MONITORING RAILWAY TRACK CONDITION USING INERTIAL SENSORS ON AN IN-SERVICE VEHICLE

by

GRAEME JAMES YEO

A thesis submitted to the University of Birmingham
for the degree of DOCTOR OF PHILOSOPHY

Birmingham Centre for Railway Research and Education

School of Engineering

College of Engineering and Physical Sciences

University of Birmingham

June 2017

Abstract

Effective maintenance of railway track is critical for the safe operation of any railway network. Efficient maintenance may also result in economic benefits for rail operators.

The work in this thesis looks into how an inexpensive measurement system could be fitted to in-service railway vehicles such as commuter trains, to provide a relatively high frequency of measurement on their routes of operation, when compared to dedicated measurement vehicles.

This thesis describes how a prototype inertial measurement system was designed and built, and fitted to a commuter train operating in the region south of London, UK. Inertial data is processed to provide a vertical profile of the track. A novel use of a modified Bryson-Frazier filter is used to produce vertical profile datasets which are repeatable to within 0.2 mm. Profiles calculated from multiple passes of the same areas of track are compared to show track degradation.

Methods of estimating track stiffness are developed using vertical geometry data from repeated passes of the same track sections at differing speeds. Some correlation to stiffness is shown through the results, but exact measurements were not possible.

Finally, two case studies are presented which show findings at a bridge approach, and through two level crossings.

Acknowledgements

I would like to thank the following for their contributions to this project:

Track 21 for an EPSRC grant (EP/H044949/1) which has funded this research.
www.track21.org.uk

Southern Railway for allowing instrumentation to be fitted to one of their trains, and for their cooperation and help maintaining the equipment.

My supervisors Clive Roberts and Paul Weston for their invaluable support in matters technical and academic.

The rest of the team at the Birmingham Centre for Railway Research and Education for their friendship and invaluable wealth of knowledge.

My dear wife Sarah, my daughter Helen, and all of my family for their love and support through the years.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF DEFINITIONS

$\boldsymbol{\theta}$ (Theta) Used to denote the angle of roll. A dot above, $\dot{\theta}$, denotes the derivative of roll with respect to time, or *roll rate*.

$\boldsymbol{\phi}$ (Phi) Used to denote the angle of pitch. A dot above, $\dot{\phi}$, denotes the derivative of pitch with respect to time, or *pitch rate*.

$\boldsymbol{\psi}$ (Psi) Used to denote the angle of yaw. A dot above, $\dot{\psi}$, denotes the derivative of yaw with respect to time, or *yaw rate*.

Accelerometer A device that measures linear acceleration along an axis. Often two or three accelerometers may be packaged together into a single device to simultaneously measure acceleration in multiple axes.

Ballasted Track Railway track comprising of rails clipped to sleepers (or ties) laid perpendicularly to the rails. This structure is supported by coarse stones known as ballast.

Bogie (*or* ***Truck***) A framework with a number of wheels attached (usually 4 or 6), mounted to the main body of a vehicle by a pivoted joint. The bogies considered in this thesis all have 4 wheels.

Diesel Multiple Unit A multiple unit train (or trainset) consisting of multiple self-propelled carriages, with diesel engines providing traction power.

DMU *Abbr.* See Diesel Multiple Unit.

Electric Multiple Unit A multiple unit train (or trainset) consisting of multiple self-propelled carriages. Electricity is used as the power source, from either overhead lines or a third rail.

EMU *Abbr.* See Electric Multiple Unit.

Gauge The distance between the two running rails of a railway track. Gauge is measured between the two inside edges of the head of the rails.

Global Positioning System A radio-based navigation system which uses a constellation of satellites to allow a receiving device to calculate its position on or above the earth's surface. Most receivers are also able to calculate their speed and heading.

GPS *Abbr.* See Global Positioning System.

Gyro *Abbr.* See Gyroscope.

Gyroscope A device used to measure the rate of rotation around an axis.

IMU *Abbr.* See Inertial Measurement Unit.

Inertial Measurement Unit A device, usually rigidly coupled to an object to be measured, which contains accelerometers and/or gyroscopes, used to measure movements experienced by the unit. In the case of this thesis, the IMU considered is mounted on a bogie of a train.

New Measurement Train A specialised measurement train operated by Network Rail in Great Britain. The train includes a track recording coach which measures and records all aspects of the track geometry.

NMT *Abbr.* See New Measurement Train.

Pitch The angle of rotation around the y axis.

Primary Suspension The suspension system between the axles and the bogie of a railway vehicle. Usually comprising of very stiff springs or hard rubber blocks.

RMS *Abbr.* See Root Mean Square.

Roll The angle of rotation around the x axis.

Root Mean Square A method of producing an absolute generalized mean of a continuously varying waveform by taking the square root of the mean of the squares of samples. This can be either a single value or discrete values for a sliding window of samples.

Secondary Suspension The suspension system between the bogie and body of a railway vehicle. On modern passenger vehicles this is often an air-based system comprising of a pair of airbags between the body and bogie.

Slab Track Railway track comprising of rails clipped to large slabs, generally constructed from moulded concrete. The slabs are laid end-to-end resulting in a continuous strip of concrete supporting the rails.

Tacho *Abbr.* See Tachometer.

Tachometer A device, usually fitted to an axle bearing of the train, which produces a signal as the train moves, which can be used to measure the speed of the train.

UGMS *Abbr.* See Unmanned Geometry Measurement System.

Unmanned Geometry Measurement System A system comprising of a number of sensors used to measure and record aspects of track geometry as a railway vehicle traverses the track. Commonly UGMS systems use optical sensors such as lasers and/or inertial sensors. The system is designed to require no human intervention during its operation.

$\boldsymbol{x}$ Used to denote distance along the axis parallel to the direction of travel. One or two dots above ($\dot{x}$ or $\ddot{x}$), denote x velocity or acceleration, respectively.

$\boldsymbol{y}$ Used to denote distance along the axis perpendicular to the direction of travel on the horizontal plane. One or two dots above ($\dot{y}$ or $\ddot{y}$), denote y velocity or acceleration, respectively.

Yaw The angle of rotation around the z axis.

$\boldsymbol{z}$ Used to denote distance along the axis perpendicular to the direction of travel on the vertical plane. One or two dots above ($\dot{z}$ or $\ddot{z}$), denote z velocity or acceleration, respectively.

CHAPTER 1

INTRODUCTION

1.1 Overview

A large number of tools are now available to railway network operators to aid in monitoring the condition of railway track. These include dedicated track measurement vehicles which have become commonplace on many railway networks over the past few decades. These vehicles are an invaluable tool for inspecting large sections of railway at high speed. Track faults identified by an inspection vehicle may also be visually inspected to gain a better understanding of the fault, but without the need to also inspect hundreds of miles of fault-free track on foot.

However, dedicated measurement vehicles have limitations. They can be expensive to run, requiring specialist crews to operate, as well as being expensive to build initially. Consequently a railway operating company may only have access to a small number of measurement vehicles. This, coupled with high traffic causing limited availability on some track, means that measurements may only be possible every one or two months at best in some areas.

The work presented in this thesis considers how a relatively inexpensive measurement system could be fitted to multiple in-service railway vehicles, to provide a much higher frequency of measurement on certain routes. The cost of running and maintaining such a system would be considerably less than operating a dedicated measurement vehicle.

During the course of the work carried out towards this thesis, an Inertial Measurement Unit (IMU), was designed and fitted to a British commuter train by the author and colleagues. The train operates on routes in the UK between London, Brighton, and Southampton. Various processing techniques are used to extract useful information from the data collected from the IMU. Multiple passes of the train over the same sections of track are used to extract information about the rates of degradation at various sites on the network.

The author has also developed processing to be performed on multiple passes of the same track in order to estimate track stiffness; an important parameter which can have significant effects on the vertical displacements experienced by a passing vehicle. This research area is still ongoing, both by the author and in other concurrent work [1].

1.2 Railway Track

Railway track consists of a number of components which must all work together to perform the track's primary role of supporting and directing railway vehicles. The work presented throughout this thesis focuses on ballasted railway track, rather than slab track or any other variations of railway track. The basic layers which compose a ballasted railway track are shown in Figure 1.1.

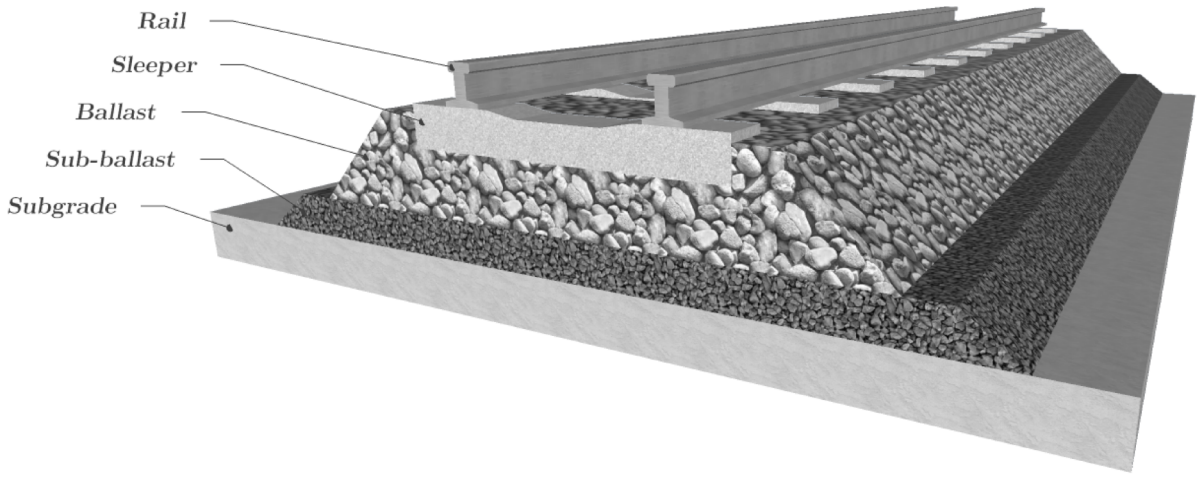


Figure 1.1: Layers forming a ballasted railway track

1.2.1 Track Geometry

Track geometry is a term used to collectively refer to the combined measurable parameters of railway track. The nature of the components of a railway track mean that its geometry varies along its length, allowing the track to follow curves and contours of the land it covers. There are a number of measurable geometry parameters. Lewis [2] provides a list of the primary geometry parameters, listed in Table 1.1.

Vertical profile	The profile of the rail head in the longitudinal vertical plane. This is usually high-pass filtered so that it measures height relative to the height of surrounding rail (usually a 35 m or 70 m length), rather than the absolute height of the rail [3], which would require a reference datum. Also known as *top*, *longitudinal level*, or *surface*.
Horizontal profile	The profile of the track centre line in the longitudinal lateral plane. This is measured as the deviation from the design profile, as many changes in lateral direction will be intentional. Also known as *alignment*, or *line*.
Gauge	The distance between the inside (gauge) faces of the two running rail heads, measured at 14 mm down from the running surface [4].
Crosslevel	The difference in height between the two running rails at any given point. This may also be referred to as *cant*, or *superelevation*, but these terms are usually specific to banked curves, with the polarity of the value matching the direction of the curve.
Curvature (horizontal)	A measure of the turning 'rate' of the track in the horizontal plane, indicating how sharp or smooth curves are. The 'rate' is measured with respect to the distance along the track rather than time. Normally the term *curvature* refers to the turn 'rate' in the horizontal plane, but it may also be measured in the vertical longitudinal plane (*vertical curvature*).

