Ambiguity Resolution of Single Frequency
GPS Measurements

by

Michael John Tandy

Thesis submitted to the University of Warwick for the degree of

Doctor of Philosophy in Engineering

WMG

April 2011

THE UNIVERSITY OF
WARWICK

Abstract

This thesis considers the design of an autonomous ride-on lawnmower, with particular attention paid to the problem of single frequency *Global Navigation Satellite System (GNSS)* ambiguity resolution.

An overall design is proposed for the modification of an existing ride-on lawnmower for autonomous operation. Ways of sensing obstacles and the vehicle's position are compared. The system's computer-to-vehicle interface, software architecture, path planning and control algorithms are all described. An overview of satellite navigation systems is presented, and it is shown that existing high precision single frequency *GNSS* receivers often require time-consuming initialisation periods to perform ambiguity resolution.

The impact of prior knowledge of the topography is analysed. A new algorithm is proposed, to deal with the situation where different areas of the map have been mapped at different levels of precision. Stationary and kinematic tests with real-world data demonstrate that when the map is sufficiently precise, substantial improvements in initialisation time are possible. Another algorithm is proposed, using a noise-detecting acceptance test taking data from multiple receivers on the same vehicle (a *GNSS compass* configuration). This allows a more demanding threshold to be used when noise levels are high, and a less demanding threshold to be used at other times. Tests of this algorithm reveal only slight performance improvements. A final algorithm is proposed, using Monte Carlo simulation to account for time-correlated noise during ambiguity resolution. The method allows a fixed failure rate configuration with variable time, meaning no ambiguities are left floating. Substantial improvements in initialisation time are demonstrated.

The overall performance of the integrated system is summarised, conclusions are drawn, further work is proposed, and limitations of the techniques and tests performed are identified.

Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my supervisor, Professor Ken Young, for his guidance and patience throughout the preparation of this thesis; and Dr Emma Rushforth and my examiners for their feedback on this thesis. I thank the administrative and technical staff of WMG and the School of Engineering for their help with and tolerance of my many manufacturing, purchase order, and other support requests; and for their company. I thank my office mates and fellow doctoral students for their suggestions and discussions, as well as their friendship.

Thanks also go to my parents, my sister, my past housemates, and to my friends both inside and outside the university, without whose support, understanding and companionship I could not have enjoyed completing this thesis as much as I did. The undergraduates I have taught also deserve a mention, for provided interesting problems, incisive questions, and trips to robot competitions. Thanks, all of you!

Declaration

Results presented in chapter 6 have been published in GPS Solutions as Tandy and Young (2012)

I declare that the material contained in this thesis is original and my own work except where otherwise indicated, cited, or commonly known. The material in this thesis is submitted to the University of Warwick for the degree of Doctor of Philosophy, and has not been submitted to any other university or for any other degree.

Acronyms

ADOP Ambiguity Dilution Of Precision

AGV Autonomous Guided Vehicle

ARP Antenna Reference Point

AS Anti-Spoofing

C-LAMBDA Constrained LAMBDA

C\A Coarse Acquisition

CDGPS Carrier Differential GPS

CDMA Code Division Multiple Access

CMU Carnegie Mellon University

COG Course Over Ground

CSIRO Commonwealth Scientific and Industrial Research Organisation

DARPA Defence Advanced Research Projects Administration

DGPS Differential GPS

DLL Delay Locked Loop

DOD Department Of Defence

DOP Dilution Of Precision

E911 Enhanced 911

EBE Epoch By Epoch

ECEF Earth Centred Earth Fixed

EKF Extended Kalman Filter

ENU East North Up

EPSRC Engineering and Physical Sciences Research Council

ETRS89 European Terrestial Reference System 1989

GDOP Geometric Dilution Of Precision

GNSS Global Navigation Satellite System

GPS Global Positioning System

HDOP Horizontal Dilution Of Precision

IA Integer Aperture

IAB Integer Aperture Bootstrapped (estimator)

ICAO International Civil Aviation Organization

IF Intermediate Frequency

IGS International GNSS Service

ILS Integer Least Squares

IMRC Innovative Manufacturing Research Centre

IMU Inertial Measuring Unit

INS Inertial Navigation System

ION Institute Of Navigation

IR Infra-Red

L1 1575.42 MHz GPS signal

L2 1227.60 MHz GPS signal

L2C Modernised 1227.60 MHz 'Consumer' GPS signal

LAMBDA Least squares AMBugiuty Decorrelation Adjustment

LIDAR LIght Direction And Ranging

LLA Longitude, Latitude, Altitude

LSAST Least Squares Ambiguity Search Technique

LTP Local Tangential Plane

MC-LAMBDA Multivariate Constrained LAMBDA

MLAMBDA Modified LAMBDA

MOSFET Metal-Oxide-Semiconductor Field-Effect Transistor

MSRP Manufacturer's Suggested Retail Price

NCO Numerically Controlled Oscillator

NED North East Down

OIA Optimal Integer Aperture (estimator)

OS Ordnance Survey

OTF On The Fly

OXTS Oxford Technical Solutions

PCB Printed Circuit Board

PCO Phase Centre Offset

PCV Phase Centre Variance

PDOP Position Dilution Of Precision

PPM Parts per million

PRN Pseudo-Random Number

P(y) Precise

RFID Radio Frequency ID

RMS Root Mean Squared

RTK Real Time Kinematic

RTK GPS Real Time Kinematic GPS

SiGe Silicon-Germanium

SIL Safety Integrity Level

SLAM Simultaneous Localization And Mapping

SV Space Vehicle

TCXO Temperature Controlled Crystal Oscillator

TDOA Time Difference Of Arrival

TDOP Time Dilution Of Precision

TEC Total Electron Content

TOF Time Of Flight

TTAR Time To Ambiguity Resolution

TTFF Time To First Fix

US United States

USB Universal Serial Bus

VDOP Vertical Dilution Of Precision

WGS84 World Geodetic System 1984

XO Crystal Oscillator

Chapter 1

Introduction

Centimetre-precision Global Positioning System (GPS) receivers - so called Real Time Kinematic GPS (RTK GPS) receivers - are used in applications such as mobile robot control (Low and Wang, 2008) and agricultural automation, where the high cost of receivers is a key influence on economic feasibility (Pedersen et al., 2006). Researchers have investigated using low-cost consumer GPS receivers to achieve similar precision, with variable results (Takasu and Yasuda, 2008) - such as requiring, under good conditions, a stationary initialisation taking on average 10 (and sometimes as much as 60) minutes.

In this contribution the causes of this variability are investigated; and several techniques to reduce the initialisation time are proposed and experimentally evaluated.

1.1   Motivating Need

Lawnmower manufacturer Ransomes Jacobsen expressed an interest in adapting of one of their lawnmowers for autonomous operation. In summer, mowing the grass of a large golf course can be a full time job - especially if decorative patterns are mowed into the grass. By saving greenskeepers' time, a team of the same size can maintain grounds to a higher standard.

A Ransomes Jacobsen E-Plex II ride-on greens mower was provided, shown in figure 1.1 on the following page. The E-Plex II is an all-electric greens mower in an inverse tricycle configuration. It has a 1.6m cut width and a top mowing speed of 6.4 km/h (4 mph).

The specification was as follows:

1. An autonomous vehicle based on a Ransomes Jacobsen E-Plex II mower.

Figure 1.1: Unmodified E-Plex II mower, parked in an engineering workshop at the University of Warwick.
Courtesy of Dr Sadiq Jaffer

2. Modifications adding less than £10,000 to the mower's £20,000 base price.

3. Capable of mowing lines accurate to 5cm, reported by Roth and Singh (2004) as required to mow decorative patterns.

4. Able to detect and avoid such obstacles as may reasonably be expected on a golf course.

1.2   Existing autonomous mowers

Robotic mowing systems have been developed in the past. Smaller systems are aimed at home users and small grass areas – such as the Belrobotics Bigmow, the Husqvarna Automower, and the Friendly Robotics Robomow. These systems require a buried wire around the area to be mown. Routes are random or dictated by dead reckoning, and obstacles are detected with bumpers and/or ultrasound. Weights vary from 10kg for the Automower to 48kg for the Bigmow.

More sophisticated systems exist in various forms – More sophisticated small robots compete in the Institute Of Navigation (ION) Autonomous Lawnmower Competition;

and larger autonomous mowers have been developed, such as McMurtry's MAS mowers, the HortiBot, CMU's autonomous mower, and CSIRO's Autonomous Tractor Project.

McMurty produces two mowers, both aimed at "the regular maintenance of sports grounds, pitches and stadia". The MAS Mower 01 uses a rotating laser to detect retroreflective passive optical beacons situated around the area to be mown. This requires 6 to 10 targets, which are 20cm wide and 1m high. Obstacle detection is accomplished with ultrasound and sensitive bumpers. The MAS Mower 05 is a larger model, controlled with Real Time Kinematic GPS (RTK GPS).

The HortiBot (Jørgensen et al., 2006) is based on a Spider ILD01 slope mower, modified to have four independently steerable wheels (as supplied it uses four synchronously motorised and steered wheels). The main sensor used is a vision module from Eco-Dan A/S, Denmark, a stereo vision system which captures colour and 3D information from horticultural and agricultural scenes. In its modified form, the mower is intended to be used as a tool carrier for plant nursing, and the vision system identifies plant rows and can precisely drive between them.

Unlike the HortiBot, Carnegie Mellon University (CMU)'s autonomous mower (Roth and Singh, 2004) retains the goal of mowing grass. It uses a scanned LIght Direction And Ranging (LIDAR) and colour stereo vision for obstacle detection, and an RTK GPS integrated with an Inertial Measuring Unit (IMU) using Kalman filtering, to determine its own position.

Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO) has also developed an autonomous mower and golf ball collector (Dunbabin et al., 2004). Sensors employed include LIDAR, omnidirectional vision, a 3-axis compass, and RTK GPS. The system has demonstrated following preplanned paths using pure pursuit control, relying only on RTK GPS and compass sensing. The same base has been used to demonstrate navigation combining LIDAR and odometry; landmark-based visual homing has also been investigated, using road cones as landmarks (Vatani et al., 2006).

Mower	Localisation sensors	Obstacle sensors
CMU	RTK GPS& Inertial navigation	LIDAR and colour stereo vision
CSIRO	Vision – recognises road cone landmarks. GPS & Inertial navigation also investigated.	Omnidirectional vision
HortiBot	Vision – recognises plant rows.	Unknown
McMurtry MAS Mower 01	Laser identification of retroreflective landmarks.	Ultrasound and bumpers
McMurtry MAS Mower 05	RTK GPS	Ultrasound and bumpers
Various smaller robots	Random movement within buried-wire boundary.	Bumpers, usually ultrasound.
Humans	Visual identification of landmarks and cut/uncut grass boundary	Vision

Table 1.1: Sensors used by existing autonomous mowers

1.3 Document Outline

This chapter provides an introduction to the work performed, and gives an overview of the integrated mower control system, to give context to the rest of the report.

In chapter 2 modifications made to the mower are summarised, along with the design decisions behind them. This provides context for later chapters, which deal with localisation using single frequency GPS.

Chapter 3 gives an overview of the Global Positioning System (GPS), including the need for ambiguity resolution, algorithms used to achieve it, and reported performance figures.

In chapter 4 an algorithm is presented to integrate topographical map data into the ambiguity resolution process. Experimental results are presented.

Chapter 5 discusses vehicles with multiple roving receivers (such as for heading or attitude detection) and describes algorithms from the literature that speed up ambiguity resolution. A method is proposed for detecting multipath noise from the difference between receivers. Experimental results are presented showing the ambiguity resolution time reductions this can offer.

Chapter 6 gives details of acceptance tests and attempts to estimate and control am-

biguity resolution failure rates found in the literature. A method is proposed for estimating the impact of multipath noise in fixed-failure-rate ambiguity resolution. Experimental results are presented.

Chapter 2

Design for Autonomous Operation

The E-Plex II is an all-electric greens mower in an inverse tricycle configuration. It has a 1.6m cut width and a top mowing speed of 6.4 km/h (4 mph). The mower's steering uses a system designed for fork lift trucks; the rear wheel is turned by a geared DC motor, with the motor's drive voltage proportional to the output of a tachogenerator connected to the mower's steering wheel. This design doesn't provide a direct mapping between steering wheel angle and rear wheel angle, but for manual operation this isn't a problem, just like a driver need not look at the steering wheel to steer a car and for the same reasons. The steering controller periodically checks the tachogenerator is present by sending high-frequency impulses and measuring its impulse response; any problems and the mower is brought to a halt.

Propulsion is controlled by a system designed for golf buggies. A 2.6 kW (3.5 Hp) 48 volt DC motor, powered by eight 6 volt lead acid batteries, drives the two front wheels by means of a 15:1 differential. The reels, reel retraction, and steering are also motorised. Manual throttle control uses a rocking foot pedal; rock forward to go forwards, backwards to reverse. A sign-magnitude signal is sent to the drive motor controller, which uses feedback control to maintain a proportional speed.

2.1   Design overview

Based on the vehicle specification, the design was divided into several subsections:

- For autonomous operation:

  - A means for a computer to control the mower's steering and throttle.

- For mowing lines accurate to 5cm:

Figure 2.1: E-Plex II mower fitted with GPS, LIDAR and control electronics.

  – A means of localising the mower relative to the desired path.

  – A path-following algorithm, which generates steering and throttle commands to keep the mower on the desired path.

• For avoiding obstacles:

  – A means of detecting permanent obstacles, such as trees, lakes and bunkers.

  – A means of detecting temporary or moving obstacles, such as people and golf equipment.

  – An algorithm to respond to detected obstacles and avoid them.

The following sections will first discuss hardware for steering and throttle control; then hardware for localisation and for obstacle detection (discussed together as some sensors can fulfil both requirements). After this software for path following and obstacle avoidance will be discussed.

2.2 Computer to vehicle interface

A control board was designed to allow for computer control of the mower. The sign-magnitude throttle pedal was emulated with two MOSFET outputs providing the sign signals, and a buffered 0-4.5v analog voltage output providing the magnitude signal. Emulating the steering tachogenerator was more complicated due to the periodic test pulses. The standard tachogenerator was retained, with the control board outputting a voltage in series with it. This voltage output was optoisolated, with a range from -4v to +4v.

A microprocessor was used to control the two analog and two digital outputs, with a USB to serial converter for interfacing to the controlling computer. The microprocessor also served as a watchdog, bringing the vehicle to a halt if data stopped arriving from the computer. Other circuits on the control board allowed reading back the mower's battery voltage, and detecting whether the emergency stop had been triggered.

Schematics for the control board can be found in appendix E on page 164.

Two multipole key switches were fitted to the mower for switching between computer and manual control - one switch for the throttle, one for the steering. This made it possible to test computerised steering control with a human still in the driving seat, operating the throttle. Emergency stop buttons controlled a safety relay between the control board and the mower's electronics, to make the emergency stop circuit as simple and reliable as possible.

2.3 Sensor selection

2.3.1 LIDAR

LIght Direction And Ranging (LIDAR) is a form of 'laser radar', where a time-of-flight laser range finder shines at a spinning mirror, the output being distance measurements at a variety of angles. This technology is common on the most advanced autonomous vehicles as it can offer fast, precise measurements - at a price. One popular sensor, the
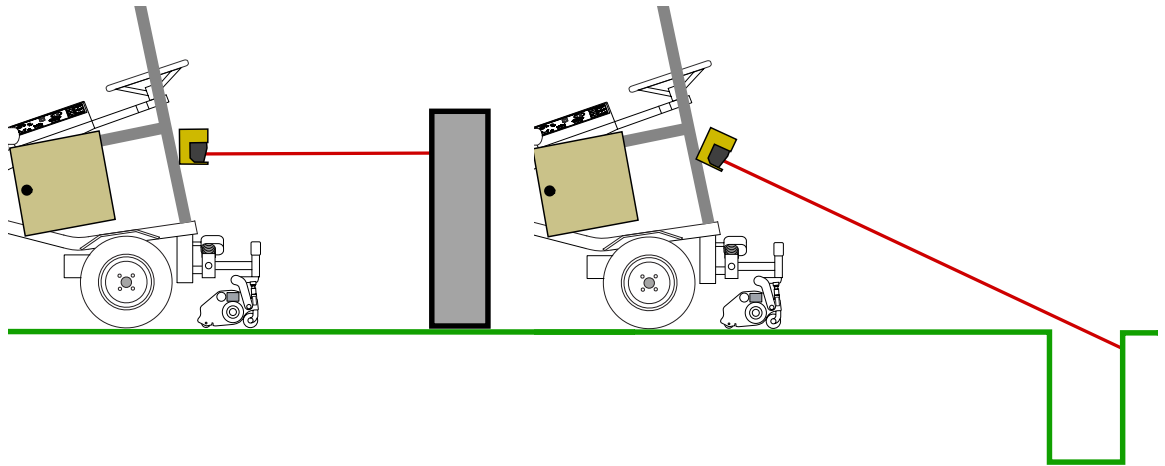
Figure 2.2: A LIDAR detecting a positive obstacle such as a person (left) and a negative obstacle such as a ditch (right)

SICK LMS 291, has a mirror spinning 75 times a second, taking 180 measurements spaced one degree apart, with a precision of a few centimetres and a maximum range of 50m(SICK, 2006). This can produce a precise profile of an obstacle, which is useful to steer around it, and in unknown terrain these sensors are invaluable to detect negative obstacles like ditches and potholes, as shown in figure 2.2.

LIDAR can also be used for localisation, using Simultaneous Localization And Mapping (SLAM). This operates by using a LIDAR to gather information about the local area and matching that information against a map. By matching features like walls and corners, vehicle location can be calculated.

There are several areas where LIDAR is weak. Costing around £4000, the SICK LMS 291 isn't cheap, especially if several of them are needed – and with a 180 degree field of view at least two are needed if you want your robot to be able to reverse. Dark-coloured objects, glass and mirrors can be hard to sense correctly. LIDARs can be dazzled by direct sunlight if the sun is low in the sky (as happened to Insight Racing during the DARPA Grand Challenge (Buehler, 2006)) or can get false readings from clouds of dust (as happened to The MITRE Meteorites). To scan an accurate terrain profile while driving needs accurate pose estimation to compensate for movement between and during scans – and pose estimation errors can create phantom obstacles (as happened to Team Cornell).
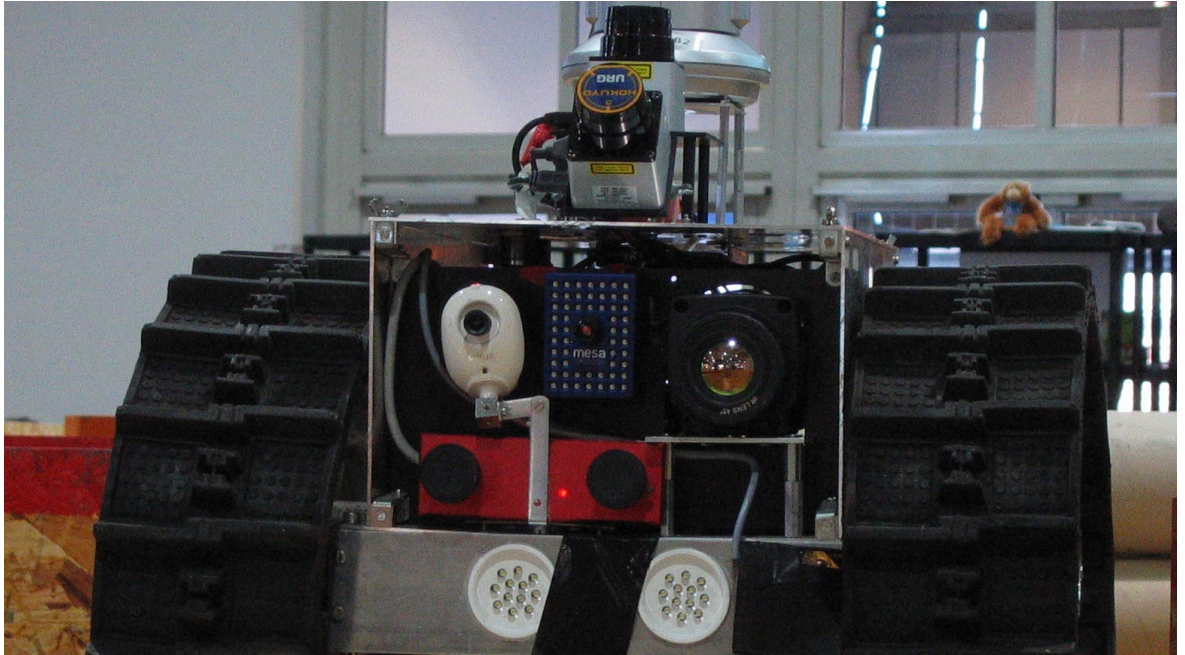
Figure 2.3: A robot fitted with several sensors, including a MESA SwissRanger Time Of Flight (TOF) camera

For SLAM navigation there have to be recognisable features within the LIDAR's range; for many golf courses, this needs a LIDAR with longer range than the LMS 291 offers. Longer range LIDAR is available, but long range LIDARs Riegl Q-240i, cost as much as £48,000(Johnson and Danis, 2006) – clearly out of our price range.

As a LIDAR was available from a previous project we fitted it to the mower. Whether such a sensor would belong on a commercial product depends on aims and future developments. A LIDAR would simplify mowing near obstacles like trees, and diverting around unexpected obstacles, but the benefits might not justify the costs at current prices. Some LIDAR manufacturers estimate that with mass production costs could drop to as little as €200 (Ax, 2006); this will change the economics substantially!

2.3.2   Time Of Flight (TOF) camera

Time Of Flight (TOF) cameras, like LIDARs, operate by generating a pulse of light and measuring the time for its reflection to return to the camera. But while LIDARs use a laser to illuminate a small point and measure a single reflection, TOF cameras illuminate an entire scene and use an array of sensors to detect the reflection. This

means tens of thousands of measurements may be taken at the same time and no moving parts are needed.

Illuminating an entire scene instead of a small point requires higher power levels, especially if the Infra-Red (IR) illumination is to be seen in sunlight. While sensors designed for indoor use, such as the MESA SwissRanger (Shown on the robot in figure 2.3 on the previous page), can be brought for as little as £2700, sensors that work in direct sunlight are more expensive - some costing as much as £10,000. Like with LIDAR this price is expected to drop as technology advances.

2.3.3 Structured light

Structured light is the operating principle of Microsoft's Kinect sensor; the scene is illuminated with a pattern of Infra-Red (IR) dots, and an IR camera offset from the light source works out the shape of objects from the way they distort the standard pattern. The distortion gets smaller as the object gets further from the sensor leading to increased measurement error. In contrast, time of flight sensors' distance measurement error is stable with increasing distance. This weakness isn't always a problem though; humans' depth perception has the same error profile but we live with it.

Unfortunately the Kinect does not work outdoors, as its source of IR is far less powerful than the sun, and in direct sunlight it quickly gets washed out. This situation may improve in the future, in which case it would be worth evaluating this technology again.

2.3.4 Ultrasound

Ultrasound is a type of sensing using sound waves too high-pitched for humans to hear, and measuring the time of flight for an echo. As sound travels at around one millionth the speed of light, ultrasound sensors require less complicated electronics than LIDAR so they can be a lot cheaper – less than £100, some short range sensors even less than £10. Lower complexity may also translate into better reliability.

As beams of sound spread out more than beams of light, it's difficult to measure the profile of an obstacle to identify or precisely steer around it. This imprecision makes ultrasound poor at spotting negative obstacles. Sound absorbent materials which produce poor echoes can be hard to detect at longer ranges. Industrial ultrasound sensors were fitted to the front of the mower; they performed well enough to stop the mower when a reasonable size obstacle was in front of it. Until LIDAR costs reduce, ultrasound is an attractive option for obstacle detection.

One means of localising robots indoors is using multiple ultrasound beacons. As ultrasound signals travel one 880,000th of the speed of light signals, measurement equipment is simpler for the same level of precision. McCarthy and Muller (2005) simulate an indoor environment, and estimated "95% accuracy of 20 cm, and a 50% accuracy of 6 cm"; while Kim et al. (2002) reports 3cm precision within a few meters of range.

Though beacons have to be installed around the area to be mown, the beacons do not need to be easily visible. However, most users of ultrasound beacons are indoors and working at a range of a few metres; operating at a range of 200m and in windy conditions would be more demanding; sensing with a square wave of wavelength 200 metres would only allow a 1.7Hz update rate, so a more sophisticated signalling system may be required. No off-the-shelf ultrasound beacon systems with the range and outdoor operation we require could be identified.

2.3.5 Machine vision

Machine vision is a means of sensing using video cameras monitored by a computer. This monitoring can range from tasks as well known as barcode scanning to sophisticated systems which try to achieve the complex 3D object recognition and tracking humans excel at. One of the strengths of machine vision is identifying objects that don't stand out by their shape. For example, LIDAR would have difficulty spotting a golf ball in golf-ball-length grass, but machine vision can pick it out by colour. The range can be as far as the eye can see, and basic cameras can be purchased for less
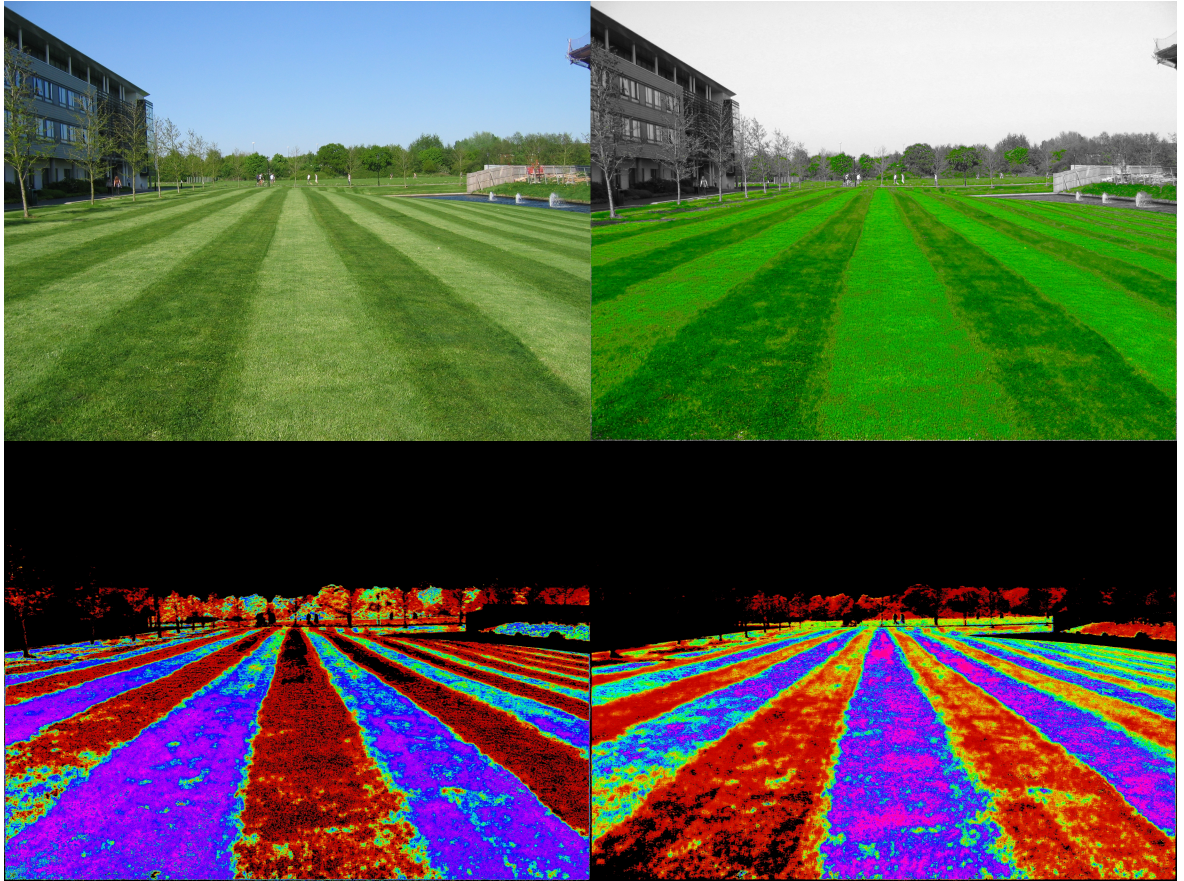
Figure 2.4: Samples of machine vision of grass. Top left: Source image. Top right: Grass identified using a 4D Bayesian filter, higher saturation indicates higher confidence of classification. Bottom left: Identification of dark grass stripes. Bottom right: Identification of light grass stripes.

than £500.

On the other hand, even if white golf balls are easy to spot, green things can be camouflaged on grass; things can look different in the amber light of dusk; and to operate in darkness extra illumination is required.

Some investigation of machine vision was performed which indicated good results from reasonably-priced equipment. Figure 2.4 shows the results of several experiments; telling grass apart from things that aren't green is fairly easy, but telling it apart from other green things and telling cut and uncut grass apart is less reliable. There was not time to perform robust series of tests in different natural lighting conditions, or a study of the type of debris found on golf courses.

Localisation using machine vision is also an option; human operators are able to see the

boundary between cut and uncut grass and control a mower along it. Accomplishing the same task with computer vision is difficult because of the obvious similarities between cut and uncut grass. Figure 2.4 on the previous page shows some examples; stripes can be identified under ideal conditions, but with quite a bit of noise. As light intensity, shadows, and suchlike cannot be controlled, it would be difficult to devise a system which is robust in dealing with changes in environmental conditions; substantially more difficult than detecting obstacles by colour.

If this vision system could be implemented reliably, it would be reasonably inexpensive. This method of navigation would be limited to certain grass patterns, such as stripes and concentric patterns, as positioning would be relative to a previous stripe. If the grass edge being followed goes out of view (such as when turning at the end of a stripe or when not mowing) alternative localisation would be needed until the edge is reacquired.

Another localisation technique humans use is visual recognition of landmarks; this will be discussed alongside other landmark recognition options, in section 2.3.10 on page 17.

Overall, using machine vision for obstacle detection would be challenging, while using it to follow cut/uncut grass boundaries would be extremely challenging. Depending on the results of field tests with ultrasound alone, machine vision could be a useful addition for detecting certain obstacles.

2.3.6 Sprung wire / bumper

Another option considered was running a spring-tensioned wire around the base of the mower (parallel to the ground and a few inches away from the floor) connected to an emergency stop switch. Obstacles the robot might hit would first be hit by the wire, pulling on it and triggering the emergency stop. Due to its simplicity, with no computer processing involved, this could be reliable as well as inexpensive. This was not adopted on our prototype vehicle as the LIDAR we made use of was sufficient for our Safety Integrity Level (SIL) 2 safety system, but it could be useful on a vehicle

without LIDAR.

Wire-based sensing is not such a good option for vehicles with long stopping distances, as the wire has to protrude further from the front of the vehicle, but with the mower's low speed this is not a major problem.

2.3.7   Dual frequency GPS

Dual frequency GPS receivers are used in surveying and other high precision applications. The operating principles are explained in more detail later in this document; a system may use a dual-frequency GPS and a subscription to a data service (£2,000 annually according to Leica (2012)), or two dual-frequency GPS receivers, one of which acts as a fixed base station.

A single surveyor's RTK GPS receiver can be hired for around £350 per week, including a data service license; a kit including two receivers can be purchased for around £15,000. Cheaper products may be available without the surveying accouterments, or second hand. Data rates are usually in the region of 2Hz.

Higher update rates are available from products like the RaceLogic VBOX III RTK, which is intended for automotive testing and offers 100Hz GPS. Costs are in the region of £24,000 including base station. GPS-only systems may lose signal when line-of-sight to satellites is lost, such as when under trees or in the shadow of tall buildings. Single GPS receivers also cannot determine a vehicle's orientation – though they can determine direction of motion, when the vehicle is moving, and with two roving receivers orientation can be determined.

As part of GPS modernisation (to be discussed in more detail later), the US government is considering phasing out semi-codeless access to L2 GPS by the end of 2020(OSC, 2008) to encourage use of the modernised L2C signal. As such current dual frequency GPS technology may only be useful for so long.

In summary, dual frequency GPS systems offer sufficient accuracy for our application, but can be quite expensive. Dual frequency GPS is also simple, in that an off-the-

shelf unit gives a precise location; as opposed to (for example) developing our own landmark recognition system, which would require a lot more programming, calibration and suchlike. Dual frequency GPS may lose signal when under large trees, in the shadow of tall buildings, or indoors.

2.3.8 Single frequency GPS

Dual frequency GPS receivers rely on several components, including: a nearby base station to offset for common errors, taking precise-but-inaccurate carrier phase measurements, and using the dual frequency signals to simplify ambiguity resolution which removes inaccuracy from the carrier phase measurements. Certain consumer receivers will also take and report these carrier phase measurements, but for a single frequency which complicates the process of ambiguity resolution. Receivers are available for around £250 (£60 in bulk) which report these measurements at a frequency of 10Hz.

Under good conditions, after ambiguity resolution these receivers can be as accurate as dual frequency receivers, but at a fraction of the cost. This was selected as worthy of further consideration; results and details are presented in later chapters.

2.3.9 RTK GPS/Inertial measurement combination

Several companies supply products integrating GPS, inertial measurement, and sometimes odometry, using Kalman filters. These products include the Trimble POS LV, and products from Oxford Technical Solutions (OXTS).

Odometry and inertial measurement allow these products to offer a higher data rate than GPS alone (100Hz and upwards); and allow continued function when GPS is unavailable (such as due to tunnels, trees, or tall buildings). However, when GPS is absent for some time, position errors can accumulate, leading to a jump in estimated position when a signal is reacquired.

Combined GPS/IMU positioning is particularly popular when using LIDAR to model terrain, as an accurate estimate of vehicle pose is required to compensate for move-

ment during and between LIDAR scans. In the DARPA Grand Challenge, all major competitors used GPS/IMU integration. Systems are even available with two GPS antennas so that, with the locations of both antennas known to within 2cm, vehicle orientation can be reliably determined. A medium-specification non-RTK GPS/IMU with 40cm advertised GPS precision costs around £22,000; a high-specification dual-frequency dual-antenna RTK model and RTK base station costs £50,000. A model with a single antenna and 20cm precision costs £33,000.

Integrated RTK GPS/IMU systems are something of a definitive standard in autonomous vehicle development. They are available as off-the-shelf units, albeit at high cost. It is uncertain how far the prices of these products will fall in the future; unlike LIDAR, the major manufacturers have not made optimistic predictions about prices. Prices may fall in the future nonetheless.

2.3.10 Landmark recognition

Landmark recognition can take various forms, depending on the landmarks and sensors being used. Landmarks can be naturally present or man-made. Sensors include lasers (both distance and angle-measuring) and machine vision.

Landmark recognition is obviously dependent on the distance from vehicle to landmark. Inspecting satellite photographs of The Warwickshire, a golf course located at 52.30947°N, 1.59392°W, it was seen that in the worst case features may be 100 to 200m away. Man-made landmarks can allow robust localisation as they can be easy to detect, and they can be specifically placed in ideal locations – for example, if the system has a maximum range of 50m, landmarks can be placed such that there is always one within that range. Disadvantages to man-made landmarks include physical appearance (road cone land marks would be incongruent with golf courses' atmosphere), the need for installation, and any maintenance required. Controllable landmarks, such as switchable LED beacons, could avoid appearance issues at the cost of system complexity.

Some autonomous mowers use a landmark-based system common in Autonomous Gui-

Figure 2.5: Artificial landmark for a factory Autonomous Guided Vehicle (AGV), comprised of a barcode of retroreflective strips.

ded Vehicle (AGV) applications. A laser reflected through a rotating mirror detects the angle (and sometimes range) to special reflectors. An example reflector is shown in figure 2.5. Laser sensors which do this include the SICK NAV200-1132 (28m range, 25mm distance accuracy, 0.1 degree angle accuracy, £5,500) and the Danaher Motion NDC8 Laser Scanner 4 (70m best-case range, 0.06 degree angle accuracy, no distance sensing). As these systems use a horizontally scanned laser, they require fairly even and level ground and targets at consistent heights, though targets can be fairly low key.

Another option is to use machine vision to identify landmarks such as trees. This would provide the angle to each landmark, with precision depending on the specifications of the vision system. To detect a vehicle's position with 5cm accuracy at a range of 200m would require an angular measurement accurate to:

$$\tan^{-1}\left(\frac{0.05}{200}\right) = 0.014^o \qquad (2.1)$$

A camera with a 360°field of view would therefore need a horizontal resolution of:

$$\frac{360^o}{0.014^o} \approx 25,000 \text{ pixels} \tag{2.2}$$

Reducing the desired specifications to 10cm accuracy at a range of 100m would reduce this to a horizontal resolution of 6,300 pixels. A commercial camera with a horizontal resolution of 5,760 pixels (the Canon EOS-5Ds Mark III) has an Manufacturer's Suggested Retail Price (MSRP) of £3,000 not including lens or optics, and not designed for the continuous high-speed operation vehicle operation would require. To achieve the full 25,000 pixels of horizontal resolution would require five such cameras. The Point Grey Ladybug2 camera, a ready-made omnidirectional video camera (3500 pixel horizontal resolution) costs around £5,000.

It may be possible to use a lower resolution camera aimed using a pan-tilt mount or rotating mirror, but aiming precise to a hundredth of a degree in a moving vehicle would be difficult.

2.3.11 Buried wire detection

Buried wires conducting a low voltage signal (frequency around 40kHz), can be detected by inductive coils when within about 30cm. Small robot lawnmowers use a 'buried wire fence' system, where the robot moves at random, detecting and staying inside the perimeter marked by the buried wire. Some AGVs used in factories follow under-floor wires, creating a 'railway' system.

It would be possible to combine these techniques, burying a wire to mark out the mower's entire path through the golf course. This would be inexpensive in hardware terms (the sensor costing less than £200) though there would be a cost to installing the buried wire, and to design new mower routes would require a new wire to be laid. Maintenance may be required due to environmental damage to the wire – although the problem may be manageable.

Satellite photos of our example golf course, The Warwickshire, show the total land area

is approximately 2 million square metres (465 acres); assuming 50% of this requires mowing with our 1.5 metre cut width mower, cutting the entire course would involve driving 667 kilometres (taking 111 hours at our 6km/h cutting speed). With the cheapest wire costing around £10 for 500m, it would cost £13,000 to wire the entire course for buried wire following.

This idea would be easy to implement with today's technology, and would be an inexpensive way to cover a smaller area of grass – 2 acres would require only £100 worth of wire. If covering an entire golf course, however, there is little cost benefit over dual frequency GPS.

2.3.12   Random movement

Small robot lawnmowers can use roomba-style random movement; with enough random mowing an entire lawn will eventually be covered. This eliminates the need for the mower to know its own location relative to the area that has been mown. This method is suitable in small gardens, where there is no penalty for covering the same area several times.

On a golf course, however, covering the same area twice halves the area one mower can maintain, increases maintenance time (potentially displacing paying customers) and prevents the creation of manicured grass patterns. And when an optimal course takes 100 hours to cover an entire golf course, a random course would take impractically long! This would also not satisfy our requirement for 5cm accuracy.

Random mowing could be practical for smaller areas, and buried wires could be used to mark especially dangerous borders like bunkers and water traps.

2.3.13   Fixed vantage point vision

By marking the mower with a distinctive pattern and fixing a video camera on a tall building, it would be possible to determine the vehicle's location with a reasonable

degree of accuracy, for small field sizes. 5cm precision location measurement for an object at the far end of a 100m field would require a camera resolution of a few thousand horizontal pixels, or a lower resolution camera and a precise pan-tilt mount.

However, like the landmark recognition described in section 2.3.10 on page 19, this would become more costly at the scale of a typical golf course, which would usually not have any high vantage points with line of sight across the entire course. A high resolution multi-camera system would rapidly increase prices, to the point of impracticality.

2.3.14 Buried RFID tags

An alternative to burying a wire would be to bury passive Radio Frequency ID (RFID) tags; depending on the technology used, such tags can have a range from a few centimetres to several metres. If 50% of The Warwickshire's 465 acres was covered with Radio Frequency ID (RFID) tags, arranged in a 5m pitch grid, 40,000 tags would be required; at a bulk cost of £0.03 per tag (Roberti, 2006) this would cost £1,200. Placing tags in a 2m pitch grid would increase this to £7,500. There would also be the cost of performing the installation; a machine would be needed to do this - although it would only be needed once per installation, so the mower manufacturer could make one and spread the cost over many installations. The costs of this are hard to estimate; they could be between £200 and £1000 per installation. A reader would cost in the region of £200.

This system would allow precise positioning, as scanning a given Radio Frequency ID (RFID) tag would yield an exact location. However, tags can only be scanned while they are in range; the mower would have to use dead reckoning or inertial measurement while moving from one tag to the next. As the system would have to get close to a tag to be able to find it at all, precise dead reckoning or inertial measurement would be needed. A tag reader would also need to provide rudimentary distance measurement if the mower is to know its position beyond 'within 30cm of location A' and off-the-shelf
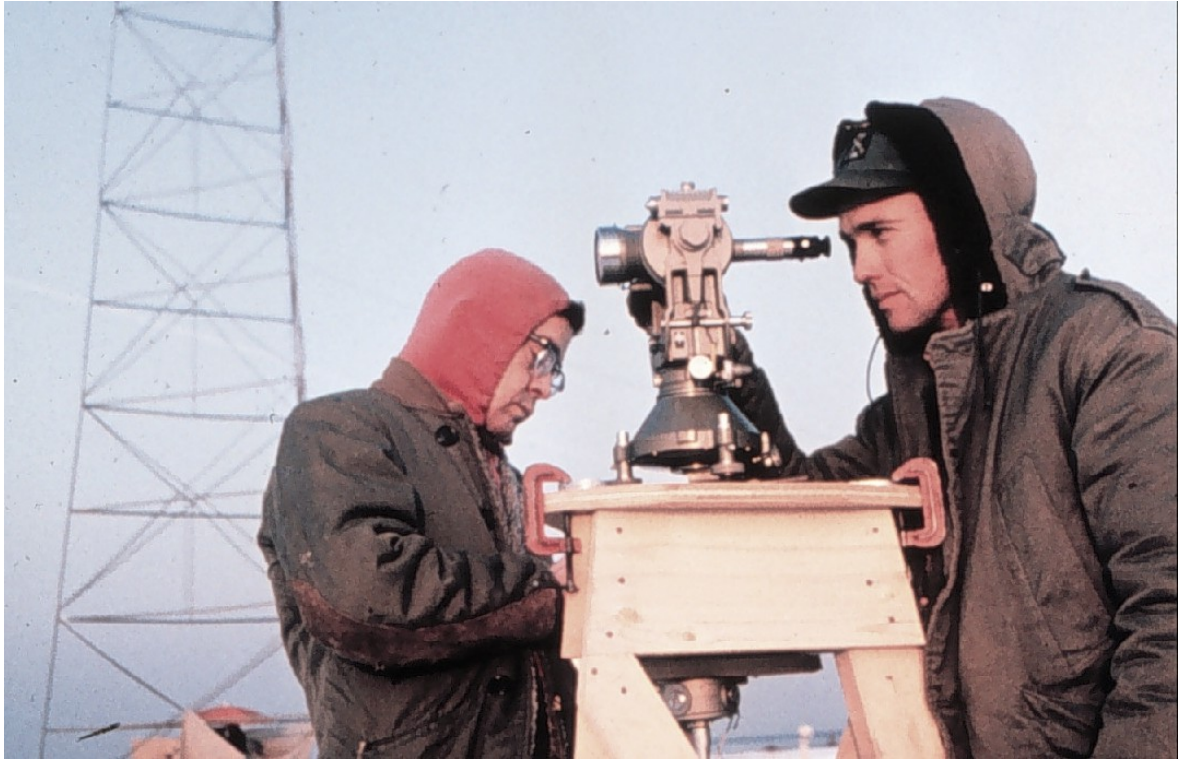
Figure 2.6: An Arctic field party takes observations with a manual theodolite.
Courtesy of National Oceanic and Atmospheric Administration; public domain license.

Radio Frequency ID (RFID) readers do not have this capability.

Overall this is an interesting idea, but would require either fairly precise dead reckoning or inertial measurement (resulting in high-cost hardware) or a high density of tags (resulting in high tag cost). For small installations it may be practical.

2.3.15 Theodolite-like sensing

The theodolite is a traditional surveying tool which combines a telescope with highly accurate angle sensing. To survey a point, the point is sighted through the telescope, and the horizontal and vertical angles to it read out. This is typically a two person task, with one person holding a target and the other operating the theodolite.

In modern times, surveyors commonly use 'total stations', theodolites with digital angle measurement, laser range finding, and data logging. The operator sights a target then presses a button and the angle and distance to the target are saved in memory for later downloading.

More sophisticated still is the 'robotic total station', a motorised total station with built in video camera and remote control. The surveyor sets up the system at a known point, then he/she walks around with the target; with the combination of camera and motors, the total station tracks the surveyor's target, and logs data when a remote control button is pressed. This allows precise point measurement by a single operator. A typical robotic total station measures angle to 0.001°and distance to 10mm at a range of 500m, at a cost around £13,000(Stakemill, 2008). This is out of our price range, but it may be practical to use a pan-tilt-zoom camera, a laser range finder, and precise angle measurement to achieve the same effects.

This could be used to measure a target on the robot from a stationary base point or to measuring targets in the environment from on board the robot. A fixed base station would require the mower to be within line of sight of a base station, while measuring from the robot would need precise, vibration compensated angle measurement. Measuring distance and angle for multiple targets would either mean stopping the vehicle; interleaving measurements and compensating for movement between them; or having multiple independent sensors. Pan-tilt-zoom cameras for landmark recognition have been seen before in the literature - such as in Li et al. (2002) which combines a landmark, a CAD model of the surrounding environment, and a pan/tilt/zoom camera and laser range finder to perform robot localisation.

A pan/tilt with 0.01°angle resolution and 4kg payload capacity can be purchased for £1,300(Directed Perception, 2008); a camera with sufficient zoom and resolution around £600(Source, 2008); and a laser range finder with basic computer connectivity, 3mm accuracy, and 150m range costs £400(Bosch, 2008). The total cost would be £2,300. However, the products priced are not designed to be mounted on vehicles; no suitably precise pan-tilt-zoom systems which are designed to be used on moving vehicles could be identified.

Overall, it may be practical to apply theodolite-like sensing for vehicle localisation at a cost of a few thousand pounds, for a vehicle-mounted camera but problems introduced by the vehicle's constant motion may be difficult to overcome. A fixed base station

would avoid this, but only work within line of sight.

2.3.16   Summary

Sensor	Summary
LIDAR	Too expensive for localisation, and fairly expensive for obstacle detection. Fitted to the test vehicle as we had spare sensors.
TOF camera	Sensors suitable for outdoor use out of our price range.
Structured light	No sensors suitable for outdoor use identified.
Ultrasound	First choice for obstacle detection due to reasonable prices. Insufficient range for localisation.
Machine vision	Second choice for obstacle detection, potentially good for detecting things ultrasound or LIDAR would miss. Difficult/complicated to use for localisation.
Sprung wire / bumper	Possible backup choice for obstacle detection; simple and reliable. Unsuitable for localisation.
Dual frequency GPS	Popular in agricultural automation, but out of our price range. Suitable for localisation only.
Single frequency GPS	Less popular than dual frequency GPS, but potentially lower cost. Fitted to the test vehicle. Suitable for localisation only.
RTK GPS with IMU	Popular in the highest performance autonomous vehicle applications, and where GPS coverage may be lost, but beyond our price range. Suitable for localisation only.
Landmark recognition	Suitable for smaller, controlled environments but costly for a golf course where landmarks aren't always close by.
Buried wire detection	Would need a lot of wire, at great cost, to cover an entire golf course. Could be used to guard around bunkers and lakes as a backup measure.
Random movement	Practical only for small areas. Not accurate to 5cm.
Fixed vantage point vision	Practical only for areas in view of a suitable vantage point, which is not the case for most golf courses. Limited ability to detect obstacles.
Buried RFID tags	Difficult/complicated to attain required accuracy. Unsuitable for obstacle detection.
Theodolite style sensing	Difficult/complicated to perform on a moving vehicle. Unsuitable for obstacle detection.

