density functions of input and output signals for a shaping filter of this kind are then given by

$$R_w(\tau) = 2\beta\sigma^2\delta(\tau) \qquad \text{(A6.11a)}$$

$$\Phi_w(\omega) = 2\beta\sigma^2 \qquad \text{(A6.11b)}$$

$$\Phi_n(\omega) = \frac{2\beta\sigma^2}{\omega^2 + \beta^2} \qquad \text{(A6.11c)}$$

$$R_n(\tau) = \sigma^2 e^{-\beta|\tau|} \qquad \text{(A6.11d)}$$

A7. Non-linear Filtering

Consider a time continuous system having its state described by the following non-linear vector differential equation

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t) + G(\tau)\mathbf{w}(\tau) \tag{A7.1}$$

In order to be able to use this state equation in the Kalman filter algorithm, the equation should be linearized. One method of doing this is by Taylor series expansion of (A7.1), resulting in

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}_0,t) + \nabla_X\left(\mathbf{x} - \mathbf{x}_0\right) = \nabla_X\mathbf{x} + \left[\mathbf{f}(\mathbf{x}_0,t) - \nabla_X\mathbf{x}_0\right] \tag{A7.2}$$

where

$$\nabla_X = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x} = \boldsymbol{x}_0}$$

The solution of (A7.1) is given by

$$\begin{aligned} x(t) = \Phi(t,t_0)\,x(t_0) &+ \int_{t_0}^{t}\Phi(t,\tau)\left[f(x^0,\tau) - \nabla_X x^0\right]d\tau \\ &+ \int_{t_0}^{t}\Phi(t,\tau)G(\tau)w(\tau)d\tau \end{aligned} \tag{A7.3}$$

where $\Phi(t,t_0)$ is the solution of the matrix differential equation

$$\dot{\Phi} = \nabla_X\Phi \qquad \Phi(t,t) = I \quad \forall t \tag{A7.4}$$

resulting in

$$\Phi(t, t_0) = e^{\int_{t_0}^{t} \nabla_x d\tau} \approx I + \int_{t_0}^{t} \nabla_x d\tau \qquad \text{(A7.5)}$$

In order to be able to use equation (A7.3) on a computer, it needs to be written in a time discrete form, which can be regarded as sampling of the state, setting $\mathbf{x}_0$ equal to the last best estimate $\hat{\mathbf{x}}_{k|k}$. This leads to

$$\boldsymbol{x}_{k+1} = \Phi_k \boldsymbol{x}_k + \Lambda_k \boldsymbol{u}_k + \Gamma_k \boldsymbol{w}_k \qquad \text{(A7.6)}$$

where

$$\Lambda_k \boldsymbol{u}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau) \left[\boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}) - \nabla_x \hat{\boldsymbol{x}}_{k|k} \right] d\tau$$

$$\Gamma_k \boldsymbol{w}_k = \int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)\, G(\tau)\, \boldsymbol{w}(\tau)\, d\tau$$

$$\Phi_k = \Phi(t_{k+1}, t_k) \approx I + \int_{t_k}^{t_{k+1}} \nabla_x d\tau$$

The propagation of the time discrete state vector between two measurements, together with its VCV matrix, is now given by

$$\hat{\boldsymbol{x}}_{k+1|k} \approx \Phi_k \hat{\boldsymbol{x}}_{k|k} + \Lambda_k \boldsymbol{u}_k \approx \hat{\boldsymbol{x}}_{k|k} + \int_{t_k}^{t_{k+1}} \boldsymbol{f}(\hat{\boldsymbol{x}}_{k|k}, \tau) + R(\hat{\boldsymbol{x}}_{k|k}, \tau) \qquad \text{(A7.7a)}$$

$$P_{k+1|k} = \Phi_k P_{k|k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T \qquad \text{(A7.7b)}$$

where $R(\hat{\underline{\mathbf{x}}}_{k|k}, \tau)$ is a rest term due to truncation of the Taylor series expansion of $\dot{\mathbf{x}}$ and series expansion of Φ.