Table 1.1: List of primary track geometry parameters

Lewis also lists secondary parameters such as twist, cyclic top, corrugation, and cant deficiency which can be derived from measurements of the primary parameters. For example twist is defined as the difference in crosslevel over a specified distance, giving a measure of the rate of change of crosslevel. Measurement of the crosslevel parameter can be used to calculate this information.

1.2.2 Track Degradation

After the initial construction of the track, the geometry of the track begins to deviate from its original geometry (the design geometry). This is a process known as track degradation. Degradation can occur for a number of reasons. The passage of railway vehicles over the track, as well as adverse weather conditions can cause the settlement and migration of ballast [5, 6]. Extremely wet or dry weather conditions, as well as extreme temperatures can also cause movement of the ballast, sub-ballast and/or subgrade layers. An example of this is an effect known as 'frost heave', where water between the ballast particles freezes, causing expansion of the ballast layer and consequent track movement [7]. Once the level of degradation of the track geometry becomes too great, corrective maintenance must be performed.

One of the main parameters of degradation is the vertical profile of the track. Degradations in the profile of a track are usually caused by uneven ballast settlement which itself can be caused by a number of issues. Variations in subgrade due to geophysical effects such as wet spots in the ground can cause some areas of ballast to move more than others during the passage of a vehicle. Similarly, the existence of man-made structures beneath the track such as culverts and bridges can cause some areas of ballast to move less than other areas. The repeated passage of railway vehicles causes settlement of the ballast as it moves. If this settlement is uneven, some of the track will become better supported than other parts resulting in a vertical profile with large variations. This can lead to a poor ride quality for passengers and freight, and if the degradation becomes particularly bad, faults such as rail cracking can occur which could ultimately lead to a fatal accident.

Uneven ballast settlement can be corrected by tamping the ballast to reset the track to an acceptable geometry. Ballast tamping is an operation performed using a specialised railway vehicle known as a tamping machine, or tamper (Figure 1.2). The tamper has finger-like 'tines' on either side of the vehicle. The vehicle first lifts the track beneath it, then thrusts its tines downwards into the ballast. The tines are then pushed towards each other, causing the ballast stones to be compacted together and upwards. Stoneblowing is another possible method of correction, where the track is lifted and ballast stones are blown underneath areas of poorly supported track. Maintenance operations such as this are expensive as they require the use of specialised machinery and crews, as well as requiring possession of the track. The ability to improve knowledge of which specific areas of track require maintenance and the frequency at which maintenance is required could have significant economic benefits for railway operators. Specifically, this thesis focuses on obtaining detail about two specific parameters of the track. Firstly the vertical geometry of the track, which tends to degrade over time, meaning that the track deviates from its design geometry (which is usually almost flat along a 70 m length of track). Secondly the track stiffness, which can itself affect the track's vertical profile, and extreme values of which can lead to track failures.

Photograph by 'Peter Broster', Source: flickr.com
Reproduced under the Creative Commons Attribution 2.0 Generic license

Figure 1.2: A tamping machine

1.2.3 Track Stiffness

Some measurement devices can measure the vertical profile of the track in an unloaded state (or a negligibly loaded state) [8–11]. However, as a train passes over the track, it will naturally deflect downwards due to the force of the train on the track, resulting in a loaded vertical profile that can differ significantly from the unloaded state. Deflection typically varies from a fraction of a millimetre to tens of millimetres in extreme scenarios. The amount of deflection depends on a number of factors. Firstly the axle load of the passing train; a heavier train will cause greater deflection, and secondly the stiffness of the track. The speed of the train can also affect the deflection, due to dynamic forces introduced to the axles. The *static axle load* of the train is the load on the axles as when the train is stationary, whereas the *dynamic axle load* of the train increases as the speed increases. A moving train can also introduce other factors affecting track deflection such as ground wave effects, but these are beyond the scope of this thesis.

A low track stiffness results in a large amount of deflection whereas a high stiffness results in little movement of the track. In fact, the relationship between stiffness and deflection can be modelled as a linear relationship [12], and can be calculated using:

$$k = \frac{F}{z} \tag{1.1}$$

Where k is the stiffness, F is the force being applied, and z is the vertical deflection. An Innotrack report [12] identifies that track degradation can be caused not only by extreme values of track stiffness (both low and high), but also by variations in stiffness.

Often rapid rates of degradation are found to exist where the stiffness of the supporting ground or structure transitions rapidly between parts of the track. For example where wet spots or culverts exist beneath the track, or on the transition from soil subgrade onto a concrete bridge or from ballasted track onto slab track. Such areas are known as transition zones or critical zones. Figure 1.3 shows the side profile of a section of track at the approach to a bridge, where a transition zone exists.

Transition zones are often smoothed using an approach slab or 'geogrid' reinforcement system, as well as using graded aggregates on the approach to larger structures [13, 14]. This typically leads to transition lengths of 20 m or more [13, 15].

Transition zones often require much more frequent maintenance than other areas of track. Regular monitoring of these areas would allow maintenance to be scheduled only when necessary, as well as providing an insight into the degradation of such areas of track. This could lead to more efficient use of maintenance resources, and ultimately, a reduction in maintenance costs.

In several recent papers [16–18] the impacts of transition zones are investigated, and consideration is given to ways of minimising stiffness variation and the formation of voids. Voids can form beneath sleepers which are regularly subjected to large displacements, and result in unloaded sleepers being supported only by the rails to which they are clipped.

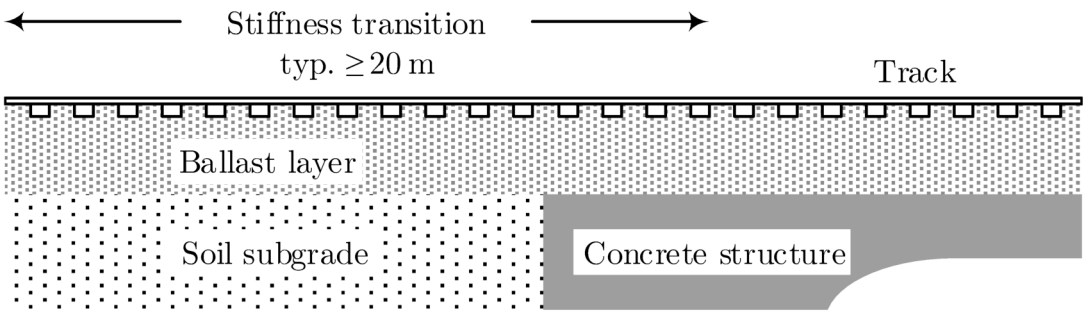


Figure 1.3: An example of a transition zone at the approach to a bridge

1.3 Measuring Track Geometry

In Section 1.2.2 it was noted how track geometry can degrade over time, and how track that has degraded significantly requires rectification. In order to determine whether a section of track requires maintenance, there exist many technologies for measuring track geometry. These technologies have been developing since the early days of railway engineering. Improvements have led to greater accuracy of measurements, and faster identification and assessment of faults.

Initially geometry was measured by hand, using simple measuring tools. Over time, instrumentation has been developed to speed up the process of geometry measurement. Many modern systems use instrumentation fitted to railway vehicles. Some of these are dedicated measurement vehicles, whilst some are standard service vehicles instrumented with various sensors. The literature review in this thesis (Chapter 2) looks in detail at some of the measurement technology developed over the last century.

The work in this thesis looks at how track measurements could be made by the instrumentation of an in-service vehicle, specifically a multiple-unit passenger train. Measurements would be made during normal operation of the train on a day-to-day basis.

The system proposed consists of a compact inertial measurement unit (IMU) capable of being mounted onto a bogie of an in-service vehicle. Equipment for receiving and storing data from the IMU would be installed on board the train itself. It would be relatively easily installed on the vehicle's bogie, and would require little or no maintenance throughout its operating life. Such a system could therefore be installed on many in-service vehicles to provide a high resolution of measurements over time for a large proportion of a rail network.

Whilst the system has a relatively low cost of manufacture compared to the cost of a complete instrumented vehicle, it is not expected that it would replace a dedicated

vehicle, which would likely provide more accurate and reliable measurements, as well as in some cases having the capability of measuring many different track parameters, rather than vertical geometry alone. The introduction of an IMU system on in-service vehicles network-wide could have some cost benefits. For example, a dedicated measurement vehicle may operate on a section of track shortly before a component begins to fail. It may be several weeks before the vehicle returns, and in the mean time, the track geometry may degrade to the point where operations must be halted, and maintenance performed. With daily measurements available from an in-service IMU system, the failure would have been detected much sooner, and maintenance could have been scheduled outside of operating hours. Ortega *et al.* [19] find that the approximate cost of a delay per minute is 2.5 times the value of travel time on average. Minimising the number of failures during revenue operation time would therefore reduce the overall operating costs.

Another key advantage of using an in-service system is that an instrumented vehicle measures track during normal operation. It may be deemed possible to reduce the frequency of operation of dedicated measurement vehicles which is of great advantage in high-traffic areas of railway networks where scheduling time to run dedicated vehicles is difficult or sometimes impossible. One example of this is the London Underground Victoria Line, where some trains are spaced apart by only 1 minute and 49 seconds [20].

1.3.1 Measuring Track Stiffness

As track stiffness variations often play a key role in the degradation of vertical track geometry, it is clear that a method of measuring track stiffness from the same measurement system would be beneficial. There are many obstacles to overcome to achieve this, especially since any measurement system mounted on a vehicle bogie will only ever be able to measure the loaded profile of the track.

An inertial geometry measurement system would need to be able to measure the track under different loading conditions. This could potentially be achieved in three different ways.

1. Two different bogies of an in-service vehicle could be instrumented, where it is known that one bogie is loaded more heavily than another. Results from the two bogie IMUs can be compared and stiffness estimated.
2. For a single bogie system, the bogie motion can be monitored when different loads are on board. Passes of the same track sections could be compared to estimate a track stiffness value.
3. A single bogie is monitored over multiple passes of the same track sections, when operating at different speeds. Speed can affect the dynamic loading of the train, and the differences could be detected by the inertial measurement system.

The specifications of the IMU system developed in this work specifically called for instrumentation of only a single bogie, which rules out the first option. The other two options rely on comparing multiple passes of the same track sections. It is likely that differences in vertical geometry perceived by the bogie under different load conditions (both static and dynamic) would be very small. This based on an approximate stiffness figure in the order of magnitude of 70 $\mathrm{MN\,m^{-1}}$ [21] and an approximate axle load of 110 kN for a Class 377 EMU, which gives a displacement of 1.6 mm. In order to measure variations in stiffness, achieving a relatively coarse resolution of 10 $\mathrm{MN\,m^{-1}}$, the system would need to be able to measure vertical displacement to within 0.2 mm repeatability from run to run.