Table 2.1: Summary of obstacle and localisation sensing options.
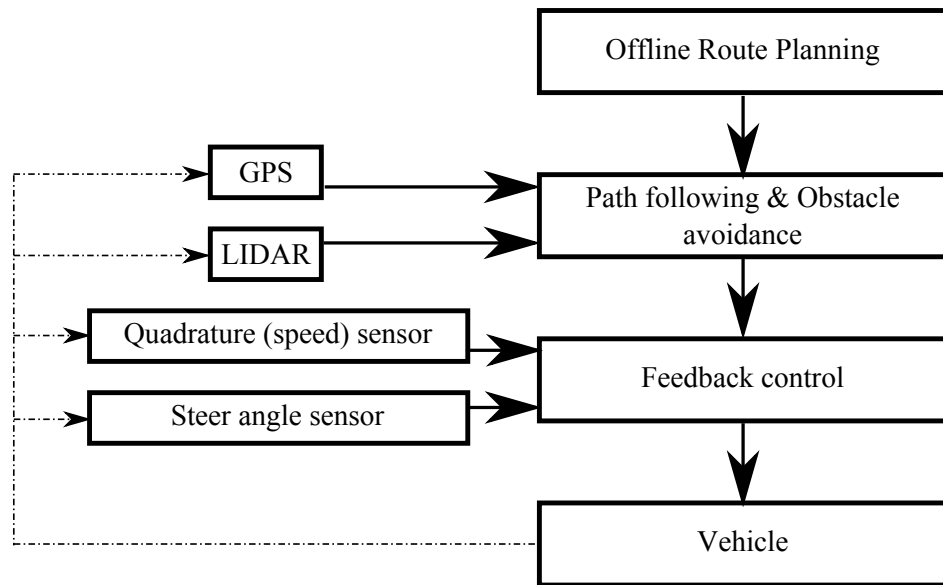
Figure 2.7: A temporal decomposition of the key subsystems used by the vehicle.

Table 2.1 on the previous page summarises the sensing options considered. For obstacle detection ultrasound sensors were fitted to the mower, as they offer sufficient performance at very reasonable prices. A LIDAR was also fitted to the mower as it is the logical next step in obstacle detection performance and we had one spare - but due to the costs LIDAR would not be recommended for the final system. For localisation a single frequency GPS system was selected as the best way to meet the cost and accuracy goals identified in section 1.1 on page 1.

One of the consumer single frequency GPS receivers identified offered some attractive features; a comparatively fast data rate of ten updates per second and a price so low it was practical to have a fixed base station and two receivers on the mower, allowing heading to be accurately determined. Reports on the performance of single frequency GPS systems are variable, sometimes reporting lengthy stationary initialisation periods; it was decided to develop a basic differential GPS system in order to investigate the performance of the single frequency GPS receiver identified, and to investigate some proposed ways to reduce the stationary initialisation period.

2.4 Software design

A temporal decomposition of the system is shown in figure 2.7 on the previous page. A planned set of waypoints is generated before the vehicle starts moving; a 10Hz control loop responds to obstacles detected by LIDAR and deviation from the planned route detected by GPS; and a 50Hz control loop actuates the vehicle's steering and throttle to achieve the commanded values.

The subsequent sections will describe these parts of the system in more detail, along with some alternatives that were considered.

2.4.1 Overall software structure

Figure 2.8 provides an overview of the main software written for the project. The computer at the fixed base station receives data from the attached GPS receiver and makes it available over the internet, using a web service. The computer on the vehicle receives this data over a 3G wireless broadband connection, and pairs it with data from the two GPSes on the vehicle. As the fixed base station is at a known location, errors common to the receivers can be cancelled out through double differencing; and because two receivers are fitted to the vehicle its heading can be determined even when stationary. The calculated vehicle position and heading are fed to control software which actuates the throttle and steering to follow the desired path.

The software was predominantly written in Java, with the separate modules run as separate threads and Java queues used for inter-thread communication. MATLAB was used in the ambiguity resolution process, as several standard algorithms are already available in MATLAB format.

2.4.2 Offline route planning

Before the vehicle starts moving, a planned set of waypoints is generated setting out the path the vehicle should take, represented as a list of latitude/longitude pairs. There are
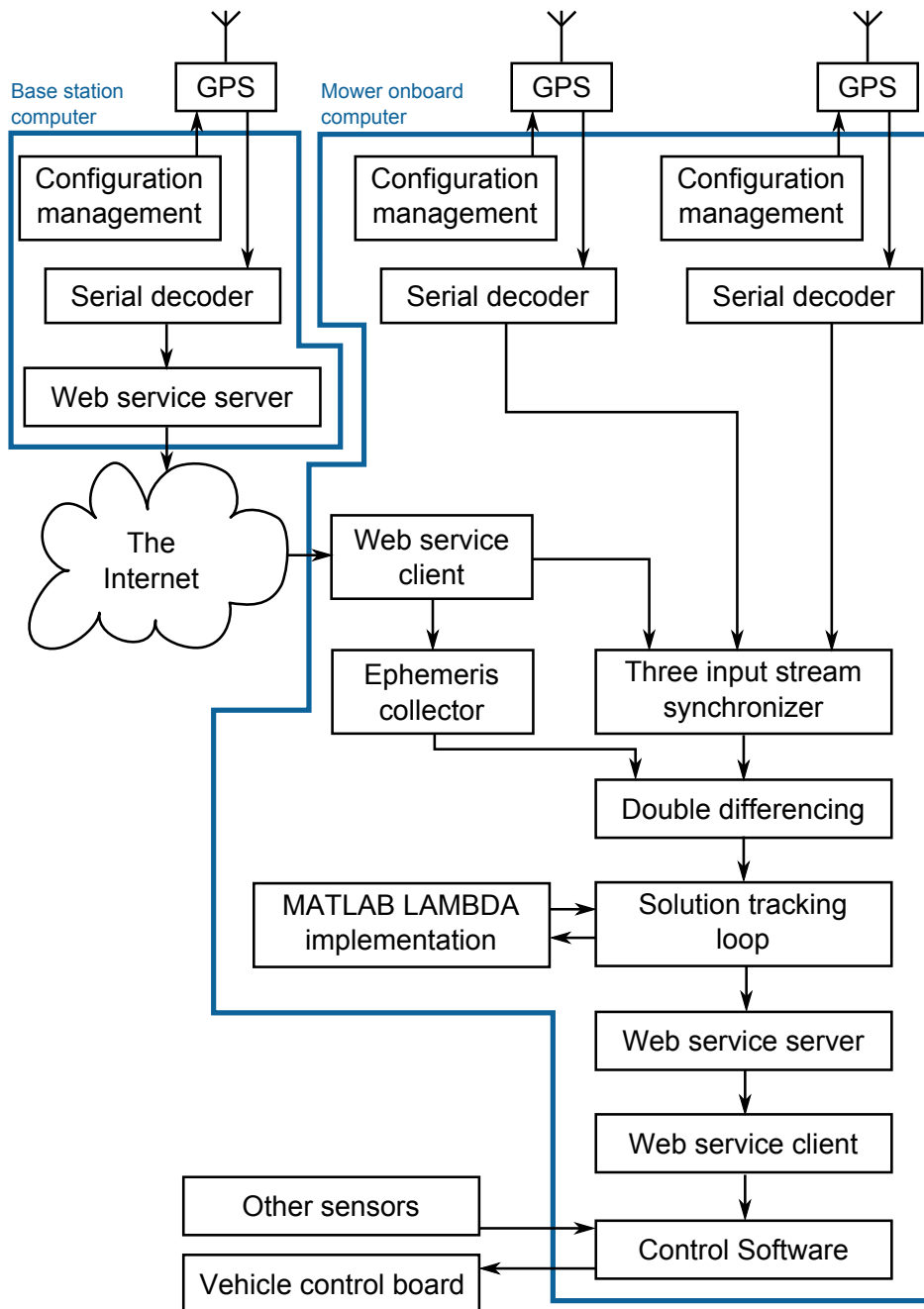
Figure 2.8: Control software overview
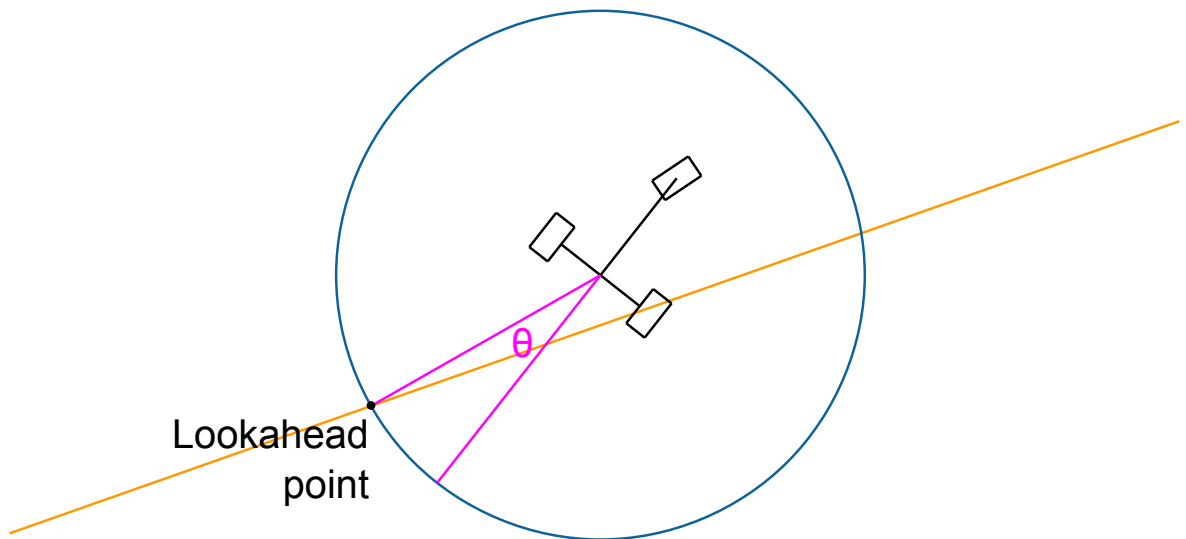
Figure 2.9: Look ahead algorithm: A look ahead distance is chosen (the radius of the blue circle) and the intersection between the circle and the path being followed (orange line) is known as the *look ahead point*. The vehicle steer angle is proportional to the angle to the look ahead point.

several ways of generating this; manually driving the outline of the area to be mowed, manually driving the entire route to be followed, or using satellite photos (i.e. Google Earth) to specify the route or outline. Outlines may be infilled using stripes or spirals.

2.4.3   Path following algorithm

Exploratory tests were performed with two path following algorithms; a simple look ahead algorithm and a more sophisticated algorithm using simulated annealing.

Look ahead is a simple control algorithm, pictured in figure 2.9. A point on the target path is chosen, a fixed distance ahead of the robot, and the angle to this point controls the vehicle steer angle - if the look ahead point is to the left we steer left, if to the right we steer right. This algorithm is simple, widely used in robotics, and works reasonably well, especially when driving the long straight segments that make up a large part of normal mowing. However, it sometimes leads to kinks exiting tight curves, because the mower's steering motor is comparatively slow.
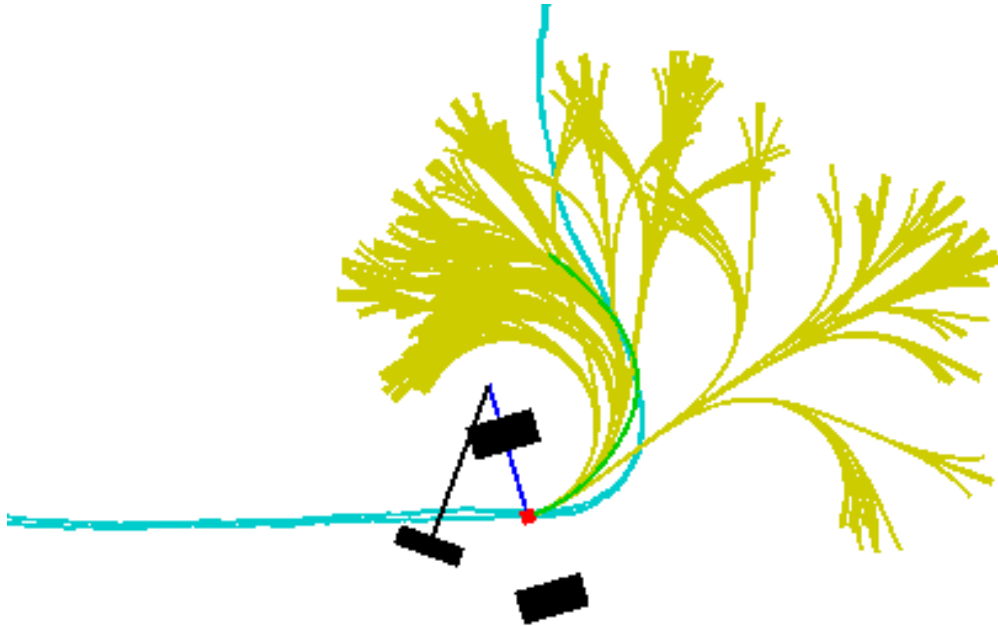
Figure 2.10: Simulated annealing algorithm: Starting from the vehicle's current position, simulate the results of different steering inputs (fast left only, fast left briefly then slow right; fast left briefly then fast right, and so on). The vehicle's location is shown by the three black squares. Various combinations of steering inputs are shown in yellow. The best choice is selected and refined further (shown in green).

Experiments were also performed with an algorithm based on simulated annealing, as shown in figure 2.10. A selection of different control inputs are generated - for example, turning the steering wheel right fast then left fast, or left fast then left slowly, and so on. Using a simple computer model of the vehicle, the inputs which would keep us closest to the planned path are selected. We choose the best of those results, randomly generate a selection of control inputs that are close to it, and evaluate those inputs. We repeat this process, adding less and less noise each time, then choose the best of all the control inputs evaluated and use it. This entire process is performed quickly by computer, as often as ten times a second as new GPS information arrives.

This algorithm offers more accurate route tracking than the look ahead algorithm, as the vehicle model used takes into account the steering wheel speed. If the vehicle's planned path requires precise tracking of turns, such as to mow a sawtooth pattern, this algorithm produces better results. On the other hand it is more complicated than

the look ahead algorithm and for conventional striped mowing patterns it's difficult to tell the difference.

2.4.4 Obstacle avoidance algorithm

The mower was fitted with a forward-facing LIDAR for obstacle detection, but no obstacle detection sensors were fitted to the rear - meaning it is not possible to safely reverse. Therefore, when an obstacle blocks the mower's path the options are to stop and wait for it to move, to swerve around it, or a combination of the two.

In its current state the mower stops and waits for the obstacle to move; this was chosen because it is simple. In the longer term, a capability to swerve around obstacles could be useful, as would rear-facing sensors and a capability to reverse.

Chapter 3

Introduction to satellite navigation

3.1 Overview

Figure 3.1: GPS is comprised of three segments; the user segment (left), the space segment (centre) and the control segment (right).

GPS is a Global Navigation Satellite System (GNSS) designed to provide precise, near-instantaneous position and time information to users anywhere on the planet, in all weather and at all times. The GPS satellites (the *space segment*) transmit a radio signal containing timing and satellite orbit information; GPS receivers (the *user segment*) with a view of at least 4 satellites can measure the arrival time of the signals, in the receiver's clock time, and compare it to the transmission time of the signals in the satellite's clock's time, and from the two the receiver's position (and the receiver's clock error) can be calculated. The final segment of the system, the *control segment*, is a network of base stations which track the satellites and upload orbit information to them (See figure 3.1).

Consumer GPS receivers, when used with a good view of the sky, are usually accurate to better than 5m, and receiver modules can cost less than £15. High precision GPS
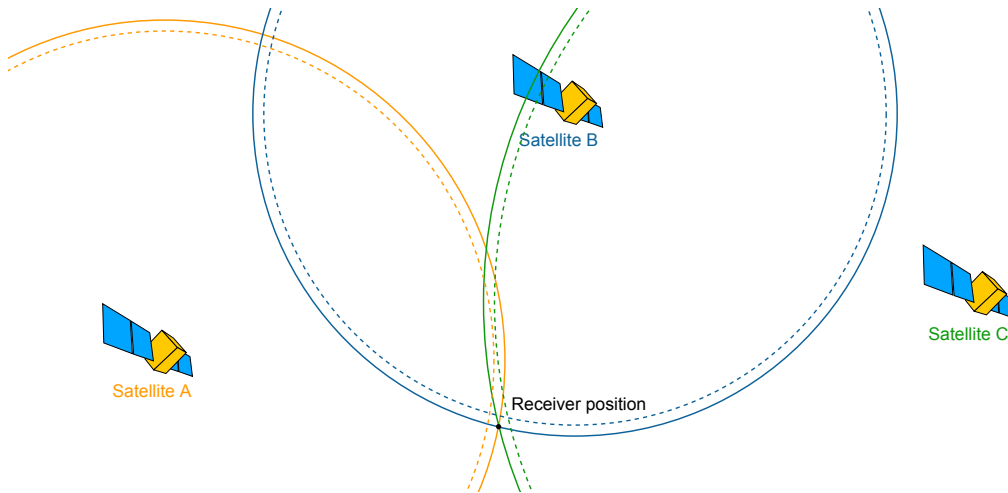
Figure 3.2: 2D example of time of flight and receiver clock error: clock error at the receiver means the measured pseudoranges don't precisely match up (dotted lines) but after correcting for receiver clock error they do (solid lines). For 3D positioning, the circles become spheres and four satellites are required.

receivers, such as those used for surveying, operate in a different way which allows them to be accurate to around 2cm; these receivers cost upwards of £10,000.

GPS is operated by the United States (US) Department Of Defence (DOD); the system was initially conceived as a military-only system, but following the shoot-down of a Korean airliner, which entered restricted soviet airspace due to a navigational error*, it was announced that the system would be opened to public use. Currently each satellite transmits two signals; one signal at 1575.42MHz (the *L1 signal*), which has an unencrypted component (used by civilians and the military) and an encrypted military-only component; and one at 1227.60MHz (the *L2 signal*) with only the encrypted military component. As part of *GPS Modernisation* efforts, some satellites also transmit a civilian component on the L2 signal (known as *L2C*), but no consumer receivers currently receive this signal.

Once the locations of four satellites and the receiver's distance from those satellites are known, the receiver's position can be calculated, like shown in figure 3.2. Because the receiver clock error is common to all the measurements, it can be included in the set

---

*The International Civil Aviation Organization (ICAO) determined the autopilot maintained a magnetic heading (used to follow air traffic control commands) instead of switching to Inertial Navigation System (INS) control (used to follow a preprogrammed flight path); either because the pilots did not switch the mode, or because when they did they were too far from the flight path for the INS control to operate.

of equations the receiver solves to determine its position.*

As the receiver needs to know the position of each satellite, and the satellites are constantly moving through space, the receivers need information about the satellites' positions and motion. This is called *ephemeris information*, and is included in the signal transmitted by each satellite. In an ideal two-body orbital system, a satellite's orbit would be a perfect ellipse (a *Kepler orbit*), which could be perfectly described using six numbers. The satellites' actual orbit is close to this ideal system, but not precisely the same, so the Keplerian parameters are supplemented with correction terms. The information broadcast is updated every 2 hours, and is valid for 4 hours; the Root Mean Squared (RMS) 3D position of the broadcast data error is 3.6m, and the radial error of the broadcast data is usually less than ±2m (Langley et al., 2000).

3.2   Signals and Reception

3.2.1   The GPS Signal

The L1 GPS signal has three components commonly used by civilian users, as depicted in figure 3.3 .

- Navigation data, broadcast at 50 bits per second, containing information such as satellite ephemeris and almanac information. The data is uploaded to the satellite by the control segment, based on satellite position measurements taken at a network of fixed base stations.

- The Coarse Acquisition (C\A) code, a *Gold code* sequence, different for each satellite and known in advance by the receiver, which allows multiple satellites to broadcast at the same frequency, and the signals to be separated later - a technology known as Code Division Multiple Access (CDMA). The C\A code is

---

*You may ask why, once the time is known, a view of 4 satellites is still required. The reason is most receivers use a Crystal Oscillator (XO) or Temperature Controlled Crystal Oscillator (TCXO) to measure time accurate to around 2 Parts per million (PPM) so in 1 second, the clock can develop an error of 1 μs - equivalent, at the speed of light, to the receiver moving 600m.
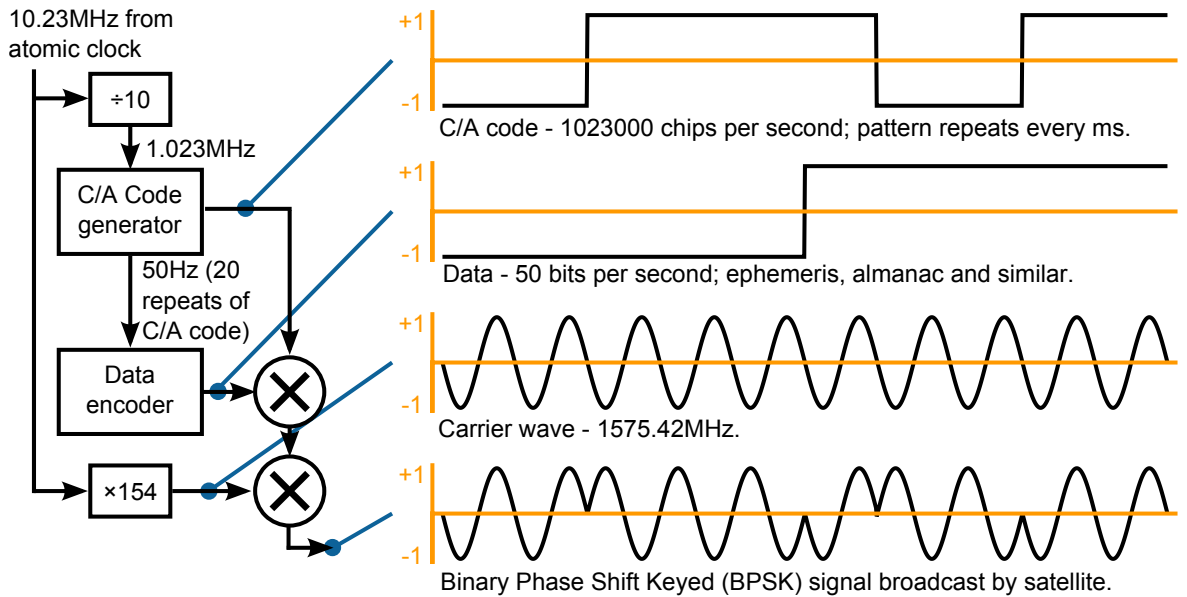
Figure 3.3: Generation of the civilian portion of the L1 GPS signal. This diagram is simplified by not showing the generation of the encrypted military signal; for a more complete diagram, see page 19 of IS-GPS-200D (2006).

Note that multiplying binary signals which are ± 1 is the same as adding, modulo 1, binary signals which are 0 or 1, which is the same as taking an exclusive or of the same signals. Hence, some sources use the symbol $\oplus$ (exclusive or) while others use the symbol $\otimes$ (multiplication).

1023 chips* long, and broadcast at 1,023,000 chips per second (with a chip length of 293.05m), so the code repeats every millisecond; there are 20 repetitions of the C\A code for each bit of the data signal. Measurements of the C\A code are accurate to a few metres.

- The L1 carrier wave, a 1575.42 MHz (wavelength of 19.0 cm) signal generated from the satellite's on board atomic clock. Due to the speed of the satellite's orbit, a receiver on the ground will receive a Doppler-shifted signal; the range of the shift is ± 5 kHz. Due to the short wavelength of the carrier wave, it can be measured very precisely - to within millimetres - but because it is a sine wave all measurements are *ambiguous* by an integer number of wavelengths - that is, a measurement of 1,000,000.00 metres looks identical to a distance of 1,000,000.19 metres. The process of working out the number of wavelengths in the *integer ambiguity* is known as *integer ambiguity resolution*.

---

*A chip is identical to a bit, but referred to as a chip because the signal carries no information (as the Gold code is already known to both the satellite and the receiver)

These three signals are combined as shown in figure 3.3 , to make up the civilian portion of the L1 GPS signal. There also are a number of other signals, received by military and high-precision civilian receivers:

- The L2 carrier wave, a 1227.60 MHz (wavelength of 24.4 cm) signal generated from the satellite's on board atomic clock. Like the L1 carrier wave, measurements can be precise, but are ambiguous by an integer number of wavelengths. Receivers with access to both the L1 and L2 signals are able to measure and compensate for certain frequency-dependent error sources, improving accuracy.

- The Precise (P(y)) code, a digital signal similar to the C\A code but with ten times the bandwidth; its shorter wavelength is designed to allow more precise positioning. The P(y) code is a publicly known signal (the P code), encrypted by modulation with the Anti-Spoofing (AS) (or Y) code. The Precise (P(y)) is broadcast on both the L1 and L2 frequencies, modulated in-phase with the carrier, while the civilian code is modulated as quadrature-phase; in other words, L1 the civilian and military signals are offset by 90°. High-precision GPS receivers can track this signal through a variety of techniques - for example, by detecting the L1 signal (which carries both C\A and P(y) codes), subtracting the C\A code to leave the P(y) code, then looking for the same code at the L2 frequency. For more information on tracking the P(y) code, see Hofmann-Wellenhof et al. (2008), figure 4.30.

- Modernised GPS signals, which are being introduced when current satellites are replaced (GPS satellites' designed lives vary from 7.5 years to 15 years (Gakstatter, 2009)). For more information on GPS modernisation, see Hofmann-Wellenhof et al. (2008).

3.2.2   GPS Signal Reception

The signals from each satellite can be separated from one another and from background noise by detecting the Gold code through a process known as cross-correlation, which
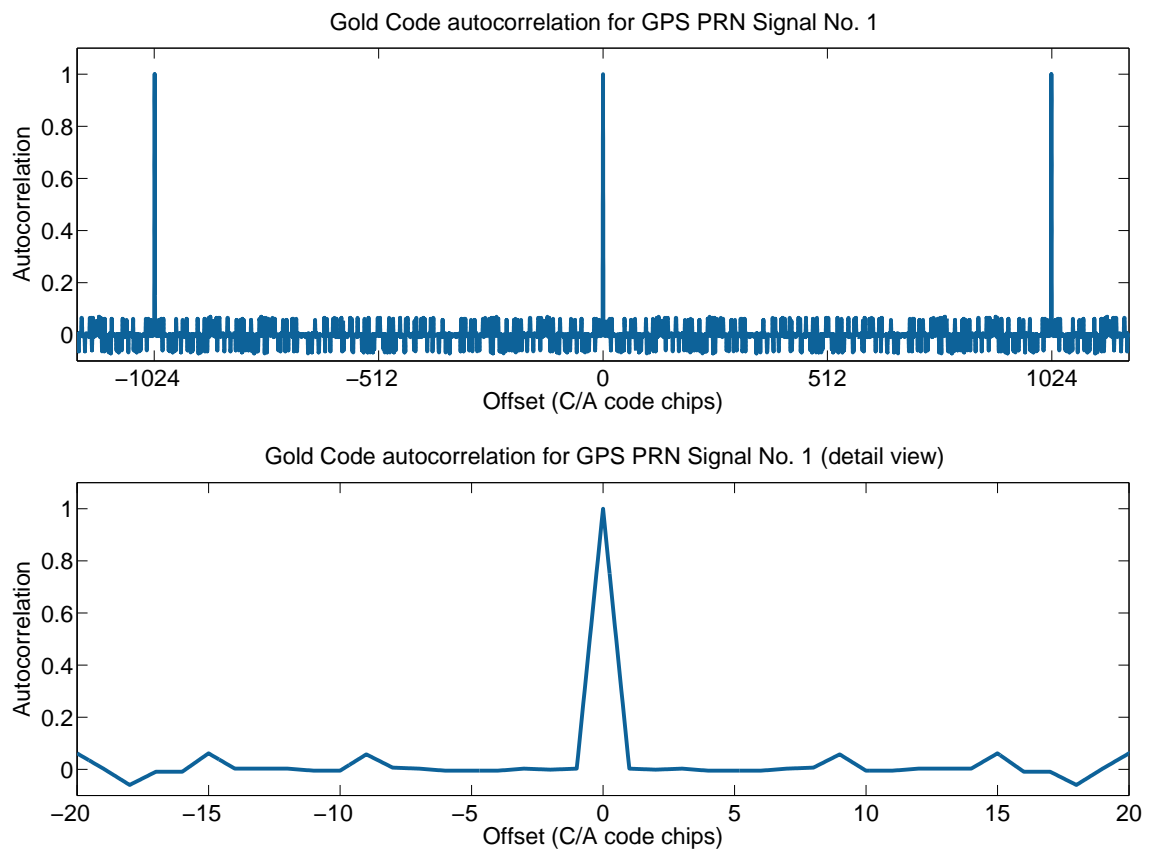
Figure 3.4: Cross-correlation between two copies of the Gold code used by SV PRN 1, shown overall (top) and zoomed to show detail (bottom). The cross-correlation of a signal with itself is called *autocorrelation*.
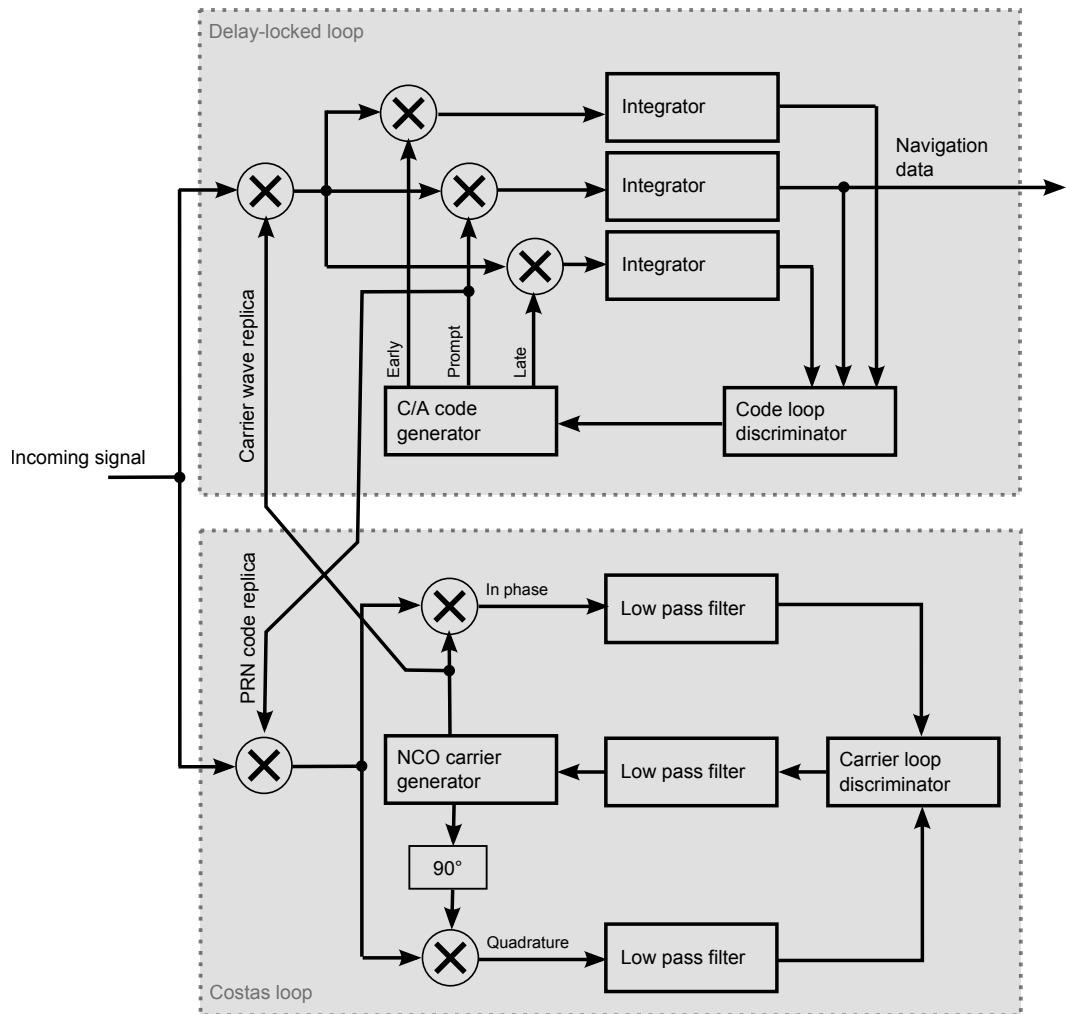
Figure 3.5: Signal tracking loop block diagram; based on a diagram from Borre et al. (2007)

uses multiplication to compare a received signal to an expected signal, with different delays applied to the expected signal. As shown in 3.4 , a peak occurs when the received and expected signals match up exactly. Gold codes (Gold, 1967) are used because several codes can be generated with low cross-correlation between them, reducing interference between signals from different satellites.

In order to track both the Gold code and the Doppler-shifted carrier signal, receivers employ two tracking loops, as shown in figure 3.5. A Delay Locked Loop (DLL) adjusts the delay on early, prompt, and late C/A code replicas, so that the received signal's cross-correlation with the early replica is equal to its cross-correlation with the late replica - with the result that the prompt code (equidistant between the early and late codes) tracks the correlation peak shown in figure 3.4 . Meanwhile, a Costas loop
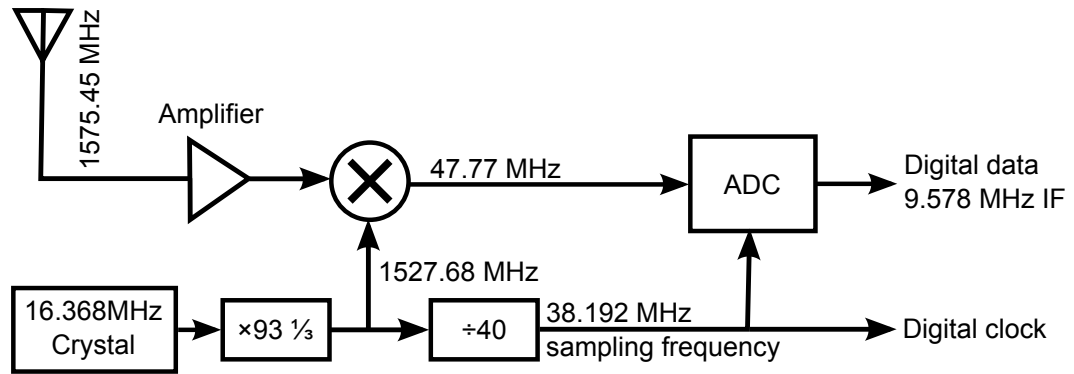
Figure 3.6: GPS receiver analog front end block diagram. It should be noted that different receivers use different crystals, intermediate frequencies, and sampling frequencies; the values shown are for the receiver described in Gromov et al. (2000).

tracks the phase and quadrature components of the incoming signal's carrier wave, and adjusts the speed of a Numerically Controlled Oscillator (NCO) to keep a carrier wave replica in phase with the received signal.

In modern receivers these tracking loops are implemented digitally, but rather than attempting to digitise the carrier signal at its full 1575.42Mhz frequency, the signal is first downmixed to an Intermediate Frequency (IF) then digitised (in the so-called *analog front end* hardware, often manufactured using a Silicon-Germanium (SiGe) fabrication process), as shown in figure 3.6. The digital data is then processed by a *baseband processor* which tracks the signals and calculates the receiver position. Some chipsets have the analog front end separate from the baseband processor, while others place them in a single chip to save space. Figure 3.7 shows the components found in a consumer bluetooth GPS receiver, such as an analog front end (3), baseband processor (2), and crystal oscillator (8).

3.3 Coordinates, Observables and Error Sources

3.3.1 Coordinates

There are a wide range of ways of describing location, from street names and house numbers to astronomical references which account for the planet's movement in the solar system. For the purposes of this document there are three systems of particular
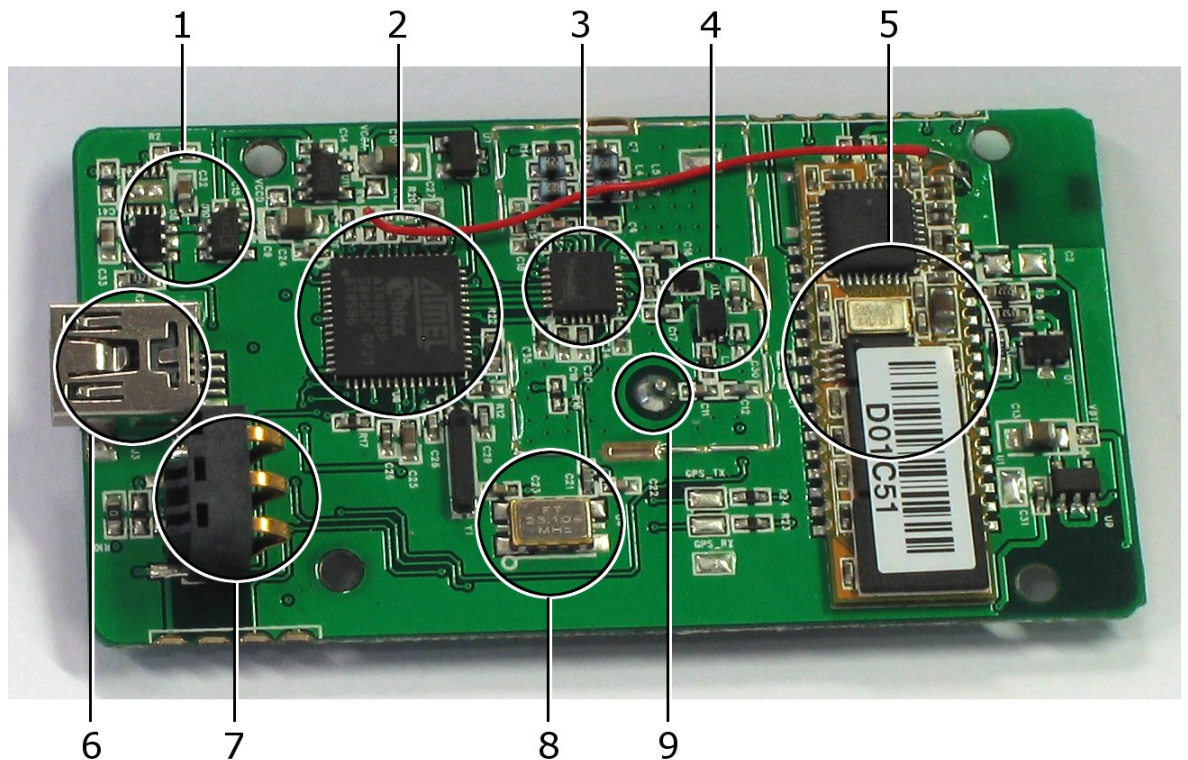
Figure 3.7: Consumer bluetooth GPS receiver (BlueNext BN-906GR) with key components labelled: (1) battery charge control electronics; (2) baseband processor, Atmel ATR0625P; (3) analog front end, Atmel ATR0601; (4) band pass filter; (5) bluetooth module, ISSC IS1005b based; (6) USB mini-b connector, used for charging; (7) battery connector; (8) crystal oscillator, 23.104 MHz; (9) connection to patch antenna on reverse of PCB.
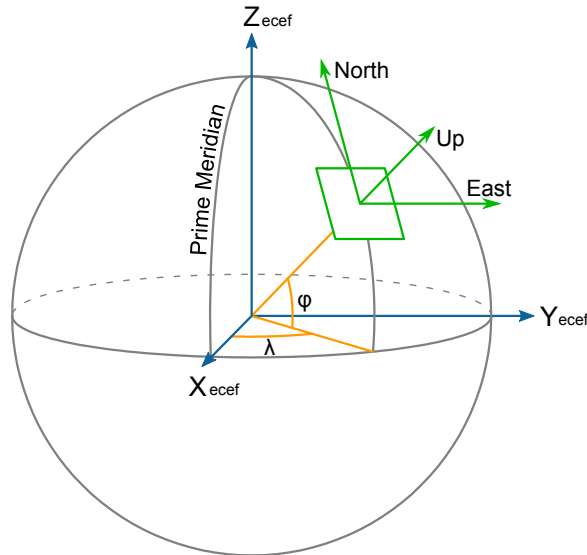
Figure 3.8: The relationship between ECEF, ENU, Longitude (λ) and Latitude (φ).
Based on a public domain licensed diagram by Wikipedia user Raffyl99, from `http://en.wikipedia.org/wiki/File:EarthTangentialPlane.png`.

interest, as they are often used in GPS calculations.

- Earth Centred Earth Fixed (ECEF) coordinates are Cartesian coordinates relative to the centre of the earth, with the Z axis to the north and the X axis to the intersection of the reference meridian with the equator. The position of the reference meridian, and the position of the centre of the earth are, for the GPS system, defined by the World Geodetic System 1984 (WGS84) datum. An example position would be $3899957.3, -106303.7, 5029020.8$. In figure 3.8, X_{ecef}, Y_{ecef} and Z_{ecef} mark the axes of the coordinate system. Other datums are also sometimes used; European Terrestial Reference System 1989 (ETRS89) is fixed to the Eurasian tectonic plate, making measurements in Europe stable in the face of 2.5cm-per-year continental drift.

- Longitude, Latitude, Altitude (LLA) coordinates are the traditional format that GPS receivers present to users. Longitude (λ) is measured in degrees east or west, and latitude (φ) is measured in degrees north or south; positions may be expressed in degrees, degrees and minutes, or degrees, minutes and seconds; and may either state whether the angles are east or west, north or south; or may use positive numbers to describe north and east, negative numbers to describe south and west. So for example a position could be described as 52°22´57.67˝ N, 1°33´40.9˝ W or
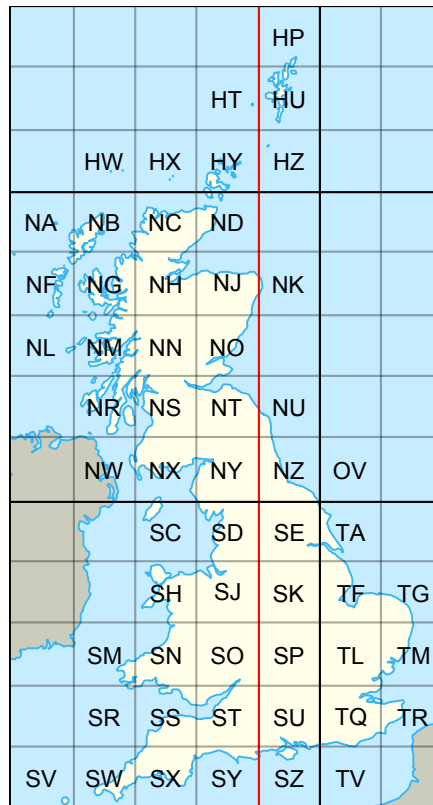
Figure 3.9: The Ordnance Survey National Grid, with 100km x 100km squares marked with their two-letter grid references, and the grid's central meridian marked in red.
Based on a diagram by Wikipedia user Nandhp, under a Creative Commons attribution-share alike license, from http://en.wikipedia.org/wiki/File:British_National_Grid.svg.

as $52.382686°, -1.561361°$. Altitudes are defined in meters, relative to a reference model of the earth defined by WGS84 (for example, 142.50m). In figure 3.8 longitude and latitude measurements are depicted, with λ marking the longitude and φ marking the latitude.

- East North Up (ENU) coordinates, also known as Local Tangential Plane (LTP) coordinates, are Cartesian coordinates defined relative to a nearby point. East North Up (ENU) coordinates may be preferred for measurements where the earth's curvature is negligible as they are simple to understand and visualise. In figure 3.8 one such system is depicted, with its axes marked *East*, *North* and *Up*.

- Ordnance Survey (OS) National Grid, a national coordinate system used in Great Britain, based on a Transverse Mercator projection. Positions may be defined as 'easting and northing', a distance in metres from the bottom-left corner of the

grid (for example 429952.5 276161.2), or by a grid square and distance from the bottom-left corner of the square (for example $SP\,2995\,276161$). Figure 3.9 shows the grid square prefixes. As more precise modern survey methods have become available it has been discovered the National Grid has offset errors in some parts of the country, so in modern usage the National Grid is defined relative to ETRS89 using a nonlinear transformation called OSTN02.

3.3.2 Observables

Range and Doppler shift measurements from the carrier and code tracking loops (described in section 3.2.2 on page 37) are known as *observables*. To make it clear that these measurements include certain errors, the range measurements are called *pseudoranges*. For example, the code pseudorange, using the terminology of Hofmann-Wellenhof et al. (2008), can be described as:

$$R_r^s = c \times (t_r(rec) - t^s(sat)) \tag{3.1}$$

where R_r^s is the code pseudorange (in metres) between satellite s and receiver r; c is the speed of light (in metres/second, 2.99792458×10^8m/s); $t_r(rec)$ is the time the signal was received, measured by the receiver's clock (in seconds); and $t^s(sat)$ is the time the signal was broadcast, measured by the satellite's clock (in seconds).

Both the satellite clock and the receiver clock can have errors; naming the satellite clock error δ^s and the receiver clock error δ_r, and by defining the signal flight time as $\Delta t_r^s = t_r - t^s$ and the combined clock error as $\Delta \delta_r^s = \delta_r - \delta^s$, equation 3.1 can be expressed as:

$$R_r^s = c \times ([t_r + \delta_r] - [t^s + \delta^s]) = c\Delta t_r^s + c\Delta \delta_r^s \tag{3.2}$$

Due to the fact that correction information is broadcast, the satellite clock error is negligibly small, making $\Delta \delta_r^s$ equal to the receiver clock bias, δ_r, which is common among all pseudoranges observed. We define the distance from the GPS receiver to the

satellite as:

$$\varrho_r^s = \|\boldsymbol{\varrho}^s - \boldsymbol{\varrho}_r\| = \sqrt{(X^s - X_r)^2 + (Y^s - Y_r)^2 + (Z^s - Z_r)^2} \qquad (3.3)$$

where $\boldsymbol{\varrho}^s$ is a vector representing the location of satellite s, which has the Cartesian coordinates $X^s\,Y^s\,Z^s$; and $\boldsymbol{\varrho}_r$ is a vector representing the location of receiver r, which has the Cartesian coordinates $X_r\,Y_r\,Z_r$. Using ν_r^s to represent all unmodelled errors, equation 3.2 can be stated as:

$$R_r^s = \varrho_r^s + c\delta_r + \nu_r^s \qquad (3.4)$$

Similarly, the phase pseudorange (which, as mentioned in section 3.2.1 on page 34, has an integer ambiguity because each sine-wave cycle looks the same), can be expressed as:

$$\lambda\Phi_r^s = \varrho_r^s + c\delta_r + \lambda N_r^s + \mu_r^s \qquad (3.5)$$

where λ is the wavelength (in metres/cycle), Φ_r^s is the carrier pseudorange (in cycles)*, N_r^s is the integer ambiguity (in cycles), and μ_r^s represents all unmodelled errors.

3.3.3 Error sources

The assumption that all error sources can be neglected is seldom an appropriate one for precise applications, but for some use cases these can be reduced. Sources of error include:

- Tropospheric refraction, where the signal is delayed while passing through the troposphere. This can be divided into a dry and a wet portion; the dry portion causes more refraction, but it's easier to model and compensate for, so the wet portion has more impact on the final position solution.

---

*Some literature measures the carrier pseudorange in metres, multiplying the integer ambiguity by the wavelength; other literature measures the carrier pseudorange in cycles, dividing the true distance and receiver clock error by the wavelength. Sources such as Chang et al. (2005) even measure the code pseudorange in wavelengths.

- Ionospheric refraction, where the signal is delayed while passing through the ionosphere. The data broadcast by GPS satellites includes Klobuchar model parameters to model the state of the ionosphere, allowing some mitigation.

- Ephemeris errors (as mentioned in section 3.1 on page 33) where, because there is only a limited bandwidth for broadcasting satellite orbit information, imperfect but succinct information is sent.

- Measurement noise, where receivers cannot measure signals with complete precision. Carrier wave signals, which have a much shorter wavelength, encounter much less measurement noise than C\A code measurements.

- Multipath, where a receiver detects a reflected signal (which has taken a longer path) rather than a direct signal from a satellite. Signals can reflect off buildings or the ground. High-cost receivers use special antennas with *choke ring* ground planes; quarter-wavelength grooves attenuate signals from low altitudes as the half-cycle-offset conducted signal interferes destructively. Attenuating signals from low altitudes helps with reflections from the ground, but doesn't offer the same benefits with signals reflected from above the ground, e.g. from buildings. The highest precision GPS applications demand an unobstructed, reflection-free view of the sky for this reason.

- Poor satellite geometry, where other errors have a magnified influence on the final position error, as shown in figure 3.10 . As this depends on the position of satellites in the sky it can be predicted in advance, making it possible to choose times when the effects will be at their lowest.

- Poor satellite visibility, where at certain times of the day fewer satellites are in view, reducing redundancy and the ability to take averages. As this depends on the position of satellites in the sky it can be predicted in advance, making it possible to choose times when the effects will be at their lowest.