A8. Computer Simulation Program

A8.1 Simulation program files

```
UVNAV    .EXE    : simulation program executable file

TO_WGS84.DAT     : geodetic datum data file
DECCA    .CHN    : Decca chain data file
UVNAV    .MNU    : menu layout data file
920917_1.RUN     : file with simulation data
```

A8.2 Program parameters default values

```
(*---------------------- Sensors and Systems -----------------------*)

Systems[1]      := 0          (* GPS                                *)
Systems[2]      := 2          (* SINS                               *)
Systems[3]      := 2          (* Omega                              *)
Systems[4]      := 2          (* Loran-C                            *)
Systems[5]      := 0          (* Decca                              *)

Sensors[ 1]     := 1          (* Gyro                               *)
Sensors[ 2]     := 1          (* Log                                *)
Sensors[ 3]     := 0          (* Rudder                             *)
Sensors[ 4]     := 2          (* Inclinometer                       *)
Sensors[ 5]     := 0          (* PressureSensor                     *)
Sensors[ 6]     := 2          (* Echosounder                        *)
Sensors[ 7]     := 2          (* For future use                     *)
Sensors[ 8]     := 2          (* For future use                     *)
Sensors[ 9]     := 2          (* For future use                     *)
Sensors[10]     := 2          (* For future use                     *)

 0  = connected but switched off
 1  = connected switched on
 2  = not connected

(*---------------------- Data files ---------------------------------*)

DatumDataFile   := 'C:\DATA\TO_WGS84.DAT'
DeccaDataFile   := 'C:\DATA\DECCA.CHN'
RunDataFile     := 'C:\DATA\'
```

```
(*------------------------ G P S Datum ------------------------------*)

name := WGS84
axis := 6378137
e2   := 0.00669437999
CM   := 0

(*------------ Default geodetic datum used for calculations ----------*)

name := ED50
axis := 6378388
e2   := 0.00672267002
CM   := 3 / 180 * pi
dX   := -87
dY   := -98
dZ   := -121

(*---------------------------- G P S --------------------------------*)

PCode           := false        (* C/A code receiver                *)
GPS.SE_SF       := 1            (* error VCV matrix scale factor    *)

(*--------------------------- Decca ---------------------------------*)

Decca.CF        := 64           (* error VCV matrix scale factor    *)
                                (* Summer's day conditions          *)

(*----------------------- Sensors default ---------------------------*)

s_weight        :=    1.026     (* specific weight of seawater      *)

(*--------------------- LSE variables setup -------------------------*)

epsilon         := 1.0          (* iteration break-off criterion    *)
s02             := 1.0          (* initial value of σ0²             *)
S2              := 0.0          (* initial value of sample variance *)

LOPReject       := 'A'          (* ask if LOPs have to be discarded *)
                                (* from set after statistical test  *)

(*------------------ statistical variables setup --------------------*)

ReliabInterval  := 5.0          (* reliability interval α = 0.05    *)
ConfRegion      := 95.0         (* confidence level     γ = 95%     *)

ScaleVCV        := false        (* no scaling of a priori VCV matrix *)

(*----------------- Kalman Filter variables setup -------------------*)

DeltaT          := 1            (* sampling interval Δt = 1 sec     *)
```

A8.3 Menu layout

- Quit
- Edit
  - Prt.Scr.
- Setup
  - Params
    - Datum
    - DR Pos
    - Current
    - Develop
    - Reset
  - Systems
    - GPS
    - SINS
    - Omega
    - Loran-C
    - Decca
  - Sensors
    - Gyro
    - Log
    - Rudder
    - Inclin.
    - Pressure
    - E/S
- Obs
  - Bearing
  - Distance
  - Radar