The ability to measure track stiffness would be useful in helping to identify the causes of track degradation, but ultimately it is the change in the loaded vertical profile itself that causes problems. Therefore it is possible that the assessment of track condition and consequent monitoring of degradation can be achieved without the need for stiffness measurements.

1.4 Hypothesis

This thesis aims to answer whether the following hypotheses are true:

(a) *An unattended IMU mounted on the bogie of an in-service vehicle could be used to assess the condition of the railway track that the vehicle operates on.*

(b) *Repeated data from multiple passes of the same track sections can be used to monitor the degradation of the vertical track geometry over time.*

(c) *Repeated data from multiple passes of the same track sections can be used to estimate the stiffness of the track.*

1.5 Thesis Structure

In Chapter 2, a thorough review of existing track measurement technologies is undertaken, and conclusions drawn from these.

Following the literature review, the development and installation of an IMU was carried out. Chapter 3 describes this work. Following this, Chapter 4 provides detail about the processing techniques used to convert raw inertial data collected from the IMU into track geometry data. Techniques used to combine data from different inertial sensors in order to enhance the accuracy of geometry measurements are presented in Chapter 5.

Once data had been collected and processed, work was done to align data from multiple passes of the IMU over the same sections of track. Chapter 6 describes this work. Aligned data was then used to develop processing to estimate track stiffness, detailed in Chapter 7. This section of development fills the remaining time available for work towards this thesis.

In Chapter 8, two case studies are included. In these studies, track condition at two sites was investigated using data collected by the developed IMU system. The first site studied is Fishbourne in Sussex, UK. Here, a track renewal took place, after which the track was

monitored over a one year period. The results of degradations and maintenance work at adjacent level crossings are inspected. The second site is Shoreham-by-Sea, also in Sussex, UK. At this site, a large steel bridge spans the river Adur. Degradation at the transition onto the bridge is monitored over a nine month period. Stiffness estimations are calculated for the start and end of this period.

Lastly, in Chapter 9, conclusions are drawn from the work done, and the hypothesis is answered. The author's novel contributions to research are also highlighted, and finally consideration is given to future research arising from this thesis.

CHAPTER 2

A REVIEW OF TRACK MEASUREMENT TECHNOLOGIES

A broad range of track measurement technologies exist. Firstly a number of methods of vehicle-based measurement technologies are reviewed, ranging from early systems through to modern instrumented vehicles.

2.1 Vehicle-Based Track Measurement

Recording the condition of track from railway vehicles is by no means a modern concept. One of the earliest examples of an automatic recording vehicle capable of detecting track faults is the Sperry Detector car [22, 23] invented by Dr. Sperry in around 1928. The self-powered electric vehicle was capable of detecting cracking and fissures in rails using an induction method. Results were recorded using a paper roll chart recorder. When faults were detected, the vehicle deposited a small amount of paint onto the track to mark the position of the defect.

Over the last century, many systems attempt to measure track geometry using a variety of sensors fitted to various railway vehicles. A doctoral thesis from the University of Birmingham [24] provides a good review of a number of track measurement systems, ranging from bespoke measurement units to commercially available measurement vehicles.

Another early vehicle-based measurement system is a mechanical feeler system from 1970 [25]. This system is able to continuously measure the static gauge of the track, as well as the alignment and cross-level. Cross-level is measured using a gyro, but is mounted on the car floor to protect it from the harsh bogie environment. Early gyros were particularly susceptible to damage from high accelerations.

Some more recent track railway systems use optical components to measure the track. One such system [26] spins a wide-profile laser so that it measures the track, as well as obstacles above and beside the track as the vehicle moves along the track. Using the laser the system is able to measure the gauge of the track. The proposed speed of operation of the vehicle is 40 $\mathrm{km\,h^{-1}}$. There are more modern variations of this system now operational in the UK [27]. Improvements over the past four decades mean that modern systems have greatly improved accuracy, and digital recordings which make it much easier to align and compare data sets recorded at different times.

Another laser-based system is described in [28] which uses two laser displacement sensors - one pointed at each rail - and a central roll-rate gyroscope. These are mounted in a single self-contained unit. The report gives examples of its installation on both a small towed wagon, and on the bogie of an inspection train. Later work by the same author [29] proposes a similar system of lasers supplemented by a longitudinally projected laser beam along a two-bogie carriage. By using this system, track inspection at a speed of 275 $\mathrm{km\,h^{-1}}$ can be achieved. An article [30] reviews an implementation of the system on the "Doctor Yellow" inspection train for the Shinkansen network in Japan.

A recent system developed in Russia [31–34] uses a combination of laser scanners pointed at each rail, and a strap-down inertial navigation system (SINS) unit mounted on the underside of a vehicle bogie. It uses the inertial system to both aid in locating the vehicle, and to provide lateral track geometry measurements, whilst the laser scanners measure the rail profile. This system is the first reviewed which uses an inertial measurement unit on a vehicle bogie to measure track geometry, although the unit is only used to measure

lateral geometry in this system.

A significant problem with laser-based measurement systems is that the optical components require frequent cleaning to ensure continued operation. In a railway environment, particularly beneath a train, components of the bogie and running gear quickly become coated in a layer of brake dust and other particulates picked up during the train's operation. Whilst the sides and top of rolling stock may be regularly cleaned, the underside of the train rarely is. This means that whilst optical components may be acceptable on a dedicated measurement vehicle, where the necessary cleaning and maintenance is performed, they are unsuitable for an in-service measurement system where the cleaning of bogie and under-carriage components is very infrequent.

Hayashi *et al.* [35] describe in detail another Japanese track fault detection system, which instead opts to use inertial sensors only, rather than the optical systems reviewed previously. Accelerometers are placed both in the carriage and on axle boxes. Exterior noise is also measured simultaneously using a microphone. The system is primarily aimed at detecting rail corrugation, and uses wavelet analysis to do so. It is proposed in the paper that a system be developed to detect corrugation from measurements obtained only from the vehicle cabin, to be installed on commercial vehicles. This system uses axle box accelerometers, rather than being bogie-mounted as proposed for the inertial measurement system in this thesis. This would allow shorter wavelength features to be detected, which whilst important for the application of measuring rail corrugation, is unnecessary for measuring features above 0.5 m in length.

One of the authors of this work [35], Tsunashima, leads a project following up this work. A conference paper [36] describes the design and development of the previously proposed device, which uses inertial measurements, that may be temporarily installed in the cab of an in-service vehicle, meaning that instrumentation is easy to install and maintain. Another paper [37] describes a later self-contained variant of the device, which further simplifies installation. A 2-dimensional model of the vehicle suspension is used to

estimate the track irregularity from accelerations experienced on-board the train. Data is supplemented with readings from a GPS receiver. The first paper acknowledges that the vehicle suspension will filter out any effects of rail corrugation so that the accelerometer will be ineffective at detecting corrugation. A microphone is added to the device to attempt to identify corrugation from the sounds experienced in the cab. The second version of the device also allows data to be reported via a cellular phone network. Two intermediate works [38, 39] show the progression of development of the system, including verification of the output of the system against data obtained from a dedicated track recording vehicle.

The combination of train body-mounted accelerometers and a microphone as used here can be used to detect long wavelength and very short wavelength features in the track geometry. It seems that some mid-wavelength information can be found by using a vehicle suspension model, however, more accurate results tend to be found with closer coupling to the track. In some scenarios, exact track geometry may not need to be determined, and only a measure of track quality may be required.

An inertial measurement system developed by Ishii *et al.* [40], shows the development of a low-cost inertial measurement system, which is also installed in the cab of an in-service vehicle. Inertial measurement is done by way of a three axis accelerometer with positional data added from a GPS receiver. Longitudinal acceleration is double-integrated and used to enhance the positional data given by the GPS. The system calculates the RMS of the measured vertical and lateral accelerations, and uses these as measures of track quality.

Another self-contained IMU system developed by Heirich *et al.* [41] is designed to measure long wavelength lateral track geometry such as curves and bank angles. The unit is mounted on an overhead luggage rack of an in-service passenger train and contains low-cost accelerometers and gyroscopes measuring acceleration and rotation in 3 axes at 100 Ss^{-1}. GPS data is also logged at 1 Ss^{-1}. The paper goes into details on the equations used to calculate the train motion. The system is still under development but is able to

determine the acceleration of the train along the track, the yaw of the train as it travels around curves, and the bank change of the train.

The last few systems reviewed have all performed inertial measurement from on-board the body of the train. The positioning of inertial sensors is significant, and must be chosen according to the application and parameters which need to be measured. Lee *et al.* [42] investigate the differences in the estimation of both vertical and lateral track geometry as measured by both bogie and axle box mounted accelerometers. Estimations are verified against a commercial laser-based geometry measurement system. The work finds that in fact the bogie-mounted accelerometer produces vertical geometry closer to that measured by the laser system. A number of reasons for this are suggested, but the paper concludes that further research is required to understandy why. A later paper by the same authors [43] uses data from the same system, but combines data from the two accelerometer locations using a Kalman filtering method in order to produce an accurate representation of the geometry. It would seem from this research that the inherent filtering of the bogie's primary suspension is in fact a benefit if only a single sensor position (bogie or axle box) is to be chosen. The work also shows that it is indeed possible to measure vertical track geometry using inertial sensors mounted on the bogie. It also shows that there is a benefit of using both bogie and axle box sensors, which is that much shorter wavelengths can be measured. However, as stated before, the system developed in this thesis does not need to be able to measure short wavelength defects such as corrugation. Addition of axle box accelerometers would mean more work would be required installing and maintaining the system, which would be disadvantageous, particularly if the system were later commercialised.

An alternative approach using a similar configuration of sensors [44], uses fuzzy logic techniques to combine measurements from bogie and axle box mounted accelerometers. The system is designed to measure track vibrations. In fact double integration of the acceleration data is done to derive displacement in both the lateral and vertical axes.

Data is supplemented with positional information from a GPS receiver, and any results deemed to be outside acceptable limits are reported to a central office using a GSM modem. Whilst combining bogie and axle box sensor outputs is not required, use of the Kalman and fuzzy logic techniques to combine sensor readings to generate better estimates of track geometry is worth considering as multiple sensors could be used on a bogie-only system to improve accuracy.

A journal paper by Bocciolone *et al.* [45] focuses on measuring rail corrugation using axle box accelerometers. The RMS of the accelerations are calculated to give an overall measure of any corrugation present. A potentially useful element of this paper is that a yaw-rate gyro is used to detect curves, which could be useful in identifying position on the track. This, along with the vehicle speed, are the only methods of positional identification in this system.