- Blunders, such as operator error and software bugs.
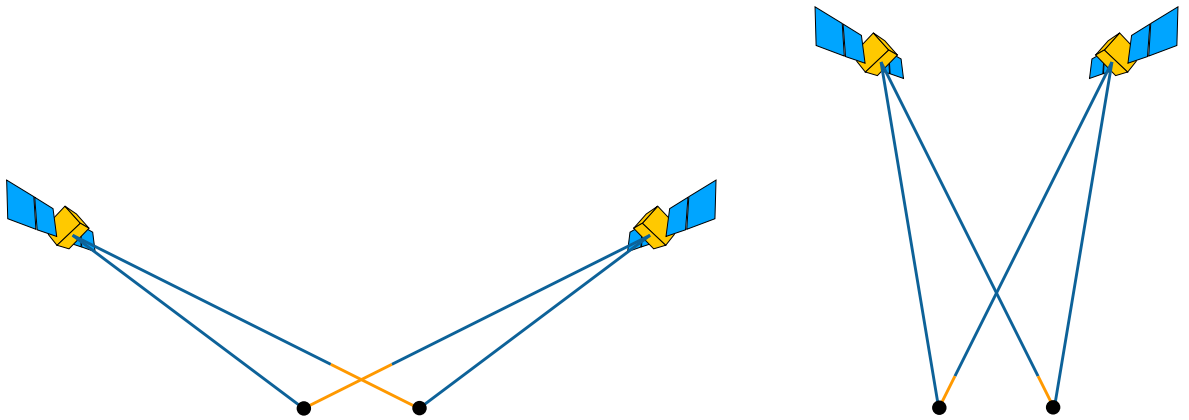
Figure 3.10: Effects of satellite position on accuracy: In the left image, with satellites spread out, receiver movement of 1 metre causes a range change of 90 centimetres; in the right image, with satellites grouped together the same movement changes the ranges by only 30 centimetres. Hence, in the former case a range measurement error of 1 metre would lead to a position error of 1.11 metres, while in the latter case the same error would lead to a position error of 3.33 metres. GPS receivers report Dilution Of Precision (DOP) to numerically represent this effect. For more information, see section 3.4.5.4 on page 73.

Error source	Means of mitigation
Tropospheric refraction	Modelling (IGS data); differential corrections.
Ionospheric refraction	Modelling (broadcast or IGS data); differential corrections.
Satellite ephemeris errors	Modelling (IGS data); differential corrections.
Measurement noise	Low-noise carrier phase measurement.
Multipath	Antenna selection and placement.
Poor satellite geometry	Select times for best satellite geometry.
Poor satellite visibility	Select times for best satellite visibility.
Operator error, software bugs	Detect and avoid!

Table 3.1: Summary of common GPS error sources

Error sources are summarised in table 3.1 . Our application allows some of these to be mitigated through differential corrections and carrier phase measurement, as noted in the table.

Chang et al. (2005) takes some of these errors into account, using this formulation for code pseudorange:

$$R_r^s = \varrho_r^s - I_r^s + T_r^s + c\delta_r + d_r - d^s + \nu_r^s \tag{3.6}$$

where I_r^s is the error caused by the ionosphere (in metres); T_r^s is the error caused by the troposphere (in metres); d_r and d^s are the receiver and satellite hardware code delays (in metres); and ν_r^s is the code measurement noise and multipath noise (in metres).

Similarly, the formulation for the carrier pseudorange is:

$$\lambda\Phi_r^s = \varrho_r^s + I_r^s + T_r^s + c\delta_r + \lambda N_r^s + D_r - D^s + \varphi_0^s + \varphi_{r0} + \mu_r^s \tag{3.7}$$

where D_r and D^s are the receiver and satellite hardware carrier delay (in metres); φ_0^s and φ_{r0} are the satellite and receiver initial phase*; and μ_r^s is the carrier measurement noise and multipath noise (in metres). You might wonder why I_r^s is negative in equation 3.6 but positive in equation 3.7. This is because the carrier is a single frequency signal (and hence propagates at the phase velocity), while the C/A code is comprised of several signals of slightly different frequencies (and hence propagates at the group velocity). Although the phase velocity and the group velocity would be the same in a vacuum, the frequency-dependent effects of the ionosphere have opposite effects on the phase velocity and the group velocity. For more information, see section 5.3 of Hofmann-Wellenhof et al. (2008).

3.3.3.1   Precision and accuracy

In discussing errors the terms *precise* and *accurate* are often used; as shown in figure 3.11 , measurements are *precise* if they are tightly grouped together and *accurate* if the grouping is centred around the true value. Geographical coordinates make this

---

*Initial phase is the carrier phase at the time the satellite (or receiver) clock turned on.

a. Not precise, not accurate

b. Accurate, not precise

c. Precise, not accurate

d. Precise and accurate
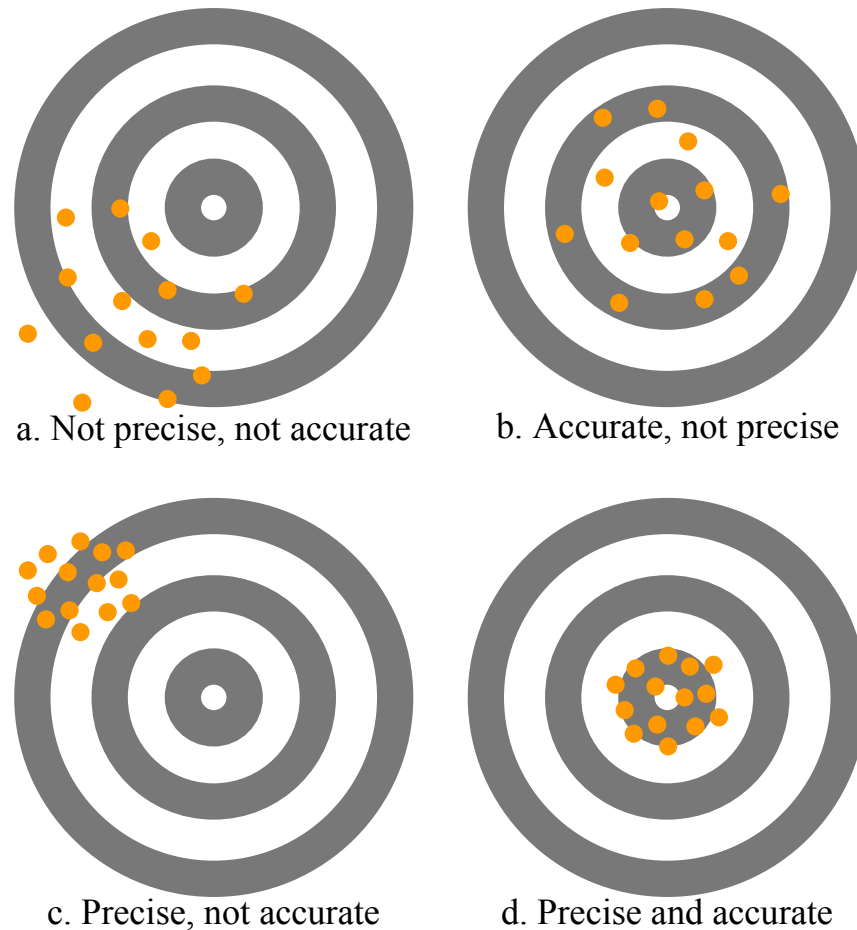
Figure 3.11: The relationship between precision and accuracy

slightly more complicated due to different frames of reference - true values change as tectonic plates move relative to one another, unless measurements are taken relative to a fixed point on the same tectonic plate. Further complicating matters, some systems aren't precise enough to notice tectonic plate motion and don't take it into account.

For the purposes of our application, if measurements are inaccurate due to a systematic error like incorrect ambiguity resolution that is a different matter to if they are inaccurate due to tectonic plate motion. For the avoidance of doubt, unless otherwise specified the term *accurate* is used to refer to accuracy relative to a nearby fixed base station.

3.3.4   Combinations of observables

3.3.4.1   Single differencing

A key method used to reduce the impact of these errors is by combining observables. The simplest form is *single differencing*, where a pseudorange from one satellite (which we will call satellite *1*) is subtracted from the pseudoranges at all other satellites, in order to eliminate receiver clock bias; the effect on equation 3.6 on page 46 is:

$$
\begin{aligned}
R_r^s - R_r^1 &= \varrho_r^s - \varrho_r^1 - I_r^s + I_r^1 + T_r^s - T_r^1 + c\Delta\delta_r - c\Delta\delta_r && (3.8) \\
&\quad + d_r - d_r - d^s + d^1 + \nu_r^s - \nu_r^1 \\
&= \varrho_r^s - \varrho_r^1 - I_r^s + I_r^1 + T_r^s - T_r^1 - d^s + d^1 + \nu_r^s - \nu_r^1 && (3.9)
\end{aligned}
$$

Using f^{s-1} to signify the difference between f for satellite s and satellite *1*, we can define $R_r^{s-1} = R_r^s - R_r^1$; $\varrho_r^{s-1} = \varrho_r^s - \varrho_r^1$; $I_r^{s-1} = I_r^s - I_r^1$; $T_r^{s-1} = T_r^s - T_r^1$; $d^{s-1} = d^s + d^1$; and $\nu_r^{s-1} = \nu_r^s - \nu_r^1$, allowing us to rewrite equation 3.9 as:

$$
R_r^{s-1} = \varrho_r^{s-1} - I_r^{s-1} + T_r^{s-1} - d^{s-1} + \nu_r^{s-1} \qquad (3.10)
$$

Likewise, by defining $N_r^{s-1} = N_r^s - N_r^1$; $D^{s-1} = D^s - D^1$; $\varphi_0^{s-1} = \varphi_0^s - \varphi_0^1$; and $\mu_r^{s-1} = \mu_r^s - \mu_r^1$, equation 3.7 on page 46 can be rewritten as:

$$
\lambda\Phi_r^{s-1} = \varrho_r^{s-1} + I_r^{s-1} + T_r^{s-1} + \lambda N_r^{s-1} - D^{s-1} + \varphi_0^{s-1} + \mu_r^{s-1} \qquad (3.11)
$$

In other words, single differencing eliminates errors common among satellites - namely receiver clock error and receiver hardware delays.

3.3.4.2   Double differencing

To eliminate errors from sources such as the ionosphere and troposphere, to allow high-precision positioning, a fixed base station at a known location takes measure-
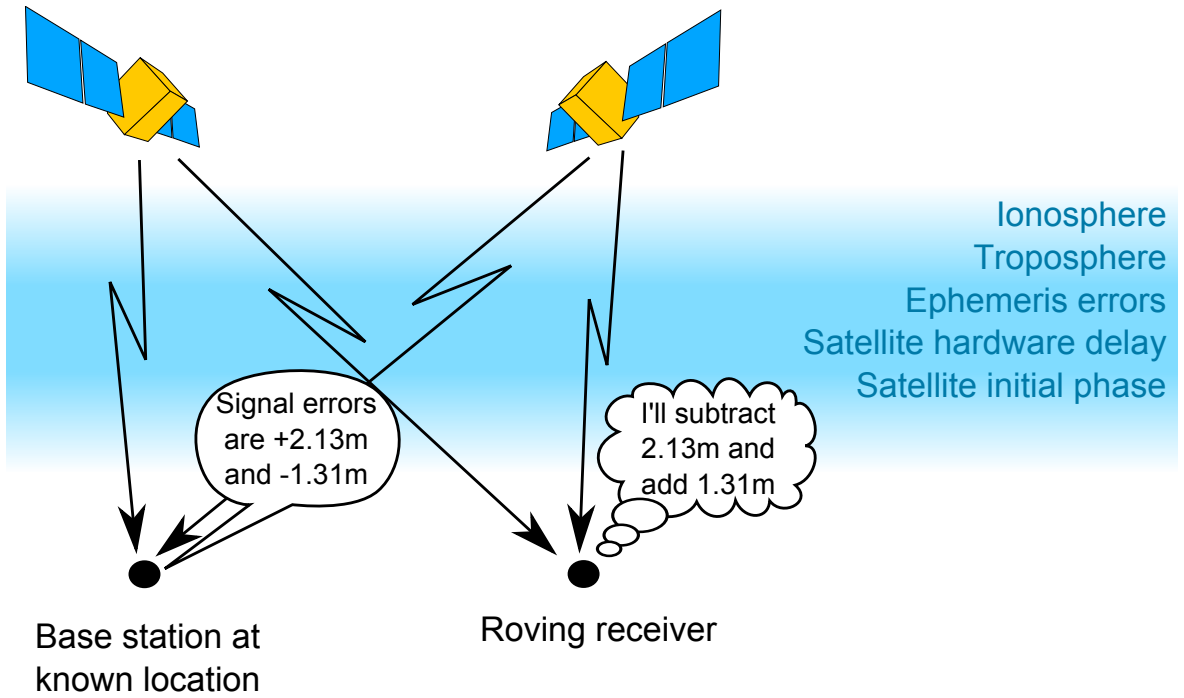
Figure 3.12: Use of a fixed base station to eliminate common errors on GPS measurements.

ments and, as its location is known, can work out the delay on each signal. This data can then be broadcast to a roving GPS receiver, where the errors can be subtracted. This is shown in figure 3.12. When the roving receiver is within a few kilometres of the base station, the errors are common, allowing them to be completely eliminated. This technique is often called Differential GPS (DGPS) when applied to the code measurement, and Carrier Differential GPS (CDGPS) when applied to the carrier measurement. When single-differenced measurements from a base station are subtracted from single-differenced measurements from a roving receiver, the results are known as *double-differenced measurements*. For example, equation 3.10, after the subtraction of the measurements at base station b, would be:

$$R_r^{s-1} - R_b^{s-1} = \varrho_r^{s-1} - \varrho_b^{s-1} - I_r^{s-1} + I_b^{s-1} + T_r^{s-1} - T_b^{s-1} \quad (3.12)$$

$$-d^{s-1} + d^{s-1} + \nu_r^{s-1} - \nu_b^{s-1}$$

$$= \varrho_r^{s-1} - \varrho_b^{s-1} + \nu_r^{s-1} - \nu_b^{s-1} \quad (3.13)$$

Using f_{r-b} to signify the difference between f for receiver r and receiver b, we can

define $R_{r-b}^{s-1} = R_r^{s-1} - R_b^{s-1}$; $\varrho_{r-b}^{s-1} = \varrho_r^{s-1} - \varrho_b^{s-1}$; and $\nu_{r-b}^{s-1} = \nu_r^{s-1} - \nu_b^{s-1}$ giving the double-differenced code measurement:

$$R_{r-b}^{s-1} = \varrho_{r-b}^{s-1} + \nu_{r-b}^{s-1} \tag{3.14}$$

Likewise, for equation 3.11 on page 48 we can define $\Phi_{r-b}^{s-1} = \Phi_r^{s-1} - \Phi_b^{s-1}$; $N_{r-b}^{s-1} = N_r^{s-1} - N_b^{s-1}$; and $\mu_{r-b}^{s-1} = \mu_r^{s-1} - \mu_b^{s-1}$ producing:

$$\lambda\Phi_{r-b}^{s-1} = \varrho_{r-b}^{s-1} + \lambda N_{r-b}^{s-1} + \mu_{r-b}^{s-1} \tag{3.15}$$

In summary, when a roving receiver is within a few kilometres of a base station, double differencing can counteract several major sources of inaccuracy.

In applications using fixed base stations there are two common configurations: Either the user owns and operates the fixed base station, or the user purchases access to a network RTK system. Network RTK uses a network of base stations around the country, interpolating between the nearest base stations to determine errors at the user's location.

For applications using dual-frequency receivers costing several thousand pounds, network RTK subscriptions costing one to two thousand pounds a year can be cost-effective, and for applications working beyond the range of a single base station, network methods are vital. One example of this arrangement is Cai et al. (2011). For golf course applications, only a single base station at the clubhouse would be required, and if a low cost single frequency receiver was used for it the savings on subscription costs would pay for the hardware within months. This was the route we chose to explore. It is possible in the future subscription costs will come down, or a special price could be negotiated; should this happen, using a reference network may become cost-effective.

3.3.4.3 Wide lane combination

When dual frequency data is available, through the use of a high precision receiver, a wide lane combination can be formed. The L1 (1575.42 MHz, 19.0 cm wavelength) signal and the L2 (1227.60 MHz, 24.4 cm wavelength) signal can be multiplied together to produce a beat frequency:

$$
\begin{aligned}
& \sin\left(1575.42 \times t\right) \times \sin\left(1227.60 \times t\right) \\
= \quad & \tfrac{1}{2}\cos\left(1575.42 \times t - 1227.60 \times t\right) \quad + \quad \tfrac{1}{2}\cos\left(1575.42 \times t + 1227.60 \times t\right) \\
= \quad & \tfrac{1}{2}\cos\left(347.82 \times t\right) \quad\quad\quad + \quad\quad\quad \tfrac{1}{2}\cos\left(2803.02 \times t\right)
\end{aligned}
\tag{3.16}
$$

Band pass filtering leaves only the 347.82 MHz component, with a wavelength of 86.2 cm. This simplifies integer ambiguity resolution because with an ambiguity distance of 86.2 cm a range error of 43cm will still be nearest to the correct ambiguity, while for a single-frequency measurement with a 19.0 cm ambiguity distance, a range error of just 9.5 cm would be required to produce the same result.

3.3.4.4 Ionosphere free combination

Ionospheric interference depends on Total Electron Content (TEC), and varies with frequency. When dual frequency data is available, first-order effects can be removed using a so-called *ionosphere free combination*. The model of ionospheric delay used for this is

$$
\Delta^{Iono} = \frac{1}{\cos z}\frac{40.3}{f^2}TVEC
\tag{3.17}
$$

where z is the zenith angle to the satellite, f is the signal frequency and $TVEC$ is the total vertical electron count, in units of 10^{16} electrons/m^2. The measurements at two frequencies may be combined like

$$
R_{Iono} = R_{L1} - \frac{f_{L1}}{f_{L2}}R_{L2}
\tag{3.18}
$$

where R_{Iono}, R_{L1} and R_{L2} are ionosphere-free, L1 and L2 pseudoranges respectively, while f_{L1} and f_{L2} are the L1 and L2 frequencies respectively. The name of the *ionosphere free combination* is not strictly correct as it doesn't entirely eliminate ionospheric errors, but it is reported to provide a substantial reduction.

3.4   Position calculation

Several of the equations above are nonlinear (such as equation 3.3 on page 43), making them difficult to solve algebraically. Several methods have been developed to solve the GPS equations.

3.4.1   Bancroft

By combining equation 3.4 on page 43 with equation 3.3 on page 43 and ignoring the combined error term ν_r^s, we get this equation relating code pseudorange to position:

$$R_r^s = \sqrt{(X^s - X_r)^2 + (Y^s - Y_r)^2 + (Z^s - Z_r)^2} + c\delta_r \qquad (3.19)$$

Bancroft (1985) proposed that it can be solved as follows: First, by subtracting $c\delta_r$ from both sides of the equation, squaring both sides, and rearranging:

$$X^{s\,2} + Y^{s\,2} + Z^{s\,2} - R_r^{s\,2} - 2\left(X^s X_r + Y^s Y_r + Z^s Z_r - c\delta_r R_r^s\right) + X_r^2 + Y_r^2 + Z_r^2 - c\delta_r^2 = 0 \qquad (3.20)$$

Bancroft defines two vectors; a vector for satellite location and pseudorange; and a vector for user location and receiver clock bias:

$$\boldsymbol{a}^s = \begin{bmatrix} X^s \\ Y^s \\ Z^s \\ R_r^s \end{bmatrix} \qquad \boldsymbol{x}_r = \begin{bmatrix} X_r \\ Y_r \\ Z_r \\ c\delta_r \end{bmatrix} \qquad (3.21)$$

Additionally, the *Minkowski functional for 4-space*, also known as the *Lorentzian inner product*, is defined as:

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3 - a_4 b_4 \tag{3.22}$$

So for example $\langle \boldsymbol{a}^s, \boldsymbol{a}^s \rangle = X^s X^s + Y^s Y^s + Z^s Z^s - R_r^s R_r^s$ and $\langle \boldsymbol{x}_r, \boldsymbol{x}_r \rangle = X_r X_r + Y_r Y_r + Z_r Z_r - c\delta_r c\delta_r$. The Lorentzian inner product is sesquilinear, i.e. $\langle \boldsymbol{a} + \boldsymbol{b}, \boldsymbol{c} + \boldsymbol{d} \rangle = \langle \boldsymbol{a}, \boldsymbol{c} \rangle + \langle \boldsymbol{a}, \boldsymbol{d} \rangle + \langle \boldsymbol{b}, \boldsymbol{c} \rangle + \langle \boldsymbol{b}, \boldsymbol{d} \rangle$. With these substitutions, equation 3.20 becomes:

$$
\begin{aligned}
0 &= \langle \boldsymbol{a}^s, \boldsymbol{a}^s \rangle - 2 \begin{bmatrix} X^s & Y^s & Z^s & -R_r^s \end{bmatrix} \boldsymbol{x}_r + \langle \boldsymbol{x}_r, \boldsymbol{x}_r \rangle \tag{3.23} \\
&= \frac{1}{2} \langle \boldsymbol{a}^s, \boldsymbol{a}^s \rangle - \begin{bmatrix} X^s & Y^s & Z^s & -R_r^s \end{bmatrix} \boldsymbol{x}_r + \frac{1}{2} \langle \boldsymbol{x}_r, \boldsymbol{x}_r \rangle \tag{3.24}
\end{aligned}
$$

When dealing with multiple satellites, this expands to:

$$
\begin{aligned}
0 &= \frac{1}{2} \begin{bmatrix} \langle \boldsymbol{a}^1, \boldsymbol{a}^1 \rangle \\ \langle \boldsymbol{a}^2, \boldsymbol{a}^2 \rangle \\ \langle \boldsymbol{a}^3, \boldsymbol{a}^3 \rangle \\ \vdots \end{bmatrix} - \begin{bmatrix} X^1 & Y^1 & Z^1 & -R_r^1 \\ X^2 & Y^2 & Z^2 & -R_r^2 \\ X^3 & Y^3 & Z^3 & -R_r^3 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \boldsymbol{x}_r + \frac{1}{2} \langle \boldsymbol{x}_r, \boldsymbol{x}_r \rangle \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} \\
&= \qquad \boldsymbol{r} \qquad - \qquad\qquad A \qquad\qquad \boldsymbol{x}_r \qquad + \lambda \qquad i_0
\end{aligned}
\tag{3.25}
$$

Ignoring, for a moment, the dependence of λ on $\boldsymbol{x}_r$, with four or more satellites visible one could find a least squares solution for $\boldsymbol{x}_r$ by solving the normal equations, i.e.

$$A \boldsymbol{x}_r = \boldsymbol{r} + \lambda i_0 \tag{3.26}$$

$$\hat{\boldsymbol{x}}_r = B \left(\boldsymbol{r} + \lambda i_0 \right) \qquad \text{where} \qquad B = \left(A^T W A \right)^{-1} A^T W \tag{3.27}$$

where W is the measurement weighting matrix (explained in section 3.4.5.1 on page 66, and in this case often the identity matrix), B is the generalised inverse of A, and $\hat{\boldsymbol{x}}_r$ is the least-squares solution for $\boldsymbol{x}_r$. However, because this solution for $\boldsymbol{x}_r$ depends on λ, which in turn depends on $\boldsymbol{x}_r$, to find a solution we have to substitute equation 3.27

back into the definition of λ - i.e.

$$\lambda \quad = \quad \frac{1}{2} \langle \boldsymbol{x}_r, \boldsymbol{x}_r \rangle \tag{3.28}$$

$$\lambda \quad = \quad \frac{1}{2} \langle B \left(\boldsymbol{r} + \lambda i_0 \right), B \left(\boldsymbol{r} + \lambda i_0 \right) \rangle \tag{3.29}$$

$$\lambda \quad = \quad \frac{1}{2} \langle B\boldsymbol{r}, B\boldsymbol{r} \rangle + \lambda \langle B\boldsymbol{r}, Bi_0 \rangle + \frac{1}{2} \lambda^2 \langle Bi_0, Bi_0 \rangle \tag{3.30}$$

$$0 \quad = \quad \frac{1}{2} \langle Bi_0, Bi_0 \rangle \lambda^2 + \left(\langle B\boldsymbol{r}, Bi_0 \rangle - 1 \right) \lambda + \frac{1}{2} \langle B\boldsymbol{r}, B\boldsymbol{r} \rangle \tag{3.31}$$

As B, $\boldsymbol{r}$, and i_0 are all known, a solution for λ can be found using the quadratic formula, then substituted into equation 3.27 to arrive at a least-squares solution.

3.4.2   Linearisation
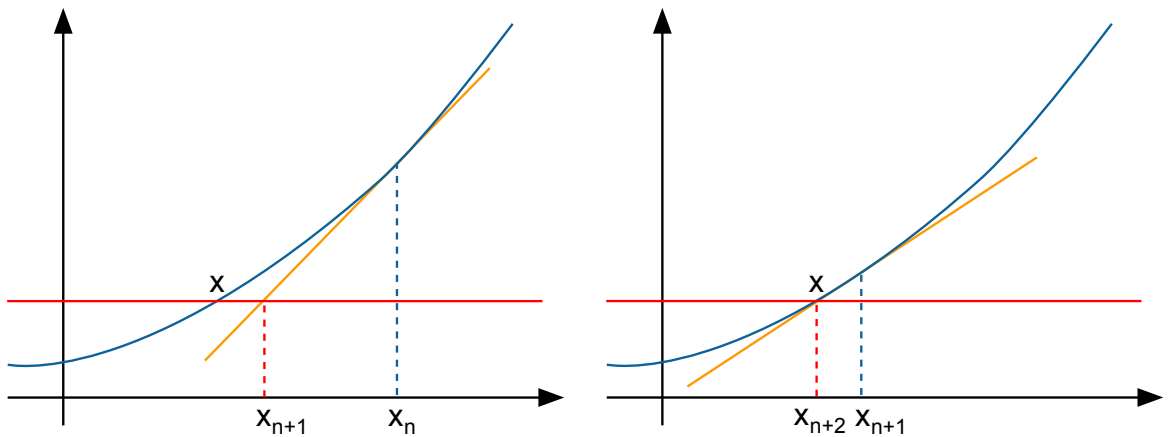
Figure 3.13: A one-dimensional example of linearisation, attempting to find the where a nonlinear equation (solid blue line) has a certain value (solid red line). A point is chosen (dotted blue line) and a linear approximation found at that point (orange line), and a solution is found for the linear approximation (dotted red line). If the dotted red line and dotted blue line are far enough apart that the linearisation may be a poor one, the process is repeated; the left graph shows iteration n, the right graph iteration n+1.

The *Newton-Raphson method* solves equations by *linearising* at a certain point (x_n), algebraically finding the solution (x_{n+1}) to the linearised system, then repeating the process, linearising around (x_{n+1}) and algebraically finding the solution (x_{n+2}), and so on. This is illustrated in figure 3.13. When adapted to solve a nonlinear least squares problem, it is known as the *Gauss-Newton algorithm*.

By combining equation 3.4 on page 43 with equation 3.3 on page 43, we get this equation relating code pseudorange to position:

$$R_r^s = \sqrt{(X^s - X_r)^2 + (Y^s - Y_r)^2 + (Z^s - Z_r)^2} + c\delta_r + \nu_r^s \tag{3.32}$$

and recalling equation 3.21 on page 52, which defined $\boldsymbol{x}_r = \begin{bmatrix} X_r & Y_r & Z_r & c\delta_r \end{bmatrix}^T$, we can define the expected pseudorange as:

$$e^s(\boldsymbol{x}_r) = \sqrt{(X^s - X_r)^2 + (Y^s - Y_r)^2 + (Z^s - Z_r)^2} + c\delta_r \tag{3.33}$$

If we perform a first-order Taylor series expansion around $\boldsymbol{x}_{ref}$, our linear approximation of $e(\boldsymbol{x}_r)$ is:

$$
\begin{aligned}
e^s(\boldsymbol{x}_r) &= e^s(\boldsymbol{x}_{ref}) + \frac{de^s(\boldsymbol{x}_{ref})}{dX_{ref}}(X_r - X_{ref}) + \frac{de^s(\boldsymbol{x}_{ref})}{dY_{ref}}(Y_r - Y_{ref}) \\
&\quad + \frac{de^s(\boldsymbol{x}_{ref})}{dZ_{ref}}(Z_r - Z_{ref}) + \frac{de^s(\boldsymbol{x}_{ref})}{dc\delta_{ref}}(c\delta_r - c\delta_{ref})
\end{aligned}
\tag{3.34}
$$

Or, in matrix form:

$$e^s(\boldsymbol{x}_r) = e^s(\boldsymbol{x}_{ref}) + \begin{bmatrix} \frac{de^s(\boldsymbol{x}_{ref})}{dX_{ref}} & \frac{de^s(\boldsymbol{x}_{ref})}{dY_{ref}} & \frac{de^s(\boldsymbol{x}_{ref})}{dZ_{ref}} & \frac{de^s(\boldsymbol{x}_{ref})}{dc\delta_{ref}} \end{bmatrix} (\boldsymbol{x}_r - \boldsymbol{x}_{ref}) \tag{3.35}$$

Substituting the expected pseudorange into equation 3.32, for multiple satellites and a best fit solution, we would get:

$$
\begin{bmatrix} R_r^1 \\ R_r^2 \\ R_r^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} e^1(\boldsymbol{x}_r) \\ e^2(\boldsymbol{x}_r) \\ e^3(\boldsymbol{x}_r) \\ \vdots \end{bmatrix} + \begin{bmatrix} \nu_r^1 \\ \nu_r^2 \\ \nu_r^3 \\ \vdots \end{bmatrix} \tag{3.36}
$$

Inserting the linearised version of $e^s(\boldsymbol{x}_r)$ from equation 3.35 , the result would be:

$$\begin{bmatrix} R_r^1 \\ R_r^2 \\ R_r^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} e^1(\boldsymbol{x}_{ref}) \\ e^2(\boldsymbol{x}_{ref}) \\ e^3(\boldsymbol{x}_{ref}) \\ \vdots \end{bmatrix} + \boldsymbol{A}(\boldsymbol{x}_r - \boldsymbol{x}_{ref}) + \begin{bmatrix} \nu_r^1 \\ \nu_r^2 \\ \nu_r^3 \\ \vdots \end{bmatrix} \tag{3.37}$$

Where

$$\boldsymbol{A} = \begin{bmatrix} \frac{de^1(\boldsymbol{x}_{ref})}{dX_{ref}} & \frac{de^1(\boldsymbol{x}_{ref})}{dY_{ref}} & \frac{de^1(\boldsymbol{x}_{ref})}{dZ_{ref}} & \frac{de^1(\boldsymbol{x}_{ref})}{dc\delta_{ref}} \\ \frac{de^2(\boldsymbol{x}_{ref})}{dX_{ref}} & \frac{de^2(\boldsymbol{x}_{ref})}{dY_{ref}} & \frac{de^2(\boldsymbol{x}_{ref})}{dZ_{ref}} & \frac{de^2(\boldsymbol{x}_{ref})}{dc\delta_{ref}} \\ \frac{de^3(\boldsymbol{x}_{ref})}{dX_{ref}} & \frac{de^3(\boldsymbol{x}_{ref})}{dY_{ref}} & \frac{de^3(\boldsymbol{x}_{ref})}{dZ_{ref}} & \frac{de^3(\boldsymbol{x}_{ref})}{dc\delta_{ref}} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{3.38}$$

Which can be arranged as:

$$\boldsymbol{A} \quad (\boldsymbol{x}_r - \boldsymbol{x}_{ref}) \;=\; \left(\begin{bmatrix} R_r^1 \\ R_r^2 \\ R_r^3 \\ \vdots \end{bmatrix} - \begin{bmatrix} e^1(\boldsymbol{x}_{ref}) \\ e^2(\boldsymbol{x}_{ref}) \\ e^3(\boldsymbol{x}_{ref}) \\ \vdots \end{bmatrix} \right) \;-\; \begin{bmatrix} \nu_r^1 \\ \nu_r^2 \\ \nu_r^3 \\ \vdots \end{bmatrix} \tag{3.39}$$

$$\boldsymbol{A} \qquad \boldsymbol{x} \qquad = \qquad\qquad \boldsymbol{b} \qquad\qquad - \qquad \boldsymbol{r}$$

The normal equations for this are given by:

$$\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A} \boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{W} \boldsymbol{b} \tag{3.40}$$

Which has the least squares solution (minimising the vector of errors, $\boldsymbol{r}$):

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A} \right)^{-1} \boldsymbol{A}^T \boldsymbol{W} \boldsymbol{b} \tag{3.41}$$

So for this iteration, the least squares solution for $\boldsymbol{x}_r$ is:

$$\boldsymbol{x}_r = \left(\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A} \right)^{-1} \boldsymbol{A}^T \boldsymbol{W} \boldsymbol{b} + \boldsymbol{x}_{ref} \tag{3.42}$$

where $\boldsymbol{W}$ is the measurement weighting matrix. This calculation can be repeatedly performed, each time replacing $\boldsymbol{x}_{ref}$ with $\boldsymbol{x}_r$, until the solution converges - i.e. the amount $\boldsymbol{x}_r$ changes per iteration drops below a certain threshold.

Other techniques for nonlinear regression are available; the methods shown in the previous sections are selected for their simplicity, rather than their computational efficiency. More information on nonlinear regression is available in Seber and Wild (1989).

3.4.3 Kalman filters

Kalman filters perform recursive estimation of linear dynamic systems' hidden states. Extended Kalman filters can operate on nonlinear dynamic systems by linearising around the most recent position estimate, and often perform well on systems amenable to this linearisation. They are commonly used in GPS applications, and especially in sensor fusion. For more information see Welch and Bishop (1995) and for an example of application to GPS see Hide et al. (2007).

3.4.4 Double-differenced carrier measurements

Sections 3.4.1 on page 52 and 3.4.2 on page 54 have given examples of how the GPS equations for C\A code pseudorange measurements can be solved. To solve the equations for the double-differenced carrier pseudorange measurements described in section 3.3.4.2 on page 48, we employ a similar but more complicated technique.

To get the best positioning accuracy, the integer ambiguities are *fixed* - their integer values are determined. One of the most popular techniques for doing this, described by Teunissen (1995), is outlined in this section. First a *float solution* is generated (a least squares solution with integer ambiguities modelled as real numbers); then the float solution is refined to produce fixed integer ambiguities; then a statistical test is applied to determine whether the solution is acceptable.

3.4.4.1   Float solution

Recall equation 3.15 on page 50, which states:

$$\lambda\Phi_{r-b}^{s-1} = \varrho_{r-b}^{s-1} + \lambda N_{r-b}^{s-1} + \mu_{r-b}^{s-1}$$

where ϱ_{r-b}^{s-1} is the double-differenced true distance from rover to satellite, i.e.

$$\varrho_{r-b}^{s-1} = \left(\varrho_r^s - \varrho_r^1\right) - \left(\varrho_b^s - \varrho_b^1\right) \tag{3.43}$$

For simplicity, we define $\varrho_{dd}^s\left(\boldsymbol{\varrho}_r\right)$ as being ϱ_{r-b}^{s-1} as a function of receiver location $\left(\boldsymbol{\varrho}_r = \left[\begin{array}{ccc} X_r & Y_r & Z_r \end{array}\right]^T\right)$. If we perform a Taylor series expansion, around reference point $\boldsymbol{\varrho}_{ref}$ (like in equation 3.35 on page 55) we get a linearised expression for $\varrho_{dd}\left(\boldsymbol{\varrho}_r\right)$, namely:

$$\varrho_{dd}^s\left(\boldsymbol{\varrho}_r\right) = \varrho_{dd}^s\left(\boldsymbol{\varrho}_{ref}\right) + \left[\begin{array}{ccc} \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dZ_{ref}} \end{array}\right]\left(\boldsymbol{\varrho}_r - \boldsymbol{\varrho}_{ref}\right) \tag{3.44}$$

Substituting this into equation 3.15 on page 50, we get the linearised equation:

$$\lambda\Phi_{r-b}^{s-1} = \varrho_{dd}^s\left(\boldsymbol{\varrho}_{ref}\right) + \left[\begin{array}{ccc} \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dZ_{ref}} \end{array}\right]\left(\boldsymbol{\varrho}_r - \boldsymbol{\varrho}_{ref}\right) + \lambda N_{r-b}^{s-1} + \mu_{r-b}^{s-1} \tag{3.45}$$

which can be rearranged to:

$$\lambda\Phi_{r-b}^{s-1} = \varrho_{dd}^s\left(\boldsymbol{\varrho}_{ref}\right) + \left[\begin{array}{ccc|c} \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^s(\boldsymbol{\varrho}_r)}{dZ_{ref}} & \lambda \end{array}\right]\left[\begin{array}{c} \boldsymbol{\varrho}_r - \boldsymbol{\varrho}_{ref} \\ \hline N_{r-b}^{s-1} \end{array}\right] + \mu_{r-b}^{s-1} \tag{3.46}$$

For multiple satellites (each of which would have their own integer ambiguity) this would be expressed as:

$$
\begin{bmatrix} \lambda\Phi_{r-b}^{2-1} \\ \lambda\Phi_{r-b}^{3-1} \\ \lambda\Phi_{r-b}^{4-1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \varrho_{dd}^2\left(\boldsymbol{\varrho}_{ref}\right) \\ \varrho_{dd}^3\left(\boldsymbol{\varrho}_{ref}\right) \\ \varrho_{dd}^4\left(\boldsymbol{\varrho}_{ref}\right) \\ \vdots \end{bmatrix} + \boldsymbol{A} \begin{bmatrix} \dfrac{\boldsymbol{\varrho}_r - \boldsymbol{\varrho}_{ref}}{N_{r-b}^{2-1}} \\ N_{r-b}^{3-1} \\ N_{r-b}^{4-1} \\ \vdots \end{bmatrix} + \begin{bmatrix} \mu_{r-b}^{2-1} \\ \mu_{r-b}^{3-1} \\ \mu_{r-b}^{4-1} \\ \vdots \end{bmatrix} \tag{3.47}
$$

where:

$$
\boldsymbol{A} = \begin{bmatrix} \dfrac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dX_{ref}} & \dfrac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dY_{ref}} & \dfrac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dZ_{ref}} & \lambda & 0 & 0 & \dots \\ \dfrac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dX_{ref}} & \dfrac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dY_{ref}} & \dfrac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dZ_{ref}} & 0 & \lambda & 0 & \dots \\ \dfrac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dX_{ref}} & \dfrac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dY_{ref}} & \dfrac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dZ_{ref}} & 0 & 0 & \lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.48}
$$

which can be rearranged to:

$$
\boldsymbol{A} \begin{bmatrix} \dfrac{\boldsymbol{\varrho}_r - \boldsymbol{\varrho}_{ref}}{N_{r-b}^{2-1}} \\ N_{r-b}^{3-1} \\ N_{r-b}^{4-1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \lambda\Phi_{r-b}^{2-1} \\ \lambda\Phi_{r-b}^{3-1} \\ \lambda\Phi_{r-b}^{4-1} \\ \vdots \end{bmatrix} - \begin{bmatrix} \varrho_{dd}^2\left(\boldsymbol{\varrho}_{ref}\right) \\ \varrho_{dd}^3\left(\boldsymbol{\varrho}_{ref}\right) \\ \varrho_{dd}^4\left(\boldsymbol{\varrho}_{ref}\right) \\ \vdots \end{bmatrix} - \begin{bmatrix} \mu_{r-b}^{2-1} \\ \mu_{r-b}^{3-1} \\ \mu_{r-b}^{4-1} \\ \vdots \end{bmatrix} \tag{3.49}
$$

$$
\boldsymbol{A} \qquad \boldsymbol{x} \qquad = \qquad \boldsymbol{b} \qquad - \qquad \boldsymbol{r}
$$

Which can be solved, to minimise the least squares residual vector, by solving the normal equations:

$$
\hat{\boldsymbol{x}} = \left(\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{W}\boldsymbol{b} \tag{3.50}
$$

where $\boldsymbol{W}$ is the measurement weighting matrix (described in section 3.4.5.1 on page 66); the top three values in the vector $\hat{\boldsymbol{x}}$ are $\left(\boldsymbol{\varrho}_r - \boldsymbol{\varrho}_{ref}\right)$ and the remaining values are double-differenced integer ambiguity estimates.

However, performing this calculation with a single set of observables presents a problem:

with seven satellites in view, there will be 6 double-differenced carrier pseudoranges, but 6 unknown integer ambiguities and 3 unknown elements in the receiver position vector. In other words, 6 equations with 9 unknowns, meaning the system has an infinite number of solutions.

One option to work around this problem is to use several consecutive sets of observables (observations over several *epochs*); as the double-differenced integer ambiguity is constant between epochs, this increases the number of equations without increasing the number of unknowns. For example, if the receiver is stationary for 2 epochs with 7 satellites in view, there will be 12 double-differenced carrier pseudoranges, but still only 9 unknowns, meaning the system is overdetermined and a least squares solution can be found.

In this case, matrices from equation 3.49 would become:

$$
\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} = \left[\begin{array}{ccc|cccc}
\frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dZ_{ref}} & \lambda & 0 & 0 & \ldots \\
\frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dZ_{ref}} & 0 & \lambda & 0 & \ldots \\
\frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dZ_{ref}} & 0 & 0 & \lambda & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\
\hline
\frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dZ_{ref}} & \lambda & 0 & 0 & \ldots \\
\frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dZ_{ref}} & 0 & \lambda & 0 & \ldots \\
\frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dX_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dY_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dZ_{ref}} & 0 & 0 & \lambda & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array} \right]
\tag{3.51}
$$

$$
\boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}
\tag{3.52}
$$

$$
\boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \end{bmatrix}
\tag{3.53}
$$

$$
\begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \end{bmatrix}
\tag{3.54}
$$

where the upper partition of matrix $\boldsymbol{A}$ contains the partial derivatives for the first epoch and the lower partition of the matrix contains the partial derivatives for the second epoch. Likewise $\boldsymbol{b}$ (the vector of double-differenced pseudoranges minus expected double-differenced pseudoranges at the reference location) and $\boldsymbol{r}$ (the vector of residuals) contain a set of data from each epoch.

In practice, for single-frequency ambiguity resolution, the data from 2 consecutive epochs may be little different, as interference may be shared and satellites will not have moved substantially, leaving the matrix of GPS equations poorly conditioned. The length of stationary initialisation required to produce reliable results will be addressed later.

3.4.4.2   Fixed solution

The most widely-used method for generating fixed ambiguities from the float solution is *Least squares AMBuqiuty Decorrelation Adjustment (LAMBDA)*, described in Teunissen (1995).

When the float solution is being calculated, a combined vector of baseline elements and integer ambiguities is used, such as in equation 3.47 on page 59. For the fixed solution, the vector is separated, allowing the baseline to be handled as real numbers and the ambiguities as integers. Teunissen (1995) states this as:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{b} + \boldsymbol{e} \tag{3.55}$$

where $\boldsymbol{y}$ is the matrix of double-differenced measurements minus expected values at ϱ_{ref}; $\boldsymbol{a}$ and $\boldsymbol{A}$ are the ambiguities and the relation between ambiguities and pseudoranges; $\boldsymbol{b}$ and $\boldsymbol{B}$ are the baseline vector and the relation between it and the pseudoranges; and $\boldsymbol{e}$ is the vector of residuals (the sum-of-squares of which will be minimised).

In other words:

$$
\boldsymbol{y} = \begin{bmatrix} \lambda\Phi_{r-b}^{2-1} \\ \lambda\Phi_{r-b}^{3-1} \\ \lambda\Phi_{r-b}^{4-1} \\ \vdots \end{bmatrix} - \begin{bmatrix} \varrho_{dd}^{2}\left(\boldsymbol{\varrho}_{ref}\right) \\ \varrho_{dd}^{3}\left(\boldsymbol{\varrho}_{ref}\right) \\ \varrho_{dd}^{4}\left(\boldsymbol{\varrho}_{ref}\right) \\ \vdots \end{bmatrix} \tag{3.56}
$$

$$
\boldsymbol{A} = \begin{bmatrix} \lambda & 0 & 0 & \dots \\ 0 & \lambda & 0 & \dots \\ 0 & 0 & \lambda & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.57}
$$

$$
\boldsymbol{a} = \begin{bmatrix} N_{r-b}^{2-1} \\ N_{r-b}^{3-1} \\ N_{r-b}^{4-1} \\ \vdots \end{bmatrix} \tag{3.58}
$$

$$
\boldsymbol{B} = \begin{bmatrix} \frac{d\varrho_{dd}^{2}(\boldsymbol{\varrho}_{r})}{dX_{ref}} & \frac{d\varrho_{dd}^{2}(\boldsymbol{\varrho}_{r})}{dY_{ref}} & \frac{d\varrho_{dd}^{2}(\boldsymbol{\varrho}_{r})}{dZ_{ref}} \\ \frac{d\varrho_{dd}^{3}(\boldsymbol{\varrho}_{r})}{dX_{ref}} & \frac{d\varrho_{dd}^{3}(\boldsymbol{\varrho}_{r})}{dY_{ref}} & \frac{d\varrho_{dd}^{3}(\boldsymbol{\varrho}_{r})}{dZ_{ref}} \\ \frac{d\varrho_{dd}^{4}(\boldsymbol{\varrho}_{r})}{dX_{ref}} & \frac{d\varrho_{dd}^{4}(\boldsymbol{\varrho}_{r})}{dY_{ref}} & \frac{d\varrho_{dd}^{4}(\boldsymbol{\varrho}_{r})}{dZ_{ref}} \\ \vdots & \vdots & \vdots \end{bmatrix} \tag{3.59}
$$

$$
\boldsymbol{b} = \boldsymbol{\varrho}_{r} - \boldsymbol{\varrho}_{ref} \tag{3.60}
$$

Finding a least squares solution, as performed in section 3.47 on page 59, is expressed (in the terminology of Teunissen (1995)) as:

$$
\min_{\boldsymbol{a},\boldsymbol{b}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|_{Q_y}^{2} \;\; \text{with } \boldsymbol{a} \in \mathbb{R}^{\mathrm{n}}, \boldsymbol{b} \in \mathbb{R}^{3} \tag{3.61}
$$

where $Q_y = W^{-1}$ (i.e. the inverse of the weighting matrix, described in more detail in section 3.4.5.1 on page 66); $\|\boldsymbol{f}\|_{Q_y}^{2} = \boldsymbol{f}^{T}Q_y^{-1}\boldsymbol{f}$; and $\boldsymbol{a} \in \mathbb{R}^{\mathrm{n}}, \boldsymbol{b} \in \mathbb{R}^{3}$ means that $\boldsymbol{a}$ is a vector of real numbers n items long; while $\boldsymbol{b}$ is a vector of real numbers 3 items long.

The solutions to this are denoted $\hat{\boldsymbol{a}}$ and $\hat{\boldsymbol{b}}$, and the minimised residual as $\hat{\boldsymbol{e}}$, i.e.

$$\hat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{A}\hat{\boldsymbol{a}} - \boldsymbol{B}\hat{\boldsymbol{b}} \tag{3.62}$$

Finding the fixed solution is a similar problem to finding the float solution, but with the constraint that ambiguity values must be not just real numbers, but integers. This can be stated as:

$$\min_{\boldsymbol{a},\boldsymbol{b}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|^2_{Q_y} \ \text{ with } \boldsymbol{a} \in \mathbb{Z}^n, \boldsymbol{b} \in \mathbb{R}^3 \tag{3.63}$$

where $\boldsymbol{a} \in \mathbb{Z}^n$ means that $\boldsymbol{a}$ is a vector of integers n items long.

Because $\mathbb{Z}^n$ is a subset of $\mathbb{R}^n$, we know that the best solution to equation 3.63 will never have a smaller residual than the best solution to equation 3.61 ; in other words, the best fixed solution will be the one closest to the float solution according to the metric Q_y. We can apply an orthogonal decomposition:

$$\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|^2_{Q_y} = \|\hat{\boldsymbol{e}}\|^2_{Q_y} + \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|^2_{Q_{\hat{a}}} + \left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|^2_{Q_{\hat{b}(a)}} \tag{3.64}$$

where $Q_{\hat{a}} = \left(\boldsymbol{A}^T Q_y^{-1} \boldsymbol{A}\right)^{-1}$ and hence 3.63 becomes

$$\min_{\boldsymbol{a},\boldsymbol{b}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|^2_{Q_y} = \min \|\hat{\boldsymbol{e}}\|^2_{Q_y} + \min_{\boldsymbol{a}} \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|^2_{Q_{\hat{a}}} + \min_{\boldsymbol{b}} \left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|^2_{Q_{\hat{b}(a)}} \tag{3.65}$$

where $\hat{\boldsymbol{e}}$ is constant and $\min_{\boldsymbol{a}} \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|^2_{Q_{\hat{a}}}$ and $\min_{\boldsymbol{b}} \left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|^2_{Q_{\hat{b}(a)}}$ are, at best, zero. As both $\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right)$ and $\boldsymbol{b}$ consist of three real values, the best solution will always be $\boldsymbol{b} = \hat{\boldsymbol{b}}\left(\boldsymbol{a}\right)$; and because $\hat{\boldsymbol{e}}$ is a constant, it has no influence on where the minimum is found. In other words, solving equation 3.63 is equivalent to solving

$$\min_{\boldsymbol{a}} \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|^2_{Q_{\hat{a}}} \ \text{ with } \boldsymbol{a} \in \mathbb{Z}^n \tag{3.66}$$

In order to solve this equation, Teunissen (1995) proposes transforming $Q_{\hat{a}}$ to minimise

the correlation between ambiguities, while preserving the problem's integer nature. This decorrelation allows a fast a search in the region of $\hat{\boldsymbol{a}}$ to find the ambiguity with the least squared error. The transform is typically defined as:

$$\boldsymbol{z} = \boldsymbol{Z}^T \boldsymbol{a}, \; \hat{\boldsymbol{z}} = \boldsymbol{Z}^T \hat{\boldsymbol{a}}, \; Q_{\hat{z}} = \boldsymbol{Z}^T Q_{\hat{a}} \boldsymbol{Z} \tag{3.67}$$

converting equation 3.66 to:

$$\min_{\boldsymbol{z}} \| \hat{\boldsymbol{z}} - \boldsymbol{z} \|_{Q_{\hat{z}}}^2 \; \text{ with } \boldsymbol{z} \in \mathbb{Z}^{\mathrm{n}} \tag{3.68}$$

where transformation matrix $\boldsymbol{Z}$ is an integer matrix similar to the $\boldsymbol{L}$ matrix in an $\boldsymbol{LDL}^T$ (Cholesky) decomposition (Press et al., 2007, p.100). More information on the determination of the $\boldsymbol{Z}$ matrix can be found in Teunissen (1995), Hofmann-Wellenhof et al. (2008), and Strang and Borre (1997). A search space is defined:

$$\Omega\left(\chi^2\right) = \left\{ \| \hat{\boldsymbol{z}} - \boldsymbol{z} \|_{Q_{\hat{z}}}^2 \leq \chi^2 \text{ with } \boldsymbol{z} \in \mathbb{Z}^{\mathrm{n}} \right\} \tag{3.69}$$

and a sequential conditional search is performed to identify all integer ambiguity vectors within the search space. The size of scalar value χ^2 dictates the size of the search space; too large and the search will be very time-consuming; too small and the search may not find any integer vectors at all. As described by Teunissen et al. (1997), a fair approximation of $\boldsymbol{z}$ (such as $\hat{\boldsymbol{z}}$ rounded to the nearest integer) can be found and used to calculate a value of χ^2 which contains at least one solution. Similarly, by rounding one integer of $\boldsymbol{z}$ to its second-nearest integer and the others to their first-nearest integers, a value of χ^2 containing at least two solutions is found.