- Kalman
  - Filtering
  - Predict.
  - Smoothing
- Statistics
  - Systems
    - GPS
    - SINS
    - Omega
    - Loran-C
    - Decca
    - Obs
  - Sensors
  - LSE
    - Stats I
    - Stats II
    - Tests
  - KF
    - StateVect
    - Tests
  - Position
    - Stats
    - Plot
    - History
  - MainPage

- Demo
  - GPS
    - Factor σ
  - Decca
    - PVC
    - Decca CF
  - Params
    - Smple.Int
    - CycleStop
    - S.Weight
  - LSE
    - It.Limit
    - AIM
    - ScaleVCV
  - Plot
    - Size
    - Scale
    - X/Y Fact
  - Recalc LS
- Data
  - Decca
  - A-prio
    - Systems
    - Varcor
  - Weights
  - Ellipsoid
- Run
  - Continue
  - Suspend
  - Start

A8.4 Layout of data files

Geodetic datum data file

```
************************************************
*
* Parameters of reference spheroids
*
* Update : 29/01/92
*
************************************************
AIRY 1830                      6377563.396   299.3249646
AIRY MODIFIED                  6377340.189   299.3249646
AUSTRALIAN NATIONAL            6378160       298.25
BESSEL 1841                    6377397.155   299.1528128
BESSEL MODIFIED                6377492.018   299.1528
CLARKE 1858                    6378235.6     294.2606768
CLARKE 1866                    6378206.4     294.9786982
CLARKE 1880                    6378249.145   293.465
CLARKE 1880 MODIFIED           6378249.145   293.4663
EVEREST                        6377276.345   300.8017
EVEREST MODIFIED               6377304.063   300.8017
FISCHER 1960                   6378166       298.3
FISCHER 1968                   6378150       298.3
HAYFORD 1909                   6378388       297
HELMERT 1906                   6378200       298.3
HOUGH                          6378270       297
INTERNATIONAL                  6378388       297
KRASSOVSKY                     6378245       298.3
MADRID 1924                    6378388       297
*
*
************************************************
*
* Transformation parameters of Datums
*
* Datum ==> WGS84
*
* Update : 08/11/91
*
************************************************
*
AUSTRALIAN GEODETIC 1966   6378160.000  298.25         -133.000  -48.000   148.000  0.00  0.00  0.00  0.00
AUSTRALIAN GEODETIC 1984   6378160.000  298.25         -134.000  -48.000   149.000  0.00  0.00  0.00  0.00
ED50                       6378388.000  297             -87.000  -98.000  -121.000  0.00  0.00  0.00  0.00
ED50 [DMA]                 6378388.000  297             -87.000  -98.000  -121.000  0.00  0.00  0.00  0.00
ED50 [UKOOA]               6378388.000  297             -86.000  -96.000  -120.000  0.00  0.00  0.00  0.00
ED79                       6378388.000  297             -86.000  -98.000  -119.000  0.00  0.00  0.00  0.00
GRS67                      6378160.000  298.247167427    59.944   -3.843    22.932  0.00  0.00  0.00  0.00
```

```
NAD 1927            6378206.400  294.9786982     -8.000   160.000  176.000  0.00  0.00  0.00  0.00
NAD 1983            6378137.000  298.257222101    0.000     0.000    0.000  0.00  0.00  0.00  0.00
OSGB 1936           6377563.396  299.3249647    375.000  -111.000  431.000  0.00  0.00  0.00  0.00
TOKYO               6377397.155  299.1528128   -128.000   481.000  664.000  0.00  0.00  0.00  0.00
WGS72               6378135.000  298.26           0.000     0.000    0.000  0.00  0.00  0.00  0.00
WGS72 [DMA]         6378135.000  298.26           0.000     0.000    4.500  0.00  0.00  0.55  0.22
WGS84               6378137.000  298.257223563    0.000     0.000    0.000  0.00  0.00  0.00  0.00
ZANDERIJ            6378388.000  297           -265.000   120.000 -358.000  0.00  0.00  0.00  0.00
*
* End of Data
```