Recent work by Molodova *et al.* [46] also uses axle box accelerometers, but uses them instead to detect rail 'squats', which are small imprints or pits on the rail surface. Left untreated, these can lead to major degradation of the rail. This system opts to use three single-axis piezoelectric accelerometers on each of four of the vehicle's axle boxes. Accelerometers on each axle box are configured to allow measurement of vertical, longitudinal and lateral accelerations. Train speed and GPS coordinates are also logged by the system. In the processing stage, wavelet analysis is used to detect the presence of squats. This is shown to be a very reliable detection method. A subsequent paper [47] presents a case study where the technology has been used again to monitor the condition of insulated rail joints. At these joints there is inherently a small gap between the adjoined rails, filled with an insulating material. Repeated passage of trains can cause deformation of the rail ends. The system uses the same configuration of accelerometers on axle boxes as in the previous paper, but detects degradation by way of spectral analysis. It should be noted that in both systems presented in these papers, it is necessary that the accelerometers are mounted on the axle boxes, because the defects being examined are of

very short wavelengths (< 50 mm). If accelerations were measured at the bogie level, the vehicle's primary suspension would have a filtering effect on the accelerations experienced at the axle boxes, which would limit the ability to detect defects.

Many geometry systems have been developed to fulfil the specific needs of rail operators. Ackroyd *et al.* [48] describe remote inertial measurement devices fitted to high-speed trains operating between Boston and Washington DC in the USA. These were developed in order to comply with U.S. federal track safety standards which mandate that for trains travelling above 125 mph, daily measurements of the body and bogie accelerations must be taken. The system uses three accelerometers. Two of these measure car body acceleration in the lateral and vertical directions, and the third measures lateral acceleration of the bogie. Signals from these sensors are collected along with GPS data onto an embedded PC. Software on the PC generates alerts when measured accelerations exceed pre-determined limits, which are then transmitted via a cellular phone network connection. This is another example of a bogie-mounted inertial measurement system, but does not measure vertical displacements from the bogie, only lateral accelerations. It is useful however, to note the methods used to collect data on-board, and to automatically generate fault alerts.

King [49] looks at AEA Technology's products used to measure track geometry on the Channel Tunnel Rail Link (CTRL). The system, which is attached to a Eurostar train, is capable of determining the vertical track profile at speeds of up to 300 $\mathrm{km\,h^{-1}}$. Measurement is performed by the use of accelerometers mounted on the bogie above axle boxes. Displacement transducers are installed vertically between the bogie and the axle boxes themselves (*i.e.* across the primary suspension). Data is processed on board, and relayed via a cellular network connection. This connection can also be used to remotely configure settings of the logging system. The system developed in this thesis does not require the short wavelength information provided by the use of displacement transducers, however, the use of four separate vertically-sensing accelerometers on different parts of the bogie highlights the fact that different areas of the bogie experience different accelerations

depending on the vertical positions of each of the wheels. This is an effect considered later on in this thesis.

Both of the previous two systems have the capability to process data on board, detect threshold exceedances, and report these using cellular network connections. This is of importance if the role of the system is to report safety critical problems in a timely manner. Systems used only for identification of maintenance requirements do not need a transmission system such as this. Simply recording the data on-board the train and transferring it off at a later time will suffice.

A number of manufacturers have been producing off-the-shelf track measurement systems for several decades. The Plasser & Theurer EM 250 Track Recording Coach and EM-SAT 120 survey car are two off-the-shelf track measurement systems, used by the ÖBB in Austria [50]. These are commercially available alternatives to some of the bespoke solutions. The main advantage of both systems are the significantly smaller number of crew required for track inspection compared to manual inspection. Data from inspections by the EM-SAT 120 may be used to automate tamping machines, which saves further manual work. The EM 250 can operate at 250 $\mathrm{km\,h^{-1}}$; significantly faster than the EM-SAT 120 which can only operate at 2.5 $\mathrm{km\,h^{-1}}$. Therefore the EM 250 is used for large area track analysis, so the EM-SAT 120 need only be operated in areas identified as needing maintenance.

A Canadian company, Andian, have produced a number of geometry measurement systems for both the North and South American markets. Their main commercial product is their SolidTrack system [51], which is fitted to a high-railer SUV. This measures a number of parameters of the rail such as curvature, gauge, cant, alignment and vertical profile, which can all be monitored in real-time from a laptop on-board the vehicle. Recorded data is tagged with GPS coordinates. Canadian and US patents for the measurement devices [52, 53] suggest that the system uses a novel setup where a gyroscope is mounted within a double-gimbal arrangement. A spin motor is used to create an inertial force.

The grade and cant of the track cause changes in the motion of the gimbals, detected by the gyroscope.

McAnaw [54] describes the use of two off-the-shelf accelerometer-based systems, Macminder and Mactrack, on the London Underground on the Central Line. These are both standalone inertial measurement devices. Macminder is designed for measuring ride quality, and Mactrack for track measurement. Mactrack is largely the same system with support for external sensors mounted on axle boxes to be connected. The systems display live data, as well as recording data which can be reviewed using software such as VTAS [55]. McAnaw also looks at other technologies which have been used on the London Underground, including a system called Automated Video Inspection (AVI). This consists of four cameras pointed at the rail, and is purely an inspection system allowing the operator to monitor the state of the track in real-time or as a video recording. For the purpose of video recordings, alignment between AVI images and Mactrack is crudely achieved by the addition of an extra camera, pointed at the live output of the Mactrack recorder. In recent years, London Underground have been using a UGMS system developed by MERMEC [56], which has the advantage of being fitted to in-service trains requiring no disruption to normal service. This is an important factor, as on many lines of the underground network, services are being operated with very little headway between trains [20] due to ever-increasing demand.

Some measurement systems used by other rail operators have been developed in-house, rather than using commercially available systems. A paper from 1984 [57] reviews the first incarnation of British Rail's High Speed Track Recording Coach (HSTRC), and the measurement techniques used on-board. It uses a combination of techniques to measure the rail. Vertical displacements are measured by accelerometers, and linescan cameras are used to monitor the horizontal profile. Additionally, gyroscopes are used to measure curvature and crosslevel. A measurement system such as this has the capability of measuring many features of the track, giving a very in-depth profile of the track structure.

Similar Track Recording Coaches are still in use today as part of Network Rail's New Measurement Train (NMT), albeit now in a much later incarnation. An article [58] provides an insight into the technologies used on the NMT. The NMT itself comprises of a pair of Class 43 power cars and a number of interchangeable specialised carriages, including a Track Recording Coaches (TRCs), generator cars, a mess carriage, and a conference carriage. Other carriages include systems for overhead power line monitoring, rail inspection cameras, and GSM-R radio network monitoring. The train can also serve as a testbed for other technologies, including a system developed at the University of Birmingham [59] for measuring air pressure at points on the vehicle for aerodynamic testing. The vehicle is capable of measuring all aspects of the track geometry at speeds of up to 200 $\mathrm{km\,h^{-1}}$. Positioning is done through a combination of GPS, wheel tachometers, and by detecting track-mounted electromagnets used in the Automatic Warning System (AWS), which is part of the British signalling system. Because of its good level of accuracy, track geometry data from the NMT have been used in this thesis as reference measurements for verification purposes.

Kolbe *et al.* [60] provide an insight into Deutsche Bahn's (DB) Railab track measurement system. This, along with other measurement technology, is installed into two carriages of the same type used in the InterCity Express (ICE) trains. The carriages are hauled by a pair of ICE TK1 locomotives, so that measurements are representative of those which would be experienced on other ICE trains on the network. A number of methods of measurement are used on the ICE-S instrumented train. Accelerometers are used to measure forces experienced by the train, and laser sensors are used to determine horizontal and vertical distances between the rail and a bogie-mounted measurement frame. Finally, a pair of cameras are used to determine the twist of the coach. The system can measure gauge, cant, curvature, Y and Z rail profiles, and can calculate the twist and cant deficiency.

The Network Rail and DB inspection vehicles, and similar used by other rail network operators, are invaluable tools for collecting very accurate measurements of the track and other railway assets. However, the use of dedicated vehicles for measurement purposes means that their operation must be specifically scheduled into operating timetables and train crews must be employed to man them. Additionally the vehicles themselves must be regularly maintained. All of this means additional cost for the network operator. As a result there is a growing focus on in-service measurement, meaning that ordinary revenue-earning trains perform the measurement themselves, saving extra train crews, track time and maintenance costs.

A system being developed concurrently with the work done for this doctorate [61] is used to instrument a DB freight locomotive, so that in-service measurement can be performed. This system is an academic research project rather than an industrial installation. Accelerometers are mounted on the bogies and axle boxes of the locomotive, and the body is instrumented with a 3-axis gyroscope. Tachometer data is obtained from one axle, and GPS data is gathered as well. A cellular phone connection is used to report defects. In this work, consideration is given to trend analysis algorithms which are used to predict the amplitudes of specific features of the track over time, allowing estimation of when maintenance of the track will be required.

Track measurement systems developed at the University of Birmingham have been evolved over a number of years. Research presented at the IEE Railway Condition Monitoring conference 2004 [62] looks at a series of experiments performed on the Tyne and Wear Metro, and also considers the effectiveness of different sensors and placements, including accelerometers, gyros, and displacement transducers between the bogie and axle boxes. This publication identifies the use of a bogie mounted pitch-rate gyro as an effective measurement source for deriving vertical displacements along a track. It also looks at the use of versines as a method of fault detection. Finally it suggests that displacement transducers could be removed from the system to create a more robust installation. A later

publication [63] investigates the possibility of reducing the number of sensors required, and describes trials performed on a Class 175 DMU trainset, which used a self-contained IMU capable of measuring three axes of acceleration and three axes of rotational velocity. No external sensors were required. The paper also examines the use of state estimation using a Kalman filter to trigger fault alarms. Two further publications [64, 65] examine in detail methods for deriving various aspects of track geometry. These methods involve the derivation of curvature in order to measure geometry in the spatial domain.

A later paper [66] considers in detail how instrumentation can be mounted on a train, including detail on mounting inertial and displacement sensors on various parts of bogies, and provides sample outputs from these sensors. This provides a good insight into the practicalities of mounting sensors, whilst considering the some of the advantages and disadvantages of using the sensors.

As seen in some of the other publications reviewed, axle-box accelerometers are useful for measuring short wavelength defects in the rail. The last paper [66] also suggests that laterally mounted axle-box accelerometers can be used to measure short wavelength lateral defects such as poor rail alignment at joins, and switches & crossings. However, axle box accelerometers can be difficult to attach, and makes maintenance of the bogie and its components more difficult, particularly if left in place permanently.

A solution providing better opportunities for mounting are bogie-mounted accelerometers and gyros. These can be used to measure vertical and lateral track geometry, as well as cross-level and twist. Whilst an installation of sensors scattered about the bogie would be difficult to maintain, a single self-contained box containing multiple sensors can be mounted unobtrusively, and can be easily removable in case bogie removal and overhaul is required.

For a temporary installation, the most convenient location for mounting inertial sensors is on board the body of the train. This can be accessed easily through normal passenger

or driving cab areas. However, this location has the greatest effect on the wavelength of track features which can be measured, allowing only long wavelength geometry to be observed.

So far, the literature review has revealed various bespoke measurement systems. Several of these have successfully used vertically-sensing accelerometers mounted on vehicle bogies to measure vertical track geometry. This answers part (a) of the hypothesis in Section 1.4. None of the literature shows studies of track degradation by using measurement systems to monitor changes in track condition over time. Part (b) of the hypothesis states that multiple measurements of the same track taken at different times could be aligned and used to assess track degradation. This appears to be an area of research where there little or no work has been done. It would be beneficial to look at rates of degradation of track to gain a better understanding of how faults develop, and potentially improve decisions about the type of maintenance required to correct such faults and to continue to maintain the track. Therefore it is worth pursuing work in this area.