Once an integer estimate ($\check{\boldsymbol{z}}$) has been found, the corresponding ambiguity is given by:

$$\check{\boldsymbol{a}} = \left(\boldsymbol{Z}^T\right)^{-1} \check{\boldsymbol{z}} \tag{3.70}$$

and the baseline solution ($\check{\boldsymbol{b}}$) for the integer estimate is given by:

$$\check{\boldsymbol{b}} = \hat{\boldsymbol{b}} - Q_{\hat{b}\hat{a}}Q_{\hat{a}}(\hat{\boldsymbol{a}} - \check{\boldsymbol{a}}) = \hat{\boldsymbol{b}} - Q_{\hat{b}\hat{z}}Q_{\hat{z}}(\hat{\boldsymbol{z}} - \check{\boldsymbol{z}}) \qquad (3.71)$$

3.4.4.3 Ambiguity validation

Under poor conditions, the least squares ambiguity solution may not be the correct one; when this happens, it is useful to gather additional observations, to ensure the ambiguity solution selected is the correct one. A test is required to determine whether this is the case. A number of methods have been developed to test this, as outlined in Verhagen (2004); one of the most popular tests is the *squared norm ratio test*, which takes the two best ambiguity solutions and calculates the ratio of their squared norms (i.e. the values minimised by equation 3.66 on page 63) which is compared to a *critical value* - i.e.

$$\frac{\|\hat{\boldsymbol{a}} - \boldsymbol{a}_2\|^2_{Q_{\hat{a}}}}{\|\hat{\boldsymbol{a}} - \boldsymbol{a}_1\|^2_{Q_{\hat{a}}}} > k \qquad (3.72)$$

where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the best and second best ambiguity solutions respectively. k, the critical value, is determined empirically; a higher threshold reduces the risk of erroneous ambiguity resolution, but increases the time taken to resolve integer ambiguities; and an excessively high threshold may preclude ambiguity acceptance all together. Hofmann-Wellenhof et al. (2008) cites papers recommending weights from 1.5 to 10. Verhagen (2004) uses geometry-free simulated data and selects a value such that there is a 1% chance of erroneous acceptance, arriving at 5.73 as a suitable value.

Another popular test is the *F-ratio test*, which takes the ratio of the squared residuals (i.e. the values minimised by equation 3.66 on page 63) and compares it to a critical value, i.e.

$$\frac{\|\hat{\boldsymbol{e}}\|^2_{Q_y} + \|\hat{\boldsymbol{a}} - \boldsymbol{a}_2\|^2_{Q_{\hat{a}}}}{\|\hat{\boldsymbol{e}}\|^2_{Q_y} + \|\hat{\boldsymbol{a}} - \boldsymbol{a}_1\|^2_{Q_{\hat{a}}}} > k \qquad (3.73)$$

which, for the same value of k is always more conservative than the squared norm ratio test, and the larger the float solution residual, the more conservative this test. This

test is used by Takasu and Yasuda (2008) with a ratio of 3.0 with data from several real receivers, achieving no erroneous acceptance with certain antennas, and less than 0.5% with any antenna. The test is also successfully used by Weisenburger (1997). Verhagen (2004) uses geometry-free simulated data and selects a value such that there is a 1% chance of erroneous acceptance, arriving at 4.23 as a suitable value.

Another means of ambiguity validation, which we will refer to as the *variance threshold test*, is described in Hatch (1991). The test calculates the estimated variance of each candidate ambiguity solution divided by the number of satellites (minus 4), and compares it to a critical value:

$$\frac{\|e\|_{Q_y}^2}{n-4} > k \tag{3.74}$$

where n is the number of satellites in view and k is the critical value. This test is performed for each candidate solution, and solutions which fail the test are discarded; if there is more than one solution which passes the test, observations are gathered until only one solution remains. Unlike the squared norm ratio test and the F-ratio test (where division cancels out the effect of the variance assumed for the range measurements), the performance of this test depends on the assumed variances for range measurements, as well as on the critical value chosen. Assuming a lower variance will lead to faster elimination of outlying candidate solutions, allowing faster ambiguity resolution; but assuming the variance to be too low will cause the correct solution to be wrongly eliminated. Also, if the best and second-best solutions are close together, but on either side of k, the best solution will be accepted even though the second-best solution is very close to it.

3.4.5 Implementation details

3.4.5.1 Selection of weighting matrices

The examples above all employ weighting matrices, denoted by W or $Q_y = W^{-1}$. When finding the least squared error for several signals, the weighting matrix allows the different errors to be scaled before they are minimised; for example if combining

Coarse Acquisition (C\A) code measurements (standard deviation 60cm) with carrier phase measurements (standard deviation 5mm) minimising the squared errors on the low-noise measurements is more important than minimising the squared errors on the high-noise measurements. When the measurements' standard deviations are taken into account properly, the least-squares solution is also the *maximum likelihood* solution; to achieve this, the weighting matrix used is the inverse of the covariance matrix of the observations (which is sometimes called the variance-covariance matrix).

The variance-covariance matrix is, as the name suggests, a matrix containing the variances of each measurement (on the diagonal), and the covariances between each measurement (elsewhere). The variance of a measurement is the square of its standard deviation; and covariances are zero if two measurements are uncorrelated.

In a system with several uncorrelated measurements, all with the same standard deviations, the variance-covariance matrix can be written as:

$$\Sigma_{\boldsymbol{c}} = \sigma_0^2 \begin{bmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.75}$$

where σ_0^2 is the variance, and the identity matrix is known as the *cofactor matrix*. In a system with several uncorrelated measurements with different standard deviations, the measurements' variances populate the diagonal of the variance-covariance matrix. For example, a system with two measurements known to have a standard deviation of 60cm and two measurements known to have a standard deviation of 5mm; and where the four measurements are uncorrelated; would have a variance-covariance matrix of:

$$\Sigma_{\boldsymbol{c}} = \begin{bmatrix} 0.6^2 & 0 & 0 & 0 \\ 0 & 0.6^2 & 0 & 0 \\ 0 & 0 & 0.005^2 & 0 \\ 0 & 0 & 0 & 0.005^2 \end{bmatrix} \tag{3.76}$$

Though GPS measurements before single- or double-differencing are uncorrelated, double-differenced GPS measurements are correlated because the same reference satellite is subtracted from all measurements; for example, a 3mm error on all double-differenced measurements could indicate the same noise affecting all satellites except the reference satellite; or it could indicate noise affecting the reference satellite only. By including the covariances between measurements into the variance-covariance matrix, this can be accurately taken into account.

We can see how double-differencing introduces covariance by considering double-differencing in matrix form. The process of double-differencing a set of measurements (as described in section 3.3.4.2 on page 48) can be represented in matrix form as:

$$
\begin{bmatrix} \lambda\Phi_{r-b}^{2-1} \\ \lambda\Phi_{r-b}^{3-1} \\ \lambda\Phi_{r-b}^{4-1} \\ \vdots \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 & \ldots & 1 & -1 & 0 & 0 & \cdots \\ -1 & 0 & 1 & 0 & \ldots & 1 & 0 & -1 & 0 & \cdots \\ -1 & 0 & 0 & 1 & \ldots & 1 & 0 & 0 & -1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \lambda\Phi_r^1 \\ \lambda\Phi_r^2 \\ \lambda\Phi_r^3 \\ \lambda\Phi_r^4 \\ \vdots \\ \lambda\Phi_b^1 \\ \lambda\Phi_b^2 \\ \lambda\Phi_b^3 \\ \lambda\Phi_b^4 \\ \vdots \end{bmatrix} \tag{3.77}
$$

$$\boldsymbol{a} \qquad = \qquad\qquad\qquad \boldsymbol{B} \qquad\qquad\qquad\qquad \boldsymbol{c}$$

Taking the variance-covariance matrix of equation 3.75 , and the law of covariance propagation, the covariance matrix of the double-differenced measurements is:

$$\Sigma_{\boldsymbol{a}} = \boldsymbol{B}\Sigma_{\boldsymbol{c}}\boldsymbol{B}^T \tag{3.78}$$

For example, with 5 satellites in view, the covariance of the double-differenced carrier phase measurements would be:

$$
\Sigma_a = \sigma_0^2
\begin{bmatrix}
-1 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 \\
-1 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
-1 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\
-1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 \\
-1 & 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1
\end{bmatrix}^{T}
$$

$$
= \sigma_0^2
\begin{bmatrix}
4 & 2 & 2 & 2 \\
2 & 4 & 2 & 2 \\
2 & 2 & 4 & 2 \\
2 & 2 & 2 & 4
\end{bmatrix}
\tag{3.79}
$$

As you may note, σ_0^2 - the covariance assumed for a single measurement - remains outside the equation as a scaling factor. When solving the equations to find a least-squares solution, scaling all weights by a constant factor does not affect the results; hence it is possible to use the cofactor matrix instead of the variance-covariance matrix. Hofmann-Wellenhof et al. (2008) use the cofactor matrix, denoting it Q_y; while Strang and Borre (1997) use the variance-covariance matrix, denoting it Σ_b.

In the event that several epochs of double-differenced data are used, matrix $\boldsymbol{B}$ from equation 3.77 on page 68 is replicated in a block-diagonal manner, i.e.

$$
\begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \\ \boldsymbol{a}_4 \end{bmatrix}
=
\begin{bmatrix} \boldsymbol{B} & 0 & 0 & \dots \\ 0 & \boldsymbol{B} & 0 & \dots \\ 0 & 0 & \boldsymbol{B} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}
\begin{bmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \boldsymbol{c}_3 \\ \boldsymbol{c}_4 \end{bmatrix}
\tag{3.80}
$$
$$
\boldsymbol{d} \quad = \quad \boldsymbol{E} \quad\quad \boldsymbol{f}
$$

Giving the covariance matrix of double-differenced measurements as:

$$
\Sigma_{\boldsymbol{d}} = \boldsymbol{E}\Sigma_{\boldsymbol{c}}\boldsymbol{E}^T
\tag{3.81}
$$

Which will itself be block diagonal - for example, with 5 satellites in view for 2 epochs, the covariance of the double-differenced measurements would be:

$$
\Sigma_{\boldsymbol{d}} = \sigma_0^2
\begin{bmatrix}
4 & 2 & 2 & 2 & 0 & 0 & 0 & 0 \\
2 & 4 & 2 & 2 & 0 & 0 & 0 & 0 \\
2 & 2 & 4 & 2 & 0 & 0 & 0 & 0 \\
2 & 2 & 2 & 4 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 4 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 2 & 4 & 2 & 2 \\
0 & 0 & 0 & 0 & 2 & 2 & 4 & 2 \\
0 & 0 & 0 & 0 & 2 & 2 & 2 & 4
\end{bmatrix}
\tag{3.82}
$$

In summary, when working with undifferenced measurements the weighting matrix would be the inverse of Σ_c, as given by equation 3.75 on page 67; for double-differenced measurements the weighting matrix would be the inverse of Σ_a, given by equation 3.78 on page 68; and for double-differenced measurements over multiple epochs the weighting matrix would be the inverse of Σ_d, given by equation 3.81 .

3.4.5.2 Combining normal equations across epochs

When combining data gathered over several epochs, as described in section 3.4.4.1 on page 60, large block-diagonal covariance matrices can result; while equation 3.82 shows a 64-element matrix to deal with 5 satellites in view for 2 epochs, the matrix size being $((\text{Number of satellites} - 1) \times \text{Number of epochs})^2$; tracking 7 satellites at 1 measurement per second for 1 hour would produce a 21600×21600 matrix, of which 99.97% of the elements would be zero.

This can be avoided because, as the covariance matrix is block-diagonal, the normal equations sum - for example, equation 3.54 on page 60 stated:

$$
\begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} - \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \end{bmatrix}
$$
$$
\boldsymbol{A} \quad \boldsymbol{x} = \boldsymbol{b} - \boldsymbol{r}
$$

which is solved from the normal equations, stated in equation 3.40 on page 56 as:

$$
\boldsymbol{A}^T \boldsymbol{W} \boldsymbol{A} \boldsymbol{x} = \boldsymbol{A}^T \boldsymbol{W} \boldsymbol{b}
$$

The block diagonal weighting matrix, is expressed as:

$$
\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_s & 0 \\ 0 & \boldsymbol{W}_s \end{bmatrix} \tag{3.83}
$$

where $\boldsymbol{W}_s$ is the weight matrix of a single epoch. Hence, the normal equations can be

expanded to give:

$$\begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{W}_s & 0 \\ 0 & \boldsymbol{W}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix}^T \begin{bmatrix} \boldsymbol{W}_s & 0 \\ 0 & \boldsymbol{W}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix} \qquad (3.84)$$

$$= \begin{bmatrix} \boldsymbol{A}_1^T & \boldsymbol{A}_2^T \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_s & 0 \\ 0 & \boldsymbol{W}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{A}_1^T & \boldsymbol{A}_2^T \end{bmatrix} \begin{bmatrix} \boldsymbol{W}_s & 0 \\ 0 & \boldsymbol{W}_s \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix}$$
$$(3.85)$$

$$= \left(\boldsymbol{A}_1^T \boldsymbol{W}_s \boldsymbol{A}_1 + \boldsymbol{A}_2^T \boldsymbol{W}_s \boldsymbol{A}_2 \right) \boldsymbol{x} = \left(\boldsymbol{A}_1^T \boldsymbol{W}_s \boldsymbol{b}_1 + \boldsymbol{A}_2^T \boldsymbol{W}_s \boldsymbol{b}_2 \right) \qquad (3.86)$$

where $\left(\boldsymbol{A}_1^T \boldsymbol{W}_s \boldsymbol{A}_1 + \boldsymbol{A}_2^T \boldsymbol{W}_s \boldsymbol{A}_2 \right)$ is known as the normal matrix and $\left(\boldsymbol{A}_1^T \boldsymbol{W}_s \boldsymbol{b}_1 + \boldsymbol{A}_2^T \boldsymbol{W}_s \boldsymbol{b}_2 \right)$ is known as the normal vector. The implication of the above is that, regardless of the number of epochs of data being processed, only the current epoch and running total normal matrix and normal vector need to be stored.

3.4.5.3 Impact of linearisation

A key technique used by the solution methods outlined previously is linearisation; replacing nonlinear equations with linear approximations generated by Taylor series expansion around a reference point. Precisely at the reference point, the nonlinear equation and its linear approximation will be perfectly matched; as the reference point gets further away, the linear approximation will become less accurate.

For example, equation 3.33 on page 55 and its linear approximation, equation 3.34 on page 55, are:

$$e^s \left(\boldsymbol{x}_r \right) = \sqrt{ \left(X^s - X_r \right)^2 + \left(Y^s - Y_r \right)^2 + \left(Z^s - Z_r \right)^2 } + c\delta_r$$

$$\begin{aligned} e^s \left(\boldsymbol{x}_r \right) =\ & e^s \left(\boldsymbol{x}_{ref} \right) + \frac{de^s \left(\boldsymbol{x}_{ref} \right)}{dX_{ref}} \left(X_r - X_{ref} \right) + \frac{de^s \left(\boldsymbol{x}_{ref} \right)}{dY_{ref}} \left(Y_r - Y_{ref} \right) \\ & + \frac{de^s \left(\boldsymbol{x}_{ref} \right)}{dZ_{ref}} \left(Z_r - Z_{ref} \right) + \frac{de^s \left(\boldsymbol{x}_{ref} \right)}{dc\delta_{ref}} \left(c\delta_r - c\delta_{ref} \right) \end{aligned}$$

Satellite	A	B	C	D	E
X position (m)	13073734.9	14853964.5	-2144251.3	24170275.7	15156391.3
Y position (m)	-9930918.1	-5969290.6	-16597147.4	1179817.3	17055129.9
Z position (m)	20475819.0	21364402.1	20578629.3	11281649.4	14151774.3

Table 3.2: ECEF positions of satellites observed by a receiver at 3899787.0, -106104.2, 5029133.6.

Satellite	A	B	C	D	E
Error at 10m	2.8685E-6	3.3155E-6	2.1942E-6	3.6918E-6	4.3027E-6
Error at 31.6m	2.8670E-5	3.3211E-5	2.1942E-5	3.6959E-5	4.2994E-5
Error at 100m	0.0002867	0.0003321	0.0002194	0.000369	0.0004299
Error at 316.2m	0.0028668	0.0033213	0.0021945	0.0036956	0.0042993
Error at 1000m	0.0286683	0.0332138	0.0219447	0.0369563	0.0429936

Table 3.3: Linearisation errors (in metres) with increasing distance from reference point.

To assess the level of error present in such approximations, the linear approximation was compared to the nonlinear equation with a number of different reference points. Representative satellite location data is shown in table 3.2 and the differences between equations 3.33 on page 55 and 3.34 on page 55, calculated by simple subtraction, are shown in table 3.3. When the system is linearised around a point within 100m of the true position, the error introduced by linearisation is less than 1mm.

3.4.5.4 Calculation of Dilution Of Precision (DOP)

As shown by figure 3.10 on page 45, a receiver movement of 1 metre may change a measured pseudorange by anything from 0 to 100 centimetres, depending on satellite positions; changes of 90 centimetres and 30 centimetres are illustrated.

When calculating receiver position from pseudoranges, the relationship is inverted; for some satellite positions, a pseudorange change of 9 centimetres implies a movement of 10 centimetres; for other positions, it implies a movement of 30 centimetres. This also applies to noise; Gaussian noise with a covariance of 9 centimetres may create position noise of covariance 10 centimetres, or 30 centimetres, depending on satellite positions. In other words, the positions of satellites in the sky can have a multiplicative effect on position noise levels.

To assess this effect, GPS receivers calculate a Dilution Of Precision (DOP) when calculating positions. There are several figures that can be calculated, including Horizontal Dilution Of Precision (HDOP), which quantifies east and north position noise levels; Vertical Dilution Of Precision (VDOP), for vertical noise levels; Position Dilution Of Precision (PDOP), which combines HDOP and VDOP; Time Dilution Of Precision (TDOP), for time noise levels; and Geometric Dilution Of Precision (GDOP), which combines PDOP and TDOP.

The relationship between receiver movement and pseudorange change is assumed to be linear over the range of likely noise levels (a reasonable assumption, given the calculations from section 3.4.5.3 on page 72) and hence can be expressed using partial derivatives; a matrix is constructed of the partial derivatives of the expected pseudorange, with respect to east, north, up, and time (similar, you may note, to 3.38 on page 56).

$$\boldsymbol{A} = \begin{bmatrix} \frac{de^1(\boldsymbol{x}_r)}{dEast} & \frac{de^1(\boldsymbol{x}_r)}{dNorth} & \frac{de^1(\boldsymbol{x}_r)}{dUp} & \frac{de^1(\boldsymbol{x}_r)}{dc\delta_r} \\ \frac{de^2(\boldsymbol{x}_r)}{dEast} & \frac{de^2(\boldsymbol{x}_r)}{dNorth} & \frac{de^2(\boldsymbol{x}_r)}{dUp} & \frac{de^2(\boldsymbol{x}_r)}{dc\delta_r} \\ \frac{de^3(\boldsymbol{x}_r)}{dEast} & \frac{de^3(\boldsymbol{x}_r)}{dNorth} & \frac{de^3(\boldsymbol{x}_r)}{dUp} & \frac{de^3(\boldsymbol{x}_r)}{dc\delta_r} \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} \tag{3.87}$$

where $e^1(\boldsymbol{x}_r)$ is the expected pseudorange from the receiver (at position $\boldsymbol{x}_r$) to satellite 1 (as defined by 3.33 on page 55); and the expressions containing dEast, dNorth, dUp and $dc\delta_r$ are that equation differentiated with respect to receiver movement east, north, up, and changing clock bias respectively.

Assuming the covariances of the observations to be 1 - making the covariance matrix of the undifferenced observations the identity matrix - then due to the law of covariance propagation, the covariance matrix of the position is given by:

$$\Sigma_{\boldsymbol{enu}} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1} \tag{3.88}$$

which is a 4-by-4 matrix, the diagonals values of which are the east, north, up, and time covariances. Naming these Σ_e, Σ_n, Σ_u and Σ_t respectively, the various values for
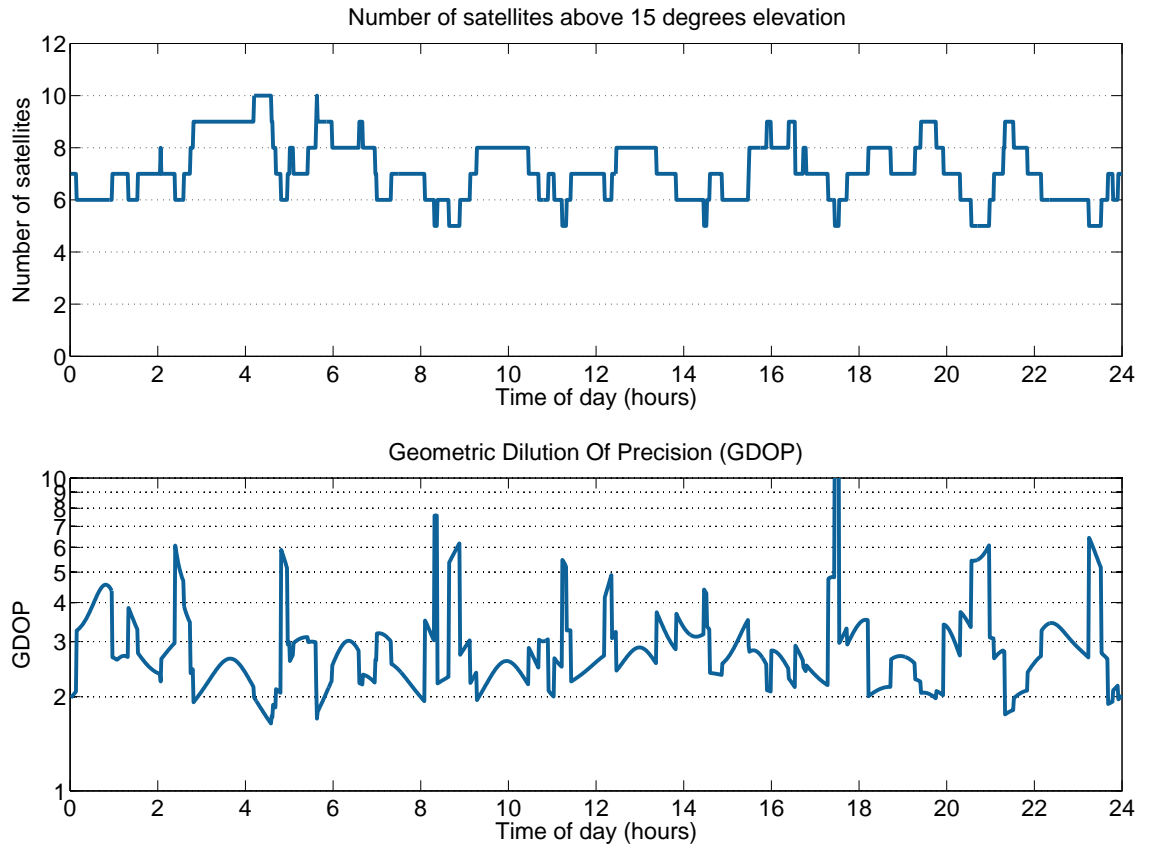
Figure 3.14: Number of satellites in view (top) and Geometric Dilution Of Precision (GDOP) (bottom) against time, over 24 hours, using a 15° elevation mask. Note the logarithmic scale of the lower graph.

DOP are calculated as follows:

$$HDOP = \sqrt{\Sigma_e + \Sigma_n} \tag{3.89}$$

$$VDOP = \sqrt{\Sigma_u} \tag{3.90}$$

$$PDOP = \sqrt{\Sigma_e + \Sigma_n + \Sigma_u} \tag{3.91}$$

$$TDOP = \sqrt{\Sigma_t} \tag{3.92}$$

$$GDOP = \sqrt{\Sigma_e + \Sigma_n + \Sigma_u + \Sigma_t} \tag{3.93}$$

A plot of GDOP against time over 24 hours is shown in figure 3.14.

3.5 Key results from the literature

Source	Receivers	Antennas	TTAR
Takasu and Yasuda (2008)	U-Blox Antaris AEK-4T	U-Blox ANN-MS	652.8 seconds
Takasu and Yasuda (2008)	U-Blox Antaris AEK-4T	NovAtel GPS-702-GG	131.2 seconds
Odijk et al. (2007)	U-Blox Antaris TIM-LL, TIM-LP	Patch antenna	227 to 660 seconds
Odijk et al. (2007)	Septentrio AsteRx1	Septentrio PolaNt	Less than 60 seconds
Lachapelle et al. (1993b)	NovAtel GPSCard	Choke ring	1090 seconds
Lachapelle et al. (1993b)	NovAtel GPSCard	Without choke ring	1826 seconds
Cannon et al. (1993)	NovAtel GPSCard	Choke ring	112 seconds
Lachapelle et al. (1993a)	NovAtel GPSCard	Choke ring	335 seconds
Lachapelle et al. (1993a)	NovAtel GPSCard	Without choke ring	810 seconds
Kim and Lee (2009)	U-Blox Antaris AEK-4T	U-Blox ANN-MS	480 seconds
Kim and Lee (2009)	U-Blox Antaris AEK-4T	NovAtel GPS-702-GG	520 seconds
Weisenburger (1997)	NovAtel MiLLenium	Choke ring	490 seconds

Table 3.4: Time To Ambiguity Resolution (TTAR) reports for single-frequency GPS positioning.

Results reported in the literature for GPS with single-frequency receivers are summarised in table 3.4.

Takasu and Yasuda (2008) evaluates the performance of a number of low-cost GPS receivers and antennas by means of a field test. Tests were performed using four low-cost receivers and one high-cost receiver, paired with a selection of low-cost antennas and one high-cost antenna. Receivers under test were mounted on a building roof (at approximately 35.87°N,138.39°E) with a high quality reference receiver 1 metre away.

Receiver	Antenna	Ambiguities resolved (%)	Mean time (s)	95% time (s)	Maximum time (s)
AEK-4T	ANN-MS	98.7 %	652.8 s	1840.0 s	3470.0 s
AEK-4T	GPS-702-GG	99.9 %	131.2 s	490.0 s	1200.0 s
OEMV-3	GPS-702-GG	99.8 %	132.7 s	630.0 s	1240.0 s

Table 3.5: Ambiguity resolution time statistics from Takasu and Yasuda (2008)

Data was logged over a period of 24 hours at a rate of 0.1 Hz; both receivers had a good view of the sky, with never less than 5 satellites in view, and 7 or more satellites in view 83% of the time. Ambiguity resolution was performed on the recorded data using an Extended Kalman Filter (EKF) and Modified LAMBDA (MLAMBDA), using the F-ratio test as the acceptance criteria (threshold 3.0) and an elevation mask of 15°. Ambiguity resolution was performed throughout the data, incrementing the start time in 10 second steps and recording time to ambiguity resolution, to gather mean time and 95% time statistics.

Results for a number of receiver/antenna combinations are presented. The best-performing low-cost receiver was a u-blox Antaris AEK-4T. Table 3.5 shows results for the AEK-4T with the antenna it is supplied with (u-blox ANN-MS) and with a high quality antenna (NovAtel GPS-702-GG); and results for single-frequency data from a high-cost dual frequency receiver (NovAtel OEMV-3) with the same high quality antenna.

For the low-cost receiver and antenna, 95% of the time ambiguities were resolved in 31 minutes or less, and the mean time to perform resolution was 11 minutes. The low-cost receiver with a high-quality antenna and the high-cost receiver with the same antenna produced similar results; 95% resolution in 8 minutes and 10 minutes respectively, both with a mean time of just over 2 minutes. In other words, the switch from a low-cost antenna to a high-quality antenna reduced the mean time by 80% and the 95% time by 73%; while the switch from a low-cost to a high-cost receiver produced a negligible (or even slightly negative) change in performance.

Odijk et al. (2007) describes two experiments using low-cost receivers (a u-blox Antaris

TIM-LL and TIM-LP); in one experiment a receiver is placed on a play area rounda-bout, less than 10 metres from a reference receiver, and data is recorded at a rate of 1Hz. With a 10° elevation mask and 6 satellites in view, a stationary initialisation resolved integer ambiguities in 320 seconds. The roundabout was rotated and the Real Time Kinematic (RTK) solutions compared to a best-fit circle; the radial standard deviation was 7.8mm and the maximum deviation 24.1mm. In a second experiment, with 7 satellites in view, a 400m baseline, and a data rate of 10Hz, integer ambiguities were resolved after 227 seconds. A third experiment was described, where a receiver on a boat, where an On The Fly (OTF) initialisation was performed with a 5 km baseline and a 15° elevation mask, and 6 or 7 satellites in view; in this case the initialisation took 660 seconds. A final test uses a more expensive receiver and antenna (Septentrio AsteRx1 and Septentrio PolaNT) in a stationary test on a roof, with a 2.7km baseline, a 1Hz data rate, and a 10° elevation mask. When 10 satellites are in view, ambiguity resolution is possible in a single epoch, while throughout the entire test (when 7 to 10 satellites are in view) ambiguity resolution never takes longer than 60 seconds.

Lachapelle et al. (1993b) describes a test on a boat equipped with two NovAtel GPS-Card receivers (two with choke ring antennas, one without) and one Ashtech P-XII (with choke ring antenna); a base station 10 to 24 km from the boat was equipped with a GPSCard and a P-XII, both with choke ring antennas. 6 satellites were in view, and data was logged at a rate of 2 Hz. The two GPSCard receivers with choke ring antennas took, on average, 1090 seconds to resolve integer ambiguities, while the GPS-Card without a choke ring antenna took on average 1825 seconds. The P-XII, on the other hand, was able to resolve integer ambiguities in 2 seconds by using dual-frequency data.

Cannon et al. (1993) performs a test with two NovAtel GPSCard receivers, both with choke ring ground planes and both with a 1 Hz data rate. One receiver was used as a base station and one receiver was mounted on the roof of a moving vehicle, within 5 km of the base station. Six satellites were in view above the 15° elevation mask. On The Fly (OTF) ambiguity resolution was performed for each, using Least Squares

Ambiguity Search Technique (LSAST) (Hatch, 1991), with ambiguity acceptance based on the variance threshold test, described in section 3.4.4.3 on page 66. Across 16 sets of data, ambiguity resolution took a mean time of 112 seconds (minimum 67 seconds, maximum 181 seconds). A second test was performed near trees, which caused higher multipath. 19 tests were performed, but ambiguity resolution was only successful in 8 of them. For those tests which were successful, ambiguity resolution took an average of 183 seconds (minimum 123 seconds, maximum 248 seconds), but the low rate of success may indicate a poor choice of threshold or range measurement variance. The paper also reports that the NovAtel GPSCard measures the C\A code accurate to 15 to 70 centimetres (depending on multipath), rather than the 2 to 3 metres reported for most receivers.

Lachapelle et al. (1993a) covers several topics, and includes results for ambiguity resolution using NovAtel GPSCard receivers with and without choke ring ground planes. The roving receivers were less than 5 kilometres from the fixed base stations, and at least 6 satellites were in view at all times. Ambiguity resolution was performed for every 60 seconds throughout the data gathered. Without choke ring ground planes ambiguity resolution took 810 seconds; with choke ring ground planes ambiguity resolution took 335 seconds. Despite using the same receivers as Cannon et al. (1993), ambiguity resolution with choke ring antennas is reported to take an average of 335 seconds, rather than 112 seconds; this may be because both tests use the variance threshold test described in section 3.4.4.3 on page 66, but Lachapelle et al. (1993a) assigns double-differenced carrier phase observations standard deviations of 15mm and 10mm in two tests, while Cannon et al. (1993) uses 7mm and 5mm. It is known that assuming a lower standard deviation allows faster ambiguity resolution at the cost of being unable to resolve the ambiguities in noisy environments - and failure to resolve ambiguities was observed in the tests of Cannon et al. (1993).

Kim and Lee (2009) gives details of a test using two low cost rovers (both u-blox AEK-4T), one with an inexpensive patch antenna (u-blox ANN-MS) and one with a high quality antenna (NovAtel GPS-702-GG). A Septentrio PolaRx2e receiver with a choke

ring antenna was used as a base station, with a 4.9 km baseline. Data was recorded at a rate of 1Hz, with a 10° elevation mask, and was processed using an on-the-fly ambiguity resolution algorithm. Ambiguities were resolved in 520 seconds ($8^2/_3$ minutes) with the high quality antenna, and in 480 seconds (8 minutes) with the patch antenna. The reduction in performance using the high quality antenna is at odds with Takasu and Yasuda (2008) - but as only one attempt at ambiguity resolution is reported, it is unclear whether this was an outlier.

Weisenburger (1997) reports on a number of tests, including single-frequency ambiguity resolution. In one test, two NovAtel MiLLenium dual frequency receivers with choke ring antennas were fitted to a boat, and another receiver was placed on the shore. The boat was within 2 kilometres of the base station at all times; data was gathered at 10Hz with a 10° elevation mask; and 6 or 7 satellites were visible throughout the 90 minute test. Ambiguity resolution was performed at different start points throughout the data. For single frequency data, the mean time to perform ambiguity resolution was 490 seconds.

Cosser et al. (2004) describes tests using single frequency GPS receivers for bridge deflection monitoring - an application where the approximate position of a receiver can be known in advance with reasonable precision. Tests were performed on two bridges, using Lecia System 500 GPS receivers with Leica AT504 choke ring antennas. Two processing techniques are described; one technique uses knowledge of the approximate position to calculate the expected distance from rover to satellite, which can be subtracted from the measured distance, then rounding performed to determine the integer ambiguity. This technique can resolve integer ambiguities instantly, as long as the receiver's position is known with an accuracy of about 3 to 5 centimetres - a reasonable assumption for short bridges. A second technique uses the approximate position to calculate an accurate float solution, which is used with the LAMBDA method to determine the ambiguity; this is applicable when the receiver's position is known less precisely (50 to 60 centimetres is suggested in the paper). The paper describes ambiguity resolution times as "greatly reduced" when using the latter technique. The paper

also states that, once ambiguities are resolved, single frequency receivers are as good as dual frequency receivers.

Saeki and Hori (2006) report on using single-frequency GPS receivers for displacement monitoring of civil structures. 2 hours of 1Hz GPS data was gathered on the roof of a building using Furuno GP8032 receivers, with both high quality antennas and patch antennas. Processing using the LAMBDA algorithm for ambiguity resolution and 5 minute fixed duration ambiguity resolution had a 98.2% success rate with two high quality antennas, and an 87.7% success rate with two patch antennas. For a 10 minute fixed duration test, these rates rose to 100.0% and 99.7% respectively. A test was also conducted with antennas placed on the ground in grass; higher multipath was encountered, reducing the patch antenna 5 minute success rate to 49.7%; the high quality antenna success rate in the test was 100.0%.

3.5.1 Choke ring antennas

Many experiments in the literature use choke ring antennas, which reduce multipath by rejecting signals arriving from near or below the horizontal - such as signals reflected from the ground. These antennas can cost several thousand pounds, ruling them out for a low-cost system. To compare results with and without these antennas it is useful to know the magnitude of benefit they can offer. Several papers have described the

Source	Antenna 1	Antenna 2	Effect on TTAR
Takasu and Yasuda (2008)	U-Blox ANN-MS	NovAtel GPS-702-GG	79.7% reduction
Lachapelle et al. (1993b)	Without choke ring	Choke ring	41.4% reduction
Kim and Lee (2009)	U-Blox ANN-MS	NovAtel GPS-702-GG	8.3% increase
Lachapelle et al. (1993a)	Without choke ring	Choke ring	58.6% reduction

Table 3.6: Impacts of choke ring antennas on Time To Ambiguity Resolution (TTAR)

interference-reducing advantages of choke ring antennas over patch antennas; reported

impacts are quoted in table 3.6 . All papers quoted are described in more detail in the previous section. The results for Kim and Lee (2009) stand out from the others as the only case where a choke ring antenna increased TTAR; the results in question were for a single ambiguity resolution using around 500 seconds of data, while Takasu and Yasuda (2008) considers a complete 24 hours of data. Ignoring Kim and Lee, the results quoted indicate choke ring antennas reduce TTAR by 40% to 80%.

Saeki and Hori (2006) compare high quality antennas and patch antennas for single-frequency applications, using a Futuno GT8032 receiver. They report that patch antennas produce results with a horizontal standard deviation of 1.2cm compared to 0.8cm for survey antennas; and that for 5 minute (fixed duration) ambiguity resolution, the success rate with survey antennas is 99% while the rate with patch antennas is 90%. The type of high quality antenna is not stated, but it's price is given as "approximately $1,000".

Chapter 4

Height Constraints

State-of-the-art ambiguity resolution algorithms are optimal in the sense that they produce a maximum likelihood estimate of receiver position based on their mathematical model of measurements and the noise thereof (Teunissen, 1999). In order to improve the success rate of ambiguity resolution, the model can be improved by integrating information from other sources. For applications such as boats on rivers with tide gauges, and ground vehicles operating on stable, rehearsed terrain, information about vehicle height is available. Depending on the accuracy and precision of the height information, ambiguity resolution times can be substantially reduced.

4.1 Previous Work

Ueno et al. (2000) reports on experiments using height constraints to improve GPS ambiguity resolution for river-depth surveys. Dual frequency receivers are used to resolve ambiguities over baselines as long as 75 km, for a ship on a river with electronic tide gauges. The GPS resolution system used first calculates an ionosphere-free combination (See 3.3.4.4 on page 51) float solution, then resolves wide-lane ambiguities, then resolves narrow-lane ambiguities; it is in the wide lane to narrow lane step that data from the tide gauges is used to improve ambiguity resolution. The paper states 5 seconds of data was used for each attempt at ambiguity resolution, and the ambiguity search is performed by varying the integer numbers in a nested loop with the calculated height solution calculated for each then compared against a predefined threshold. The best results were obtained by assuming the tide gauges were accurate to ± 15 cm; in this case the success rate rose from 70% to 90%. When the measurements were assumed to be more accurate correct solutions could be excluded, while a ± 30 cm range was equivalent to not using height information; as the wide-lane GPS ambiguities are

Source	Receivers	Antennas	Test type	Without height constraint	With height constraint
Ueno et al. (2000)	Dual frequency	Unknown	Dual frequency, 5 seconds	70% correct	90% correct
Zhu et al. (2005)	Dual frequency	Unknown	Dual frequency, 5 seconds	69% and 61% correct	93% and 95% correct
Zhu and Santerre (2002)	Dual frequency	Unknown	Dual frequency, 5 seconds	69% and 61% correct	95% and 96% correct
Weisenburger (1997)	NovAtel MiLLe-nium	Choke ring	Single Fre-quency, TTAR	490 seconds	140 seconds

Table 4.1: Previous results applying height constraints to GPS positioning.

resolved, the receiver's location is already known with ±30 cm accuracy even without height information.

Zhu et al. (2005) uses the same data gathered by Ueno et al. (2000). In four experiments, dual frequency GPS measurements were gathered on a boat, along with tide gauge measurements; in two tests the GPS base station was within 7 kilometres, and in two other tests the distances were 35 kilometres and 45 kilometres respectively. The data is then divided into five second blocks, to simulate a 5 second initialisation, and the ambiguity resolution success rate is evaluated. To perform ambiguity resolution, possible ambiguity solutions are enumerated using LSAST (Hatch, 1991), then subjected to Bayesian evaluation under the assumption that the probability is zero outside of the height boundaries. The mean value of the resulting distribution is used as the chosen solution, and is compared to the solution for the entire (120 minute) test. For the two tests with a 7 kilometre baseline, the success rates without height information were 69% and 61%; while with a height window of ±20 cm, the success rates increased to 93% and 95% (an increase of 24 points and 34 points respectively). As the success criteria for the tests was matching a known ambiguity, rather than matching while also passing a validation test (such as a squared norm ratio test), resolution may take

longer in real situations, where the ambiguity is not known in advance and hence a va-
lidation test is required. The paper does not mention why the mean of the probability
distribution is used in preference to the mode, when the mode would be a maximum
likelihood estimate.

Zhu and Santerre (2002) discusses the use of height information, using the same dual-
frequency data and test procedure as Zhu et al. (2005); but instead of using the height
as a filter, it is used as a 'quasi-observation' - that is, it is integrated into the observation
equations as if there were a satellite at the centre of the earth. A least-squared-error
solution to the observation equations therefore is the best solution for both height and
pseudorange information, assuming all measurements have only zero-mean Gaussian
noise and an accurate a priori standard deviation is known. Assuming a 20 cm standard
deviation on height measurements, the success rates for the two tests with a 7 kilometre
baseline rose from 69% and 61% to 95% and 96% respectively. This is a greater
improvement than using height constraints. However, assuming a height measurement
to have zero-mean Gaussian noise may not always be a safe assumption - for example,
when there is a bias due to distance from the tide gauge measuring station. Also, in
common with Zhu et al. (2005), no validation test was performed.

Weisenburger (1997) examines the use of height information, with constraints applied
as observations to a Kalman filter. In one test, two NovAtel MiLLenium dual frequency
receivers with choke ring antennas were fitted to a boat, and another receiver was placed
on the shore. The boat was within 2 kilometres of the base station at all times, data was
gathered at 10Hz with a 10° elevation mask, and 6 or 7 satellites were visible throughout
the 90 minute test. The recorded data was analysed to determine a mean height, then
ambiguity resolution was performed at different start points throughout the data. This
was done once for dual-frequency data; and once with only single-frequency data. With
single frequency data, without height constraints, the mean time to perform ambiguity
resolution was 490 seconds; with a height constraint (based on the mean height, with
a standard deviation of 50 centimetres) the mean time dropped to 140 seconds; a
reduction of 71%.
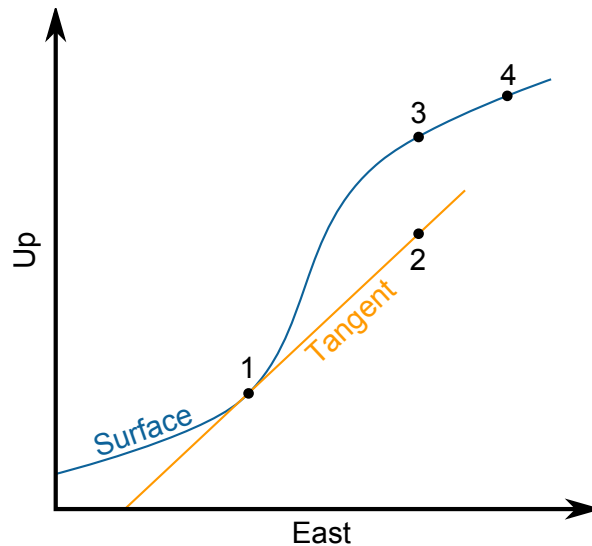
Figure 4.1: Integrating surface topography into a pseudolite-based measuring system, using method 5 from Amt and Raquet (2006). Point 1 is the initial position; point 2 the updated position, based on measurements and the tangential plane constraint; point 3 is the updated position moved vertically to match the topographical map; point 4 is the true position. Based on figure 7 from Amt and Raquet (2006).

Amt and Raquet (2006) report on using surface topography with a pseudolite-based system to improve vertical accuracy. Pseudolites are ground based transmitter beacons which broadcast GPS-like signals; they can be used either alone, or in conjunction with normal GPS satellites. Applications include improving GPS coverage and reliability (such as in open-pit mining and near airports); key disadvantages include lack of vertical observability (as the receiver and all transmitters are at approximately the same height, leading to poor vertical Dilution Of Precision (DOP); see section 3.3.3 on page 44 for more details), near-far problems (as the signal strength is subject to an inverse-square drop off with range), multipath problems (due to obstacles and reflective surfaces at ground level), and traditionally high costs. For more information on pseudolites, see Cobb (1997).

The paper proposes several means of integrating height from a topographic map into an iterative least-squares algorithm, alongside carrier phase measurements from pseudo-lites. The method identified as best (on the grounds of numerical precision and number of iterations to converge) uses an iterative process based on a tangential approximation of the surface at the estimated position; this is illustrated in figure 4.1. Evaluation of

the processes was performed through two tests. The first test was conducted on a road, comparing results from a pseudolite-only based position (using five Locata LocLite pseudolites) to solutions from two high-precision GPS receivers. Topographical information accurate to 1mm was available for the road. Comparing the pseudolite-based position to the GPS-based position, the pseudolite solution had a mean error of 2.3 centimetres and a peak error of 9 centimetres. However, due to the relatively flat road, the same results were gathered for all five of the proposed height-integration algorithms. A second test was performed, simulating positioning on a hill. Perfect measurements were used, to simplify comparison of the different algorithms. It was found that the tangent-based algorithm produced the most accurate results (although, with perfect measurements, the errors never exceeded 50 micrometres), and required the lowest number of iterations. Pseudolites themselves have not been investigated in detail in this contribution; the algorithms used by Amt and Raquet are described to give a complete picture of the literature.

4.2 Application to ground vehicles

As seen in the previous section, height constraints can improve ambiguity resolution rates on boats. However, the application of similar techniques on land presents certain problems.

4.2.1 Dealing with limited map precision

Unlike boats on still waters, ground vehicles can encounter hills, kerbs, and uneven ground. Hills can be accounted for using a topographical map, but precisely accounting for uneven ground is difficult; a very detailed map and a model of vehicle-ground interaction would be required to achieve consistently high precision. It is simpler to use a limited-precision map and to take the map's imprecision into account.

For a stationary vehicle, map imprecision would take the form of an offset, rather than zero-mean Gaussian noise. Similar to Zhu et al. (2005), map imprecision can be modelled using a minimum and maximum admissible height.
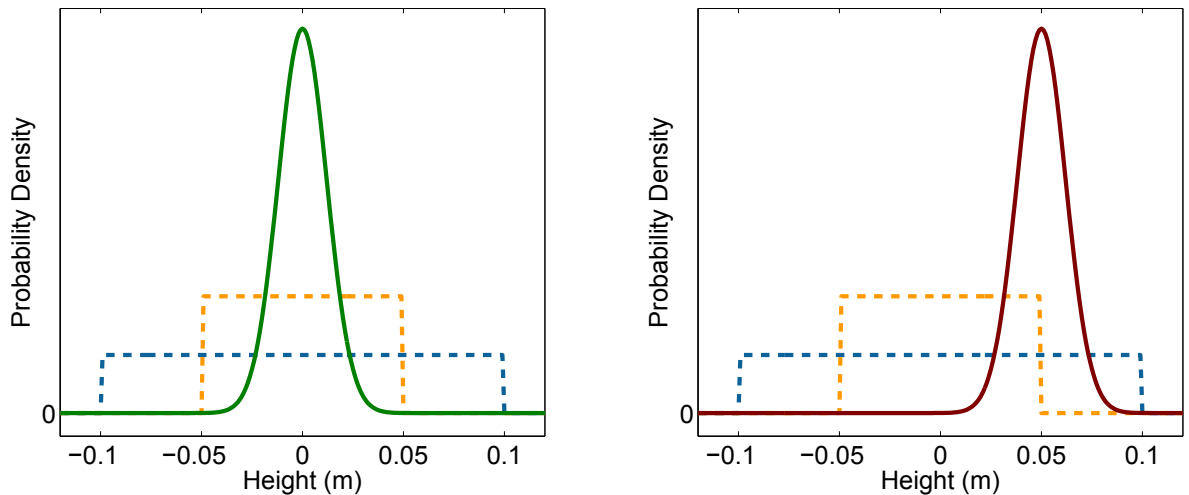
Figure 4.2: Probability densities of height measurements and threshold functions. The green and red normal distributions are probability densities for height measurements assuming true heights of 0cm and +5cm respectively, and a 1.2cm standard deviation on height measurements. The yellow box marks an acceptance threshold of $\pm$5cm and the blue an acceptance threshold of $\pm 5 + 4*1.2$cm, i.e. 4 standard deviations.