Decca chain data file

```
* Decca chain data
*
* Frisian Islands Chain (chain 9B)
*
9B
53 12 12.388 N 007 06 00.590 E   85.7200   * Master : Finsterwolde
55 01 07.110 N 008 41 38.420 E  114.2930   * Red    : Hoyer
52 35 30.089 N 004 43 47.118 E  128.5800   * Green  : Heiloo
53 17 07.219 N 009 15 48.858 E   71.4333   * Purple : Zeven
ED50                                       * Geodetic Datum
0                                          * Min. lanecount - Red
0                                          * Min. lanecount - Green
0                                          * Min. Lanecount - Purple
299550.000                                 * Assumed signal prop. speed
*
* Holland Chain (chain 2E)
*
2E
51 36 36.629 N 004 55 36.521 E   84.5500   * Master : Gilze-Rijen
52 35 24.325 N 004 44 37.086 E  112.7333   * Red    : Heiloo
51 13 27.506 N 003 51 41.203 E  126.8250   * Green  : Sas v. Gent
52 11 11.350 N 001 35 46.272 E   70.4583   * Purple : Thorpeness
ED50                                       * Geodetic Datum
-3.920                                     * Min. lanecount - Red
3.205                                      * Min. lanecount - Green
0                                          * Min. lanecount - Purple
299550.000                                 * Assumed signal prop. speed
*
*
* End of Data
```

Simulation data file

Date	Time	GPS lat	GPS lon	Decca	Heading	Speed	Rudder
920917	083538	52 50 33.020 N	004 31 51.134 E	.862 .325 .109	232.4	12.7	-.2
920917	083539	52 50 32.781 N	004 31 50.672 E	.862 .337 .111	232.3	12.7	-.2
920917	083540	52 50 32.706 N	004 31 50.497 E	.866 .343 .115	232.2	12.7	-.2
920917	083541	52 50 32.584 N	004 31 50.125 E	.873 .347 .118	232.1	12.7	-.4
920917	083542	52 50 32.509 N	004 31 49.950 E	.877 .353 .118	231.8	12.6	-.5
920917	083543	52 50 32.359 N	004 31 49.663 E	.879 .357 .116	231.7	12.6	.2
920917	083544	52 50 32.269 N	004 31 49.489 E	.872 .361 .112	231.8	12.8	1.2
920917	083545	52 50 32.223 N	004 31 49.289 E	.872 .368 .107	232.3	12.9	1.5
920917	083546	52 50 32.073 N	004 31 48.973 E	.870 .372 .103	232.9	13.0	-.1
920917	083547	52 50 31.997 N	004 31 48.798 E	.869 .376 .104	233.0	13.0	-2.4

etc.

A8.5 Output screen pages layout

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 100
________________________________________________________________________________

| Suspended... | NPP Datum : ED50              | LSE Datum : ED50              |
|              |     Lat   :  52°50'32.373" N  |     Lat   :  52°50'32.373" N  |
|              |     Lon   : 004°31'50.958" E  |     Lon   : 004°31'50.958" E  |
|              |     Hg    :    0.0 m          |     Hg    :    0.0 m          |
|              |                               |                               |
| GYRO  232.3° |     RMS   :     6.35m         |     RMS   :     6.35m         |
| LOG   12.7kts|     a / b :   74.87m   21.37m |     a / b :   74.87m   21.37m |
| RUDR      OFF|     Az    :  112.0°           |     Az    :  112.0°           |
| INCL      N/A|                               |                               |
| PRESS     OFF|                               |                               |

|     GPS       |    SINS     |    OMEGA    |   LORAN-C   |    DECCA       |
|               |             |             |             |                |
| X  3848703.06 |             |             |             | Chain 9B       |
| Y   304893.45 |     not     |     not     |     not     |                |
| Z  5059951.55 |  connected  |  connected  |  connected  | Red     B20.9 *|
| Lat 52°50.50'N|             |             |             | Green   F34.3 *|
| Lon004°31.77'E|             |             |             | Purple  A51.1  |
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 611
________________________________________________________________________________