2.2 Stiffness Measurement

Part (c) of the hypothesis in this thesis states that an in-service inertial measurement system could be used to measure track stiffness. A number of methods of stiffness measurement technologies exist. As stiffness measurement is usually performed by measuring the deflection of the track between an unloaded and loaded state, many stiffness measurement systems are trackside-based in order to be able to measure the level of the track with no train present.

A system used by the University of Southampton involves the use of geophones, which are temporarily glued to sleeper ends in order to measure the vertical movements of sleepers during the passage of a train. Priest *et al.* [67] describe how these measurements can be used to determine track stiffness. Some collaborative work was later done by the author

with Southampton to investigate vertical track movements at a level crossing approach [8]. This journal paper includes some of the work done towards this thesis.

An earlier system developed by the University of Southampton [68], uses Digital Image Correlation (DIC) where a video camera is used to monitor the positions of 'targets' affixed to components of the track as a vehicle passes by. This system is also used as part of the study in [8]. By detecting the positions of specific patterns on the targets in each frame of the video, the system is able to measure the displacement of the track over time as a train passes. A potential problem with a system like this is camera shake caused by a passing train, causing the system to perceive that all of the targets are moving.

A comparable system described in [9] also uses a DIC video inspection technique, but without the need for the visual targets. In fact it is claimed that the system works better using the natural patterns occurring in the materials making up the track, such as the shape of a rail clip. The system identifies patterns, and detects their movement down to 1/100th of a pixel. This means that the resolution of movement detectable by the system is dependent on the resolution of the camera, and the distance from the track being examined. In the given example using a 1024×768 pixel camera observing a 1.1 m span of track, produces a resolution of 0.011 mm. Recent developments in this system allow tracking of components in three dimensions.

Another trackside system developed at the University of Birmingham [10, 11] uses a fixed line laser and a number of interconnected node units attached to sleeper ends. Each node contains a vertically mounted position sensitive detector (PSD) which measures the height of the laser beam relative to the sleeper. This allows high-resolution monitoring of the displacements of up to 20 sleepers simultaneously. This can suffer from the laser source being shaken by a passing train, but this is compensated for in software.

The systems reviewed so far have all been trackside-based. These provide a useful insight into the techniques used to derive track stiffness, but are unable to provide continuous

stiffness measurements along the length of a track. This would require the use of a vehicle-based stiffness measurement system.

Arguably the most common vehicle-based stiffness measurement device is the Falling Weight Deflectometer (FWD). Originally a device developed for testing the stiffness of roads and pavement, the design has been adapted for use on the railway. Often packaged in a trailer towed behind a road vehicle, the device raises and drops a heavy weight whilst measuring the vertical displacement experienced by a sleeper. The rails are unclipped from the sleepers before the measurement takes place. A presentation given at the Railway Track Science & Engineering Workshop [69] provides a good example of the use of a FWD, whilst a paper [70] looks at some of the calculations used to determine the stiffness of the track from the measurements taken. A FWD must be moved to one location and stopped, then a test is done, and it is then moved to the next location. The device used in [69] is able to perform one test per minute. This means that it can be slow to measure stiffness along a long section of track, and continuous measurement is not possible.

An Innotrack report [12] considers a number of vehicle-based track stiffness measurement devices. Two of these are based on the same principles as the FWD, using an oscillating mass to create varying force on the track, whilst displacement is measured. The Rolling Stiffness Measurement Vehicle (RSMV) [71, 72] is a rebuilt 2-axle freight wagon. A mass is oscillated above one of the axles, and the resultant track deflection is measured. As the mass is continuously oscillated, measurements are also continuous, so the RSMV is able to operate continuously at speeds up to 50 $\mathrm{km\,h^{-1}}$ for basic measurements. More detailed analysis is possible below 10 $\mathrm{km\,h^{-1}}$. The second system, called a Railway Portancemetre [73, 74], works on a similar principle, except a heavy wheelset is used, and the axle itself is oscillated whilst displacement measurements are taken. This is a more compact system capable of operating at 6 $\mathrm{km\,h^{-1}}$.

Another approach to the continuous measurement of track stiffness was developed by the China Academy of Railway Sciences, and was one of the first to achieve continuous

stiffness measurement. The system uses a pair of instrumented railway cars, one lightly loaded and one heavily loaded. Each car has small non-load-bearing wheels either side of a load-bearing wheel. It uses the displacements of the smaller wheels relative to the car body to measure the track deflection using a chord measurement technique. The system is effective at speeds up to 60 km h^{-1}. This system achieves continuous vehicle-based stiffness measurement, but requires a dedicated train to do so.

A system by TTCI in the USA, uses a similar principle to the Chinese system, instead using a laser and camera based method of measuring deflection. This system is mounted on both heavy and light wagons, allowing the track stiffness to be determined.

A third system based on this principle was developed by the Swiss railways (SBB) [75]. This uses a box wagon as the heavy vehicle, and a lightweight frame with small wheels as the light vehicle.

Systems such as these are effective at measuring stiffness continuously from a moving vehicle. However, whilst it is possible to fit a measurement system to an in-service vehicle, it is highly unlikely that a lightweight instrumented wagon would be approved to be towed as part of an in-service train. There are many reasons for this, such as safety, the speed limit that would have to be imposed on a train towing the wagon, and the impracticality of having to shunt the wagon onto the rear of the train when required.

A system which has been used for stiffness measurement on the Swedish railways [76] is adapted from a commercially available track recording vehicle, the Plasser & Theurer EM 80. The vehicle has a mechanical three-point chord measurement system. The EM 80 vehicle was rebuilt and renamed the Infranord Measurement Vehicle (IMV) 100, and a modern inertial measurement system was installed to measure track geometry, whilst the existing chord measurement system was kept. Using a combination of the two systems, the IMV can be used to determine track stiffness.

All of the vehicle-based solutions reviewed so far have required dedicated measurement vehicles to be used.

Another different method of stiffness measurement was developed at the University of Nebraska at Lincoln (UNL) [77, 78]. This uses a pair of crossed flat-beam lasers mounted on an arm attached to the bogie of an in-service freight wagon. A video camera observes the lines projected by the lasers onto the rail head. Through measurement of the distance between the laser lines, the deflection of the track can be inferred. This system is capable of determining changes in track stiffness as the wagon traverses the track, but the accuracy of absolute values determined by the system is unknown. The system is capable of working at around 65 $\mathrm{km\,h^{-1}}$. This seems to be a good approach to stiffness measurement, but requires a lot of space underneath the wagon, making it unsuitable for instrumenting commuter trains, which have limited space underneath them.

A conference paper by Network Rail and TRL, UK [79] addresses the struggle by Network Rail to find a method of measuring track stiffness at line speed. It also highlights the fact that when poor track geometry detected at a site where there is poor track stiffness, the measurement is of the symptom of the fault, not the cause. Consequently, if maintenance is carried out to correct the geometry fault, it does not solve the problem. The paper finds two systems with potential for use. Firstly a laser-based measurement of vertical velocity response, which is used when measuring the stiffness of roads in the UK. The second method would use laser displacement sensors which are used on the NMT and other measurement vehicles, to calculate track stiffness. The research opts to investigate the first method and does not consider the second.

The research reviewed in this section has shown several different methods of measuring stiffness from a rail vehicle, but none which use inertial measurements and which are capable of fitting to an in-service train. One of the systems reviewed [77, 78] was fitted to an in-service freight wagon, but is very bulky and would not be possible to fit to a UK passenger train, where under-carriage space is limited. The system also uses optical

sensors which would likely require frequent cleaning in the under-train environment. The hypothesis in Section 1.4 states that repeated inertial measurements of a track would be able to be used to provide a measurement of track stiffness. This has not been found in the available research. An in-service stiffness measurement system would allow problems with track support to be identified before they become significant geometry faults, which could stop railway traffic during service hours. Knowing the root cause of a geometry fault could also potentially aid in a more appropriate choice of maintenance. It is therefore worth pursuing further research in this area.

2.3 Conclusions from Literature Review

Many of the track measurement systems reviewed in Section 2.1 gather basic measures of the quality of the track using accelerometers or other motion sensors mounted on-board the vehicle body. This is an acceptable method of track quality measurement, where the exact geometry of the track does not need to be reconstructed. This is also the best way to gauge "ride quality", as the sensors measure movements similar to those experienced by passengers on board the train.

In order to reconstruct the geometry of the track, a better approach when using inertial sensors seems to be to install the instrumentation on the bogie or axle box of the vehicle. This moves the measurement point to below the secondary suspension. Often the primary suspension is very stiff, meaning that measurements taken on the bogie are close to those experienced at axle box level. Instrumentation of the axle box appears to be the best way of detecting very short wavelength features such as corrugation and squats. Bogie instrumentation is appropriate for providing measurements mid-wavelength geometry features, as the primary suspension and bogie itself provide a natural low-pass filtering of the accelerations experienced. Part (a) of the hypothesis states that an IMU on the bogie of an in-service vehicle could be used to assess track condition. The literature reviewed proves that this is possible.

Some of the systems reviewed have optical measurement systems, using either laser measurement devices or a combination of lasers and video cameras. Such systems can provide very accurate measurements (to within 0.1 mm) of the track geometry, but could be more expensive to maintain in an in-service scenario because they require frequent cleaning in order to achieve continuous operation.

Part (b) of the hypothesis states that IMU data from repeated passes of the same track sections could be used to monitor track degradation. No research was found where this has been done. The ability to do this would allow degradation rates to be monitored, and could provide a better understanding of how track faults develop over time. It was decided to continue with research to answer this part of the hypothesis.

Most of the vehicle-based stiffness measurement systems reviewed in Section 2.2 require the use of dedicated measurement vehicles. Because of the need to use bulky oscillating masses, or specific wagons of differing masses, these systems would not be possible for fitment to in-service vehicles, unless specifically designed wagons were hauled as part of a freight train. This could consequently reduce the maximum cargo load of the train. A system reviewed from UNL is fitted to an in-service freight wagon, but due to the use of optical components would require frequent cleaning due to the buildup of dirt from the track, and brake dust produced by the train under braking. The system also uses a bracket mounted from the bogie of the wagon. For many British and European trains, components of rolling stock undercarriages tend to be situated very close together in order to maximise the use of space. It could therefore be difficult or impossible to fit such a device. The scope of use of in-service measurement systems would ultimately to be to fit them to vehicles network-wide, meaning that measurement systems would need to be small enough to fit to locations on most multiple-unit vehicles.

Part (c) of the hypothesis states that data from multiple passes of the same track could be used to determine track stiffness. A method of measuring track stiffness using inertial sensors on an in-service train would be beneficial because it would remove the need for

bulky stiffness measurement devices, and lessen the need to run dedicated vehicles to measure these parameters. Ultimately this would lead to cost savings, both directly from reduced running time of dedicated measurement vehicles, and from increased availability of the railway to revenue-earning vehicles and potentially through more effective maintenance as a result of continuous stiffness measurements being available for entire routes of the railway network.