When applying a minimum and maximum admissible height, noise must be taken into account because even if the true height of a receiver falls within a given range, noise may result in measured heights outside that range. Figure 4.2 illustrates this; both the green and the red distributions have their means within the $\pm$5cm range (yellow dotted line), but half the measurements following the red distribution would fall outside of the threshold.

One way of addressing this is to set the acceptance thresholds further apart; the blue dotted line shows the height threshold expanded by 4 standard deviations, to account for 99.99% of noise. Another way is, when a solution falls outside of the acceptance threshold, to evaluate it when constrained to lie on that threshold. This document deals with the latter approach.

4.2.2 Integer ambiguity search algorithm

Section 3.4.4.2 on page 61 describes the LAMBDA method, a popular means for calculating least-squared-error ambiguity solutions. The calculations presented in that section use Earth Centred Earth Fixed (ECEF) coordinates relative to a reference point, so we can describe our height boundaries as:

$$\min_{\boldsymbol{a},\boldsymbol{b}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|^2_{Q_y} \text{ with } \boldsymbol{a} \in \mathbb{Z}^{\mathrm{n}}, \boldsymbol{b} \in \mathbb{R}^3, minup \leq \boldsymbol{d}^T\boldsymbol{b} \leq maxup \qquad (4.1)$$

where $\boldsymbol{d}^T = \begin{bmatrix} cos\varphi cos\lambda & cos\varphi sin\lambda & sin\varphi \end{bmatrix}$ (which converts the relative ECEF coordinates to a height in the local tangential plane), φ and λ are the reference location's latitude and longitude respectively, and *minup* and *maxup* are the minimum and maximum admissible heights respectively. There are several ways we can go about calculating this solution.

4.2.2.1 Triple integer search method

If the LAMBDA method is used without height constraints, then a height constraint is applied, good solutions fall into three categories: solutions which fall between the minimum and maximum heights; solutions which fall above the maximum height but which would still look good if they were exactly at the maximum height; and solutions which fall below the minimum height but which would still look good if they were exactly at the minimum height.

The triple integer search method calculates the best solutions for each of the three categories separately, then selects the best solution overall. This is depicted in figure 4.3 .

The calculations presented in section 3.4.4.2 on page 61 use relative Earth Centred Earth Fixed (ECEF) coordinates; these can be transformed to relative East North Up (ENU) coordinates with a simple rotation matrix (see, e.g. Hofmann-Wellenhof et al. (2008) p.281). Hence, equation 3.55 on page 61 can be rewritten as:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{a} + \tilde{\boldsymbol{B}}\tilde{\boldsymbol{b}} + \boldsymbol{e} \qquad (4.2)$$
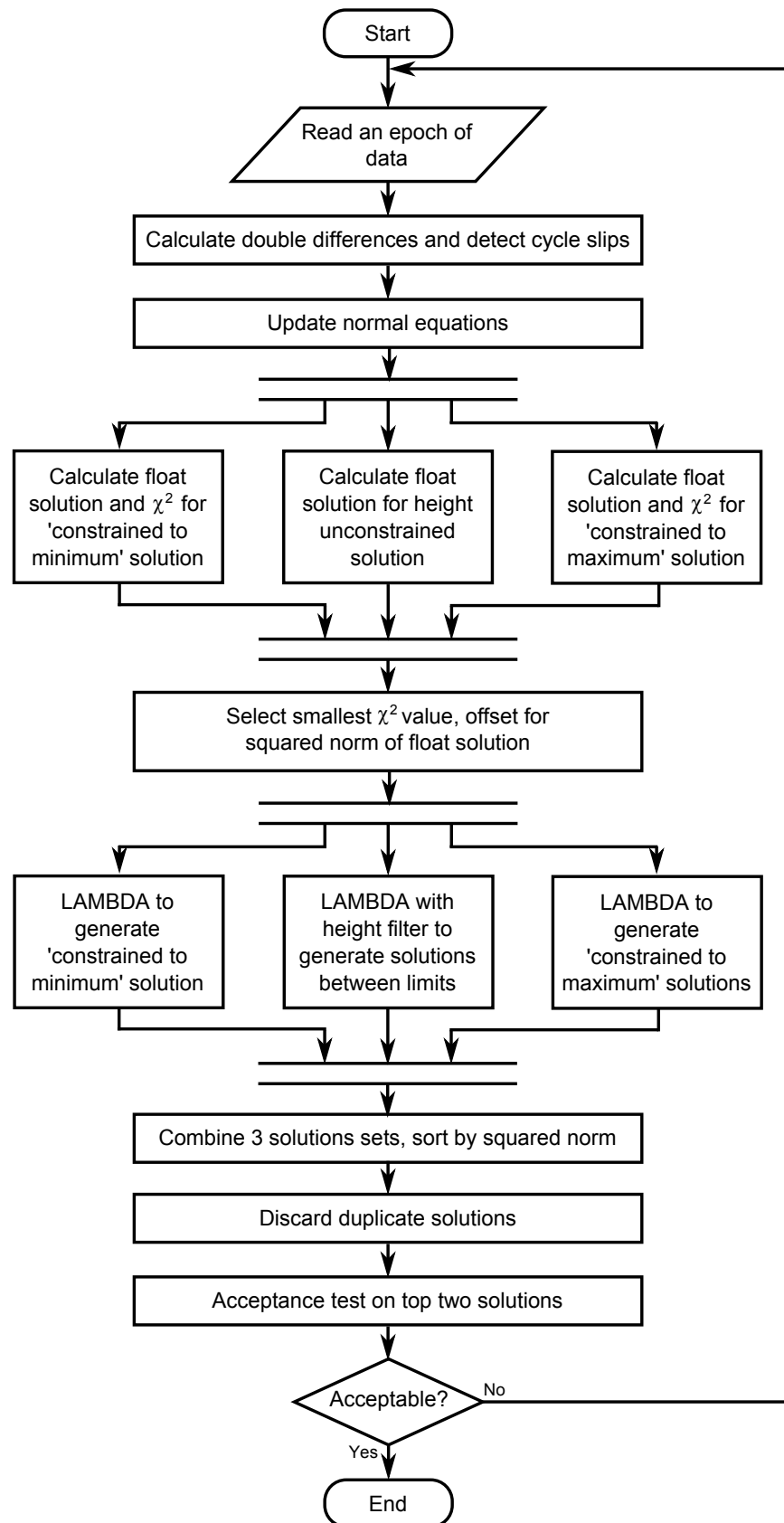
Figure 4.3: Flow chart depicting triple integer search ambiguity resolution process.

where

$$\tilde{\boldsymbol{B}} = \begin{bmatrix} \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dN_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dU_{ref}} \\[6pt] \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dN_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dU_{ref}} \\[6pt] \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dN_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dU_{ref}} \\[6pt] \vdots & \vdots & \vdots \end{bmatrix} = \boldsymbol{B}\boldsymbol{D}^{-1} \tag{4.3}$$

$$\boldsymbol{D} = \begin{bmatrix} -sin\lambda & cos\lambda & 0 \\ -sin\varphi cos\lambda & -sin\varphi sin\lambda & cos\varphi \\ cos\varphi cos\lambda & cos\varphi sin\lambda & sin\varphi \end{bmatrix} \tag{4.4}$$

$$\tilde{\boldsymbol{b}} = \begin{bmatrix} e \\ n \\ u \end{bmatrix} = \boldsymbol{D}\boldsymbol{b} \tag{4.5}$$

where φ and λ are the reference location's latitude and longitude and e, n and u are the east, north, and up distances of the solution relative to the reference location. As this is simply a rotated form of equation 3.55 on page 61, many of the simplifications used in section 3.4.4.2 on page 61 can still be applied.

As normally implemented, the LAMBDA technique minimises $\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2$, the squared residual of the fixed solution compared to the float solution, because the float solution is a constant offset so this also minimises the squared residual of the fixed solution.

As the triple integer search uses three different float solutions, the fixed residual compared to the float solution ($\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2$) is not so useful for comparison purposes; the residual of the float solution ($\|\hat{\boldsymbol{e}}\|_{Q_y}^2$) must be added to give the total residual of the fixed solution.

Solving for a fixed height   To find the best solutions lying precisely on the top and bottom height thresholds, we can perform ambiguity resolution with a two-dimensional rather than a three-dimensional geometry matrix. Equation 4.2 on page 89 can be

converted from:

$$
\boldsymbol{y} = \boldsymbol{Aa} +
\begin{bmatrix}
\frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dN_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dU_{ref}} \\
\frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dN_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dU_{ref}} \\
\frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dN_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dU_{ref}} \\
\vdots & \vdots & \vdots
\end{bmatrix}
\begin{bmatrix}
e \\
n \\
u
\end{bmatrix}
+ \boldsymbol{e}
\tag{4.6}
$$

to:

$$
\boldsymbol{y} -
\begin{bmatrix}
\frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dU_{ref}} \\
\frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dU_{ref}} \\
\frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dU_{ref}} \\
\vdots
\end{bmatrix}
u = \boldsymbol{Aa} +
\begin{bmatrix}
\frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^2(\boldsymbol{\varrho}_r)}{dN_{ref}} \\
\frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^3(\boldsymbol{\varrho}_r)}{dN_{ref}} \\
\frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dE_{ref}} & \frac{d\varrho_{dd}^4(\boldsymbol{\varrho}_r)}{dN_{ref}} \\
\vdots & \vdots
\end{bmatrix}
\begin{bmatrix}
e \\
n
\end{bmatrix}
+ \boldsymbol{e}
\tag{4.7}
$$

which can be solved by the same methods used for 3D problems, such as the LAMBDA method; it is equivalent to subtracting u multiplied by the 'up' column of the normal matrix from the normal vector, then discarding the 'up' row and column from the normal matrix and the 'up' row from the normal vector (as defined in section 3.4.5.2 on page 72).

Solving for a position within height constraints   As section 3.4.4.2 on page 64 describes, the LAMBDA technique searches for integer ambiguity vectors with sum-of-squares residuals close to those of the float solution. Most implementations of the search algorithm allow the user to request n results, and return the n best (i.e. lowest sum-of-squares residual) results. We can modify the search algorithm to return the n best solutions which fall within our height constraints by applying the following logic. It is known from equation 3.71 on page 65 and the East North Up (ENU) LAMBDA formulation of equation 4.2 on page 89 that:

$$
\begin{bmatrix}
\check{e} \\
\check{n} \\
\check{u}
\end{bmatrix}
=
\begin{bmatrix}
\hat{e} \\
\hat{n} \\
\hat{u}
\end{bmatrix}
- Q_{\hat{b}\hat{a}} Q_{\hat{a}} \left(\boldsymbol{Z}^T \right)^{-1} (\hat{\boldsymbol{z}} - \check{\boldsymbol{z}})
\tag{4.8}
$$

$$
\begin{bmatrix} \check{e} \\ \check{n} \\ \check{u} \end{bmatrix} = \begin{bmatrix} \hat{e} \\ \hat{n} \\ \hat{u} \end{bmatrix} - \boldsymbol{U}(\hat{\boldsymbol{z}} - \check{\boldsymbol{z}}) \tag{4.9}
$$

and defining $\boldsymbol{u}^T$ as the bottom row of $\boldsymbol{U}$ we can formulate

$$
\check{u} = \hat{u} - \boldsymbol{u}^T(\hat{\boldsymbol{z}} - \check{\boldsymbol{z}}) \tag{4.10}
$$

$$
\check{u} = \hat{u} - \boldsymbol{u}^T\hat{\boldsymbol{z}} + \boldsymbol{u}^T\check{\boldsymbol{z}} \tag{4.11}
$$

$$
\check{u} - \hat{u} + \boldsymbol{u}^T\hat{\boldsymbol{z}} = \boldsymbol{u}^T\check{\boldsymbol{z}} \tag{4.12}
$$

and hence the height acceptance criteria

$$
minup \leq \check{u} \leq maxup \tag{4.13}
$$

can be expressed as

$$
\begin{aligned}
minup - \hat{u} + \boldsymbol{u}^T\hat{\boldsymbol{z}} &\leq \boldsymbol{u}^T\check{\boldsymbol{z}} \leq maxup - \hat{u} + \boldsymbol{u}^T\hat{\boldsymbol{z}} \\
minuz &\leq \boldsymbol{u}^T\check{\boldsymbol{z}} \leq maxuz
\end{aligned} \tag{4.14}
$$

so, during the search for an acceptable ambiguity vector, checking that a given ambiguity vector is of an acceptable height can be performed by multiplying by $\boldsymbol{u}^T$ and comparing the result to the thresholds $minuz$ and $maxuz$.

Selection of search space size   As described in section 3.4.4.2 on page 64, LAMBDA searches for ambiguities in a search space around the float solution, the size of which is defined by χ^2. For the solutions which are constrained to lie on the minimum or maximum heights, the technique described in section 3.4.4.2 on page 64 can be used to define a search space containing at least two ambiguity vectors.

Defining the search space size is not so simple for the height-constrained search, as simply rounding the float solution may give a result which lies outside the height thresholds, leading to too low a value for χ^2 being selected.

The process being used performs three LAMBDA searches, combines the results, then selects only the two best solutions. Therefore, if it is known that one search can produce two solutions with squared norms below a given threshold, another search need only search for solutions with squared norms below that threshold as any found above that threshold would be discarded as not being among the two best solutions. Hence, the χ^2 values chosen for solutions constrained to lie on the height thresholds can be used to select a value of χ^2 for the search between the height thresholds by selecting the lower of the two and offsetting for the difference in float solution squared norms.

4.2.2.2 Search-and-correct method

Equivalently, if we recall from equation 3.63 on page 63 that the standard LAMBDA technique minimises

$$\min_{\boldsymbol{a},\boldsymbol{b}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|_{Q_y}^2 \text{ with } \boldsymbol{a} \in \mathbb{Z}^{\mathrm{n}}, \boldsymbol{b} \in \mathbb{R}^3$$

using the orthogonal decomposition from equation 3.64 on page 63, i.e.

$$\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|_{Q_y}^2 = \|\hat{\boldsymbol{e}}\|_{Q_y}^2 + \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2 + \left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|_{Q_{\hat{b}(a)}}^2$$

by minimising $\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2$, which minimises the former term because $\|\hat{\boldsymbol{e}}\|_{Q_y}^2$ is a constant and $\left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|_{Q_{\hat{b}(a)}}^2$ can always be zero. On the other hand, if $\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right)$ is unconstrained while the *up* portion of $\boldsymbol{b}$ is constrained to lie between minimum and maximum height thresholds, the assumption that the term will be zero does not hold true. Instead, a height-constrained value of $\boldsymbol{b}$ must be found - call it $\breve{\boldsymbol{b}}$ - which minimises $\left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \breve{\boldsymbol{b}}\right\|_{Q_{\hat{b}(a)}}^2$, and $\left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \breve{\boldsymbol{b}}\right\|_{Q_{\hat{b}(a)}}^2$ must be added to the value of $\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2$.

When $\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right)$ falls between the height thresholds, $\breve{\boldsymbol{b}} = \hat{\boldsymbol{b}}\left(\boldsymbol{a}\right)$; otherwise, the height-constrained value of $\breve{\boldsymbol{b}}$ is found as the conditional least squares solution, given the

height and $\boldsymbol{b}$. The float solution variance-covariance matrix is

$$
\left[\begin{array}{c|c}
Q_{\hat{b}} & Q_{\hat{a}\hat{b}} \\
\hline
Q_{\hat{b}\hat{a}} & Q_{\hat{a}}
\end{array}\right]
=
\left[\begin{array}{ccc|c}
Q_{\hat{e}} & Q_{\hat{e}\hat{n}} & Q_{\hat{e}\hat{u}} & Q_{\hat{e}\hat{a}} \\
Q_{\hat{n}\hat{e}} & Q_{\hat{n}} & Q_{\hat{n}\hat{u}} & Q_{\hat{n}\hat{a}} \\
Q_{\hat{u}\hat{e}} & Q_{\hat{u}\hat{n}} & Q_{\hat{u}} & Q_{\hat{u}\hat{a}} \\
\hline
Q_{\hat{a}\hat{e}} & Q_{\hat{a}\hat{n}} & Q_{\hat{a}\hat{u}} & Q_{\hat{a}}
\end{array}\right]
\tag{4.15}
$$

And we can define

$$
\left[\begin{array}{c|c}
Q_{\hat{e}\hat{n}} & Q_{\hat{u}\hat{a}\hat{e}\hat{n}} \\
\hline
Q_{\hat{e}\hat{n}\hat{u}\hat{a}} & Q_{\hat{u}\hat{a}}
\end{array}\right]
=
\left[\begin{array}{cc|cc}
Q_{\hat{e}} & Q_{\hat{e}\hat{n}} & Q_{\hat{e}\hat{u}} & Q_{\hat{e}\hat{a}} \\
Q_{\hat{n}\hat{e}} & Q_{\hat{n}} & Q_{\hat{n}\hat{u}} & Q_{\hat{n}\hat{a}} \\
\hline
Q_{\hat{u}\hat{e}} & Q_{\hat{u}\hat{n}} & Q_{\hat{u}} & Q_{\hat{u}\hat{a}} \\
Q_{\hat{a}\hat{e}} & Q_{\hat{a}\hat{n}} & Q_{\hat{a}\hat{u}} & Q_{\hat{a}}
\end{array}\right]
\tag{4.16}
$$

then, for a known ambiguity and height, the constrained baseline is given by

$$
\breve{\boldsymbol{b}} =
\left[\begin{array}{c}
e \\
n \\
u
\end{array}\right]
\quad \text{where} \quad
\left[\begin{array}{c}
e \\
n
\end{array}\right]
=
\left[\begin{array}{c}
\hat{e} \\
\hat{n}
\end{array}\right]
- Q_{\hat{e}\hat{n}\hat{u}\hat{a}} Q_{\hat{u}\hat{a}}^{-1}
\left[\begin{array}{c}
\hat{u} - u \\
\hat{\boldsymbol{a}} - \boldsymbol{a}
\end{array}\right]
\tag{4.17}
$$

where u is the constrained height value and $\hat{e}$ $\hat{n}$ and $\hat{u}$ are the east, north and up values from the float solution. Equivalently

$$
\hat{\boldsymbol{b}}(\boldsymbol{a}) - \breve{\boldsymbol{b}} =
\left[\begin{array}{c}
\hat{e} \\
\hat{n} \\
\hat{u}
\end{array}\right]
- Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} [\hat{\boldsymbol{a}} - \boldsymbol{a}] -
\left[\begin{array}{c}
\hat{e} \\
\hat{n} \\
\hat{u}
\end{array}\right]
+
\left[\begin{array}{c}
Q_{\hat{e}\hat{n}\hat{u}\hat{a}} Q_{\hat{u}\hat{a}}^{-1} \\
0
\end{array}\right]
\left[\begin{array}{c}
\hat{u} - u \\
\hat{\boldsymbol{a}} - \boldsymbol{a}
\end{array}\right]
+
\left[\begin{array}{c}
0 \\
0 \\
1
\end{array}\right]
[\hat{u} - u]
\tag{4.18}
$$

$$
\hat{\boldsymbol{b}}(\boldsymbol{a}) - \breve{\boldsymbol{b}} = -
\left[\begin{array}{cc}
0 & Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1}
\end{array}\right]
\left[\begin{array}{c}
\hat{u} - u \\
\hat{\boldsymbol{a}} - \boldsymbol{a}
\end{array}\right]
+
\left[\begin{array}{c}
Q_{\hat{e}\hat{n}\hat{u}\hat{a}} Q_{\hat{u}\hat{a}}^{-1} \\
0
\end{array}\right]
\left[\begin{array}{c}
\hat{u} - u \\
\hat{\boldsymbol{a}} - \boldsymbol{a}
\end{array}\right]
+
\left[\begin{array}{c|c}
0 & \\
0 & 0 \\
1 &
\end{array}\right]
\left[\begin{array}{c}
\hat{u} - u \\
\hat{\boldsymbol{a}} - \boldsymbol{a}
\end{array}\right]
\tag{4.19}
$$

$$\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \check{\boldsymbol{b}} = \left(\begin{bmatrix} Q_{\hat{e}\hat{n}\hat{u}\hat{a}} Q_{\hat{u}\hat{a}}^{-1} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & Q_{\hat{b}\hat{a}} Q_{\hat{a}}^{-1} \end{bmatrix} + \left[\begin{array}{c|c} \begin{matrix} 0 \\ 0 \\ 1 \end{matrix} & 0 \end{array} \right] \right) \begin{bmatrix} \hat{u} - u \\ \hat{\boldsymbol{a}} - \boldsymbol{a} \end{bmatrix} \quad (4.20)$$

hence, the value of $\left\| \hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b} \right\|_{Q_{\hat{b}(a)}}^{2}$ can be found with a small number of matrix multiplications. The calculation and addition of this to the squared norm can be included as part of the ambiguity search process.

Selection of search space   The value of $\left\| \hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \check{\boldsymbol{b}} \right\|_{Q_{\hat{b}(a)}}^{2}$ can be anything from zero to substantially larger than $\left\| \hat{\boldsymbol{a}} - \boldsymbol{a} \right\|_{Q_{\hat{a}}}^{2}$ making it difficult to select a search space size. One means of choosing the search space size is given in section 4.2.2.1 on page 93. Another technique starts with a small size (such as using the technique from section 3.4.4.2 on page 64), performing a search, then if fewer than two acceptable values are found, expanding the search space size and repeating the search.

4.2.3   Ambiguity validation

To produce results for comparison with other findings in the literature, tests were performed with a number of ambiguity acceptance criteria:

1. Accept the solution with the lowest squared residual immediately.

2. Accept the solution with the lowest squared residual after a certain time (5 seconds is the criteria used by Zhu et al. (2005) and Zhu and Santerre (2002))

3. Accept the solution with the lowest squared residual if the F-ratio (see section 3.4.4.3 on page 65) is greater than a threshold (A ratio of 4.0 is used by Weisenburger (1997)).

4. Accept the solution with the lowest squared residual if it is maintained as the best candidate for a set duration, with that duration depending on satellite visibility and measured noise levels.
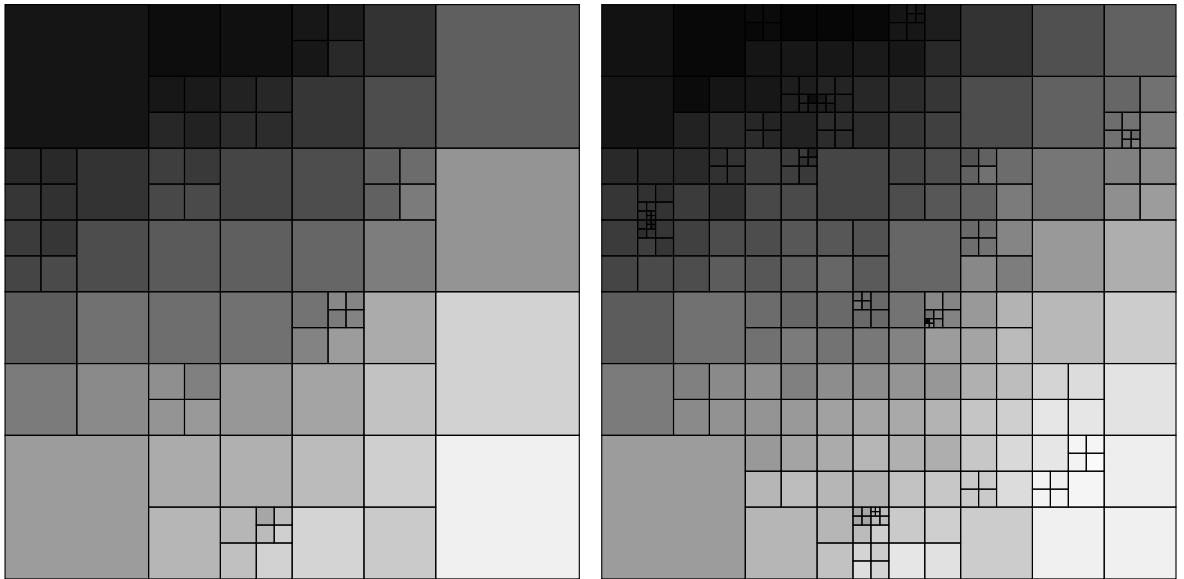
Figure 4.4: Two quadtrees storing height information for a 60m x 60m test area; the left quadtree stores height information precise to ±5cm, while the right quadtree stores height information precise to ±2.5cm.

4.2.4   Storage of topographical information

To supply the minimum and maximum admissible heights for given positions, a topographical map must be stored. A quadtree with leaves comprised of planar approximations was used as an efficient means of height information storage. Figure 4.4 shows two quadtrees of planar approximations, based on height data from the field pictured in figure 4.6 on page 100. The left image describes the test area precise to ±5cm, while the right image describes the same area precise to ±2.5cm.

Figure 4.5 illustrates the detail of the quadtree data structure. The structure starts with a top level *root node* (blue outline) and each node may either contain data (if it is a *leaf node*) or be subdivided into four (if it is a *branch node*). In areas where the ground is flat, a single large node might suffice to represent a large area (such as the yellow squares in the figure). In areas where the topography is more complicated, a larger number of smaller nodes can be used (such as the purple squares in the figure). For each datum added, one starts at the root node and descends to the leaf node corresponding to the latitude and longitude. It is added to the stored data for that node, then that stored data is assessed and if appropriate the node can be split and its
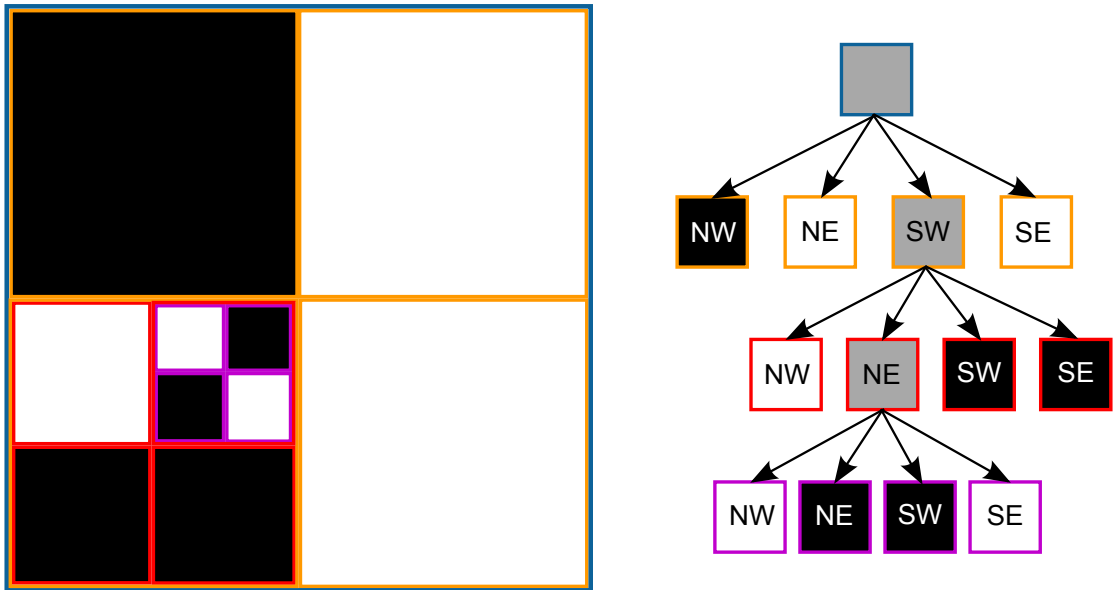
Figure 4.5: Illustration of quadtree structure; in this diagram grey-filled boxes are *branches* which have been sub-divided, while black and white filled boxes are *leaves* which have not been subdivided.

data distributed among the new leaf nodes.

4.2.5   Sources of topographical information

Several methods are available to gather topographical information in the base station reference frame, depending on the level of vertical and horizontal precision required. Obviously, it's easier to gather precise data for a very flat area, as fewer measurements are needed to describe its shape fully!

One option would be a survey using conventional theodolites or total stations - a traditional means of height surveying, this technique is accurate but would be time-consuming over an entire golf course. Another option is a GPS survey, using high-precision receivers and low speed or stop/start movement. This is more amenable to automation, and was the means we used to gather the test data shown in figure 4.4 . Other surveying techniques like photogrammetry are also available, although better suited to roads than large areas of grass!

The minimum and maximum height bounds must be set to take into account the stability of the ground, and for the highest precision the map may need to be updated

from time to time. It may be possible for the vehicle to update its own map where inaccuracies can be spotted, but this is beyond the scope of the work performed.

4.2.5.1 Varying precision across the map

As the quadtree minimum and maximum map has variable horizontal and vertical precision, locations of interest can be mapped with higher precision. For example, ambiguity resolution is likely to be needed at route start points and near signal-loss-inducing obstacles, so a higher level of detail could be employed at those locations. With manual operator intervention, the mower could even be guided to specific precisely surveyed landmarks (such as tees or greens) where ambiguity resolution could be particularly fast.

4.3 Experimental Configuration

4.3.1 Stationary Experiment

An experiment was performed to compare ambiguity resolution performance with and without height constraints.

To reflect real-world use, a test was devised as shown in figure 4.6 . One U-Blox Antaris AEK-4T GPS receiver with U-Blox ANN-MS patch antenna was placed on a building roof (at approximately 52.3827°N, 1.5613°W), while an identical receiver and antenna were fitted to a ground vehicle (at approximately 52.3846°N, 1.5585°W, for a 280m baseline length). Using the software described in section 2.4.1 on page 26, 24 hours of single-frequency data was logged at the receivers' maximum measurement rate of 10Hz. Satellite visibility was fair to good, with 5 or more satellites in view 97.7% of the time, and 7 or more in view 53.6% of the time. A reference location and series of ambiguity solutions was determined by using the standard LAMBDA method over the entire 24 hours of data.

The recorded files were post-processed using the software described in section 2.4.1 on page 26. The proposed algorithm was used to perform ambiguity resolution through
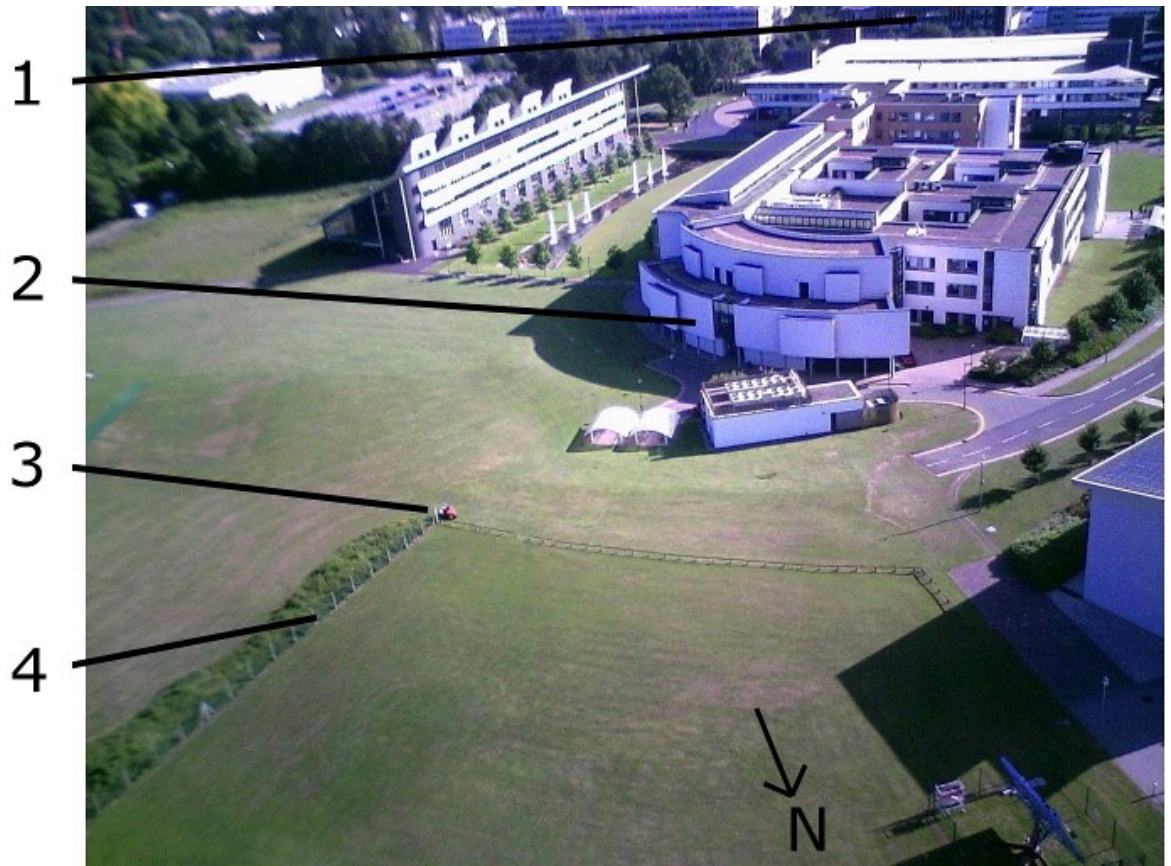
Figure 4.6: Photograph showing the test configuration; (1) base station antenna on building, 280m baseline; (2) tall buildings within 30m; (3) test vehicle; (4) chain link fence within 1m.

the entire 24 hours of data, with the test start time incrementing in 30 second steps. A solution was considered correct if it matched the reference ambiguity solution for that time. For each start time, the time to achieve ambiguity resolution was recorded, along with whether the solution was correct.

The data was processed using height constraints of ± 5cm and ± 2.5cm; and the acceptance criteria:

1. Accept immediately.

2. Accept after 5 seconds.

3. Accept when F-ratio exceeds 3.0.

4. Accept when F-ratio exceeds 4.0.

In all tests, if the acceptance criteria was not met within 1800 seconds (30 minutes) the best candidate result was selected. This simplified handling of cases where an acceptance criteria was unattainable. Tests indicated this introduced no erroneous acceptances. However, in tests where the median or 95% time is reported 1800 seconds, the true value may have been higher.

To determine the impact of the F-ratio test acceptance threshold on speed and reliability, a series of tests were performed using an F-ratio test as the acceptance criteria, with a threshold varying between 1.0 and 5.0. This was performed without height constraints, and with constraints of ±25cm, ±10cm, ±5cm and ±2.5cm.

For comparison purposes, the data was processed integrating a single height (instead of a range) using a pseudo-observation based algorithm based on Zhu and Santerre (2002). The acceptance criteria used was a F-ratio of 4.0. Finally, to simulate the impact of inaccurate height information, tests were performed with 2.5cm and 5cm biases on the a priori height.

4.3.2 Kinematic experiment

Although for an autonomous vehicle it would be acceptable (or even desirable) to come to a stop when an RTK position was not available, having the option to perform ambiguity resolution while in motion enables a broader range of applications. To investigate this, a kinematic test was performed using the same equipment and environment as the stationary test of section 4.3.1 on page 99.

A real-time version of the software used in these experiments was run on the test vehicle. After a stationary initialisation the test vehicle was set to operate autonomously, following a path comprised of several dense spirals at a speed of 6 km/h for ninety minutes. Single frequency GPS observables were recorded at 10 Hz. Between four and seven satellites were in view throughout the entire test, allowing a continuous RTK fix to provide a series of reference solutions. To simulate poor satellite visibility, two data sets were generated with certain satellites excluded; the first with four to six satellites

in view and the second with four to five satellites in view.

The proposed method was used to perform on-the-fly ambiguity resolution throughout the ninety minute data sets, with the resolution start time incrementing in five minute steps. Height constraints from ±25 cm to ±2.5 cm were used and a test without height constraints was performed for comparison. Solution acceptance used the F-ratio test, and as in section 4.3.1 a range of thresholds were tested.

4.4 Results

4.4.1 Stationary Experiment

Acceptance	Height	Solved	Time to ambiguity resolution		
criteria	precision	right (%)	Mean (s)	Median (s)	95% (s)
Accept	None	*17.38%*	*0.1*	*0.1*	*0.1*
immediately	±5.0cm	*63.58%*	*0.1*	*0.1*	*0.1*
	±2.5cm	*69.33%*	*0.1*	*0.1*	*0.1*
Accept	None	*17.94%*	*5.0*	*5.0*	*5.0*
after	±5.0cm	*64.40%*	*5.0*	*5.0*	*5.0*
5 seconds	±2.5cm	*69.89%*	*5.0*	*5.0*	*5.0*
Accept if	None	99.89%	679.6	604.0	1800.0
F-ratio	±5.0cm	*99.36%*	*265.2*	*182.0*	*786.0*
exceeds 3.0	±2.5cm	99.68%	224.9	142.0	708.0
Accept if	None	100.00%	969.5	919.0	1800.0
F-ratio	±5.0cm	99.96%	421.7	330.0	1140.0
exceeds 4.0	±2.5cm	99.93%	374.3	264.0	1130.0

Table 4.2: Results of stationary height-constrained tests using various acceptance criteria. Results with success rates below 99.5% in *italics*.

The results of experiments with various acceptance criteria are shown in table 4.2. When no height constraints are used, results are comparable to the results reported in Takasu and Yasuda (2008) and Odijk et al. (2007) when using a low-cost receiver, patch antenna and 3.0 F-ratio threshold; Takasu and Yasuda (2008) reports on an AEK-4T receiver with ANN-MS antenna correctly resolving 98.7 % of ambiguities with a mean time of 652.8 seconds and a 95% time of 1840.0 seconds, while the 3.0 ratio test in the table above correctly resolves 99.89 % of ambiguities with a mean time of 679.6 seconds and a 95% time of 1800.0 seconds. Applying a ±5.0cm height constraint reduces the

mean ambiguity resolution time by 61%, while a ±2.5cm height constraint results in a 67% reduction.

Zhu et al. (2005) and Zhu and Santerre (2002) present results using 5 seconds of data, with dual-frequency receivers and choke-ring antennas: 61 % to 69 % success without height information and 93 % to 96 % success with height information - a 25 to 30 point improvement. With single-frequency receivers and patch antennas, we observed 17.94 % success without height information and 64.49 % / 69.89 % success with ±5.0cm / ±2.5cm height constraints - a 46 to 52 point improvement. In other words, the single frequency/patch antenna test *with* height information was as successful as the dual-frequency/choke ring antenna test *without* height information. However, the dual-frequency test was conducted with a longer baseline, and a 70% success rate is insufficient for most practical applications.

Results with an F-ratio threshold of 4.0 show reliability of 99.9% and mean ambiguity resolution times of 969.5, 421.7 and 374.3 seconds, height information reducing times by 56% and 61%. These results are a lot less impressive than the 140 second initialisation times reported by Weisenburger (1997), and may be due to Weisenburger's use of choke ring antennas; this would fit with the report in Takasu and Yasuda (2008) that mean initialisation times were 131.2 seconds with a choke ring antenna and 652.8 seconds with a patch antenna.

Results of tests with various F-ratio thresholds are shown in figure 4.7 . Higher F-ratio thresholds increase the reliability of ambiguity resolution, at the cost of increasing time to ambiguity resolution. Height constraints increase the reliability and decrease the mean resolution time; for example, for a success rate of 99.5 % the ±5.0cm calculations have a 47% lower mean time than the calculations without height constraints while the ±2.5cm calculations are 61% lower. For ±10.0cm and ±25.0cm the reductions are 39% and 16% respectively.

Results of the test using pseudo-observations in the presence of a priori height biases are shown in table 4.3 on page 105. In every case, the ambiguity resolution success rate was at least 99.5%. Only mean ambiguity resolution times are displayed, for purposes
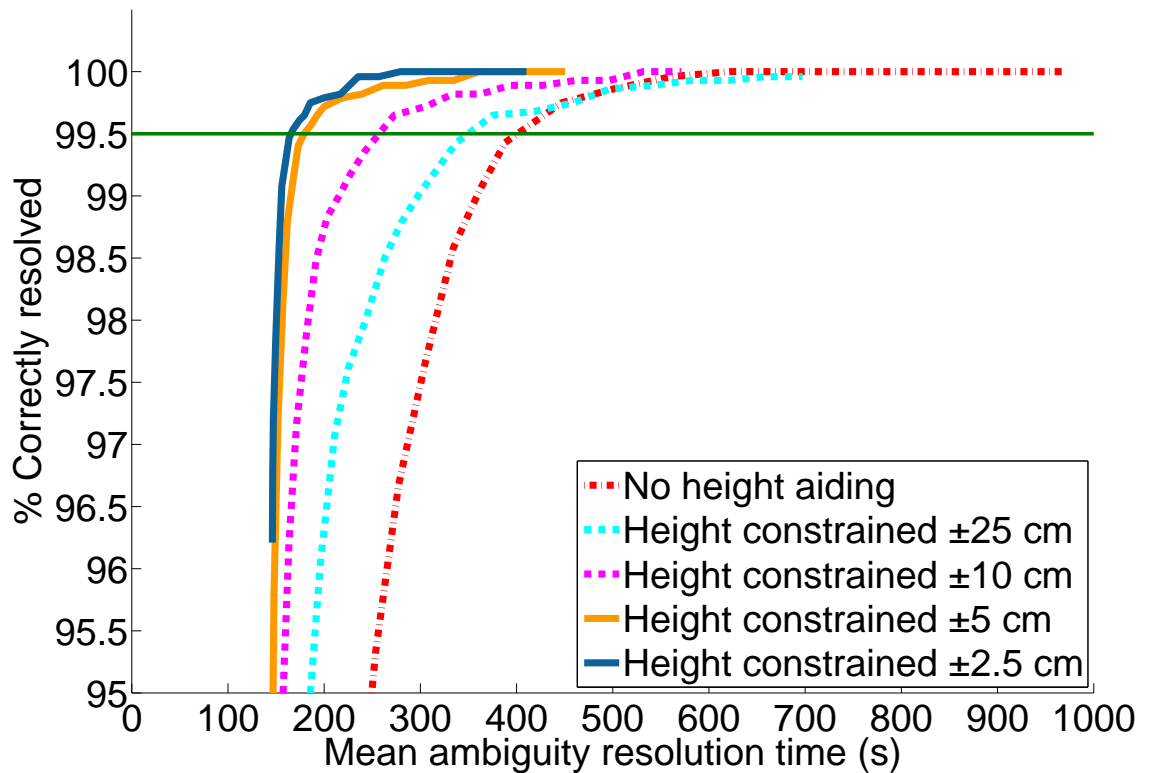
Figure 4.7: Ambiguity resolution time and success rate for various F-ratio thresholds, with height constraints between of ± 2.5cm and ± 25.0cm

of table clarity; for tables with median and 95% times, see appendix A on page 152.

When the a priori height has no bias the pseudo-observation based algorithm performs well. When there is a bias on the a priori height, the pseudo-observation based algorithm performs less well; with a bias, the pseudo-observation based algorithm never outperforms the ± 5.0cm minimum and maximum algorithm, and only outperforms the ± 2.5cm algorithm when the height bias is greater than 2.5cm.

4.4.2 Kinematic Experiment

The results reported are for a comparatively short baseline, which strengthens the assumption that errors such as ionospheric and tropospheric effects will be common at the receiver and the fixed base station. For golf course mowing applications this assumption remains reasonable, as most golf courses are only so large. For this work to be generalised to applications covering larger areas, further work could look at longer

A priori	Assumed height accuracy						
height bias	No a priori height	Minimum & maximum		Pseudo-observation standard deviation			
		±2.5cm	±5.0cm	2.5cm	5cm	10cm	20cm
Accurate	969.5	374.3	421.7	431.2	463.9	566.0	731.0
2.5cm offset	969.5	424.6	441.7	732.7	640.3	644.7	761.0
5.0cm offset	969.5	908.3	507.9	1390.3	1109.9	866.9	1050.4

Table 4.3: Mean ambiguity resolution times (in seconds) for several algorithms, when there is a bias on a priori height information. This shows how different algorithms handle topographical map errors.

baselines, and interpolating errors between multiple base stations.

Satellites in view	Height precision	F-ratio threshold	Solved right	Time to ambiguity resolution	
				Mean (s)	Median (s)
	None	3.6	16/18	619	480
	±25.0cm	3.2	17/18	528	402
4 - 7	±10.0cm	3.2	17/18	365	358
	±5.0cm	2.8	17/18	253	242
	±2.5cm	2.5	17/18	222	198
	None	4.0	15/18	746	640
	±25.0cm	3.8	16/18	640	548
4 - 6	±10.0cm	3.6	16/18	535	432
	±5.0cm	3.0	16/18	343	290
	±2.5cm	2.8	17/18	306	258
	None	4.0	10/18	1063	738
	±25.0cm	4.0	11/18	939	644
4 - 5	±10.0cm	4.0	12/18	735	548
	±5.0cm	4.0	12/18	599	516
	±2.5cm	4.0	13/18	512	334

Table 4.4: Kinematic ambiguity resolution times with various acceptance criteria.

As shown in table 4.4, as only ninety minutes of data were available and only 18 attempts at ambiguity resolution were performed per data set, the success rate is hard to report precisely. As a test that starts five minutes from the end of the data set can only succeed if it does so in five minutes or less, failures are reported that may have been successes had a longer data set been available.

The on-the-fly resolution times are greater than the stationary test times, reflecting the greater noise and higher frequency of cycle slips found in the kinematic data. However, the benefits of using height information can still be observed, in the form of a 12% to 64% reduction in mean ambiguity resolution time and a modest improvement in

success rate. As one would expect, the greatest improvements in performance were with the most precise height information; with ±2.5 cm precision height information the mean improvement was 58%; with ±5 cm, ±10 cm and ±25 cm data the mean improvements were 52%, 33% and 13% respectively.

4.5   Conclusions

In this chapter an algorithm and experimental results have been presented for the integration of a priori height information in the integer ambiguity resolution process. The algorithm differs from prior work by:

- Representing the a priori height not as a single height but as an acceptable range (which is better suited to encoding limited-precision height information) while producing a maximum likelihood solution.

- Employing the computationally efficient LAMBDA algorithm.

The experimental results differ from prior work by:

- Using the new algorithm

- Using low-cost single frequency receivers and patch antennas

- Using 24 hours of data to account for the varying satellite constellation

- By being conducted on ground in moderate-multipath conditions rather than at sea.

Experiments without a height constraint were found to be comparable to typical results in the literature using single-frequency receivers and patch antennas. Height information accurate to ±5.0cm reduced mean ambiguity resolution times by 47% to 56%, while information accurate to ±2.5cm reduced times by 61% to 67%. This is roughly comparable to the ambiguity resolution time reductions of 40% to 80% reported from the use of choke ring antennas, as described in section 3.5.1 on page 81. Less precise

height information yielded less impressive improvements, with ±10.0cm reducing the mean time by 39%, and ±25.0cm data offering only a 16% performance improvement.

Kinematic experiments showed lower overall performance, but similar percentage improvements from the proposed algorithm. Success rate was also improved, but as only 90 minutes of data was available, generalisations about the success rate are difficult to make.

Experiments integrating height information using a pseudo-observation produce fair results when there is no bias on the a priori height, but poor results when there is a bias on the a priori height. The proposed algorithm outperforms the pseudo-observation based algorithm from the literature when a height bias is present.

Chapter 5

Multiple receiver configurations
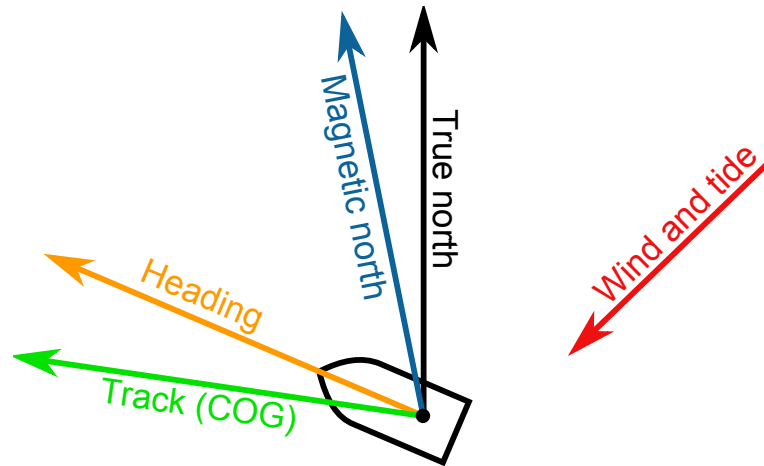
Figure 5.1: Heading (the direction a vehicle is pointing) and track (the direction a vehicle is travelling, also known as Course Over Ground or COG) for a boat. Were a compass fitted to the boat, it would measure the angle between the heading and magnetic north (which is different from true north due to *magnetic variation*), and may be influenced by the boat's own magnetic field. As the track is the direction a vehicle is travelling, it is undefined if the vehicle is stationary.