| Suspended... | GPS statistical data :                                        |
|              |                                                               |
|              | datum        lat                lon               Hg          |
|              | WGS84        52°50'30.000" N 004°31'46.200" E   39.5          |
|              | ED50         52°50'32.777" N 004°31'51.052" E   0.00          |
| Time 08:35:39Z|                                                              |
| Systems: G...D|       obs            calc                                    |
| LOPs   :   4 | x :   -31613.554    -31615.400                                |
| Filter : N/A | y :  5857312.625   5857300.165                                |
|              |                                                               |
| CM 005°00.00'E| a priori VCV matrix :  2500.000      0.000                   |
|              |                            0.000   2500.000                    |
| NPP          |                                                               |
| Lat 52°50.54'N|                                                              |
| Lon004°31.85'E|                                                              |
| a     74.87 m|                                                               |
| b     21.37 m|                                                               |
| Az   112.0°  |                                                               |
| RMS    6.35 m|                                                               |
```

Quit Edit Setup Obs Kalman Statistics Demo Data Run Pg 615

Suspended...

Time 08:35:39Z
Systems: G...D
LOPs : 4
Filter : N/A

CM 005°00.00'E

MPP
Lat 52°50.54'N
Lon004°31.85'E
a 74.87 m
b 21.37 m
Az 112.0°
RMS 6.35 m

Decca statistical data :

Chain : 9B

	Obs	C-O	Calc	LEF	Az
Red :	44.86	-0.05	44.81	3.98609	61.3
Green :	274.34	-0.05	274.29	1.58335	295.0
Purple :	1.11	-0.05	1.01	33.45517	257.6

a priori VCV matrix (SD) :

916.877	18.050	0.000
18.050	83.585	0.000
0.000	0.000	1.000

Quit Edit Setup Obs Kalman Statistics Demo Data Run Pg 631

Suspended...

Time 08:35:39Z
Systems: G...D
LOPs : 4
Filter : N/A

CM 005°00.00'E

MPP
Lat 52°50.54'N
Lon004°31.85'E
a 74.87 m
b 21.37 m
Az 112.0°
RMS 6.35 m

Syst	Pat	Read	C-O	Calc	v	Az	Sel
GPS	Lat	52°50.55'N	0.00	52°50.54'N	-12.5	90.0	*
GPS	Lon	004°31.85'E	0.00	004°31.85'E	-1.8	180.0	*
Decca	Red	44.86	-0.05	44.81	1.6	61.3	*
Decca	Grn	274.34	-0.05	274.29	-0.3	295.0	*
Decca	Prp	1.11	-0.05	1.01		257.6	

Datum : ED50 Proj.: TM CM : 005°00'00.000" E

MPP : 52°50'32.373" N 004°31'50.958" E
a/b/c : 74.87m 21.37m 0.00m
Az/elev. : 112.0° 0.0°

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 633

Suspended...          v        σd        w        τ     sel

                   -12.5     52.09     0.26     1.04    *
                    -1.8     60.90     0.04     1.22    *
                     1.6     50.89    -0.09     1.68    *
Time 08:35:39Z      -0.3     44.25     0.17     4.84    *
Systems: G...D       0.0    999.99   999.99   999.99
LOPs   :   4
Filter : N/A       S² :  0.034
                   σ² :  0.021    (4)
CM 005°00.00'E

MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run            Pg 641

Suspended...     Datum      : ED50
                 Projection : TM
                 CM         :  005°00'00.000" E

                              Lat                Lon                  x           y
Time 08:35:39Z   LSE :  52°50'32.373" N 004°31'50.958" E  -31615.40  5857300.16
Systems: G...D   KF  :  52°49'59.742" N 004°29'59.447" E  -33709.65  5856305.50