In Section 1.3.1 three potential methods of using inertial sensors to determine track stiffness were identified. Of these, two fitted the specifications of the system developed in this thesis. Inertial data from repeated passes of the train over the same track sections could be compared when the train has different loads on board, or when the train is operating at differing speeds, affecting the dynamic load on the track.

Research into existing vehicle-borne stiffness measurement systems in this chapter has not revealed any in-service systems using inertial methods. This, coupled with the benefits discussed previously, makes this an area which would benefit from further research.

CHAPTER 3

IN-SERVICE INERTIAL MEASUREMENT SYSTEM

3.1 System Specification and Practicalities

The Birmingham Centre for Railway Research and Education (BCRRE) at the University of Birmingham has an existing measurement system [80, 81], which is installed on one of Southern Railway's Class 377 'Electrostar' EMUs (Figure 3.1). The existing system measures the height of conductor rails using laser measurement devices. It was agreed that this system could be expanded to include an inertial measurement system for the purposes of track condition monitoring and ride quality measurement.

3.1.1 Inertial Measurement System Requirements

Many of the requirements for the inertial measurement system were based on knowledge gained from previous projects at the University of Birmingham of a similar nature. A joint paper with other UK universities [66] considers in detail the possibilities for instrumentation of vehicle bogies. A University of Birmingham thesis [24, ch. 5 & 6] provides an insight to two inertial measurement systems on both a Class 375, and on a Merseyrail Class 508. From these, some of the requirements for this system were drawn. Other requirements were stipulated by the train operating company. Two systems are

Figure 3.1: A Southern Class 377 'Electrostar' EMU

described in a pair of journal papers [64,65]. These were fitted to a Tyne and Wear metro vehicle, and a Class 175 DMU.

The list of requirements for the system to be fitted to the Southern Electrostar are listed in Table 3.1.

Requirement	**Justification**
One IMU to be mounted on the vehicle bogie	Previous work has identified this as an acceptable location for measuring track geometry. This was agreed with Southern as a non-obtrusive position for operation and maintenance.
One IMU to be mounted beneath a passenger seat	This is an acceptable place to measure ride quality. Agreed with Southern to be a non-obtrusive position.

Both IMUs to measure six degrees of freedom	As this is an experimental system, three degrees of acceleration and three degrees of rotation should be measured, so that geometry in each direction can be reconstructed if necessary. If the system were to be commercialised, only a subset of these degrees of freedom may need to be measured.
Accelerometers must have a range greater than ±100 ms^{-2} for bogie, and ±10 ms^{-2} for body	These are the maximum expected range of accelerations for these locations [82].
Gyros must have a range greater than ±10 $°s^{-1}$ for bogie, and ±1.25 $°s^{-1}$ for body	Maximum values observed from gyro data collected during previous research work [65]. For body rotations, the vertical velocities of the two bogies supporting the body are ~20 m apart rather than ~2.6 m for the bogie wheelbase. This means that the body rotations are around 8 times lower than the bogie (before accounting for secondary suspension). Off-the-shelf gyros usually measure in ranges of at least ±100 $°s^{-1}$, so these specifications lie well within expected availability.
IMU to output via RS-422 at 115200 baud	RS-422 is a suitable transmission protocol for noisy environments, and a baud of 115200 will allow transmission along a cable of up to 10 m (from bogie to body end cupboard) without degradation of the signal.

IMU to output samples at 250 Ss^{-1}	When data is processed in the spatial domain, a sample 'rate' of 8 Sm^{-1} is desired for accurate representation of geometry features as small as 1 m long. 250 Ss^{-1} gives 8 Sm^{-1} at approximately 70 mph (31 ms^{-1}), only requiring upsampling above this speed. This fits within the bandwidth (8 channels × (16 bits/sample/channel + 4 for RS-422 overhead) × 250 Ss^{-1} = 40000) and allows for packaging bytes to be added.
IMU sensors to be sampled at 8000 Ss^{-1} and downsampled internally	To minimise aliasing effects of high-frequency components of the sensor signal, including the gyro resonating ring frequency of 28 kHz, whilst allowing the output to fit within the available bandwidth.
GPS location to be measured	To allow geometry features to be located to within the accuracy of the GPS.
GPS receiver to provide samples at at least 1 Ss^{-1}	More frequent samples are not required, as position is approximate. 1 Ss^{-1} is most common among off-the-shelf GPS receivers.
Vehicle wheel tachometer to be measured	Wheel tachometer data to be used to measure distance travelled to improve inertial data positioning.
Tachometer to be sampled synchronously with inertial measurements	Allows precise alignment of position for every sample.
Data to be stored to removable media for retrieval	Large amounts of data will need to be retrieved from the system regularly.
Bogie IMU must be able to withstand the harsh bogie environment	The bogie is subject to high accelerations (> 10 g), water, dust, and dirt which should not affect the IMU operation.

Bogie IMU to have 1 kV electrical isolation from carriage body	A fault causing indirect contact between the bogie chassis and the 750 V DC third rail must not allow on-board parts to reach a high potential.
No sensors are to be fitted to the bogie outside of the IMU enclosure	To allow easy maintenance of the bogie itself by Southern staff.
On-board logging equipment to fit into pantograph transformer cupboard on the EMU	Equipment cannot be in passenger area for safety. The Class 377/4 does not have a pantograph, so the cupboard is empty and may be used for instrumentation.
The system must operate from the 240 V AC power supply	This is the supply available on-board the train.
IMU system repeatability to be within 0.2 mm from run to run	This was the value identified from initial calculations in Section 1.3.1. This will provide a possibly of estimating track stiffness.
System to be verified against an existing known-working measurement system	It is most important that the system provides repeatable results, but it should be verified against another system to check that its output is valid, and that the results can be relied upon for assessing track condition.

Table 3.1: Inertial measurement system specifications

A bespoke designed system was opted for in order to satisfy all of the requirements. The IMU designed at the University of Birmingham for the Merseyrail project [65] had to meet stringent safety requirements, meaning it already meets several of the requirements, however it has several limitations. The first of these is that it stores inertial data to an SD card inside the bogie IMU, making retrieval of stored data difficult. Secondly, the GPS antenna is also located in the IMU unit, meaning that visibility of GPS satellites is very limited, resulting in poor estimates of position. For these reasons, a new design was

produced, based on the original, but with modifications where necessary. These are listed in Section 3.2.1.

Many of the systems reviewed in Chapter 2 use accelerometers mounted on axle boxes [42–44], which has the advantage of providing accurate short-wavelength geometry results. However, the use of many external sensors such as these mean that the vehicle bogie becomes more difficult to maintain. Mounting a single IMU directly onto the bogie frame results in a much simpler installation process, as well as making it easier to maintain the running gear in the long run. A bogie-mounted IMU gives acceptable measurement results, particularly when the exact geometry does not need to be reconstructed [65]. It is suitable for measurement of mid-wavelength (> 1 m) geometry features. Instrumentation mounted in this way is better suited to in-service installation, as it is less obtrusive and easier to remove and re-attach when the bogie requires maintenance.

Using a bogie-mounted IMU, the very short wavelength geometry features are filtered out by the primary suspension, however, in this scenario, the very short wavelength features are not required. Some short wavelength defects are filtered out by the dynamic effects of the bogie, which could result in some loss of fidelity. However, processing allows some of these features to be reconstructed (Chapter 5). The point of measurement is before the secondary airbag suspension between the bogie and the coach of the train, allowing much shorter wavelength measurements to be recorded than if the IMU were mounted on board the coach. The ride quality, however, is best measured from on board the car itself, as this is where passengers would be sitting or standing. The ride quality IMU is therefore mounted on board in a void beneath one of the seats.

There is a European standard for measuring ride quality [83] which includes the positioning of sensors. However, as this part of the measurement system was purely to give an approximate ride quality value, and was to be a temporary installation on an existing vehicle, the positioning was chosen as a compromise between the ideal measurement position, and ease of installation. The ride quality measurement system uses only a relatively

simple calculation of the RMS value of the accelerations to indicate the 'roughness' of the ride. The measurements from the body mounted IMU are not used in further work in this thesis, so greater detail is not provided here.

3.2 Inertial Measurement Unit Design

3.2.1 System Design

As mentioned before, the design of the IMU for this research was based on a previous IMU designed at the University of Birmingham for a Merseyrail train [80,81], rather than designing the unit from scratch. The primary advantage of this is that many of the system specifications are identical to the ones listed in Table 3.1. The Merseyrail system also had to undergo safety testing before installation, and it is therefore already known that the mechanical and electrical designs are suitable for the railway environment.

The new design does not require the use of an SD card, as used in the Merseyrail design. Instead the inertial measurement data is transmitted from the IMU directly to an industrial PC on-board the train. Here, incoming data will be logged and timestamped along with positional data acquired from the GPS and other instrumentation.

The Merseyrail system transmits tachometer data from systems on-board the train to the bogie mounted IMU. The tachometer data is stored with the inertial data on the SD card. The new system, however, will accept an input directly from a sensorised wheel bearing. Information from the tacho will be transmitted to the on-board PC along with the inertial data.

The overall system diagram is shown in Figure 3.2. Components belonging solely to the third-rail measurement system are greyed out for clarity.