Single-antenna consumer GPS receivers are able to determine the *track* (also known as Course Over Ground or COG) of the receiver by measuring the Doppler shift on satellite signals. However, when a vehicle is stationary its track is undefined; when a vehicle is moving slowly its track may be difficult to measure precisely; and in some applications vehicles have a heading which is distinct from their track, as illustrated in figure 5.1. For our application we do not anticipate big differences between course and heading, but we do anticipate slow speeds making the vehicle's track hard to measure precisely using Doppler shift.

To address these issues, some high-precision GPS systems will fit a vehicle with two or three GPS receivers; with high precision GPS measurements taken at the front and rear of a vehicle, the vehicle's heading and pitch can be precisely determined even

when it is stationary; with three receivers, the vehicle's heading, pitch and roll can all be determined (See, for example, the Oxford Technical Solutions RT-3003(OXTS, 2009)). There also exist *GPS compass* products which, instead of positioning two roving receivers relative to a fixed base station, position them relative to one another only, allowing them to determine a heading with high precision when no base station coverage is available (See, for example, the JRC JLR-10 (JRC, 2002)).

In applications where there are multiple rover receivers fitted to one vehicle and the distance between the antennas is known, ambiguity resolution times can be improved by using that information.

5.1 Previous work

Teunissen (2006) describes an algorithm for single-epoch, single-frequency ambiguity resolution between two receivers a known distance apart. The algorithm takes the LAMBDA orthogonal decomposition from equation 3.63 on page 63, i.e.

$$\min_{\boldsymbol{a},\boldsymbol{b}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|_{Q_y}^2 \ \text{ with } \boldsymbol{a} \in \mathbb{Z}^{\mathrm{n}}, \boldsymbol{b} \in \mathbb{R}^3$$

applying the decomposition from equation 3.64 on page 63, i.e.

$$\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{a} - \boldsymbol{B}\boldsymbol{b}\|_{Q_y}^2 = \|\hat{\boldsymbol{e}}\|_{Q_y}^2 + \|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2 + \left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|_{Q_{\hat{b}(a)}}^2$$

but with the additional constraint that $\|\boldsymbol{b}\|_{I_3}^2 = l^2$ where l is the distance between the antennas. Because this violates the assumption that the third term of equation 3.64 on page 63 can always be minimised to zero, a search is performed to find the value of $\boldsymbol{b}$ which minimises $\left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|_{Q_{\hat{b}(a)}}^2$ subject to the $\|\boldsymbol{b}\|_{I_3}^2 = l^2$ constraint. Because the constraint is a sphere of radius l centred at the origin while $\left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|_{Q_{\hat{b}(a)}}^2 = c^2$ is an ellipse centred at $\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right)$, the minimum value is found when the ellipse just touches the sphere, and hence their normal equations are equal. The minimised value of $\left\|\hat{\boldsymbol{b}}\left(\boldsymbol{a}\right) - \boldsymbol{b}\right\|_{Q_{\hat{b}(a)}}^2$ is then added to $\|\hat{\boldsymbol{a}} - \boldsymbol{a}\|_{Q_{\hat{a}}}^2$. Teunissen (2006) calls this a quadratically-constrained least squares problem, and sources such as Buist et al. (2010) call the

Source	Receivers	Antennas	Test type	Without length constraint	With length constraint
Teunissen et al. (2010)	Sokkia GSR 2700 ISX	Internal pinwheel	Heading, single epoch	75% to 82% right	99% to 100% right
Giorgi et al. (2010a)	Trimble 5700	Trimble Zephyr	Heading, single epoch	15% to 20% right	50% to 70% right
Giorgi et al. (2010b)	Trimble R7 and SSi	Trimble Zephyr Geodetic and Geodetic W	Attitude, single epoch	40.1% right	97.1% right
Lachapelle et al. (1993a)	NovAtel GPSCard	Choke ring	Position, TTAR	335 seconds	181 seconds
Lachapelle et al. (1993a)	NovAtel GPSCard	No choke ring	Position, TTAR	810 seconds	471 seconds
Zheng and Gebre-Egziabher (2009)	NovAtel Superstar II	Patch antennas	Position, TTAR	71 to 390 seconds	22 to 139 seconds
Weisenburger (1997)	NovAtel MiLLe-nium	Choke ring	Position, TTAR	490 seconds	220 seconds
Buist et al. (2009)	Simulated	Simulated	Position, single epoch	24% to 97% right	53% to 100% right
Pinchin et al. (2008)	NovAtel Superstar II	Trimble Zephyr	Position, single epoch	0% right	31.6% right
Buist et al. (2010)	Simulated	Simulated	Position, single epoch	19% right	42% right

Table 5.1: Summary of key results from papers regarding multiple roving receivers, and the Time To Ambiguity Resolution (TTAR) or success rate they report.

algorithm Constrained LAMBDA (C-LAMBDA). Further information is available in Giorgi and Teunissen (2009). In Teunissen et al. (2010) experimental results with this algorithm are described. In one experiment, six Sokkia GSR 270 ISX receivers with high quality antennas were placed in a straight line, with separation distances from 60 centimetres to 10 metres. 35 minutes of 1Hz data was logged, with satellite numbers varying from 6 to 8. The single-epoch single-frequency ambiguity resolution algorithm was tested on the logged data, and tests were performed with certain satellites excluded to simulate the effects of reduced satellite visibility. The ambiguity resolution success rate was then compared with and without the baseline length constraint. Without the baseline constraint and with six satellites in view, the success rate varied between 75% and 82% for different baselines; with the baseline constraint applied, the success rate rose to 99% to 100%. When 8 satellites were in view, the success rate without baseline length constraint was 85% to 94%, and with the constraint was 100% for all baseline lengths. Somewhat unexpectedly, there was a high level of multipath on the seventh satellite at certain times, so the success rate for seven satellites was less than that for six satellites. In another test on a boat, (using a Leica SR530 and an Ashtech Z12, both with high quality antennas, and a 2 metre baseline) the success rate with baseline length constraint was 99.5% during a 2.5 hour test on a boat with satellite visibility varying between 6 and 8 satellites.

Luo and Lachapelle (1999) report on simulated on-the-fly ambiguity resolution between multiple moving platforms, with a number of baselines and receiver configurations. With 8 satellites in view and a 1.5 km baseline between moving platforms, mean times to resolve the ambiguities between the moving platforms vary between 48 seconds and 300 seconds, depending on the level of receiver and multipath noise simulated. The paper suggests the ambiguity resolution procedure can be sped up by adding the constraint that, for three moving platforms, the three ambiguities must form a loop; this is implemented by finding the top n candidates for each baseline, considering the n^3 possible combinations, discarding those solutions which do not form loops; and evaluating the remaining solutions based on the residuals of the three baselines.
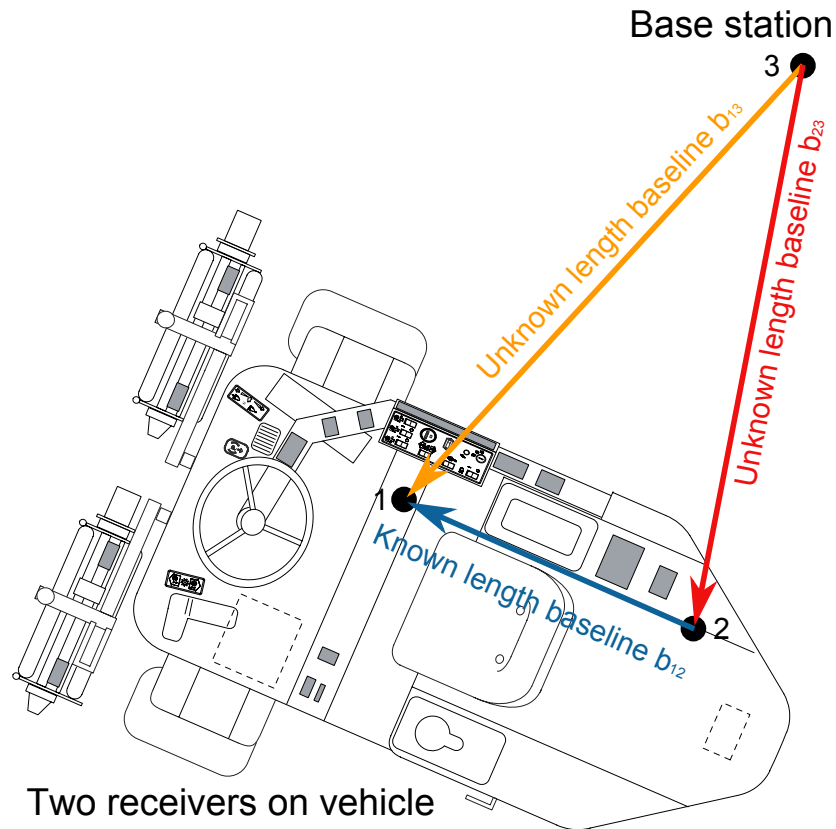
Figure 5.2: A common multiple roving receiver arrangement involves two receivers on a vehicle, a known distance apart, and one fixed base station at a known location.

For the test cases mentioned above, the mean times to resolve ambiguities are 40 to 290 seconds (dropping by 16.4% and 8.4% respectively). Other simulated situations (such as for very low noise or dual frequency data) present greater improvements, with time reductions as high as 58%. The simulation with an initialisation time averaging 300 seconds is in line with Odijk et al. (2007); the faster simulations with shorter initialisation times may be based on overly-optimistic assumptions about noise.

Lachapelle et al. (1993a) investigates a similar means of improving ambiguity resolution; with two base stations and two roving receivers mounted on the same vehicle, ambiguity resolution being simpler for known-length baselines. Known-length baselines are resolved "using a few seconds of observations", candidate ambiguities are generated for all the unknown-length baselines, and then those solutions which do not form a loop given the already-fixed ambiguities are rejected. An experiment was performed with four NovAtel GPSCards, and repeated with and without choke ring ground planes. The roving receivers were less than 5 kilometres from the fixed base stations, and at

Source	Algorithm
Lachapelle et al. (1993a)	Best n solutions found for known and unknown length baselines; solutions where three ambiguities do not form a loop discarded.
Buist et al. (2009)	Known-length baseline resolved; known ambiguity added to float solution for one rover, to give equivalent at other rover; two float solutions for other rover averaged; standard LAMBDA with averaged float solution.
Pinchin et al. (2008)	Known length baseline resolved; LAMBDA gives top 500 solutions for one rover; known length baseline added to give corresponding solution for other rover; residuals at both rovers summed.
Weisenburger (1997)	Kalman filter, solution constrained with linear approximation of baseline length constraint.

Table 5.2: Summary of multiple roving receiver positioning algorithms from the literature

least 6 satellites were in view at all times. Ambiguity resolution was performed for every 60 seconds throughout the data gathered. Without choke ring ground planes or quadruple receivers, ambiguity resolution took 810 seconds; with choke ring ground planes but without quadruple receivers ambiguity resolution took 335 seconds; without choke ring ground planes but with quadruple receivers, ambiguity resolution took 471 seconds; and with both choke ring ground planes and quadruple receivers, ambiguity resolution took 181 seconds. In other words, the use of choke ring ground planes reduced initialisation time by approximately 60%, while the use of quadruple receivers reduced initialisation time by approximately 45%.

Buist et al. (2009) also investigates improving ambiguity resolution with multiple receivers, simulating a system with two single-frequency base stations and a vehicle fitted with two single-frequency roving receivers a known distance apart. The paper attempts to resolve integer ambiguities instantaneously, by combining C\A code and carrier phase measurements. Two techniques are described, a sub-optimal technique where the known-length baseline between receivers on the same vehicle is resolved using the algorithm from Teunissen (2006), then the unknown-length baseline is resolved using combined observations from all receivers; and an optimal technique which resolves both the known-length and unknown-length baselines at the same time. In the suboptimal

algorithm first the fixed-length baseline ambiguity (a_{12}) is resolved, solving

$$\check{a}_{12} = \arg\min_{a_{12}} \left(\|\hat{a}_{12} - a_{12}\|^2_{Q_a} + \min_{b_{12}} \left(\left\|\hat{b}_{12}\left(a_{12}\right) - b_{12}\right\|^2_{Q_{b(a)}} \right) \right) \qquad a_{12} \in \mathbb{Z}^n \ \|b_{12}\| = l$$
$$(5.1)$$

where l is the known length. Float solutions ($\hat{a}_{23}$ and $\hat{a}_{13}$) are found between the base station and the two rovers; the unknown-length baseline is found with:

$$\check{a}_{23} = \arg\min_{a_{23}} \left(\|\hat{a}_{23}\left(a_{12}\right) - a_{23}\|^2_{\frac{3}{4}Q_a} \right) \qquad (5.2)$$

where the conditional float solution is given by:

$$\hat{a}_{23}\left(a_{12}\right) = \frac{\hat{a}_{23} + \left(\hat{a}_{13} - a_{12}\right)}{2} \qquad (5.3)$$

in other words, the solution between the receivers is used to transform the solution at receiver 1 to its equivalent at receiver 2, and the mean of those two float solutions is taken. The paper states the performance difference between the optimal and suboptimal algorithms is negligible. The solution with the lowest squared residual is selected; in the 1 base/2 rover configuration this is not simply the sum of the squared norms for each rover; a weighting matrix takes into account the fact that both rovers' double-differenced observations are relative to the same base station. Simulations were performed for several configurations; 1 base/1 rover, 1 base/2 rovers and 2 bases/2 rovers; from 5 to 8 satellites in view; C\A code measurements with standard deviations from 5 cm to 30 cm; and code phase measurements with standard deviations from 1 mm to 3 mm. Information is then provided on the success rate of instantaneous ambiguity resolution. Results vary substantially depending on configuration assumptions; with a single rover and single base station, a 5cm standard deviation on C\A code measurements and with 6 satellites in view, the success rate is 97%; on the other hand, with a 30cm standard deviation, the success rate drops to 24%. With two roving receivers and two base stations, those rates increase to 100% and 53% respectively. Hence, in the former case a 4-receiver configuration improves the success rate by 3%,

while in the latter case the rate is improved by 19%. However, all these results depend on how reasonable the simulation assumptions are; even a 30cm standard deviation on C\A code measurements seems an optimistic assumption when Cannon et al. (1993) states the (high performance) NovAtel GPSCard with a choke ring antenna has a code accuracy of 15 to 70 cm; and that for standard C\A code receivers 2 to 3 m is typical.

Pinchin et al. (2008) also investigates instantaneous ambiguity resolution with multiple single-frequency receivers; a vehicle was fitted with two low-cost NovAtel Superstar II receivers with high quality Trimble Zephyr antennas; and a base station 2.7 km from the test site was operated with a high quality antenna and receiver (Trimble NetR5 and Trimble Zephyr Geodetic Antennae). 28 minutes of data was gathered at a rate of 1Hz, with the vehicle in motion. An algorithm is used which, for each epoch, first determines the ambiguity for the (known length) baseline between the receivers on the vehicle; then finds the top 500 ambiguity solutions between the base station and one receiver on the vehicle; then uses the ambiguity already found to find the equivalent solutions at the second receiver on the vehicle. Each solution is evaluated based on the sum of the squared norms at the two receivers on the vehicle, i.e.

$$\check{a}_{23} = \arg\min_{a_{23}} \left(\|\hat{a}_{23} - a_{23}\|^2_{Q_a} + \|(\hat{a}_{13} - a_{12}) - a_{23}\|^2_{Q_a} \right) \tag{5.4}$$

In the experiment, the results from the single epoch algorithm were compared to the results for an algorithm run over all epochs; single epoch algorithm results were classed as correct if they matched the all-epoch algorithm. The single epoch algorithm chose the same vehicle position as the all-epoch algorithm in 31.6% of epochs; however, at no time in the test did the squared norm ratio exceed 1.3 (as mentioned in section 3.4.4.3 on page 65, it is common to require a higher ratio to be confident that a given solution is a correct one).

Another demonstration of instantaneous ambiguity resolution can be found in Giorgi et al. (2010a), which describes experiments using a combination of dual-frequency and single-frequency receivers to determine the position and attitude of ships entering and

leaving the harbour of Hong Kong. While conventionally this would be done with a dual-frequency base station and three dual-frequency receivers fitted to the ship, instead one dual-frequency receiver and two single-frequency receivers were fitted to the boats tested, with baseline length constraints used to improve ambiguity resolution on the two single-frequency baselines. Single epoch ambiguity resolution was attempted on these baselines. The success rate of resolving ambiguities between the receivers on the ship was 15-20% without the length constraint and 50-70% with the length constraint. The paper states the inclusion of two or three more measurements would achieve a close-to-100% success rate in resolving the fixed-length baseline ambiguities although no explicit results are given.

Zheng and Gebre-Egziabher (2009) also investigates use of multiple receivers and a fixed-length baseline. Three roving receivers and one fixed receiver are used; the known-length baselines between the three roving receivers (e.g. a_{12}) are first resolved (using the baseline length constrained algorithm of Teunissen (2006)), then the double-differenced measurements from each of the three roving receivers can be offset to have a single set of ambiguities relative to the base station. The squared norms at the three roving receivers are summed, i.e.

$$\check{a}_{23} = \arg\min_{a_{23}} \left(\|\hat{a}_{23} - a_{23}\|_{Q_a}^2 + \|(\hat{a}_{13} - a_{12}) - a_{23}\|_{Q_a}^2 + \|(\hat{a}_{43} - a_{42}) - a_{23}\|_{Q_a}^2 \right) \quad (5.5)$$

which is to say, a similar criteria to Pinchin et al. (2008). The LAMBDA technique was used to minimise the sum of squared residuals, and the ambiguity acceptance criteria were a ratio test with a critical value of 2.0 and an estimate of the success rate, minimum 99.99%. An experiment was conducted where three NovAtel SuperStar II receivers with patch antennas were fitted to a model aeroplane; tests were performed in stationary, taxiing, and flying conditions, within a few hundred metres of a fixed base station. Times to perform ambiguity resolution were reported for one roving receiver and three roving receivers. During the static test, with 7 satellites visible, ambiguity resolution for a single roving receiver took 327 seconds, while with three receivers the

time dropped to 139 seconds. During the test while the plane was taxiing, 7 satellites were in view, and the times were 390 seconds and 116 seconds with one and three receivers respectively. In the in-flight test, 9 satellites were in view, and the times were 71 seconds and 22 seconds. In other words, for the three tests, the triple-receiver configuration reduced ambiguity resolution times by 57%, 70%, and 69%. Results are also reported with two rather than three roving receivers; depending on the test, the double-receiver configuration reduced ambiguity resolution times by 28% to 69%.

Weisenburger (1997) examines the use of dual roving receivers, with constraints applied as observations to a Kalman filter. Unlike Teunissen (2006)'s quadratic constraints, Weisenburger implements the baseline length constraint by linearising the baseline relationship; because a vehicle-sized sphere is poorly represented by a linear approximation, accurate approximate coordinates are required to linearise around; Weisenburger waits several epochs for the filter to warm up to address this. Experimental results are presented; in one test, two NovAtel MiLLenium dual frequency receivers with choke ring antennas were fitted to a boat, and another receiver was placed on the shore. The boat was within 2 kilometres of the base station at all times; data was gathered at 10Hz with a 10° elevation mask; and 6 or 7 satellites were visible throughout the 90 minute test. Ambiguity resolution was performed at different start points throughout the data, once for dual-frequency data; and once with only single-frequency data. Without a constraint on the distance between roving receivers, the mean time to perform ambiguity resolution was 490 seconds; with the constraint the time dropped to 220 seconds; in other words, the baseline constraint reduced ambiguity resolution times by 55%.

Giorgi et al. (2010b) gave details of an algorithm named Multivariate Constrained LAMBDA (MC-LAMBDA), which resolves multiple known-length baselines - such as when there are three receivers on a vehicle, to measure the vehicle's pitch, yaw, and roll. When there is only one known-length baseline the algorithm is equivalent to Constrained LAMBDA (C-LAMBDA) from Teunissen (2006); MC-LAMBDA is specifically for situations where there are two or more known length baselines. The relative positions

of the receivers on a vehicle are measured in the vehicle frame, then an orthonormal rotation matrix gives the receivers' positions in ECEF coordinates. Where conventional LAMBDA searches for a position baseline vector, MC-LAMBDA performs a search for the constituent values of the rotation matrix, subject to the constraint that the matrix must be orthonormal. Similar to C-LAMBDA, a LAMBDA search is performed with a term added to solution squared norms, to represent the difference between the unconstrained and constrained solutions. The paper reports on experiments applying the algorithm to approximately 3 hours of single-frequency data, gathered using a Trimble R7 and two Trimble SSi receivers, with Trimble Zephyr Geodetic L1/L2 and Trimble Geodetic W ground plane antennas held stationary at ground level. 9 satellites were tracked, and performance with fewer satellites was tested by discarding data for some satellites. For single-epoch, single frequency ambiguity resolution of one fixed length baseline (a case for which MC-LAMBDA is equivalent to C-LAMBDA algorithm) with six satellites in view, a length constraint improved the ambiguity resolution success rate from 40.1% to 97.1%; with two fixed length baselines (using the MC-LAMBDA algorithm) the success rate for both baselines was 100.0%. Another test, to determine amount of data required to achieve a 99% success rate, found that with 6 satellites in view the time without constraints was greater than 20 seconds; the time with a single constrained baseline was 4 seconds; and the time with two constrained baselines was instant. Tests on an aircraft in flight demonstrated similar results.

Buist et al. (2010) reports on simulated experiments, using the MC-LAMBDA algorithm and a combination of known-length and unknown-length baselines. First MC-LAMBDA is used to resolve known-length baselines, then standard LAMBDA is used to resolve the unknown-length baseline, bootstrapped from the MC-LAMBDA solutions. Tests perform single-epoch, single-frequency ambiguity resolution, with the number of known-length baselines varying between 0 (standard LAMBDA) and 6 (MC-LAMBDA with 3 receivers per vehicle). The benefits of MC-LAMBDA are clearest when the code noise is simulated with a standard deviation of 15cm; with no known length baselines, the success rate is 19%, which rises to 42%, 51% and 56% with

2, 4 and 6 known-length baselines respectively.

5.2 Adaption to the considered application

Preliminary experiments were performed with the algorithms of Buist et al. (2009) and Pinchin et al. (2008). Both Buist et al. and Pinchin et al. report success rates for single-epoch, single-frequency ambiguity resolution (one entirely simulated and one with high quality antennas) but in our experiments with patch antennas the success rate for this was low; multi-epoch ambiguity resolution was investigated instead, with ambiguity resolution times and success rates as measures of algorithm quality.

5.2.1 Float solution difference test

It was found that multipath was a common cause of incorrect solutions, but that when present it is rarely identical at both receivers, leading to the two roving receivers' float solutions not matching up. The following test was proposed (using the terminology of equation 5.3 on page 114) to detect this situation:

$$\max \text{abs} \left(\hat{a}_{23} - (\hat{a}_{13} + a_{12}) \right) < k \tag{5.6}$$

where $\max \text{abs} (a)$ means to take the largest absolute value from the vector a; and k is a threshold value (1.5 L1 cycles in our experiments). When the largest difference between float solutions at the two roving receivers exceeds k it is taken as a sign of multipath, and the solution is not accepted until the difference has reduced to below k.

5.2.2 Fixed length baseline residual in ratio test

In line with previous chapters, an F-ratio test (See section 3.4.4.3 on page 65) is used during multi-epoch ambiguity resolution, in order to choose whether to accept a solution or gather more data before deciding. However, instead of a ratio of the squared

residuals of the two best solutions for the fixed length baseline, it was proposed to add to them the squared residual of the best solution for the known-length baseline. When multipath at the roving receivers is low, the residual between the two roving receivers should be low, making the new ratio test less conservative, while higher multipath will lead to a greater residual, making the new ratio test more conservative.

In other words, while the standard F-ratio test is defined by equation 3.73 on page 65 as:

$$\frac{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - a_2\|_{Q_{\hat{a}}}^2}{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - a_1\|_{Q_{\hat{a}}}^2} > k$$

the proposed test

$$\frac{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - a_2\|_{Q_{\hat{a}}}^2 + \|\hat{e}_{fixed}\|_{Q_y}^2 + \|\hat{a}_{fixed} - \check{a}_{fixed}\|_{Q_{\hat{a}}}^2}{\|\hat{e}\|_{Q_y}^2 + \|\hat{a} - a_1\|_{Q_{\hat{a}}}^2 + \|\hat{e}_{fixed}\|_{Q_y}^2 + \|\hat{a}_{fixed} - \check{a}_{fixed}\|_{Q_{\hat{a}}}^2} > k \qquad (5.7)$$

where $\|\hat{e}_{fixed}\|_{Q_y}^2$ is the float solution residual for the fixed-length baseline and similarly $\|\hat{a}_{fixed} - \check{a}_{fixed}\|_{Q_{\hat{a}}}^2$ is the squared norm between the float solution and the chosen fixed solution.

5.2.3 Performance comparison criteria

As both the modifications outlined in previous sections will lead to variations in both success rate and ambiguity resolution time, it was decided to gather data for a range of F-ratio thresholds, and to make comparisons between ambiguity resolution times with a success rate of 99.5% or higher.

5.3 Experimental Configuration

Section 4.3 on page 99 describes the collection of 24 hours of 10Hz GPS data from a stationary vehicle in a moderate multipath environment. Two single-frequency GPS receivers were present on the vehicle, and a third receiver, 280 metres away, served as a fixed base station. The distance between the two antennas on the vehicle was 1.9 metres, as shown in figure 5.3 . The baseline between the roving receivers ran from

Figure 5.3: Test vehicle fitted with two single frequency GPS receivers 1.9 metres apart.

north to south.

The recorded data was processed by performing ambiguity resolution starting every 30 seconds throughout the data, taking as long as needed to meet the acceptance threshold, and recording the ambiguity resolution times and success rates. This processing was performed using several different algorithms:

- A single receiver, and an F-ratio threshold between 1.0 and 3.0

- Dual receivers using Buist et al. (2009)'s suboptimal algorithm, an F-ratio threshold between 1.0 and 3.0, with and without the float difference test of equation 5.6 on page 119.

- Dual receivers using Pinchin et al. (2008)'s algorithm, an F-ratio threshold between 1.0 and 3.0, with and without the float difference test.

- Dual receivers using Zheng and Gebre-Egziabher (2009)'s algorithm, a squared
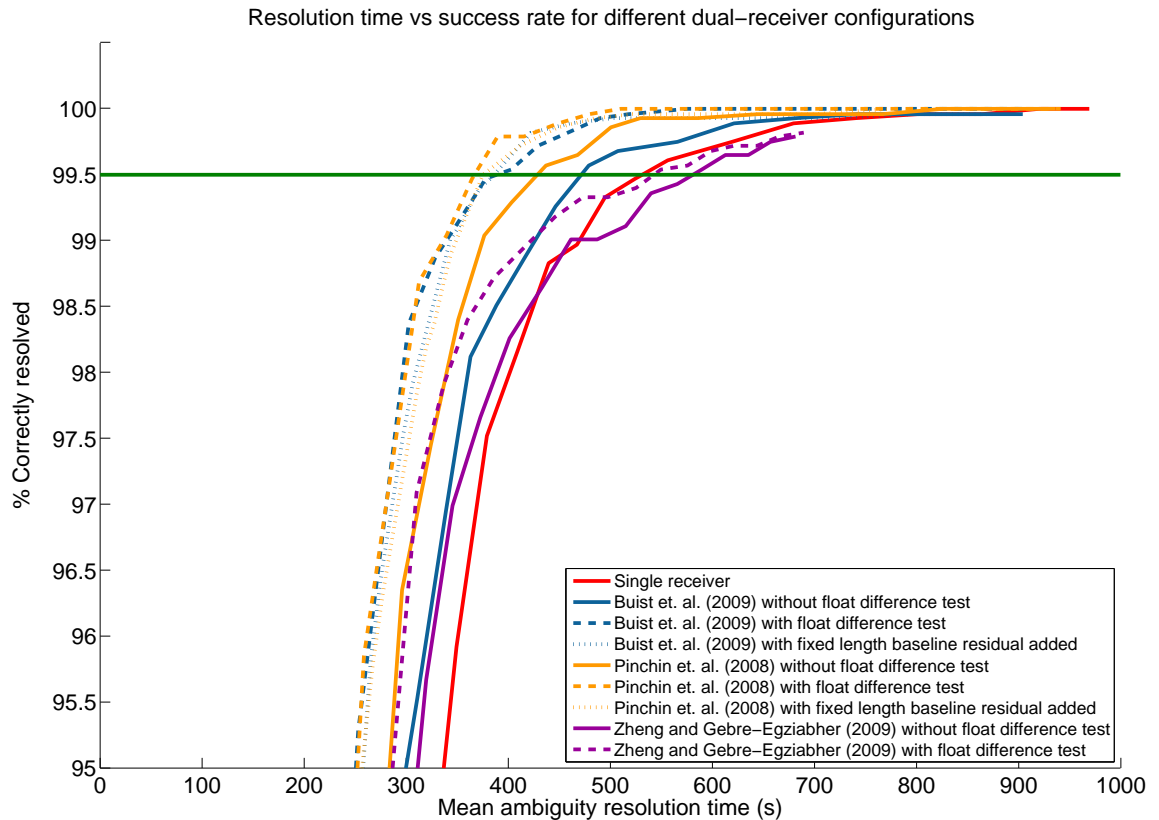
Figure 5.4: Graph of resolution time vs. success rate as F-ratio threshold is varied, for several algorithms.

norm ratio threshold between 1.0 and 6.0, with and without the float difference test.

- Dual receivers using Buist et al. (2009)'s suboptimal algorithm with the ratio test with fixed length baseline residual added, thresholds between 1.0 and 3.0.

- Dual receivers using Pinchin et al. (2008)'s algorithm with the ratio test with fixed length baseline residual added, thresholds between 1.0 and 3.0.

Results graphs and tables are presented in the following section.

5.4 Results

The results shown in table 5.3 show a modest reduction in ambiguity resolution time when dual-receiver algorithms are used; with an F-ratio of 2.5, the two dual receiver

Algorithm used	Ratio threshold	Solved right (%)	Time to ambiguity resolution		
			Mean (s)	Median (s)	95% (s)
Single receiver	*2.0*	*97.52%*	*379.1*	*320.0*	*968.0*
	2.5	99.47%	524.4	460.0	1364.0
	3.0	99.89%	679.6	604.0	1800.0
Buist et al. (2009) without float difference test	*2.0*	*96.95%*	*338.9*	*292.0*	*832.0*
	2.5	99.57%	478.9	422.0	1262.0
	3.0	99.89%	621.4	560.0	1800.0
Pinchin et al. (2008) without float difference test	*2.0*	*98.40%*	*351.0*	*304.0*	*852.0*
	2.5	99.86%	500.7	440.0	1348.0
	3.0	99.96%	644.7	578.0	1800.0
Zheng and Gebre-Egziabher (2009) without float difference test	*2.0*	*89.04%*	*262.6*	*220.0*	*636.0*
	2.5	*95.67%*	*319.6*	*264.0*	*786.0*
	3.0	*97.66%*	*372.6*	*312.0*	*892.0*
Buist et al. (2009) with float difference test	*2.0*	*99.18%*	*354.3*	*304.00*	*848.0*
	2.5	99.93%	490.2	430.0	1262.0
	3.0	100.00%	629.6	566.0	1800.0
Pinchin et al. (2008) with float difference test	2.0	99.50%	364.6	312.0	862.0
	2.5	100.00%	510.3	448.0	1342.0
	3.0	100.00%	651.8	584.0	1800.0
Zheng and Gebre-Egziabher (2009) with float difference test	*2.0*	*94.82%*	*284.6*	*237.0*	*664.0*
	2.5	*97.91%*	*336.9*	*280.0*	*818.0*
	3.0	*98.72%*	*386.1*	*324.0*	*914.0*
Buist et al. (2009) with fixed length baseline residual added	2.0	99.89%	461.1	412.0	1150.0
	2.5	99.93%	670.3	614.0	1800.0
	3.0	99.96%	875.9	820.0	1800.0
Pinchin et al. (2008) with fixed length baseline residual added	2.0	99.72%	411.6	364.0	976.0
	2.5	99.96%	589.0	534.0	1684.0
	3.0	100.00%	775.1	716.0	1800.0

Table 5.3: Results of dual receiver tests using various algorithms and F-ratio thresholds. Results with a success rate below 99.50% in italics.

algorithms reduce ambiguity resolution time by 8.6% and 4.5% respectively - substantially less than the 60% reductions reported by Zheng and Gebre-Egziabher (2009) with three receivers. Success rates also vary with F-ratio threshold; a 99.5% success rate can be achieved in 364.6 seconds (mean) with a 2.0 ratio using dual receivers and a float difference test, while the same success rate with a single receiver requires a ratio of 3.0 and takes 679.6 seconds (mean).

The relationship between success rate and ambiguity resolution time is illustrated in figure 5.4 on page 122. At the 99.5% confidence level (green horizontal line), the implementation of Zheng and Gebre-Egziabher (2009)'s algorithm, which uses a squared norm ratio test, takes 581 seconds without the float difference test and 543 seconds with it - longer than any of the algorithms which use the F-ratio test. It is due to this finding that other experiments use the F-ratio test.

Among experiments using the F-ratio test, at a 99.5% confidence level ambiguity resolution with a single receiver takes a mean time of 531 seconds; with dual receivers that reduces to 428 and 471 seconds; and with dual receivers and the float difference test the mean time is further reduced, to 370 and 388 seconds. In other words, for a constant success rate, dual roving receivers without a float difference test reduced mean ambiguity resolution time by 11% and 19%, while with the float difference test the reduction is 30%.

The impact of using the fixed length baseline residual is almost identical to the impact of the float difference test, reducing the ambiguity resolution time by 30%.

This compares favourably to results from Lachapelle et al. (1993a) without a choke ring, being a 21% improvement on the 471 second time they report. However, Lachapelle et al.'s results represent a 42% time reduction, greater than the 30% time reduction found above. A 30% reduction in initialisation time is also not as great an improvement as reported by Zheng and Gebre-Egziabher (2009), where 57% to 70% time reductions were found with three rover receivers, and 28% to 69% improvements with two rover receivers. This may be because Zheng and Gebre-Egziabher perform only three tests rather than testing throughout 24 hours of data.

Another possible difference is multipath at the base station; multiple roving receivers can average out multipath at the rovers, but can do nothing for multipath at the base station. If the level of multipath at the base station is low this is not a problem, but should it occur it would slow ambiguity resolution regardless of how many rover receivers were used. Further investigation, using multiple base station receivers, could address this. Needless to say, the assumption multipath at the base station is low is often a reasonable one, as base stations are usually positioned with a view to reducing multipath.

5.5   Conclusions

In this chapter, several algorithms for ambiguity resolution with dual roving receivers have been compared, and an additional acceptance test has been proposed and examined. Key differences to prior work include:

- Proposal of an acceptance test based on the difference between receivers' float solutions.

- Proposal of an acceptance test adding the residual of the fixed-length baseline to the residual of the unknown-length baseline when performing an F-ratio comparison.

- Comparison of algorithms by considering mean ambiguity resolution time when configured to achieve a success rate of 99.5% - rather than comparing single epoch ambiguity resolution success rates.

- Using 24 hours of experimental data, to account for the varying satellite constellation.

- Using data from low cost single frequency receivers fitted with patch antennas.

Compared to ambiguity resolution with a single receiver, the dual receiver algorithms alone improve mean ambiguity resolution time by 11% to 19%; with the float difference test or the fixed length baseline residual added the reduction is 30%.

Chapter 6

Alternative acceptance tests

Conventional ambiguity resolution finds best and second-best ambiguity solutions, then applies an acceptance test to decide whether to accept the best ambiguity solution, or to gather more data to be more confident the chosen solution is the correct one. The performance of this test is critical to minimising ambiguity resolution times; a test which is too conservative prolongs ambiguity resolution needlessly - but a test which is not conservative enough can allow the acceptance of incorrect solutions.

6.1 Previous Work

6.1.1 Tests based on comparison between best and second-best solutions

Verhagen (2004) gives details of a number of integer ambiguity validation algorithms; recall equation 3.55 on page 61:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{b} + \boldsymbol{e}$$

where $\boldsymbol{y}$ is the vector of double-differenced measurements, $\boldsymbol{a}$ and $\boldsymbol{A}$ are the vector of ambiguities and the relationship between ambiguities and measurements, $\boldsymbol{b}$ and $\boldsymbol{B}$ are the baseline vector and the relationship between baseline and measurements, and $\boldsymbol{e}$ is the vector of residuals. The variance-covariance matrix of the measurements is given by Q_y. Verhagen also defines $Q_y = \sigma^2 G_y$, where G_y is the cofactor matrix of Q_y and σ^2 is the variance factor of unit weight.

$\hat{\boldsymbol{a}}$ is the float ambiguity solution, $\breve{\boldsymbol{a}}$ the integer least squares ambiguity solution, and $\breve{\boldsymbol{a}}_S$ the second-best integer least squares ambiguity solution, allowing the following

definitions:

$$\hat{\Omega} = \hat{\boldsymbol{e}}^T G_y^{-1} \hat{\boldsymbol{e}} \tag{6.1}$$

$$\breve{\Omega} = \breve{\boldsymbol{e}}^T G_y^{-1} \breve{\boldsymbol{e}} \tag{6.2}$$

$$R = \breve{\Omega} - \hat{\Omega} = (\hat{\boldsymbol{a}} - \breve{\boldsymbol{a}})^T G_a^{-1} (\hat{\boldsymbol{a}} - \breve{\boldsymbol{a}}) \tag{6.3}$$

Verhagen examines six different acceptance tests. The three tests found most effective are: The *ratio test*, attributed to Euler and Schaffrin (1990):

$$\frac{R_S}{R} > k \tag{6.4}$$

the *difference test*, credited to Tiberius and De Jonge (1995) for:

$$R_S - R > k \tag{6.5}$$

and the *F-ratio test*, described by Frei and Beutler (1990):

$$\frac{\breve{\Omega}_S}{\breve{\Omega}} > k \tag{6.6}$$

where in each case k is a critical value which is empirically calibrated.

Verhagen generated 100,000 random samples of dual-frequency GPS data, then found single epoch float and fixed solutions for each sample, and the different acceptance tests were performed. For each sample there are four possible outcomes; accepting a correct solution (success), rejecting a correct solution (undecided), rejecting an incorrect solution (undecided), and accepting an incorrect solution (failure). The critical value (k) for each test was calibrated so that the chance of accepting an incorrect solution was 1%. The tests were then evaluated based on the number of correct acceptances; given that each test has the same failure rate, tests with higher success rates are preferable. The ratio test correctly accepted 48.5% of samples, the difference test 48.2%, and the F-ratio test 43.8%.

6.1.2 Monte Carlo based failure rate estimates

Teunissen and Verhagen (2004) discusses the ratio test and gives an algorithm for a *ratio test with fixed failure rate*; the software is configured with a failure rate, P_f such as 0.01. For each epoch of GPS data, LAMBDA is used to find the best and second-best solutions, and the squared norm ratio (k) is calculated. Simulated data is then generated, using the epoch's satellite geometry, and ambiguity resolution is performed using the k as the acceptance ratio. The failure rate for simulated data is tracked, and if simulation with a ratio of k produces a failure rate greater than P_f then k is too low a ratio, meaning the acceptance test has not been passed this epoch (A later paper, also by Teunissen and Verhagen (2009), suggests using a look-up table to avoid repeating simulation each epoch).

In a test with simulated data, a ratio test with fixed failure rate (0.5%) is compared to a ratio test with a fixed ratio (3.0), testing the probability of acceptance against the number of epochs of data used. The fixed ratio test only resolved ambiguities faster than the fixed failure rate test in one instance, and in that instance had a higher failure rate than the ratio test with fixed failure rate. However, when the fixed ratio test was slower than the fixed failure rate test, it did offer failure rates below the fixed failure rate.

Teunissen and Verhagen (2004) also discusses a similar test, named the Optimal Integer Aperture (OIA) estimator, which relies on the same metric as integer least squares, and it is optimal in the sense that it maximises the probability of success for a given probability of failure, assuming measurement noise is zero-mean and normally distributed. While the standard ratio test compares the squared norm of the best solution to the squared norm of the second-best solution, the OIA estimator compares the squared norm of the best solution to the sum of the squared norms of all other solutions. Comparing the OIA estimator with fixed failure rate to the ratio test with fixed failure rate, using experiments with simulated data, the difference in performance between the ratio test and the OIA estimator is small - on the order of 2% in some cases. The differences

between the OIA and the ratio test are most pronounced when the third-best solution is nearly as good as the first two solutions.

Teunissen and Verhagen (2009) further discusses the ratio test with fixed failure rate. An experiment is reported using 1 hour of 1Hz single-frequency data gathered with two receivers with a 2.5 to 4.5km baseline, using a Leica SR530 and a Trimble 4700 receiver. Stored data was processed with a fixed failure rate of 0.1%, and with fixed squared norm ratios of 1.5, 2.0 and 3.0. This was performed once with ambiguities assumed to be fixed throughout the experiment, and once with the ambiguities recalculated every epoch.

When integer ambiguities were considered constant, none of the acceptance tests accepted a wrong ambiguity. With a fixed squared norm ratio, in some epochs the ratio was not met despite the solution being the right one - a so-called *false alarm*. This happened in 3.9% of epochs with a ratio of 1.5 and 12.3% of epochs with a ratio of 3.0. It never happened when the 0.1% fixed failure rate was used.

When Epoch By Epoch (EBE) processing was used (where ambiguities are not assumed constant between epochs) wrong ambiguities were accepted more often; fixed ratios of 1.5, 2.0 and 3.0 produced failure rates of 7.4%, 2.7% and 0.7% respectively. No wrong acceptances occurred with the fixed failure rate algorithm, but the fixed failure rate algorithm rejected 58.9% of solutions even though they were correct - more false alarms than any of the fixed ratio approaches.

6.1.2.1   Look up table of ratio thresholds

Teunissen and Verhagen (2009) also suggests instead of performing a Monte Carlo simulation every epoch to determine the ratio threshold to produce a fixed failure rate, it is faster to use a look up table, as long as any imprecision in the look up table errs on the side of lower failure rates. Simulation results are presented indicating this would result in only fractionally lower rates of correct ambiguity fixing, while substantially reducing the real-time computational demands.

An equation such as 6.10 is used to estimate the *Integer Least Squares (ILS) failure rate* - the failure rate if no acceptance test was used - which makes up the first axis of the lookup table. The second axis of the table is the number of satellites in view. The body of the table contains ratio test thresholds sufficient to produce a target failure rate (such as 0.5%) when the ratio test is used. These values are generated by Monte Carlo simulation.

6.1.3   Non-Monte Carlo based failure rate estimates

Verhagen (2005a) discusses several ways of estimating the reliability of integer ambiguity estimation without relying on Monte Carlo methods. The paper reports on simulations comparing various success rate upper-bound and lower-bound functions to the success rates found by Monte Carlo simulation. The best-performing lower bound on success rates is based on integer bootstrapping. The paper also outlines a success rate approximation known as the Ambiguity Dilution Of Precision (ADOP), which is not a lower bound but which is expected to be a good approximation of the success rate.

6.1.3.1   Integer bootstrapping success rates

Integer bootstrapping is an ambiguity resolution method which does not perform quite as well as LAMBDA, but it offers close performance when the decorrelating Z-transformation of the LAMBDA method is used, and allows a closed form expression for its success rate. Teunissen (2001) gives this expression as:

$$P_{s,B} = \prod_{i=1}^{n} \left(2\Phi \left(\frac{1}{2\sigma_{i|I}} \right) - 1 \right) \tag{6.7}$$

where $\sigma_{i|I}$ is the standard deviation of ambiguity i conditioned on ambiguities 1 to $i-1$ (which is to say the diagonal elements of D in an LDL^T decomposition of the

variance-covariance matrix $Q_{\hat{a}}$) and

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2}v^2\right) dv \qquad (6.8)$$

This is the cumulative distribution function for a normally-distributed variable with mean 0 and variance 1, and is equivalent to

$$\Phi(x) = \frac{1}{2}\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) + \frac{1}{2} \qquad (6.9)$$

$$P_{s,B} = \prod_{i=1}^{n} \mathrm{erf}\left(\frac{1/2}{\sqrt{2}\sigma_{i|I}}\right) \qquad (6.10)$$

where erf is the Gauss error function. As integer bootstrapping is known to have a lower success rate than LAMBDA, the success rate for LAMBDA will be at least as high as the success rate for integer bootstrapping.

6.1.3.2 Integer bootstrapping with scaled acceptance regions

Teunissen (2005) describes a fixed failure rate algorithm based on integer bootstrapping, called the Integer Aperture Bootstrapped (IAB) estimator; by using a scaled-down version of the integer aperture pull-in region, an *acceptance region* with a known failure rate can be defined. The paper gives equations for the success, failure, and undecided rates for a chosen scaling factor. Integer bootstrapping is performed by taking a first float ambiguity and rounding it to the nearest integer, updating the float solution given that integer, then rounding a second integer, and so on. Because the calculation performed to update the float solution involves only multiplication, this can be represented by a triangular matrix - L from an LDL^T decomposition of the variance-covariance matrix $Q_{\hat{a}}$. In other words, if integer solution $\check{a}$ is the integer bootstrapped solution for float solution $\hat{a}$, then

$$\left|c_i^T L^{-1}(\hat{a} - \check{a})\right| \leq 1/2, \quad i = 1, \ldots, n \qquad (6.11)$$

where c_i^T is an n element column vector with the ith element 1 and all others 0. In other words, all elements of $L^{-1}(\hat{a} - \check{a})$ must be in the range from $\frac{-1}{2}$ to $\frac{1}{2}$, which is to say the range that would round to zero. For the Integer Aperture Bootstrapped (IAB) estimator a scaled down version of the acceptance criteria is used; instead of using the range $\frac{-1}{2}$ to $\frac{1}{2}$, a smaller range can be used, and solutions falling outside that smaller range (i.e. cases closer to the borderline) can use the float solution in preference to the integer solution. The smaller acceptance range means that confidence in those ambiguities which are accepted is higher. For example, integer solution $\check{a}$ is accepted as the solution for float solution $\hat{a}$ if it meets the criteria:

$$\left| c_i^T L^{-1}(\hat{a} - \check{a}) \right| \leq \beta \times \nicefrac{1}{2}, \quad i = 1, \ldots, n \tag{6.12}$$

where β is between 0 and 1.

In this configuration there are three possible outcomes to an attempt at ambiguity resolution with a given set of data; correct acceptance (success), incorrect acceptance (failure), and indecision. The probability of success is similar to equation 6.10 :

$$P_s = \prod_{i=1}^{n} \mathrm{erf}\left(\frac{\beta \times \nicefrac{1}{2}}{\sqrt{2}\sigma_{i|I}} \right) \tag{6.13}$$

while the probability of failure is given by considering every other solution's probability, i.e.

$$P_f = \sum_{z \in Z^n \setminus \{0\}} \prod_{i=1}^{n} \left[\frac{1}{2}\mathrm{erf}\left(\frac{c_i^T L^{-1} z + \beta \times \nicefrac{1}{2}}{\sqrt{2}\sigma_{i|I}} \right) - \frac{1}{2}\mathrm{erf}\left(\frac{c_i^T L^{-1} z - \beta \times \nicefrac{1}{2}}{\sqrt{2}\sigma_{i|I}} \right) \right] \tag{6.14}$$

and the probability of indecision is:

$$P_u = 1 - P_s - P_f \tag{6.15}$$

Summing 6.14 across all integers is not possible; instead a finite set of integers can be chosen, containing all non-negligible integers - such as the thousand ambiguity solutions

with the lowest squared norms.   Verhagen (2005b) uses search spaces varying from a few integers to a few tens of thousands of integers.

6.1.3.3   Integer bootstrapping success rates based on Ambiguity Dilution Of Precision (ADOP)

The Ambiguity Dilution Of Precision (ADOP) relies on the determinant of the variance-covariance matrix of ambiguities.   Unlike the integer bootstrapping lower bound, it does not vary depending on the order in which rounding is performed, or whether the ambiguities have been decorrelated before calculation. It can be used to calculate what the success rate would be if the ambiguities were fully decorrelated. The ADOP is defined as:

$$ADOP = \sqrt{|Q_{\hat{a}}|}^{\frac{1}{n}} \tag{6.16}$$

where n is the number of satellites, $Q_{\hat{a}}$ is the variance-covariance matrix of ambiguities, and $|Q_{\hat{a}}|$ is the determinant of $Q_{\hat{a}}$. Teunissen (2003) gives the following upper bound for the integer bootstrapping success rate:

$$P_s \approx \mathrm{erf}\left(\frac{1/2}{\sqrt{2}ADOP}\right)^n \tag{6.17}$$

This is a very similar equation to 6.10 on page 131; it is less conservative, but faster to calculate as it can be calculated before decorrelation is performed.