LOPs   :   4     MPP :  52°50'32.373" N 004°31'50.958" E  -31615.40  5857300.16
Filter : N/A
                 a posteriori VCV matrix of MPP :
CM 005°00.00'E
                    814.936  -298.408
MPP                -298.408   196.755
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m   Standard  :     30.59     8.73     0.00
b      21.37 m   95.0%     :     74.87    21.37     0.00
Az   112.0°      Azimuth   :     112.0
RMS     6.35 m   Elevation :       0.0
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run              Pg 810

Suspended...      Datum  : ED50
                  Chain  : 9B
                  Master :  53°12'12.388" N 007°06'00.590" E
                  Red    :  55°01'07.110" N 008°41'38.420" E
                  Green  :  52°35'30.089" N 004°43'47.118" E
Time 08:35:39Z    Purple :  53°17'07.219" N 009°15'48.858" E
Systems: G...D
LOPs   :   4        44.862  274.337    1.111
Filter : N/A
                   -31615.400 5857300.165
CM 005°00.00'E     140333.462 5899445.769
                   236240.537 6105681.577
MPP                -18315.065 5829339.403
Lat 52°50.54'N     284292.608 5914986.604
Lon004°31.85'E
a      74.87 m       177026.82               1.32392
b      21.37 m       365220.90   227344.07   0.81675   3.98609   1.07034
Az   112.0°           30962.69   173435.71   2.69107   1.58335   5.14909
RMS     6.35 m       321036.13   144712.47   1.38371  33.45517   4.49540
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run              Pg 821

Suspended...        0.00000   1.00000          12.45880
                    1.00000   0.00000           1.84520
                   -0.47409   0.88046          -1.60637
                   -0.42885  -0.90336           0.28186
                    0.00000   0.00000           0.00000
Time 08:35:39Z
Systems: G...D     2500.000     0.000     0.000     0.000     0.000
LOPs   :   4          0.000  2500.000     0.000     0.000     0.000
Filter : N/A          0.000     0.000   916.877    18.050     0.000
                      0.000     0.000    18.050    83.585     0.000
CM 005°00.00'E        0.000     0.000     0.000     0.000     1.000

MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run              Pg 830

Suspended...      0.00040  0.00000  0.00000  0.00000  0.00000
                  0.00000  0.00040  0.00000  0.00000  0.00000
                  0.00000  0.00000  0.00110 -0.00024  0.00000
                  0.00000  0.00000 -0.00024  0.01202  0.00000
                  0.00000  0.00000  0.00000  0.00000  0.00000
Time 08:35:39Z
Systems: G...D    0.00000  0.00000  0.00000  0.00000  0.00000
LOPs    :   4     0.00000  0.00000  0.00000  0.00000  0.00000
Filter : N/A      0.00000  0.00000  0.00000  0.00000  0.00000
                  0.00000  0.00000  0.00000  0.00000  0.00000
CM 005°00.00'E    0.00000  0.00000  0.00000  0.00000  0.00000

MPP
Lat 52°50.54'N
Lon004°31.85'E
a      74.87 m
b      21.37 m
Az   112.0°
RMS     6.35 m
```

```
Quit  Edit  Setup  Obs  Kalman  Statistics  Demo  Data  Run              Pg 840

Suspended...      LSE :  935.48      0.92721    0.37455
                          76.21     -0.37455    0.92721

                         112.00     0.00
                           0.00     0.00   -22.00
Time 08:35:39Z
Systems: G...D    KF  :
LOPs    :   4
Filter : N/A

CM 005°00.00'E

MPP               MPP :  935.48      0.92721    0.37455
Lat 52°50.54'N            76.21     -0.37455    0.92721
Lon004°31.85'E
a      74.87 m           112.00     0.00
b      21.37 m             0.00     0.00   -22.00
Az   112.0°
RMS     6.35 m
```