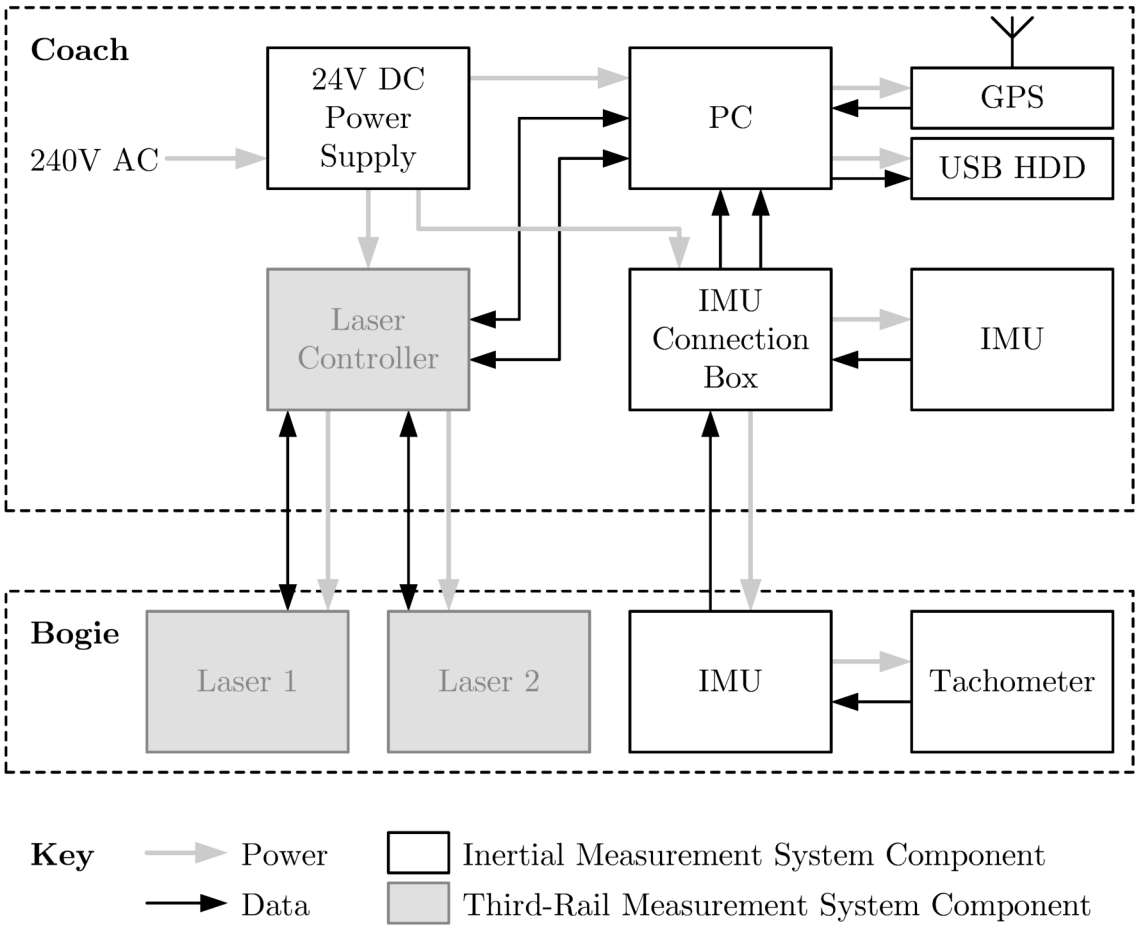


Figure 3.2: System diagram for inertial measurement system

3.2.2 Electronic Design

The IMU electronics is split into two printed circuit boards (PCBs). The analogue board provides power to the sensors, and conditions the outputs of the sensors before passing the signals on to the digital board. This board samples the analogue sensor signals via an analogue-to-digital converter (ADC) onto a microcontroller, where they are buffered and then transmitted to the PC on-board the train. Figure 3.3 shows the functional sub-sections of the IMU.

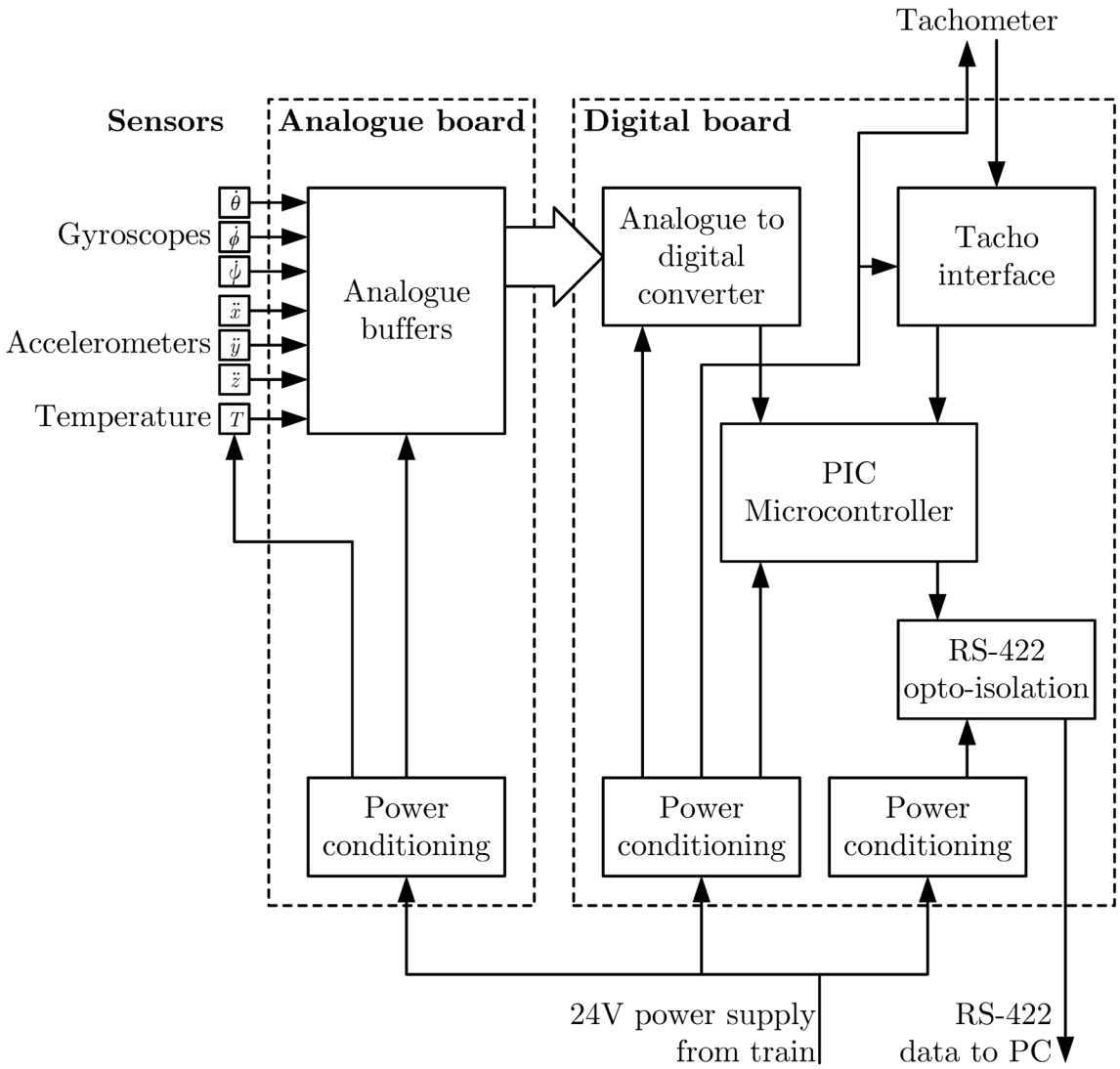


Figure 3.3: Functional block diagram of the IMU

It was decided to use the on-board PC's RS-422 ports to receive data from the IMUs. The digital board generates a serial data stream containing the sampled data. The signal is passed through a level converter, which converts it to an RS-422 compatible signal. RS-422 is suitable to be transmitted over a reasonably long distance and in harsh environments (*i.e.* where there is a lot of electro-magnetic (EM) noise) because it is a differential line system using a twisted pair [84]. This makes it suitable for communications between the IMU on the bogie of the train and the PC in the car of the train.

After completing a prototype IMU, a small test application was written in C#, which was used to receive and interpret the serial data transmitted by the IMU. Tests were performed to prove the functionality of the device. Some minor modifications were made to the prototype to correct its operation before a corrected PCB design was created, and two copies to be used on board the train were produced. These boards were mounted into frames with the sensors mounted in appropriate orientations to measure the three axes of rotation and acceleration. This is shown in Figure 3.4.

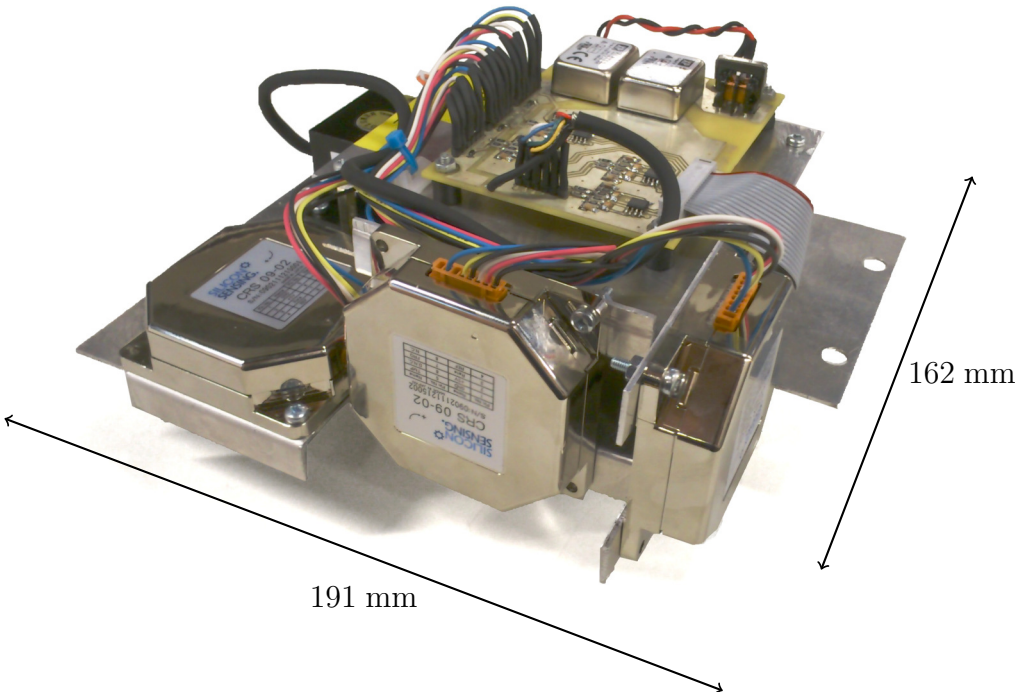


Figure 3.4: A completed IMU sensor frame

The IMU samples its sensors at 8000 samples per second per channel. These samples are averaged together in blocks of 32 samples, resulting in a downsampled output rate of 250 samples per second. The sensors also report their temperatures to allow compensation for variations in their outputs due to temperature changes. Only one of the sensors' temperatures is recorded, as other temperatures are likely to be very close to the recorded one. Pulses generated by the tacho are constantly counted by the microcontroller, and the tacho count is transmitted with each sample from the sensors. This results in a total of 8 channels of data being transmitted at a rate of 250 samples per second. The ADC samples each of the seven analogue channels (3 accelerometers, 3 gyros, 1 temperature) at a 16-bit resolution, and the tacho count is recorded as a 16-bit integer. The tacho counter wraps around to zero after it reaches its maximum value of 65535. This wrap-around transition is dealt with in interpreting software.

The tachometer itself (see Figure 3.5) is an SKF sensorised wheel bearing, which generates 108 pulses per revolution of the axle of the train, using a magnetic disc. Assuming a standard wheel circumference of approx 2.5 m, if the train travels at its top speed of 100 mph (44.69 ms^{-1}), the tacho will generate just under 2000 pulses per second. It should be noted that the wheel circumference can vary depending on factors such as wheel wear and maintenance which can affect its size. This effect is considered in Section 4.2.3.

3.2.3 Logging Software

Data arrives at the on-board PC from multiple sources. These are two IMUs and a GPS receiver, as well as data from other instrumentation used for the third-rail monitoring system mentioned previously.

Figure 3.5: Sensorised wheel bearing used on a Class 377 EMU

The PC runs a Linux-based operating system. Software was developed to run continuously on the on-board PC, which acquires data from the two IMUs, GPS receiver, and other equipment. Data is timestamped before being stored to disk. Data is stored both locally on the PC (as a backup), and onto an external USB hard disk which allows retrieval of the data.

The software starts once the PC has booted up, and continues to run until the PC is powered off. During the boot up stage, the PC's clock is synchronised with the clock provided by the GPS receiver, which ensures that timestamps inserted into the data files remain accurate.