6.1.4   Failure rate estimates assuming coloured noise

Petovello et al. (2005) and O'Keefe et al. (2006) point out that when time-correlated noise is present, failure rate estimates which do not take this into account will be overly optimistic. It is proposed that in the conventional linearised model of equation 3.55 on page 61

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{a} + \boldsymbol{B}\boldsymbol{b} + \boldsymbol{e}$$

the noise parameter $\boldsymbol{e}$, should be assumed to be a first-order Gauss-Markov process, i.e.

$$\boldsymbol{e} = \boldsymbol{u}_k + \boldsymbol{n}_k \tag{6.18}$$

$$\boldsymbol{u}_{k+1} = \boldsymbol{S}_{k+1}\boldsymbol{u}_k + \boldsymbol{\varepsilon}_k \tag{6.19}$$

where $\boldsymbol{u}_k$ is the coloured noise, $\boldsymbol{n}_k$ and $\boldsymbol{\varepsilon}_k$ are vectors of white noise, and $\boldsymbol{S}_{k+1}$ is the measurement error transition matrix - a diagonal matrix with elements $\exp\left(-\beta\left(t_{k+1} - t_k\right)\right)$ where β^{-1} is the correlation time and $(t_{k+1} - t_k)$ is the time between observations.

The papers give results of processing using a Kalman filter where the inputs are differenced between epochs, i.e.

$$\boldsymbol{y}_k = \boldsymbol{A}_k\boldsymbol{a}_k + \boldsymbol{B}_k\boldsymbol{b}_k + \boldsymbol{u}_k + \boldsymbol{n}_k \tag{6.20}$$

$$\begin{aligned}
\boldsymbol{y}_k - \boldsymbol{S}_k\boldsymbol{y}_{k-1} &= \boldsymbol{A}_k\boldsymbol{a}_k + \boldsymbol{B}_k\boldsymbol{b}_k + \boldsymbol{u}_k + \boldsymbol{n}_k - \ldots \\
&\quad \boldsymbol{S}_k\left(\boldsymbol{A}_{k-1}\boldsymbol{a}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{b}_{k-1} + \boldsymbol{u}_{k-1} + \boldsymbol{n}_{k-1}\right)
\end{aligned} \tag{6.21}$$

$$\begin{aligned}
\boldsymbol{y}_k - \boldsymbol{S}_k\boldsymbol{y}_{k-1} &= \boldsymbol{A}_k\boldsymbol{a}_k + \boldsymbol{B}_k\boldsymbol{b}_k + \left(\boldsymbol{S}_k\boldsymbol{u}_{k-1} + \boldsymbol{\varepsilon}_{k-1}\right) + \boldsymbol{n}_k - \ldots \\
&\quad \boldsymbol{S}_k\left(\boldsymbol{A}_{k-1}\boldsymbol{a}_{k-1} + \boldsymbol{B}_{k-1}\boldsymbol{b}_{k-1}\right) - \boldsymbol{S}_k\boldsymbol{u}_{k-1} - \boldsymbol{S}_k\boldsymbol{n}_{k-1}
\end{aligned} \tag{6.22}$$

$$\boldsymbol{y}_k - \boldsymbol{S}_k\boldsymbol{y}_{k-1} = \boldsymbol{A}_k\boldsymbol{a}_k + \boldsymbol{B}_k\boldsymbol{b}_k + \boldsymbol{\varepsilon}_{k-1} + \boldsymbol{n}_k - \boldsymbol{S}_k\boldsymbol{A}_{k-1}\boldsymbol{a}_{k-1} - \boldsymbol{S}_k\boldsymbol{B}_{k-1}\boldsymbol{b}_{k-1} - \boldsymbol{S}_k\boldsymbol{n}_{k-1} \tag{6.23}$$

i.e. the correlated noise terms are removed, leaving only white noise. This can be simplified to a linear set of equations - but because this introduces correlation to the Kalman filter measurement and process noise vectors, which must be accounted for by modifying the conventional Kalman filter state and covariance update equations.

The eventual outcome of this is a Kalman filter that produces a substantially higher

standard deviation estimate than a conventional filter which assumes only white noise is present in the input measurements.

Petovello et al. (2009) presents test results for the algorithm, with both simulated and real data. Simulated noise included both white noise and Gauss-Markov noise, and results are presented for the simulated period showing that when the time-correlated noise is not accounted for the Kalman filter underestimates the results' standard deviations - while when the noise is accounted for in the filter, the standard deviation matches or exceeds the results' standard deviations - depending on whether the filter's time constant matches or exceeds that of the simulated noise.

Data was recorded for 6 hours on a building roof, using NovAtel OEM4 GPS receivers with NovAtel 700 antennas, 10m apart. The first hour of data was used to estimate Gauss-Markov parameters, and the subsequent 5 hours were processed once without accounting for time correlation between measurements, and once while accounting for it using the new algorithm. The resulting positions were compared to the true receiver position to give an RMS error, and the RMS error was compared to the estimated covariance; for a perfect covariance estimate, they would be equal. The results for the east axis show the standard filter results have a standard deviation of about half what it should be, while the proposed filter has a standard deviation of about five times what it should be, making it substantially more conservative.

Khanafseh et al. (2010) takes a similar approach, but instead of creating a differenced measurement, the measurement error transition matrix is added to the system state transition matrix. The paper performs 10,000 simulations to assess the impact when the correlation time assumed by the filter does not precisely match the correlation time of the noise. For example, when the time constant of the simulated noise is 100 seconds and the filter assumes a time constant of 20 seconds, after 60 seconds the filter underestimates the variance of the ambiguity estimate. With the same simulated noise but a filter time constant of 400 seconds, the filter underestimates the ambiguity variance for the first 170 seconds. Only when the filter correlation time matches the simulated data's correlation time are the estimates consistently accurate.
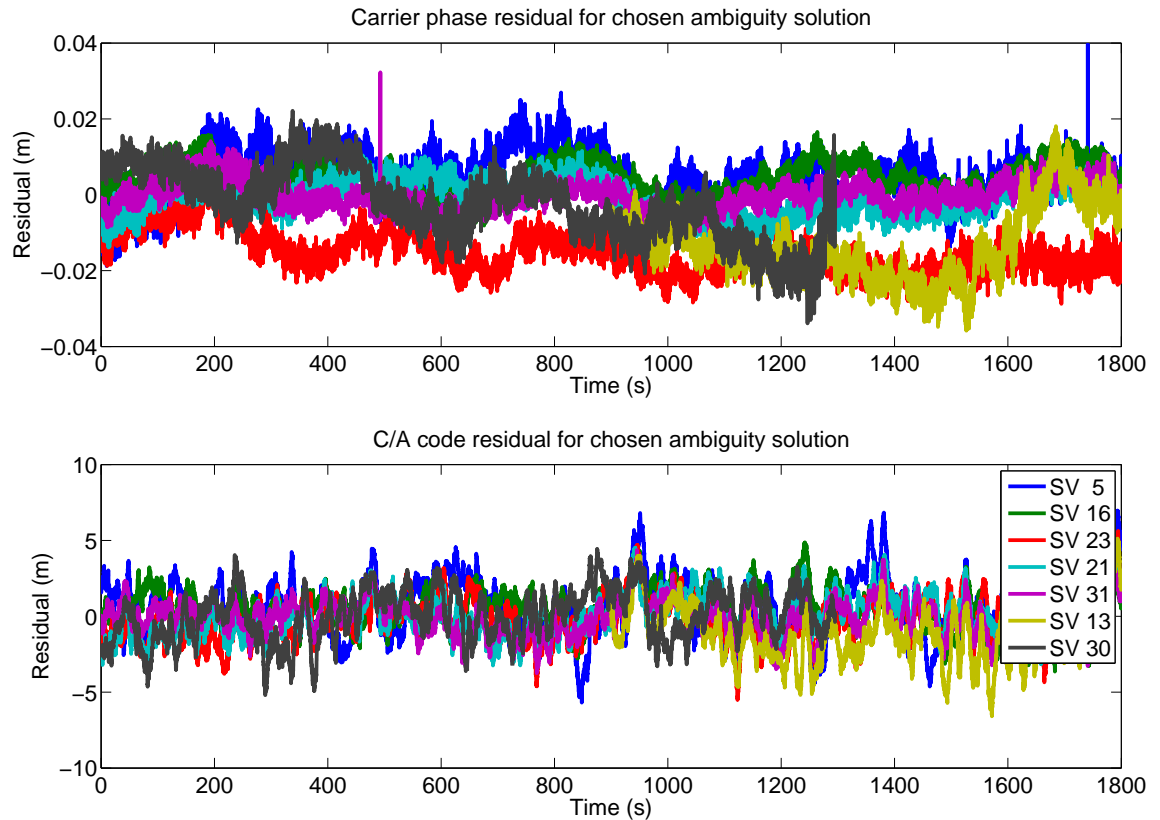
Figure 6.1: Carrier phase and C\A code residuals after successful ambiguity resolution (taking 972 seconds, to achieve an F-ratio of 3.5). Note the difference in vertical scales between top and bottom graphs.

The paper goes on to propose that, to avoid variances being underestimated, a "matrix of Kalman filters" covering a range of time constants could be used.

6.2 Adaption to the considered application

Teunissen and Verhagen (2004) and Teunissen (2005) propose algorithms that, given a variance-covariance matrix and a target failure rate, can calculate a ratio test threshold that will produce no more than the target rate of failures. Ambiguities are resolved anew each epoch, and in epochs when the ratio test is not passed the float solution is used instead.

On the other hand to reliably perform ambiguity resolution on single frequency data when multipath noise is present it is typically necessary to use several epochs of data. The number of epochs required depends on the number and position of satellites in
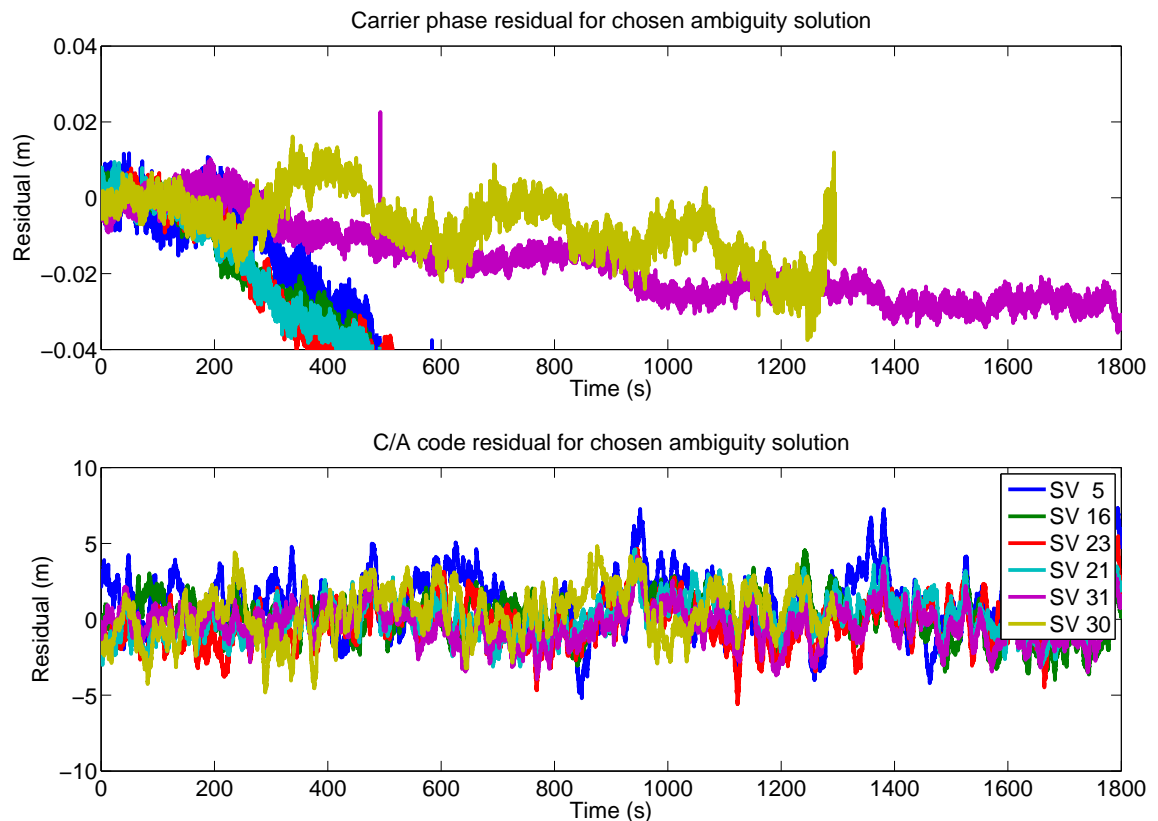
Figure 6.2: Carrier phase and C\A code residuals after unsuccessful ambiguity resolution (taking 148 seconds, to achieve an F-ratio of 2.5). Note the difference in vertical scales between top and bottom graphs.

view (represented by the variance-covariance matrix) and on the quality of the data (usually judged using a test such as the ratio test or F-ratio test). Ambiguities are considered constant between epochs, except where cycle slips are detected.

In addition preliminary experiments indicated a time-correlated element to multipath noise, which is not fully modelled by the assumption of zero mean Gaussian noise used by Teunissen and Verhagen. For any success rate model to be reliable, its modelling assumptions must be accurate; one common assumption is that the noise on observables is independent, zero mean, and normally distributed. Figure 6.1 and figure 6.2 show residuals after ambiguity resolution for an example set of data. When ambiguities are correctly resolved (e.g. figure 6.1) it can be seen that there is a time-correlated element to the noise, meaning statistical techniques expecting independent noise can overestimate success rates. For more details about the noise observed and zero baseline

test results, see appendix C on page 156.

Two new features are required to use a fixed failure rate approach under the test conditions used; use of variable resolution times instead of undecided results, and a noise model representative of typical multipath noise. Two means of achieving this were investigated.

6.2.1 Modified statistical tests

One means of accounting for time correlated noise is by scaling a priori standard deviations up to make success rate estimates more conservative, and combining this with other acceptance tests such as the F-ratio test. By requiring that multiple tests be passed for acceptance, including equation 6.10 on page 131, the failure rate will be no higher than it would be using equation 6.10 on page 131 alone.

This technique benefits from speed and simplicity, and the fact it is at least as statistically justified as using equation 6.10 on page 131 alone. However, using this technique is an admission that equation 6.10 on page 131 alone is insufficient to achieve a given success rate, and that faster performance may be available using more complicated models.

6.2.2 Monte Carlo simulation based tests

Another method is the use of Monte Carlo simulation to estimate success rates, using noise models which include time-correlated noise; the success rate estimate found by a Monte Carlo simulation with an accurate noise model will be more reliable than the success rate estimate found by statistical tests based on an inaccurate noise model. Monte Carlo models can not just be used to estimate the success rate in the absence of other acceptance tests; they can be used to determine success rates and mean ambiguity resolution times for different combinations of acceptance tests.

For example, if the acceptance criteria uses the format "F-ratio greater than a and time greater than b" a search of possible values of a and b (sometimes called the

configuration space) can be performed to find the value pair predicted to offer the lowest mean ambiguity resolution times with a failure rate at or below a target value. When the right data from the simulations is retained, multiple acceptance criteria can be evaluated without the need to repeat simulations for every acceptance criteria, allowing more efficient searches of the configuration space.

Teunissen and Verhagen (2009) demonstrated simulations can be conducted in advance and results stored in a lookup table. As the lookup table is calculated once and then stored, it is practical to conduct a more detailed search of the configuration space than would be possible in real time, or to use a more complicated configuration space.

6.2.2.1 Accurate noise modelling

For the results of the Monte Carlo simulation to be useful, an accurate noise model is required. A Gauss-Markov noise model was chosen, to represent time-correlated noise. To calibrate the noise model the real data was analysed to find the distribution of residuals' means, standard deviations and Gauss-Markov time constants (measured via autocorrelation). At the same time samples of satellite geometry were taken from the real data; samples of poorer satellite geometry were generated by discarding visible satellites.

An alternative approach which would guarantee accurate noise modelling would be to gather a very large sample of real-world data and use it directly, or to find the residuals on it and add the real-world residuals to simulated data. However, this would require a large sample of real world data, and that data would have to be gathered carefully to avoid over-fitting to the noise characteristics of one particular test location or base station.

6.2.2.2 Generation of lookup table

Data was simulated in 30 minute blocks with a 10Hz data rate. 100,000 simulations were performed, with ambiguity resolution performed every two seconds. A record was
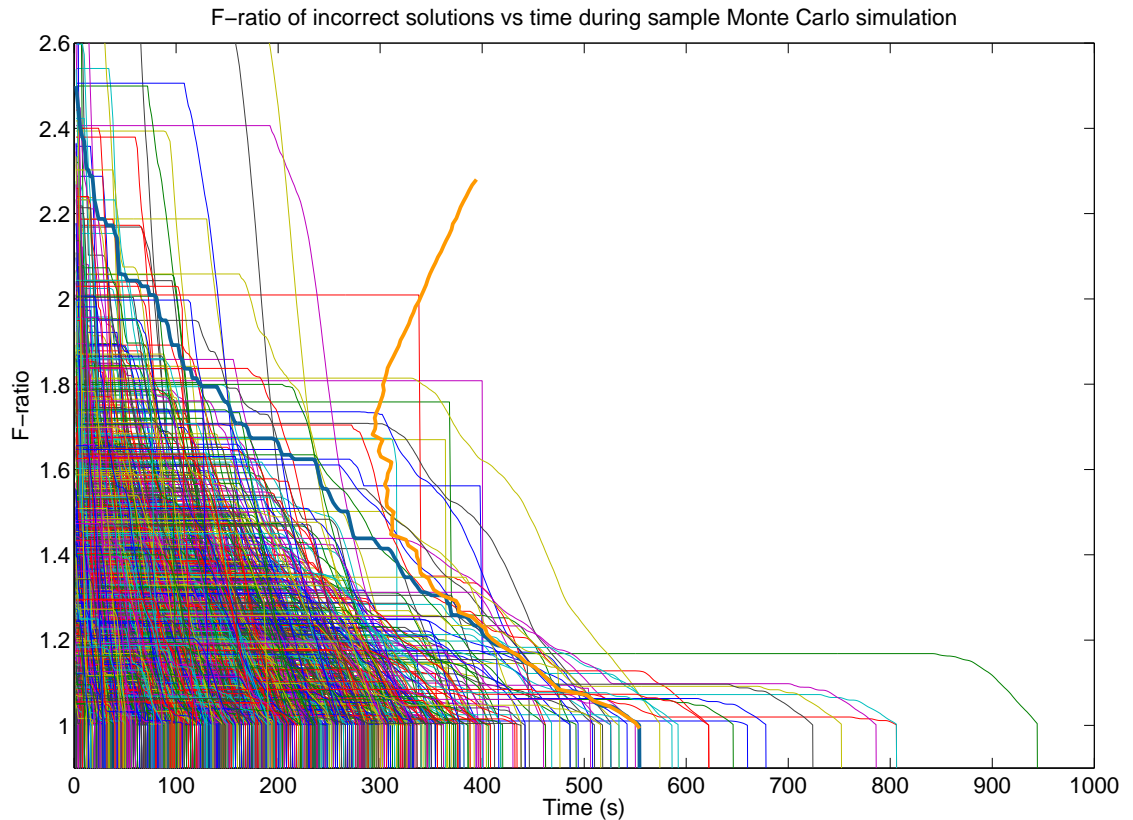
Figure 6.3: F-ratio thresholds that could result in acceptance of an incorrect solution at different times (thin lines) along with minimum acceptance time for a 99.5% success rate different F-ratio values (thick blue line) and mean ambiguity resolution time for these F-ratio values (thick yellow line).

kept of the resulting F-ratio values and whether the top solution was the chosen one, alongside the number of satellites in view and the success probability after ten seconds, as estimated by equation 6.17 on page 133.

Records are then divided up by number of satellites in view and success probability estimate, into bins corresponding to lookup table cells. The records in each bin are then used to choose the acceptance test thresholds for the corresponding lookup table cell.

With knowledge of the maximum F-ratios achieved by incorrect solutions at different times it is possible to identify the time after which an F-ratio will produce a success rate above a given threshold. This is illustrated in figure 6.3 by the thick blue line; for example an F-ratio threshold of 2.5 can be applied almost immediately and a 99.5%

success rate will be obtained - whereas if a 1.0 ratio threshold is used results should not be accepted in the first 554 seconds if that success rate is to be achieved.

With knowledge of the F-ratios achieved by correct solutions and the times they are achieved, the mean ambiguity resolution time can simply be calculated for a given F-ratio and minimum time. This is shown by the thick yellow line in figure 6.3 ; the lowest mean ambiguity resolution time is 292.5 seconds, for an F-ratio of 1.68 and a minimum time of 172 seconds.

For comparison a graph showing the impact of F-ratio on median and 95% ambiguity resolution times can be found in appendix B on page 154.

Lookup table generation in summary:

- Gather experimental data to generate a representative pool of noise model parameters.

- For each cell in the lookup table, perform 100,000 simulations with noise samples drawn from the noise model pool, covering 1000 seconds and attempting ambiguity resolution every second. For each attempt record the F-ratio and whether the solution is correct or not.

- From the simulation results, determine for each F-ratio value the time threshold required for a failure rate below 0.5% and the mean ambiguity resolution time for these parameters.

- Select the F-ratio/time threshold pair with the lowest mean ambiguity resolution time.

6.2.2.3 Structure of lookup table

Teunissen and Verhagen (2009) structure their lookup table based on the number of satellites in view and the estimated success rate, with that success rate estimate based on a lower bound approximation such as equation 6.10 on page 131. Given our findings that even lower bound approximations may overestimate the success rate, we could use

Number of satellites	$0.0 < P_{s10} \leq 0.1$	$0.1 < P_{s10} \leq 0.2$	$0.2 < P_{s10} \leq 0.3$	$0.3 < P_{s10} \leq 0.4$	$0.4 < P_{s10} \leq 0.5$	
⋮	⋮	⋮	⋮	⋮	⋮	
6	1.71 @ 56.0s	1.74 @ 22.0s	1.14 @ 147.0s	1.12 @ 33.0s	1.10 @ 40.0s	⋯
7	1.27 @ 141.0s	1.69 @ 26.0s	1.39 @ 131.0s	1.36 @ 109.0s	1.13 @ 92.0s	⋯
8	1.30 @ 146.0s	1.36 @ 55.0s	1.64 @ 50.0s	1.25 @ 95.0s	1.48 @ 106.0s	⋯
9	1.30 @ 45.0s	1.26 @ 36.0s	1.23 @ 48.0s	1.26 @ 60.0s	1.26 @ 29.0s	⋯
⋮	⋮	⋮	⋮	⋮	⋮	

Table 6.1: Example lookup table giving F-ratio threshold and acceptance start time. Satellite counts make up the vertical axis and statistical success rate estimates after ten seconds make up the horizontal axis.

faster but less conservative success rate estimates, such as equation 6.17 on page 133, but the time benefit of doing so is marginal.

As the success probability increases with the amount of data gathered, the results of a simulation would move between table cells as it progressed were the success probability recalculated every epoch. Instead, the lookup table value is selected based on the success probability calculated after 10 seconds of data. This is denoted P_{s10} in table 6.1.

6.2.2.4   Alternative lookup table designs

As the use of a precalculated lookup table makes it possible to search larger configuration spaces in advance, it is possible to set more complicated limits than an F-ratio threshold and a minimum time.

For example, just as equation 6.12 on page 132 assesses a solution based on the distance in cycles between fixed and float solutions, we can also measure the euclidean distance between fixed and float solution positions - the *distance from float*. This can be used for an acceptance test by applying a threshold to it.

A lookup table was generated looking not just at F-ratio threshold values but also at F-ratio threshold/distance from float threshold combinations. The table was genera-

ted using a local search of F-ratio thresholds from 1.0 to 2.0 and distance from float thresholds from one third of a metre to ten metres.

A third design of lookup table was also generated, with two F-ratio thresholds and start times. This was generated by testing different F-ratio pairs, determining the start times required to achieve the target failure rate as one would with a single threshold, then selecting the pair that resulted in the lowest mean ambiguity resolution times. Depending on the application, it would also be possible to set the two thresholds to optimise different things, such as an early threshold minimising mean ambiguity resolution time while a later, lower threshold minimises 95% ambiguity resolution time.

6.3 Experimental Configuration

Tests were then carried out using real data; section 4.3 on page 99 describes the collection of 24 hours of 10Hz GPS data from a stationary vehicle in a moderate multipath environment. A single-frequency GPS receiver was used (U-Blox AEK-4T with ANN-MS patch antenna), with an identical receiver acting as a fixed base station located 280 metres away. The collection of 90 minutes of kinematic data under similar conditions is also described.

6.3.1 Modified statistical acceptance tests

The stationary recorded data was processed by performing ambiguity resolution starting every 30 seconds throughout the data, taking as long as needed to pass the acceptance test, and recording the ambiguity resolution times and whether resolution was successful.

Ambiguity resolution was performed with the acceptance test described in section 6.2.1 on page 138, requiring that the success rate estimated by equation 6.10 on page 131 is greater than 99.5%. This was repeated with scale factors on the a priori standard deviation of $1\times$, $50\times$, $100\times$, $200\times$ and $300\times$; tests were also performed with and without an F-ratio test.

From the recorded ambiguity resolution times and successfulness, mean, median and 95% ambiguity resolution times can be calculated, as can the overall success rate.

6.3.2 Monte Carlo simulation based tests

Monte Carlo simulations were performed to generate the three lookup tables described in section 6.2.2 on page 138.

The real stationary test data was then processed, starting ambiguity resolution every 30 seconds throughout the data and using acceptance criteria from the lookup table cell corresponding to the test situation. The ambiguity resolution time and successfullness were recorded and the mean, median and 95% ambiguity resolution times calculated. The process was then performed every 5 minutes throughout the 90 minutes of kinematic data, with the same statistics calculated.

6.4 Results

6.4.1 Modified statistical acceptance tests

Results for tests with real data are presented in table 6.2 . As can be seen, when applying the constraint that the integer bootstrapping success rate (determined with equation 6.10 on page 131) should exceed 99.5%, but applying no other constraints, the success rate seen is below the 99.5% expected, at 24%. To confirm this stark difference was not due to a programming error, the test was repeated using simulated data with zero-mean white noise matching the a priori standard deviations; when noise matching the noise model was used the success exceeded 99.5%. This shows that, in this test circumstance, equation 6.10 on page 131 over-estimates the success rate. Scaling up the a priori standard deviations $\times 300$ allows a success rate above 99.5% with no F-ratio test - but the resulting mean ambiguity resolution time is high compared with some other solutions with comparable success rates. For example, standard deviations scaled $\times 300$ and an F-ratio threshold of 1.0 has a mean time of 548.5 seconds and a success

Algorithm used	Ratio threshold	Solved right (%)	Time to ambiguity resolution		
			Mean (s)	Median (s)	95% (s)
Single receiver	*1.0*	*17.38%*	*2.0*	*2.0*	*2.0*
No success statistic used	*2.0*	*97.52%*	*379.1*	*320.0*	*968.0*
	2.5	*99.47%*	*524.4*	*460.0*	*1364.0*
	3.0	99.89%	679.6	604.0	1800.0
Single receiver	*1.0*	*23.69%*	*15.8*	8.0	86.0
equation 6.10 on page 131	*2.0*	*98.19%*	*379.4*	*320.0*	*1800.0*
$\geq$99.5%	2.5	99.65%	495.0	430.0	1800.0
weights as measured	3.0	99.96%	679.7	604.0	1800.0
Single receiver	*1.0*	*85.07%*	*210.3*	*194.0*	*384.0*
equation 6.10 on page 131	2.0	100.00%	367.5	320.0	806.0
$\geq$99.5%	2.5	100.00%	487.6	424.0	1120.0
weights $\times$50	3.0	100.00%	627.3	556.0	1678.0
Single receiver	*1.0*	*94.43%*	*302.5*	*284.0*	*512.0*
equation 6.10 on page 131	2.0	100.00%	405.6	362.0	808.0
$\geq$99.5%	2.5	100.00%	511.7	452.0	1124.0
weights $\times$100	3.0	100.00%	642.1	562.0	1678.0
Single receiver	*1.0*	*99.40%*	*435.0*	*408.0*	*900.0*
equation 6.10 on page 131	2.0	100.00%	490.8	446.0	1800.0
$\geq$99.5%	2.5	100.00%	572.6	510.0	1800.0
weights $\times$200	3.0	100.00%	680.9	610.0	1800.0
Single receiver	1.0	99.93%	548.5	510.0	1800.0
equation 6.10 on page 131	2.0	100.00%	579.8	534.0	1800.0
$\geq$99.5%	2.5	100.00%	643.5	574.0	1800.0
weights $\times$300	3.0	100.00%	730.3	648.0	1800.0

Table 6.2: Results using real data with various acceptance tests and F-ratio thresholds. Results with a success rate below 99.50% in italics.

rate of 99.93% - whereas standard deviations scaled $\times$50 with an F-ratio threshold of 2.0 offers 100.00% success and a mean time of just 367.5 seconds.

Results with a variety of F-ratio thresholds are shown in figure 6.4 ; for a 99.5% success rate using F-ratio alone, the mean ambiguity resolution time is 530 seconds. With the addition of equation 6.10 on page 131 this drops to 476 seconds with accurate a priori standard deviations, 280 seconds with them scaled by 50$\times$, and 320 seconds with a scale factor of 100$\times$. With a 200$\times$ scale factor 99.5% success is achieved in 435 seconds, and with 300$\times$in 550 seconds.

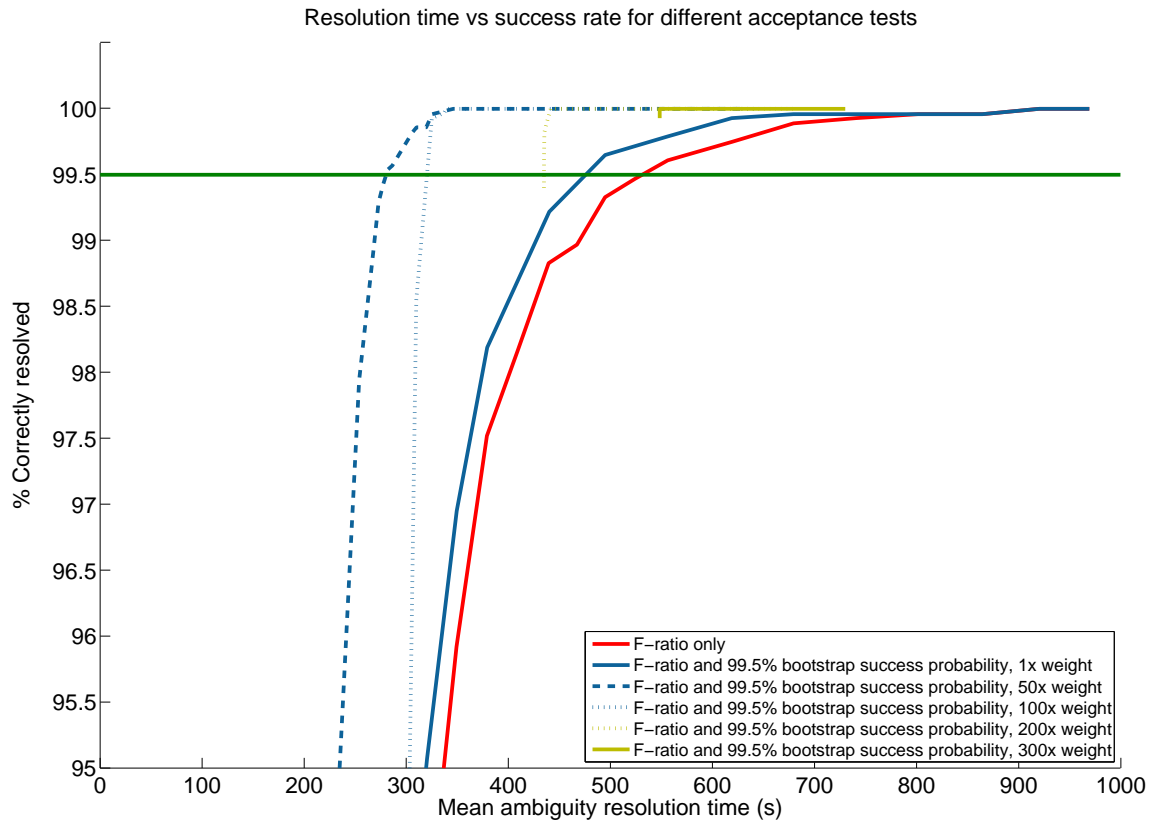The best configuration represents a reduction in mean ambiguity resolution times of 47%.

Figure 6.4: Resolution time vs. success rate for various F-ratio thresholds, for several acceptance tests.

6.4.2 Monte Carlo simulation based tests

Results for the stationary test with a Monte Carlo lookup table are shown in table 6.3 . Compared to using a fixed F-ratio, a lookup table for F-ratio threshold and start time reduces mean ambiguity resolution time by 52%. A lookup table for F-ratio threshold, distance from float threshold and start time reduces mean ambiguity resolution time by 55%; and a lookup table with two F-ratio thresholds with different start times reduces ambiguity resolution time by 58%.

Results for the kinematic test using the Monte Carlo lookup table are shown in table 6.4 . Mean ambiguity resolution times drop by 33% and 39% using the single F-ratio and dual F-ratio tables respectively. Given the extra complexity of the dual F-ratio table technique, the improvement over the single F-ratio table is limited.

Algorithm used	Solved right (%)	Time to ambiguity resolution		
		Mean (s)	Median (s)	95% (s)
Fixed F-ratio of 2.5 only	*99.47%*	*524.4*	*460.0*	*1364.0*
Lookup table for F-ratio threshold and start time.	99.79%	248.8	222.0	842.0
Lookup table for F-ratio threshold, distance from float threshold, and start time.	100.0%	233.2	217.0	858.0
Lookup table for two F-ratio thresholds and start times.	100.0%	219.8	203.0	684.0

Table 6.3: Results using stationary real data with acceptance criteria from lookup table generated using Monte Carlo simulation.

Algorithm used	Solved right	Time to ambiguity resolution	
		Mean (s)	Median (s)
Fixed F-ratio of 2.5 only	16/18	619	480
Lookup table for F-ratio threshold and start time.	17/18	512	328
Lookup table for two F-ratio thresholds and start times.	17/18	380	286

Table 6.4: Results using kinematic real data with acceptance criteria from lookup table generated using Monte Carlo simulation.

When an ambiguity resolution attempt begins near the end of the data set, if the end of the data is reached before a correct ambiguity is accepted, we consider this a failure. As the kinematic test was shorter than the stationary test, the final minutes of the test have a proportionally greater impact; this explains the increase in failure rates compared to the stationary tests; in every test the final attempt failed. A longer duration kinematic test would be ideal, in order to confirm the overall failure rate is as expected.

6.5 Conclusions

In this chapter a number of modifications to conventional ambiguity solution acceptance tests have been discussed. Acceptance tests from the literature have been shown to not always perform well when using single-frequency data from consumer grade receivers with patch antennas; unmodelled time-correlated noise has been implicated as the cause of this. Two means were proposed by which this problem can be overcome.

The first new proposal was to scale up the a priori standard deviations assigned to measurements until success rate estimates from conventional algorithms match empirical success rates. With a scale factor of $\times 300$ the empirical success rate exceeded the estimated success rate, although this did not help ambiguity resolution times. It was also observed that combining a scale factor of $\times 50$ with an appropriately chosen F-ratio could reduce mean ambiguity resolution times by 47%, from 530 seconds to 280 seconds.

The second new proposal was to use Monte Carlo simulation with a realistic noise model to generate a lookup table of acceptance thresholds to produce a specified failure rate. Previous works similar to this have not been designed to account for time-correlated noise or variable-length multi-epoch ambiguity resolution. As the lookup table is precalculated it proved possible to evaluate a wider range of possible configurations than would be practical in real time, such as applying different F-ratio thresholds at different times. The best performing configuration offered a mean ambiguity resolution time of 219.8 seconds, a 58% improvement compared to a single fixed F-ratio. In another test with 90 minutes of kinematic data, the improvements were 33% and 39% respectively.

Conclusions

In this contribution several methods to reduce mean GNSS ambiguity resolution times have been proposed, and the results of empirical evaluations using 24 hours of single frequency data have been presented.

For GPS applications involving vehicles operating in rehearsed areas, a method was demonstrated using a topographic map in the ambiguity resolution process. Unlike to other algorithms from the literature, the method uses maximum and minimum height constraints and produces a maximum likelihood estimate. Tests demonstrated a map accurate to ±5cm reduced mean ambiguity resolution times by 47 to 56%, while a map accurate to ±2.5cm reduced times by 61 to 67%. As well as being computationally efficient, compared to other methods from the literature the minimum-and-maximum constraint deals with prior height uncertainty effectively and in a statistically justifiable way, making it particularly well suited to the type of errors introduced by topographic maps.

For GPS applications involving multiple receivers on one vehicle (in a *GNSS compass* arrangement) two algorithms from the literature were examined, offering mean ambiguity resolution time reductions of 11% to 19%. It was proposed that multipath interference could be detected by comparing results at the two receivers, and that this multipath detection would allow faster resolution when multipath was low. Used alongside the algorithms from the literature allowed mean ambiguity resolution time reductions of 30% with both algorithms, an 11-percentage-point improvement on the better algorithm.

Finally, tests showed problems due to time-correlated noise were noted, and it was demonstrated this causes some conventional statistical techniques to over-estimate success rates. An algorithm was presented using Monte Carlo simulation with realistic noise modeling to compensate for the difference between statistical and empirical

results. This provided a justification for the configuration of ambiguity acceptance test thresholds. It was also demonstrated that pre-calculated lookup tables allow complicated acceptance criteria to be evaluated, offering improvements to mean ambiguity resolution times. Compared to using a fixed F-ratio, a lookup table for F-ratio threshold and start time reduces mean ambiguity resolution time by 52%. The best performing acceptance criteria tested was a lookup table with two F-ratio thresholds with different start times; this reduces ambiguity resolution time by 58%.

The methods and results represent an improvement in single-frequency ambiguity resolution performance, making the use of such systems more attractive and improving the feasibility of using high precision GNSS in applications such as mobile robot control and agricultural automation.

Compared to the requirements identified in section 1.1 on page 1, the single frequency GPS system met the accuracy requirements, although availability reduced when satellite visibility is poor. The cost and obstacle detection requirements were also met by the combined system described in chapter 2 on page 6.

Future work

For the work on GPS which comprises the majority of this thesis, the three sets of results presented in this work are for distinct techniques; no results have yet been produced using multiple techniques in combination.

The data used in this work was from low cost single frequency GPS receivers with patch antennas, which were stationary in an environment with some multipath interference. Further work could cover a wider range of sample data, covering different receivers and antennas, dual frequency data, kinematic tests, longer baselines, higher multipath levels such as urban canyons, other GNSS systems such as Galileo. As well as demonstrating the system performance in different situations, the tests could be chosen to simplify comparison with other results from the literature as the present single-frequency work is difficult to compare with (for example) dual-frequency work from the literature.

The noise shown in figure 6.1 on page 136 is attributed to multipath on the basis

of elimination of the other causes found in common models in the literature, but no detailed antenna analysis has been performed. A more in-depth analysis may be able to verify this and devise mitigation techniques.

Just as benefits were found from using multiple roving receivers, benefits may be found to using multiple base stations; this work addressed multiple roving receivers first, making the assumption the good visibility afforded to the base station antenna would make noise small compared to the noise at the roving receivers; a test with multiple base stations could verify this assumption. Users of high quality receivers often make use of commercial network RTK; our experiments have not looked at this, but further experiments could do.

Work on the autonomous mower system more broadly could improve on operation in situations of poor satellite visibility, such as by using sensor fusion with odometry and inertial sensors. The obstacle response system could also be more sophisticated, such as driving around stationary obstacles.

Appendix A

Additional pseudo-observation results

Results of the test using pseudo-observations in the presence of a priori height biases are shown in tables A.1 on the following page and A.2 on the next page. Median and 95% ambiguity resolution times are reported, to supplement the means reported in table 4.3 on page 105. In every case, the ambiguity resolution success rate was at least 99.5%.

Note that the median and 95% times of 1800 seconds may be due to the 1800 second acceptance criteria described in section 4.3.1 on page 100. Were this criteria not used the reported values would likely be higher.

These median and 95% performance figures are in line with the mean figures reported in table 4.3 on page 105.

A priori height bias	Assumed height accuracy						
	No height	Minimum & maximum		Pseudo-observation standard deviation			
		±2.5cm	±5.0cm	2.5cm	5cm	10cm	20cm
Accurate	919.0	264.0	330.0	292.0	327.0	425.0	603.5
2.5cm offset	919.0	322.0	344.0	590.0	535.5	553.0	645.0
5.0cm offset	919.0	852.0	414.0	1800.0	1141.0	827.5	1060.0

Table A.1: Median ambiguity resolution times (in seconds) for several algorithms, when there is a bias on a priori height information. This shows how different algorithms handle topographical map errors.

A priori height bias	Assumed height accuracy						
	No height	Minimum & maximum		Pseudo-observation standard deviation			
		±2.5cm	±5.0cm	2.5cm	5cm	10cm	20cm
Accurate	1800.0	1130.0	1140.0	1800.0	1800.0	1800.0	1800.0
2.5cm offset	1800.0	1180.0	1260.0	1800.0	1800.0	1800.0	1800.0
5.0cm offset	1800.0	1800.0	1422.0	1800.0	1800.0	1800.0	1800.0

Table A.2: 95% ambiguity resolution times (in seconds) for several algorithms, when there is a bias on a priori height information. This shows how different algorithms handle topographical map errors.

Appendix B
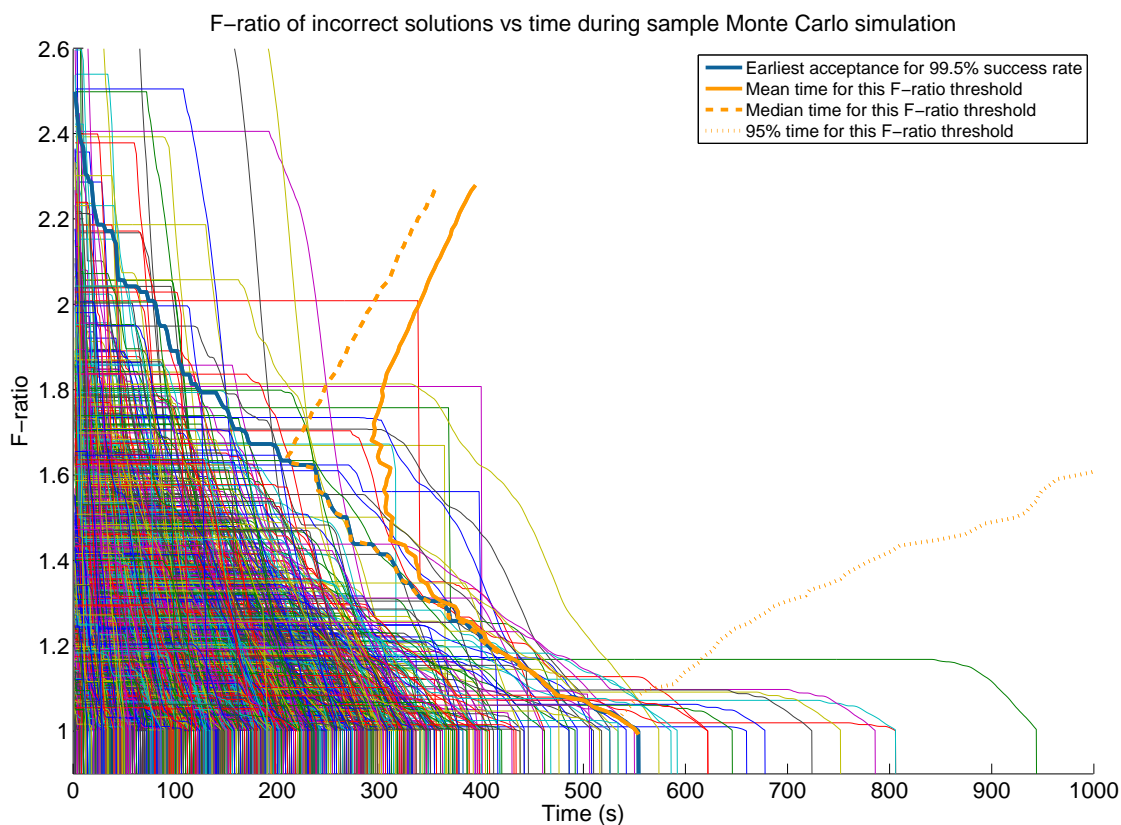
Monte Carlo for median and 95% times

Figure B.1: F-ratio thresholds that could result in acceptance of an incorrect solution at different times (thin lines) along with minimum acceptance time for a 99.5% success rate different F-ratio values (thick blue line) and mean, median and 95% ambiguity resolution time for these F-ratio values (thick yellow solid, dashed and dotted lines respectively).

Figure B.1 shows F-ratio test thresholds and minimum acceptance times that would result in a 99.5% success rate, along with the mean, median and 95% ambiguity resolution times for those thresholds. In this example, an F-ratio of 1.63 minimises the median time, 1.68 the mean time, and 1.03 the 95% time.

In this contribution settings are selected to minimise the mean ambiguity resolution

time, as it is a performance measure often provided in the literature, simplifying comparisons between techniques. In an application where minimising median or 95% time were deemed more important, it could be achieved by recalculating the lookup table.

Appendix C
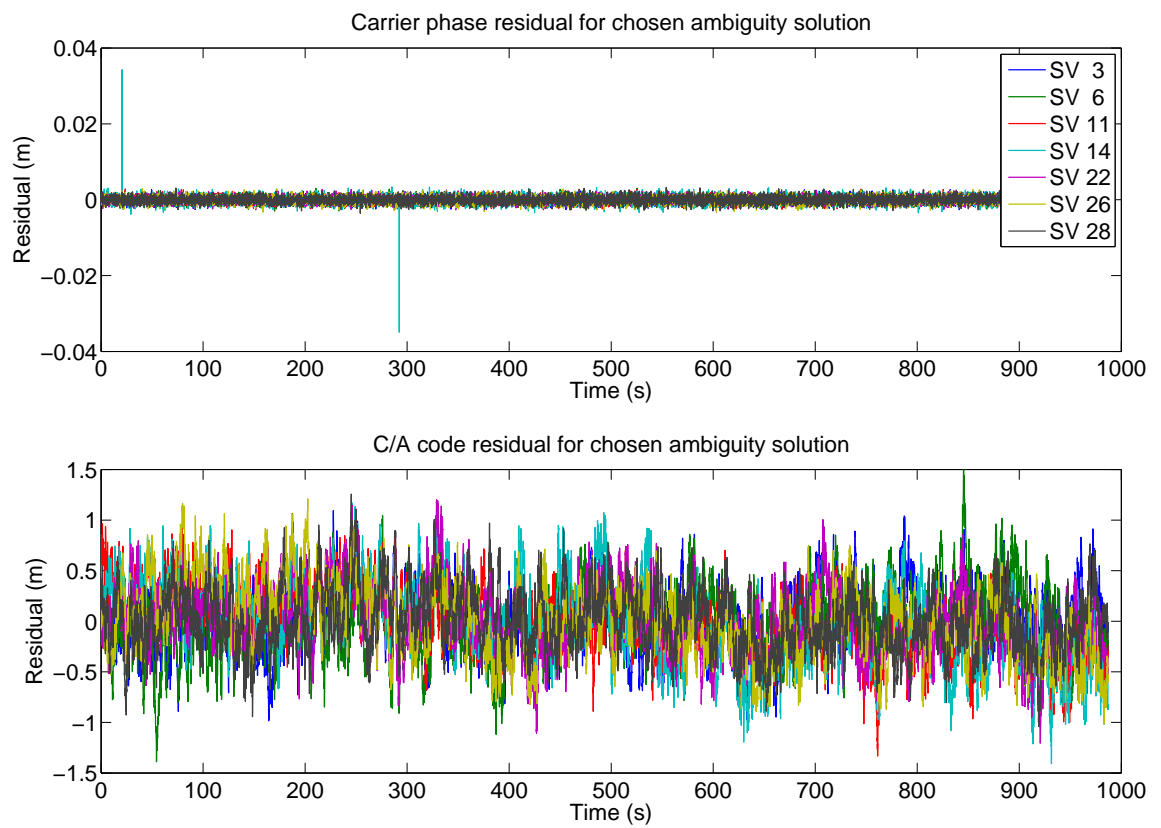
Measurement noise and zero baseline test results
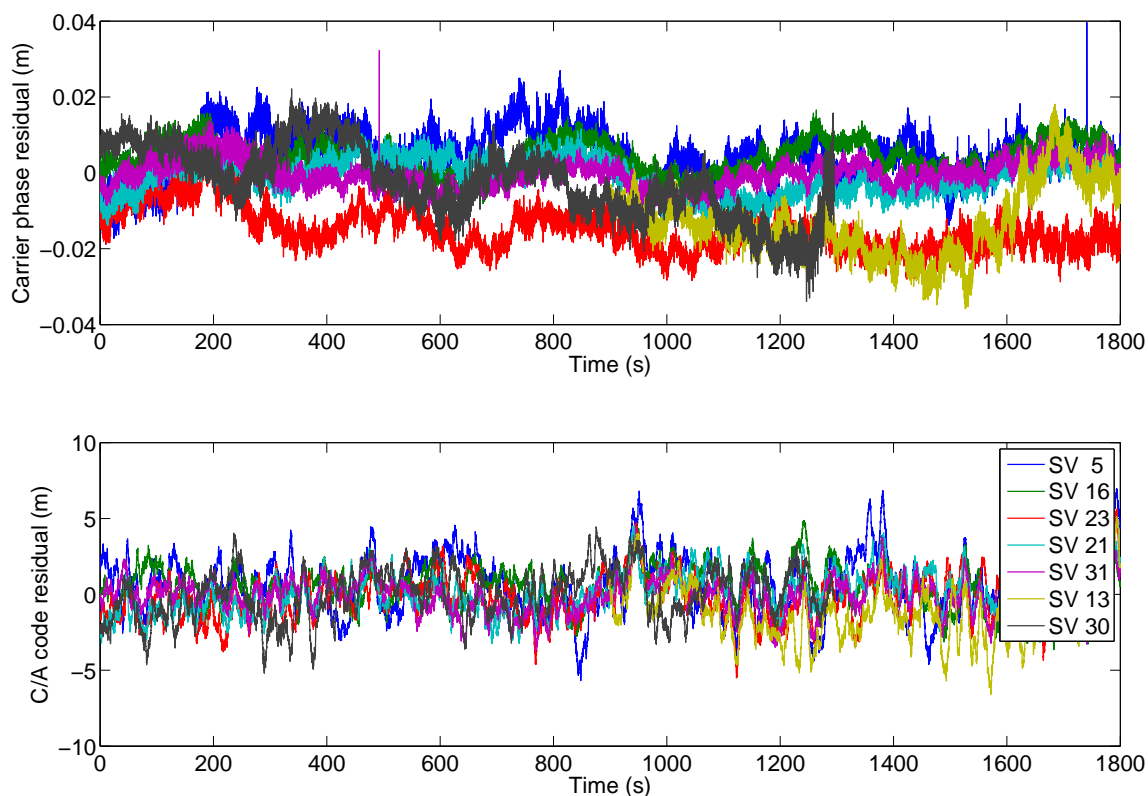
Figure C.1: Carrier phase and C\A code residuals for a stationary antenna, 280m baseline. Note the difference in vertical scales between top and bottom graphs.

Having noticed the impact of noise in some experiments, zero baseline tests were performed to investigate the source of this noise. In a zero baseline test, a signal splitter is used to feed one antenna's signal into two receivers. Noise which is common to both receivers must have happened at or before the antenna (for example, multipath noise), while noise which differs between receivers must have happened in the receivers

Figure C.2: Carrier phase and C\A code residuals for a stationary antenna, zero baseline. Note the difference in vertical scales between top and bottom graphs.
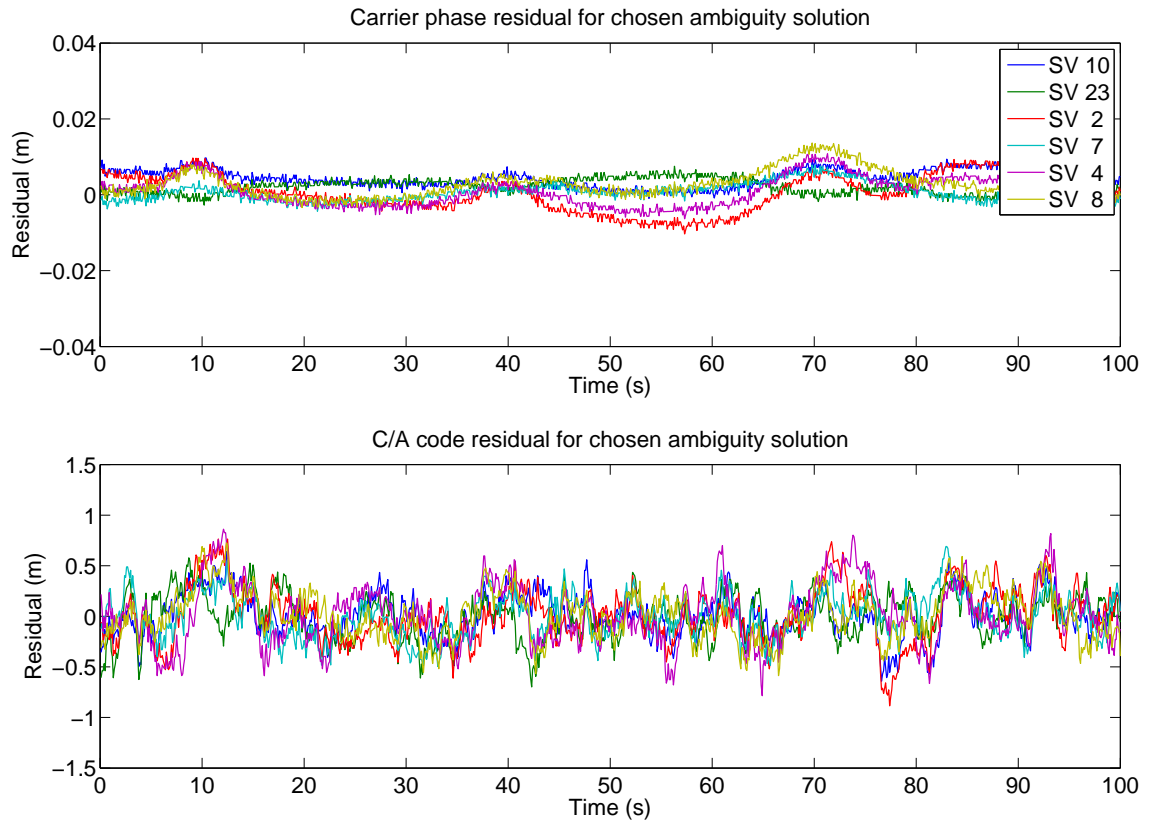
Figure C.3: Carrier phase and C\A code residuals for a moving antenna, zero baseline. Note the difference in vertical scales between top and bottom graphs.

themselves (for example, measurement noise).

The first test was a stationary one. A Mini-Circuits ZAPDQ-2-S power splitter was used, along with and the receivers and antenna used in section 4.3 on page 99. Figure C.2 on the previous page shows the difference in measurements between the two receivers; with the exception of two single-epoch excursions, the measurements are within 3mm of identical, with a standard deviation of 0.8mm. The noise level for a 280m baseline, shown in figure C.1 on page 156, is substantially higher; the noise, then, mostly occurs at or before the antenna. Multipath noise is a possible cause. It is reported by Blake et al. (2008) that the measurement noise levels of the U-Blox Antaris AEK-4T GPS receiver increase with increased motion. To assess the impact of this a zero-baseline test was performed with the antenna on the roof of a car travelling from 51.7624°N, 0.2433°W and 51.7701°N, 0.2412°W at normal road speeds. As shown in

Figure C.3 on the previous page, the noise level increases to within 10mm of correct, and the standard deviation increases to 3.2mm. Errors peak at 10, 40 and 70 seconds corresponding to three roundabouts crossed in the journey. The measurement noise level is higher than the stationary zero baseline test, corroborating Blake et al. (2008), but is still low compared to the stationary 280m baseline test indicating that multipath is remains the largest noise source.

Appendix D

Other Results

D.1   Kinematic relative accuracy experiment

To determine whether the system met the accuracy requirements, it was programmed to follow a path marked out with rope.

First, we marked out a line from north to south and programmed the mower with a path that made four passes; southbound with the rope on the left, northbound with the rope on the left, southbound with the rope on the right, and northbound with the rope on the right. The difference in error between the north and south passes let us calibrate for the offset of the antennas relative to the centre of the vehicle, and the difference in error between the right and left passes let us calibrate for the width of the vehicle and pass-to-pass overlap, while rope placement errors had predictable symptoms.

After calibrating the system, the same tests were performed with video taken showing the deviation from the rope. The video was manually evaluated frame-by-frame, samples are shown in figure D.1 on the following page. The rope is always between the fourth and sixth ribs on the cutter's front roller, indicating ± 2cm peak cross-track error.

It should be noted that this is not an ideal test; the dynamics of the mower naturally dampen the effects of GPS noise. Indeed, once the mower is going in a straight line the steering control system could completely stop and the mower would keep on going straight just because of its mechanical configuration. There was also some movement of the rope caused by the mower's tyres going over it. Nonetheless, this test indicates the system offers the required level of accuracy.
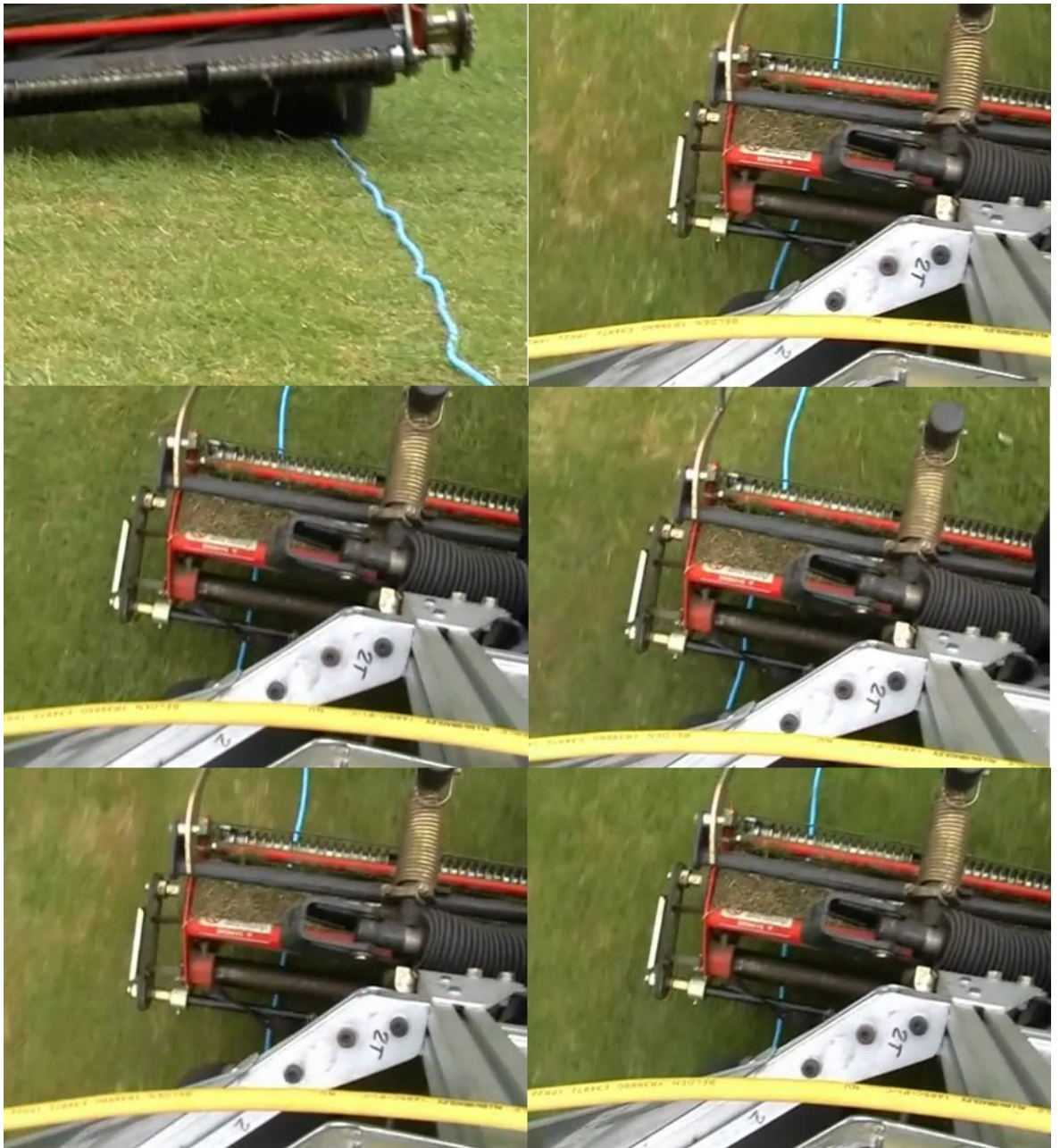
Figure D.1: Line-following precision test results. Note that the rope is consistently between the fourth and sixth ribs on the front roller of the cutter.
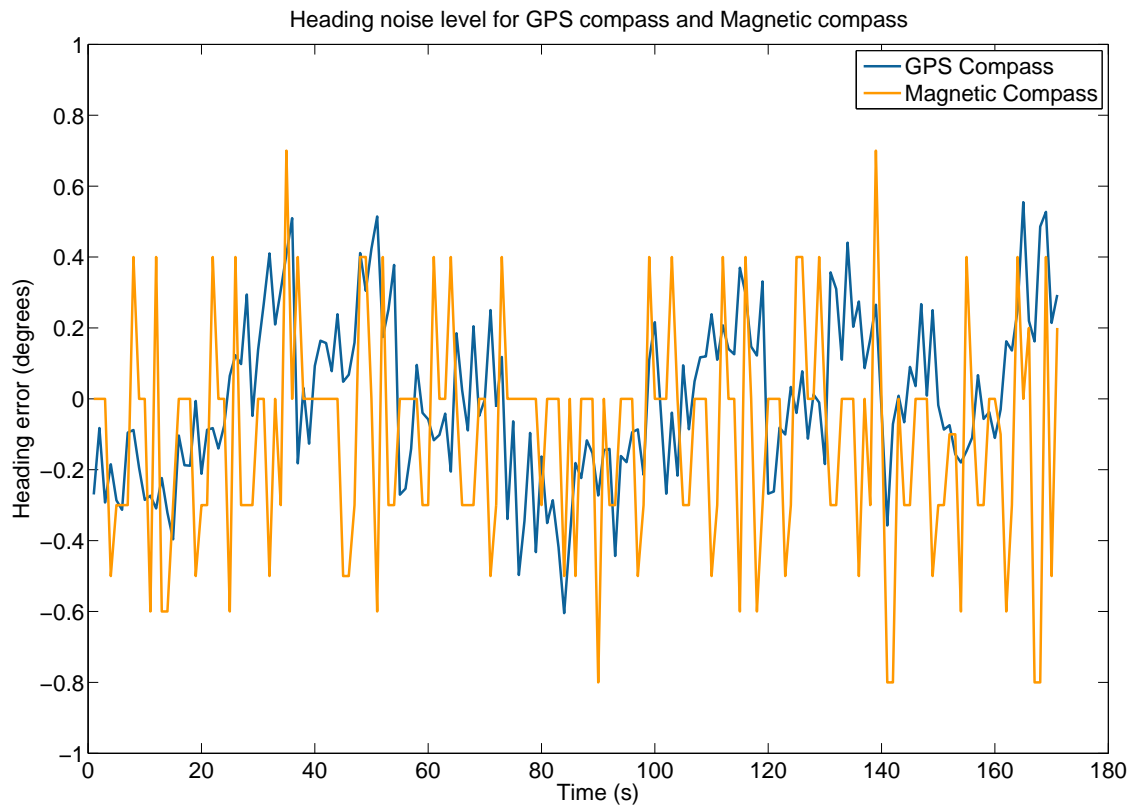
Figure D.2: Heading error for GPS and magnetic compass.

D.2 Heading measurement precision experiment

To determine how accurate heading measurement was, measurements were taken with the vehicle stationary.

In this project we did not directly use a magnetic compass, preferring to use two GPS receivers, but we did fit a magnetic compass to the vehicle to gather data for other researchers. We found the magnetic compass quite hard to calibrate due to the metal structures and magnetic fields of our test vehicle and engineering workshop. For the purposes of this test we parked the vehicle stationary facing north and ran the compass manufacturer's calibration procedure. We then gathered data from both the dual GPS system and the magnetic compass.

As shown in figure D.2, the GPS system less than ±0.6°with a standard deviation of 0.2°. This test shows the magnetic compass in a good light - its performance is

comparable to the GPS compass system - but once the vehicle was set in motion, the magnetic compass error levels were higher - sometimes above ten degrees, while the GPS system error rose only slightly. A different magnetic compass or a more complicated calibration procedure might have improved performance, but this was beyond the scope of this project.

Appendix E

Control board schematics

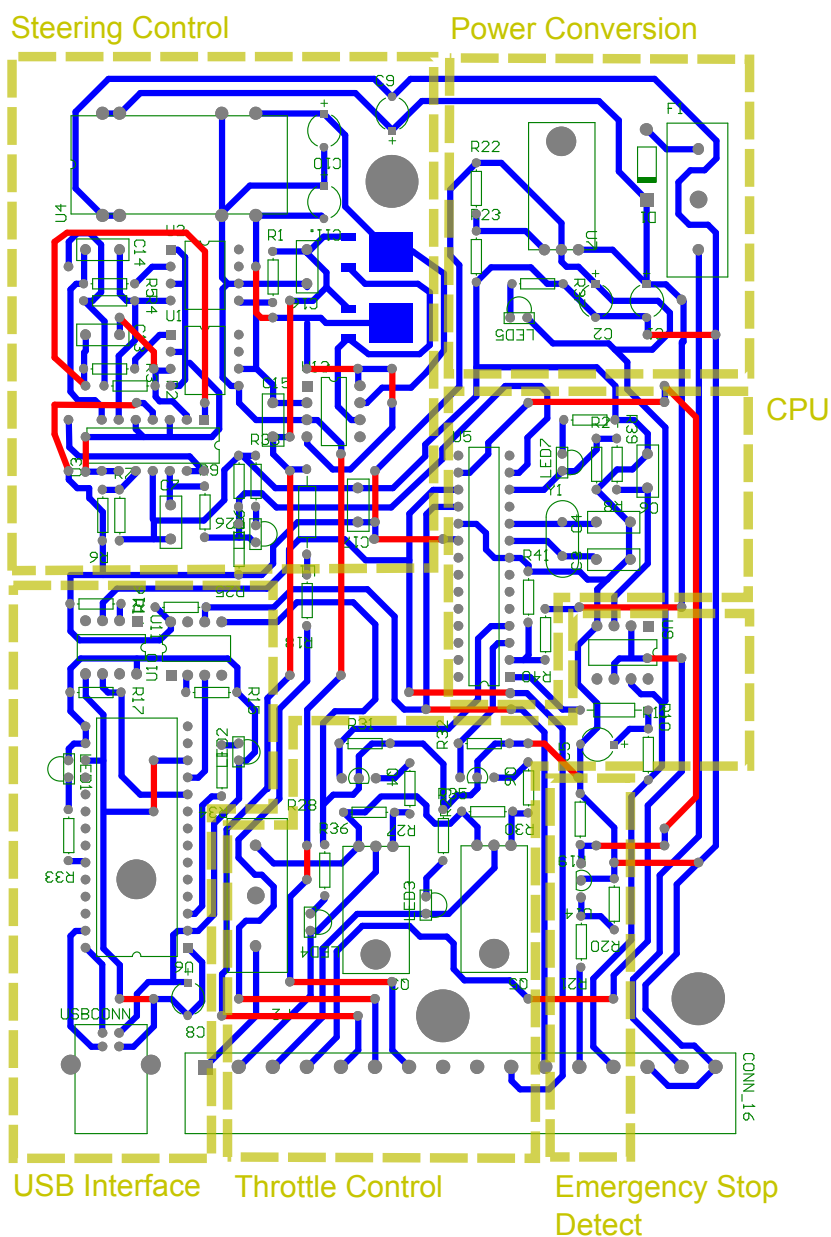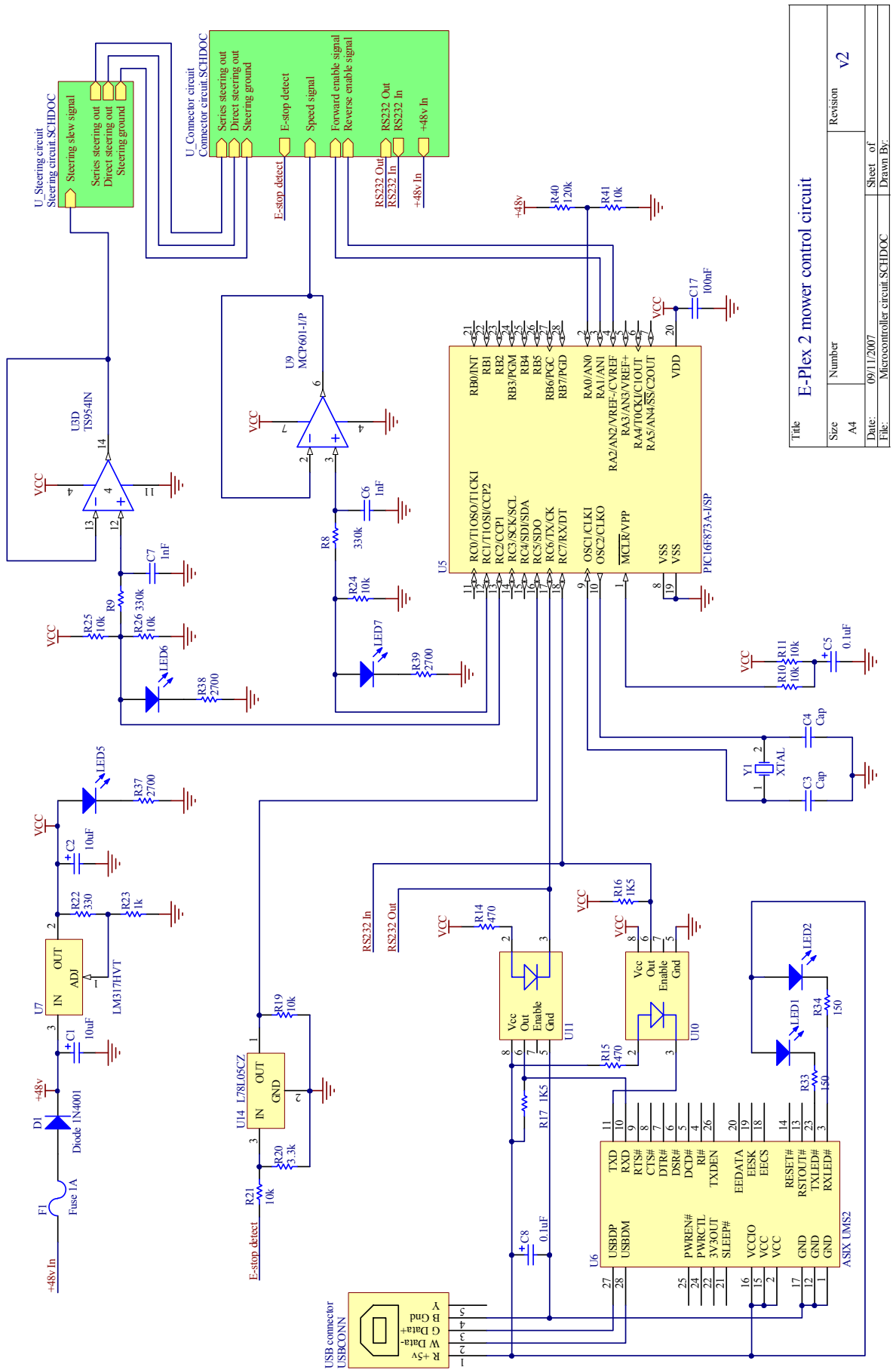The following pages show the PCB layout and schematic design of the mower-to-computer interface board.

Figure E.1: Control PCB layout: Blue bottom layer, red top layer (or wire links)

Figure E.2: Control PCB schematic page 1
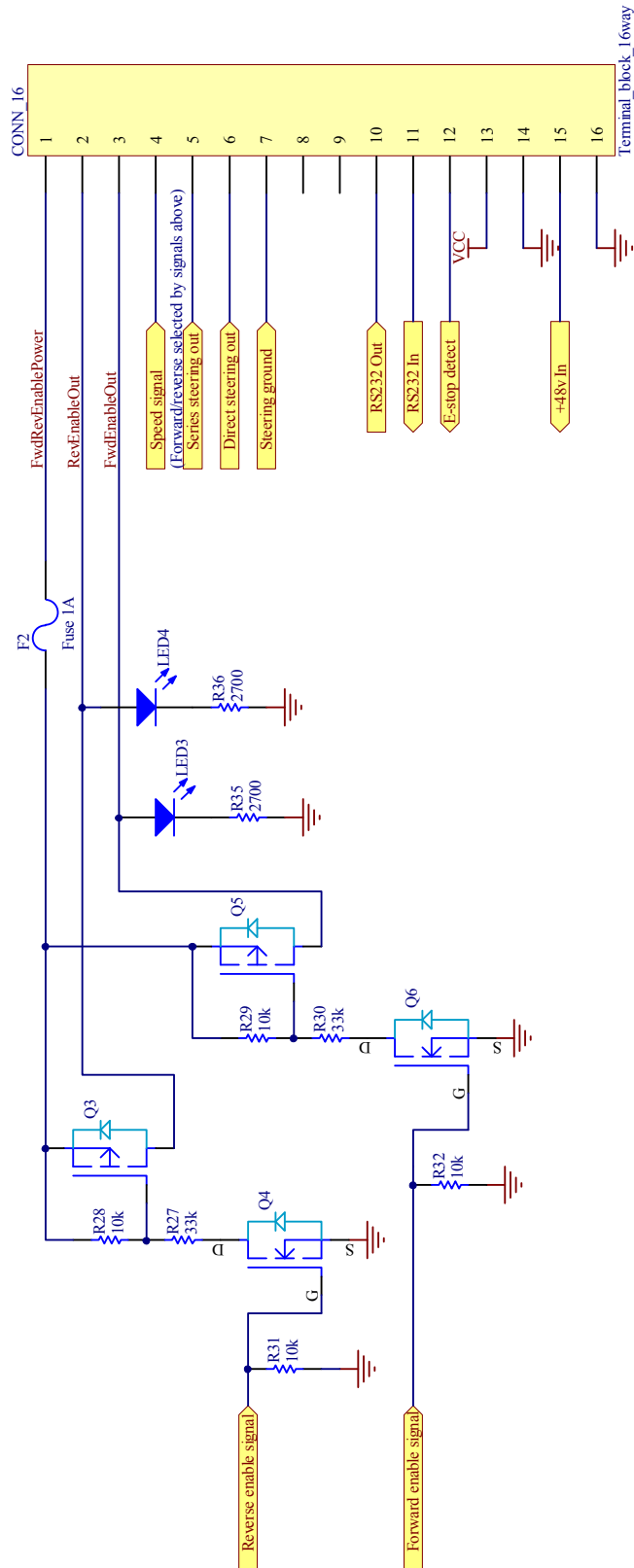
Figure E.3: Control PCB schematic page 2
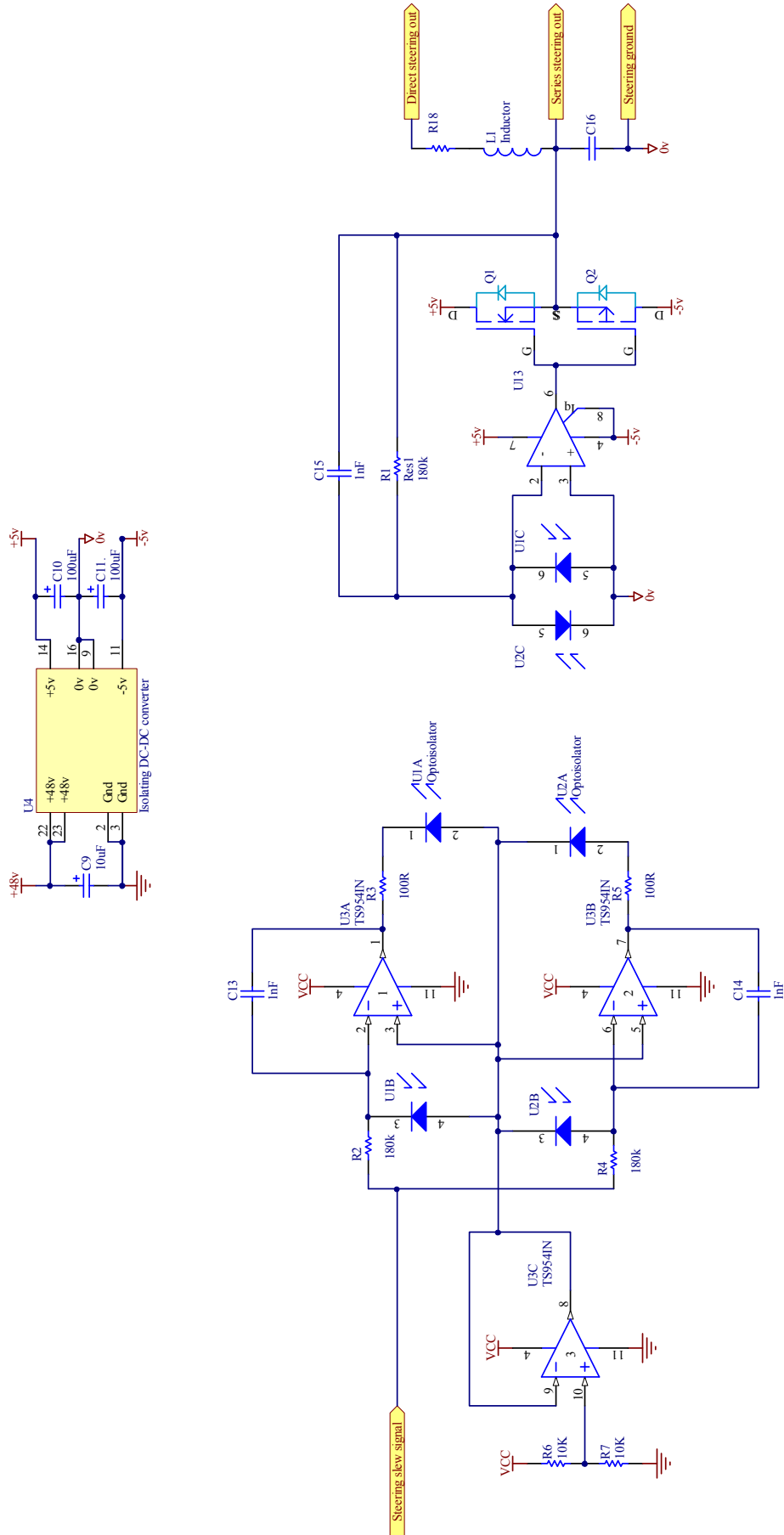
Figure E.4: Control PCB schematic page 3

References

John H.R. Amt and John F. Raquet. Positioning for Range-Based Land Navigation Systems Using Surface Topography. In *Proc. 19th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2006)*, pages 1494–1505, Fort Worth, Texas, United States, September 2006. 85, 86, 87

Martin Ax. Wir bauen das vollautomatische Auto. *Welt Online*, October 2006. URL `http://www.welt.de/print-welt/article156695/Wir_bauen_das_vollautomatische_Auto.html`. 10

Stephen Bancroft. An Algebraic Solution of the GPS Equations. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-21(1):56–59, January 1985. ISSN 0018-9251. doi: 10.1109/TAES.1985.310538. 52

Stephen Blake, Chris Hill, Terry Moore, Chris Hide, and David Park. A Heave Compensation Algorithm Based on Low Cost GPS Receivers. *The Journal of Navigation*, 61(02):291–305, 2008. ISSN 1469-7785. doi: 10.1017/S0373463307004651. 158, 159

Kai Borre, Dennis M. Akos, Nicolaj Bertelsen, Peter Rinder, and Soren Holdt Jensen. *A Software-Defined GPS and Galileo Receiver: A Single-Frequency Approach*. Springer, 2007. ISBN 978-0-8176-4390-4. 37

Bosch. DLE 150 Connect Overview, 2008. 23

Martin Buehler. Summary of Darpa Grand Challenge 2005 results. *Journal of Field Robotics*, 23(8):465–466, 2006. ISSN 1556-4967. doi: 10.1002/rob.20145. 9

Peter J. Buist, Peter J. G. Teunissen, Gabriele Giorgi, and Sandra Verhagen. Multiplatform Instantaneous GNSS Ambiguity Resolution for Triple-and Quadruple-Antenna Configurations with Constraints. *International Journal of Navigation and Observation*, 2009. doi: 10.1155/2009/565426. 110, 113, 119, 121, 122, 123

Peter J Buist, Peter J. G. Teunissen, Gabriele Giorgi, and Sandra Verhagen. Attitude bootstrapped multi-frequency relative positioning. In *Proc. Position Location and Navigation Symposium (PLANS) 2010*, pages 585–591. IEEE/ION, May 2010. doi: 10.1109/PLANS.2010.5507315. 109, 110, 118

Y. Cai, P. Cheng, X. Meng, W. Tang, and C. Shi. Using network rtk corrections and low-cost gps receiver for precise mass market positioning and navigation applications. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 345–349, June 2011. doi: 10.1109/IVS.2011.5940570. 50

M. Elizabeth Cannon, Gérard Lachapelle, and Gang Lu. Kinematic Ambiguity Resolution With a High Precision C/A Code Receiver. *Journal of Surveying Engineering, Amer. Soc. Civil Eng*, 119(4):147–155, 1993. 76, 78, 79, 115

Xiao-Wen Chang, Christopher C. Paige, and Lan Yin. Code and carrier phase based short baseline GPS positioning: computational aspects. *GPS Solutions*, 9(1):72–83, April 2005. ISSN 1521-1886. doi: 10.1007/s10291-004-0112-8. 43, 46

H. Stewart Cobb. *GPS pseudolites: Theory, design, and applications.* PhD thesis, Stanford University, September 1997. 86

Emily Cosser, Gethin W. Roberts, Xiaolin Meng, and Alan H. Dodson. Single Frequency GPS for Bridge Deflection Monitoring: Progress and Results. In *1st FIG International Symposium on Engineering Surveys for Construction Works and Structural Engineering*, Nottingham, England, July 2004. 80

Directed Perception. PTU-D46-70 Datasheet, 2008. 23

M. Dunbabin, J. Roberts, K. Usher, and P. Corke. In the rough: in-field evaluation of an autonomous vehicle for golf course maintenance. In *Proc. 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, volume 4, pages 3339–3344, 2004. 3

HJ Euler and B. Schaffrin. On a measure for the discernibility between different ambiguity solutions in the static-kinematic GPS-mode. In *Kinematic Systems in Geodesy,*

Surveying, and Remote Sensing: Symposium, No. 107, page 285, Banff, Alberta, Canada, September 1990. 127

E. Frei and G. Beutler. Rapid static positioning based on the fast ambiguity resolution approach FARA: theory and first results. *Manuscripta Geodaetica*, 15(6):325–356, 1990. 127

Eric Gakstatter. End of an Era: Final Block IIR-M Satellite Launched. *GPS World*, August 2009. 35

Gabriele Giorgi and Peter J. G. Teunissen. On the Time-To-Fix for Single-Frequency GNSS-Based Attitude Determination. In *Proc. IGNSS Symposium 2009*, Queensland, Australia, 2009. 111

Gabriele Giorgi, Tim P. Gourlay, Peter J. G. Teunissen, Lennard Huisman, and Kim Klaka. Carrier Phase Ambiguity Resolution for Ship Attitude Determination and Dynamic Draught. In *Proc. FIG Congress 2010*, Sydney, Australia, April 2010a. 110, 115

Gabriele Giorgi, Peter J. G. Teunissen, Dennis Odijk, and Peter J. Buist. Enhancing the Time-To-Fix for the unaided single-frequency integer ambiguity resolution in GNSS attitude determination applications. In *Proc. Position Location and Navigation Symposium (PLANS) 2010*, pages 236–244. IEEE/ION, May 2010b. doi: 10.1109/PLANS.2010.5507206. 110, 117

Robert Gold. Optimal binary sequences for spread spectrum multiplexing (corresp.). *Information Theory, IEEE Transactions on*, 13(4):619–621, October 1967. ISSN 0018-9448. 37

Konstantin Gromov, Dennis Akos, Sam Pullen, Per Enge, and Bradford Parkinson. GIDL: Generalized Interference Detection and Localization System. In *Proc. ION GPS 2000*, Salt Lake City, UT, United States of America, September 2000. 38

Ron Hatch. Instantaneous ambiguity resolution. In *Proc. IAG International Symposium*

107 on Kinematic Systems in Geodesy, Surveying and Remote Sensing, volume 107, pages 299–308, Banff, Canada, September 1991. 66, 79, 84

Chris Hide, Terry Moore, and Chris Hill. A Multi-Sensor Navigation Filter for High Accuracy Positioning in all Environments. *The Journal of Navigation*, 60(03):409–425, 2007. ISSN 1469-7785. 57

Bernhard Hofmann-Wellenhof, Herbert Lichtenegger, and Elmar Wasle. *GNSS - Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and more.* Springer, 2008. ISBN 978-3-211-73012-6. 35, 42, 46, 64, 65, 70, 89

IS-GPS-200D. NAVSTAR GPS Space Segment and Navigation User Interfaces. Technical report, ARINC Engineering Services LLC, March 2006. 34

Peter Johnson and Marek Danis. Unmanned aerial vehicle as the platform for light-weight laser sensing to produce sub-meter accuracy terrain maps for less than 5 dollars per square km. *Columbia Undergraduate Science Journal*, 2006. 10

JRC. JLR-10 GPS Compass Datasheet, 2002. 109

R.N. Jørgensen, C.G. Sørensen, J.M. Pedersen, I. Havn, H.J. Olsen, and H.T. Søgaard. Hortibot: An accessory kit transforming a slope mower into a robotic tool carrier for high-tech plant nursing - part i. *ASABE Annual International Meeting*, 2006. 3

Samer Khanafseh, Steven Langel, and Boris Pervan. Overbounding position errors in the presence of carrier phase multipath error model uncertainty. In *Position Location and Navigation Symposium (PLANS), 2010 IEEE/ION*, pages 575–584. IEEE, 2010. doi: 10.1109/PLANS.2010.5507313. 135

Hee-Sung Kim and Hyung Keun Lee. Real-Time Implementation of L1 RTK System Based on Position-Domain Hatch filter. In *Proc. European Navigation Conference - Global Navigation Satellite Systems*, Naples, Italy, May 2009. 76, 79, 81, 82

Young-gie Kim, Jong-Suk Choi, Mun Sang Kim, and Jin-Oh Kim. A new localization system for mobile robots using radio frequency and ultrasound. In *2002 Internatio-*

nal Conference on Control, Automation and Systems (ICCAS2002), pages 148–152, 2002. 12

Gérard Lachapelle, Congyu Liu, and Gang Lu. Quadruple Single Frequency Receiver System for Ambiguity Resolution on the Fly. In *Proc. ION GPS93*, pages 1167–1172, Salt Lake City, Utah, United States, 1993a. 76, 79, 81, 110, 112, 113, 124

Gérard Lachapelle, Congyu Liu, Gang Lu, Bryan Townsend, M. Elizabeth Cannon, and Rob Hare. Precise marine DGPS positioning using P code and high performance C/A code technologies. In *Proc. ION National Technical Meeting*, pages 241–250, San Francisco, United States, January 1993b. 76, 78, 81

R.B. Langley, H. Jannasch, J. Peeters, and S. Bisnath. The GPS broadcast orbits: an accuracy analysis. In *Proc. 33rd COSPAR Scientific Assembly*, Warsaw, Poland, July 2000. 33

Leica. Leica Geosystems SmartNet website, 2012. URL `http://www.leica-geosystems.com/uk/en/2346_Smart_Net_Brochure_.indd.pdf`. 15

X. J. Li, A. T. P. So, and S. K. Tso. Cad-vision-range-based self-localization for mobile robot using one landmark. *Journal of Intelligent and Robotic Systems*, 35:61–81, 2002. ISSN 0921-0296. doi: 10.1023/A:1020240026070. 23

Chang Boon Low and Danwei Wang. GPS-Based Path Following Control for a Car-Like Wheeled Mobile Robot With Skidding and Slipping. *Control Systems Technology, IEEE Transactions on*, 16(2):340–347, March 2008. ISSN 1063-6536. doi: 10.1109/TCST.2007.903100. 1

Ning Luo and Gérard Lachapelle. Centimeter-level Relative Positioning of Distributed Moving Platforms Using GPS Carrier Phase Observables. In *Proc. 55th Annual Meeting of the Institute of Navigation*, pages 307–313, June 1999. 111

Michael McCarthy and Henk L. Muller. Positioning with independent ultrasonic beacons. Technical Report CSTR-05-005, Department of Computer Science, Univer-

sity of Bristol, September 2005. URL `http://www.cs.bris.ac.uk/Publications/Papers/2000430.pdf`. 12

Dennis Odijk, Johannes Traugott, Gottfried Sachs, Oliver Montenbruck, and Christian C.J.M. Tiberius. Two approaches to precise kinematic GPS positioning with miniaturized L1 receivers. In *Proc. of ION 20th International Techical Meeting of the Satellite Division*, pages 827–838, Fort Worth, Texas, United States, September 2007. 76, 77, 102, 112

Kyle O'Keefe, Mark Petovello, Gérard Lachapelle, and M. Elizabeth Cannon. Assessing probability of correct ambiguity resolution in the presence of time-correlated errors. *Navigation*, 53(4):269–282, 2006. 133

OSC. Request for Public Comments on Codeless and Semi-Codeless Access to GPS, 2008. URL `http://www.space.commerce.gov/gps/semicodeless/`. 15

OXTS. Oxford Technical Solutions RT3003 Inertial and Dual GPS Navigation System Datasheet, 2009. 109

S.M. Pedersen, S. Fountas, H. Have, and B.S. Blackmore. Agricultural robots-system analysis and economic feasibility. *Precision Agriculture*, 7(4):295–308, September 2006. doi: 10.1007/s11119-006-9014-9. 1

Mark G. Petovello, Kyle O'Keefe, Gérard Lachapelle, and M. Elizabeth Cannon. Quantifying ambiguity resolution performance in the presence of time-correlated measurement errors using geometric-based techniques. In *Proc. 61st Institute of Navigation Annual Meeting*, pages 1073–1085, Cambridge, Massachusetts, United States of America, June 2005. 133

Mark G. Petovello, Kyle O'Keefe, Gérard Lachapelle, and M. Elizabeth Cannon. Consideration of time-correlated errors in a Kalman filter applicable to GNSS. *Journal of Geodesy*, 83(1):51–56, 2009. ISSN 1432-1394. doi: 10.1007/s00190-008-0231-z. 135

James Pinchin, Chris Hide, David Park, and XiaoQi Chen. Precise kinematic positioning using single frequency gps receivers and an integer ambiguity constraint. In

Proc. IEEE/ION Position, Location and Navigation Symposium 2008, pages 600–605, May 2008. doi: 10.1109/PLANS.2008.4570103. 110, 113, 115, 116, 119, 121, 122, 123

William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes: the art of scientific computing, Third Edition.* Cambridge University Press, 2007. ISBN 978-0-521-88068-8. 64

M Roberti. A 5-cent breakthrough. *RFID Journal*, May 2006. URL `http://www.rfidjournal.com/article/articleview/2295/1/128/`. 21

Stephan Roth and Sanjiv Singh. Application of robust, high-accuracy positioning for autonomous ground vehicles. In *Proc. AUVSI Unmanned Systems North America*, August 2004. 2, 3

Masayuki Saeki and Muneo Hori. Development of an accurate positioning system using low-cost L1 GPS receivers. *Computer-aided civil and infrastructure engineering*, 21 (4):258–267, 2006. ISSN 1093-9687. doi: 10.1111/j.1467-8667.2006.00433.x. 81, 82

G.A.F. Seber and C.J. Wild. *Nonlinear regression.* Wiley, 1989. ISBN 0471617601. doi: 10.1002/0471725315. 57

SICK. LMS200/211/221/291 Technical Manual, 2006. 9

Imaging Source. DBK 31BF03-Z Overview, 2008. 23

Stakemill. Sokkia SRX5 Overview, 2008. 23

Gilbert Strang and Kai Borre. *Linear Algebra, Geodesy, and GPS.* Wellesley Cambridge Press, 1997. ISBN 0-9614088-6-3. 64, 70

Tomoji Takasu and Akio Yasuda. Evaluation of RTK-GPS Performance with Low-Cost Single-Frequency GPS Receivers. In *Proc. International Symposium on GPS/GNSS 2008*, Odaiba, Tokyo, Japan, November 2008. xi, 1, 66, 76, 77, 80, 81, 82, 102, 103

Michael Tandy and Ken Young. Variable duration fixed failure rate ambiguity resolution. *GPS Solutions Online First*, pages 1–7, 2012. ISSN 1080-5370. doi: 10.1007/s10291-012-0265-9. xiii

Peter J. G. Teunissen. The least-squares ambiguity decorrelation adjustment: a method for fast GPS integer ambiguity estimation. *Journal of Geodesy*, 70(1):65–82, November 1995. ISSN 1432-1394. doi: 10.1007/BF00863419. 57, 61, 62, 63, 64

Peter J. G. Teunissen. An optimality property of the integer least-squares estimator. *Journal of Geodesy*, December 1999. ISSN 1432-1394. doi: 10.1007/s001900050269. 83

Peter J. G. Teunissen. GNSS ambiguity bootstrapping: theory and application. In *Proc. Int. Symp. on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS2001)*, pages 246–254, 2001. 130

Peter J. G. Teunissen. An invariant upperbound for the GNSS bootstrappend ambiguity success-rate. *Journal of Global Positioning Systems*, 2(1):13–17, 2003. 133

Peter J. G. Teunissen. Integer aperture bootstrapping: a new GNSS ambiguity estimator with controllable fail-rate. *Journal of Geodesy*, 79:389–397, 2005. ISSN 1432-1394. doi: 10.1007/s00190-005-0481-y. 131, 136

Peter J. G. Teunissen. The LAMBDA method for the GNSS compass. *Artificial Satellites*, 41(3):89–103, 2006. ISSN 0208-841X. doi: 10.2478/v10018-007-0009-1. 109, 113, 116, 117

Peter J. G. Teunissen and Sandra Verhagen. On the foundation of the popular ratio test for GNSS ambiguity resolution. In *Proc. ION GNSS 17th International Technical Meetings of the Satellite Division*, pages 2529–2540, Long Beach, California, September 2004. 128, 136, 137

Peter J. G. Teunissen and Sandra Verhagen. The GNSS ambiguity ratio-test revisited: a better way of using it. *Survey Review*, 41(312):138–151, 2009. ISSN 0039-6265. doi: 10.1179/003962609X390058. 128, 129, 139, 141

Peter J. G. Teunissen, Paul J. De Jonge, and Christian C.J.M. Tiberius. The least-squares ambiguity decorrelation adjustment: its performance on short GPS baselines and short observation spans. *Journal of Geodesy*, 71(10):589–602, 1997. ISSN 1432-1394. 64

Peter J. G. Teunissen, Gabriele Giorgi, and Peter J. Buist. Testing of a new single-frequency GNSS carrier phase attitude determination method: land, ship and aircraft experiments. *GPS Solutions*, March 2010. ISSN 1521-1886. doi: 10.1007/s10291-010-0164-x. 110, 111

Christian Tiberius and Paul J. De Jonge. Fast positioning using the LAMBDA method. In *Proc. 4th International Symposium on Differential Satellite Navigation Systems DSNS*, volume 95, pages 24–28, 1995. 127

Mami Ueno, Rock Santerre, Daniel Langelier, and Guy Marceau. Improvement of GPS ambiguity resolution using height constraint for the support of bathymetric surveys. In *Proc. IAIN/ION Conference*, volume 850, San Diego, California, June 2000. 83, 84

N.N. Vatani, M Bosse, J Roberts, and M. Dunbabin. Practical path planning and obstacle avoidance for autonomous mowing. In *Proc. Australasian Conference on Robotics and Automation (ACRA 2006)*, 2006. 3

Sandra Verhagen. Integer ambiguity validation: an open problem? *GPS Solutions*, 8 (1):36–43, 2004. ISSN 1521-1886. doi: 10.1007/s10291-004-0087-5. 65, 66, 126, 127

Sandra Verhagen. On the reliability of integer ambiguity resolution. *Navigation*, 52(2): 99–110, 2005a. ISSN 0028-1522. 130

Sandra Verhagen. *The GNSS integer ambiguities: estimation and validation*. PhD thesis, 2005b. 133

Shawn D. Weisenburger. Effect of constraints and multiple receivers for on-the-fly ambiguity resolution. *UCGE Reports No*, 20109, April 1997. 66, 76, 80, 84, 85, 96, 103, 110, 113, 117

Greg Welch and Gary Bishop. An introduction to the Kalman filter. *University of North Carolina at Chapel Hill, Chapel Hill, NC*, 7(1), 1995. 57

Guijin Zheng and Demoz Gebre-Egziabher. Using Carrier Phase Attitude Determination Constraints to Enhance Single Frequency RTK Integer Ambiguity Resolution Performance. In *Proc. 22nd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2009)*, pages 2156–2168, Savannah, Georgia, United States, September 2009. 110, 116, 121, 123, 124

Jian-jun Zhu and Rock Santerre. Improvement of GPS phase ambiguity resolution using prior height information as a quasi-observation. *Journal of the Canadian Institute of Geomatics*, 56(3):211–221, 2002. 84, 85, 96, 101, 103

Jian-jun Zhu, Rock Santerre, and Xiao-Wen Chang. A Bayesian method for linear, inequality-constrained adjustment and its application to GPS positioning. *Journal of Geodesy*, 78(9):528–534, 2005. ISSN 1432-1394. doi: 10.1007/s00190-004-0425-y. 84, 85, 87, 96, 103