Every few months, the external USB disk is removed from the train by depot staff and swapped for a blank disk. The disk from the train is then sent to the University of Birmingham for post-processing.

3.3 Installation

3.3.1 Equipment On-Board the Carriage

A rack frame was created by University of Birmingham staff, to be installed in the instrumentation cupboard aboard the train. This was designed to fit all of the on-board equipment on it.

The completed rack is shown in Figures 3.6 and 3.7 bolted in place inside the cupboard, with the instrumentation installed on it. The first figure shows the 240 V AC to 24 V DC power supply, the external USB hard disk in a mounting frame, the GPS receiver unit, and a connection box for interfacing the IMUs with the PC connections and the power supply. Other equipment shown in the figure is used for the third-rail monitoring system, including a laser control unit and components of a pressurised air system. The antenna for the GPS is attached to the underside of the roof of the train, above the rack and out of the camera shot. The on-board PC is installed on the lower shelf (Figure 3.7), and connections through the floor of the train to the undercarriage can also be seen in the figure.

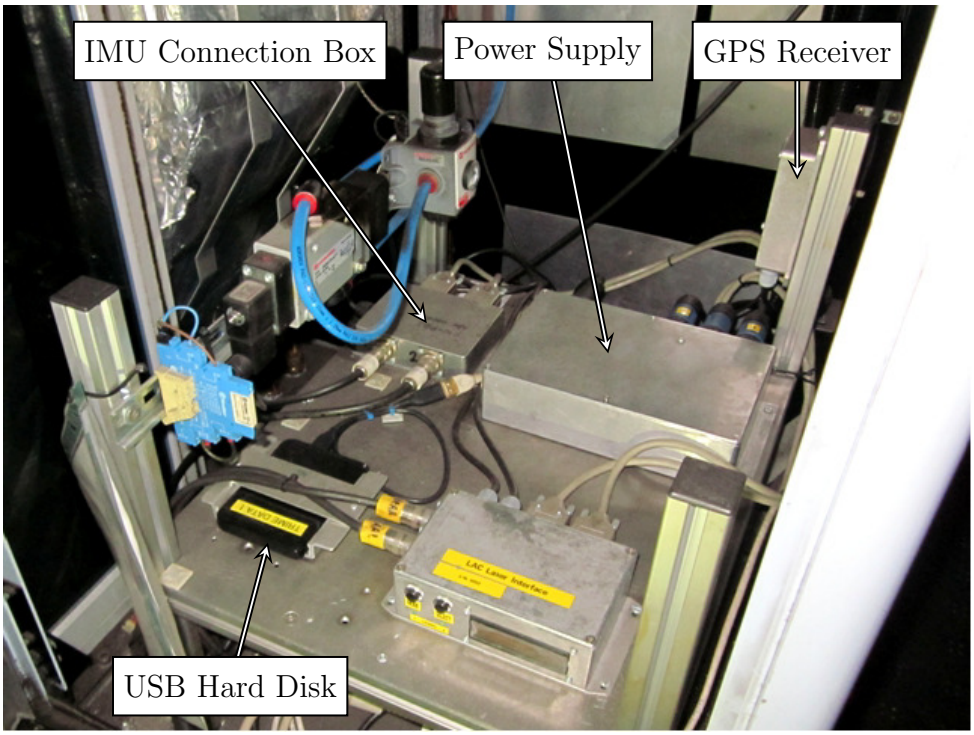


Figure 3.6: Equipment installed on-board the carriage - top shelf

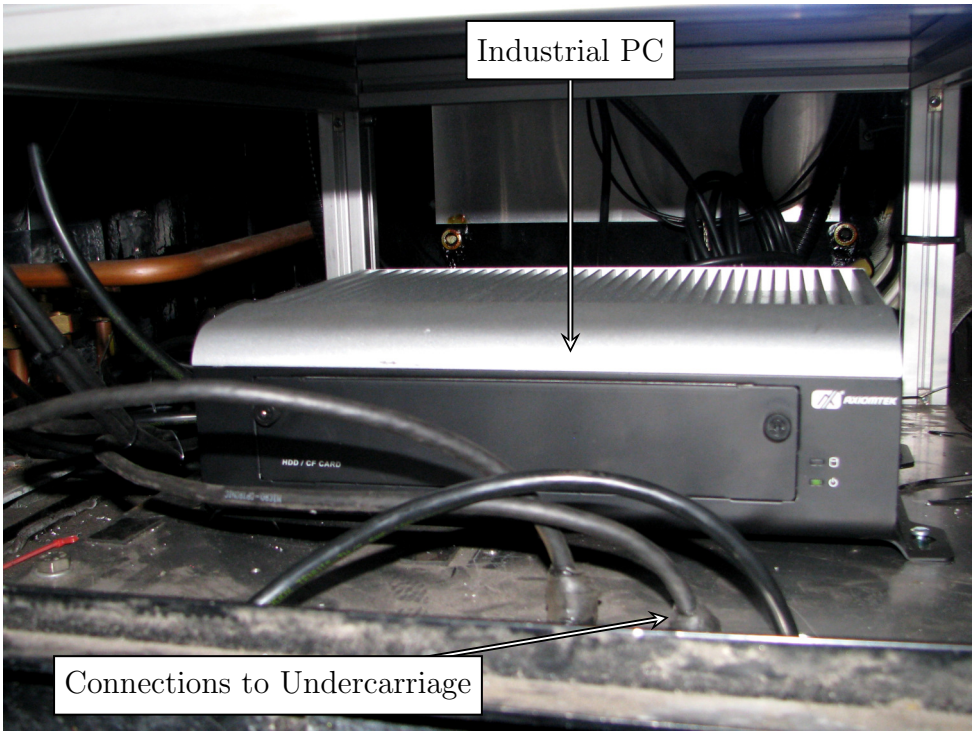


Figure 3.7: Equipment installed on-board the carriage - bottom shelf

3.3.2 Equipment on the Bogie

The IMU itself is mounted on top of the bogie, using four tapped holes provided for the addition of extra equipment to the bogie. The positioning of the IMU on the bogie is shown in Figure 3.8.

A conduit from the coach above to the IMU contains a twisted-pair cable with two pairs. These carry 24 V DC power to the IMU, and the RS-422 signal back to the coach. A second conduit connects the IMU to the tacho. These connections are shown in Figure 3.9.

Small cable-mounted connectors are located inside the conduits so as to protect against vibrations. Using panel-mounted connectors on the IMU could cause failure due to the high forces being exerted on the IMU.

Once all equipment was installed on the train, the system was commissioned and the train was returned to service.

Figure 3.8: IMU positioning on the bogie

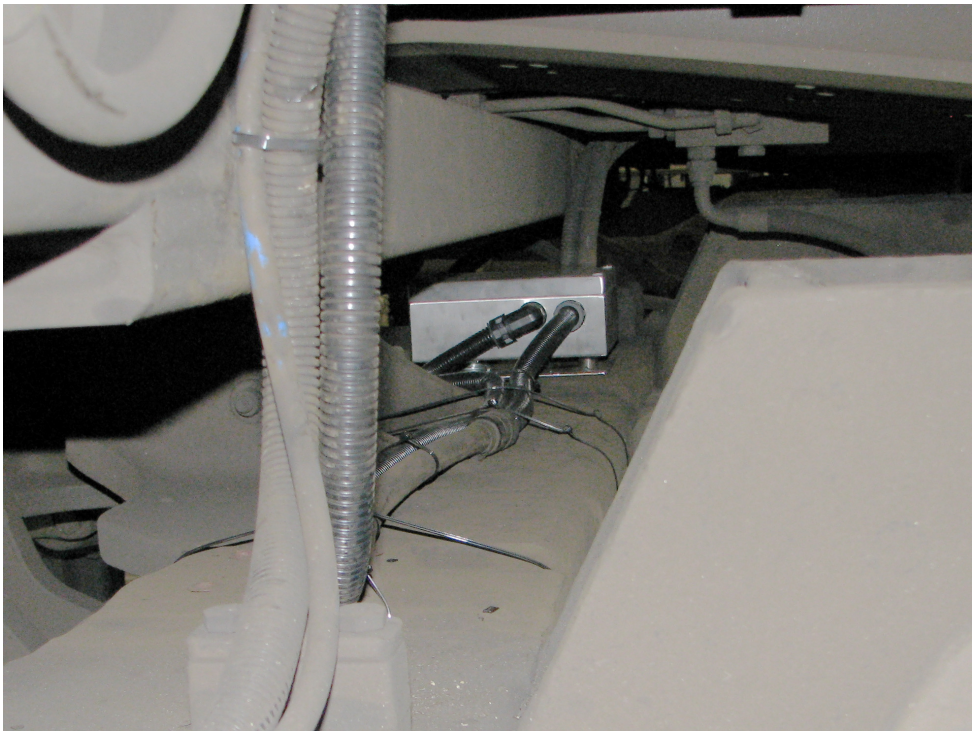


Figure 3.9: Conduit connections to the IMU

CHAPTER 4

INITIAL PROCESSING OF INERTIAL DATA

Data processing is required to transform recorded inertial data, consisting of accelerations and rotational velocities recorded at a constant rate over time, into meaningful information which can be used to assess the condition of the track, which can in turn be used to make decisions regarding maintenance.

For the purposes of this research, all data processing is performed away from the train. The computer on board the train acquires and stores sampled data onto a removable disk. This data is downloaded and processed remotely. The long-term aim of the project is to perform the processing on-board the train as data is accumulated (close to real-time), and relay degradation reports using a cellular connection. Some consideration has been given to this possibility [85], although it is beyond the scope of this thesis.

4.1 Data Management

The data logging system on board the train logs a substantial amount of data. A single IMU samples 8 channels 250 times per second, at 16 bits per sample per channel. Inertial processing also requires data from the GPS which samples once per second at approximately 500 bytes per sample. Typically the train may operate around 18 hours per day. This results in approximately 280 MB of data being produced per day for a single IMU with GPS. A method of extracting relevant data from these recordings was

required, as it would be impractical to attempt to consider such a large quantity of data at once.

4.1.1 Data Management Software

An application was developed in C#, to allow recorded data to be imported and managed, and data of interest to be exported for processing and inspection.

The application allows the following functions to be performed:

- **Import data** Allows data from a number of sources to be added to the database, including the Southern system, and historic data from the University of Birmingham's Merseyrail inertial monitoring system. At this stage calibrations are applied to the raw IMU data so that all data stored in the database use S.I. units.
- **View inertial data** Plots a graph of selected inertial data streams against time.
- **View GPS data** Plots GPS points from selected data over a map of the rail network.
- **Search by time** Find where the train was at a particular time. This is useful for fault-finding, for example when it is known that a sensor stopped working at a specific time.
- **Search by location** This is one of the most used features of the program, which allows a search by GPS location. A result is created for every pass of the train within a specified radius of a given location.
- **Data export** Selected data from the inertial plot or from selected location search results, can be exported to standard Comma Separated Variable (CSV) files or directly to Mathworks' Matlab native file format (MAT).

Figures 4.1 and 4.2 show screenshots of the software, detailing GPS data plotted on a map, and the search by location function, respectively.

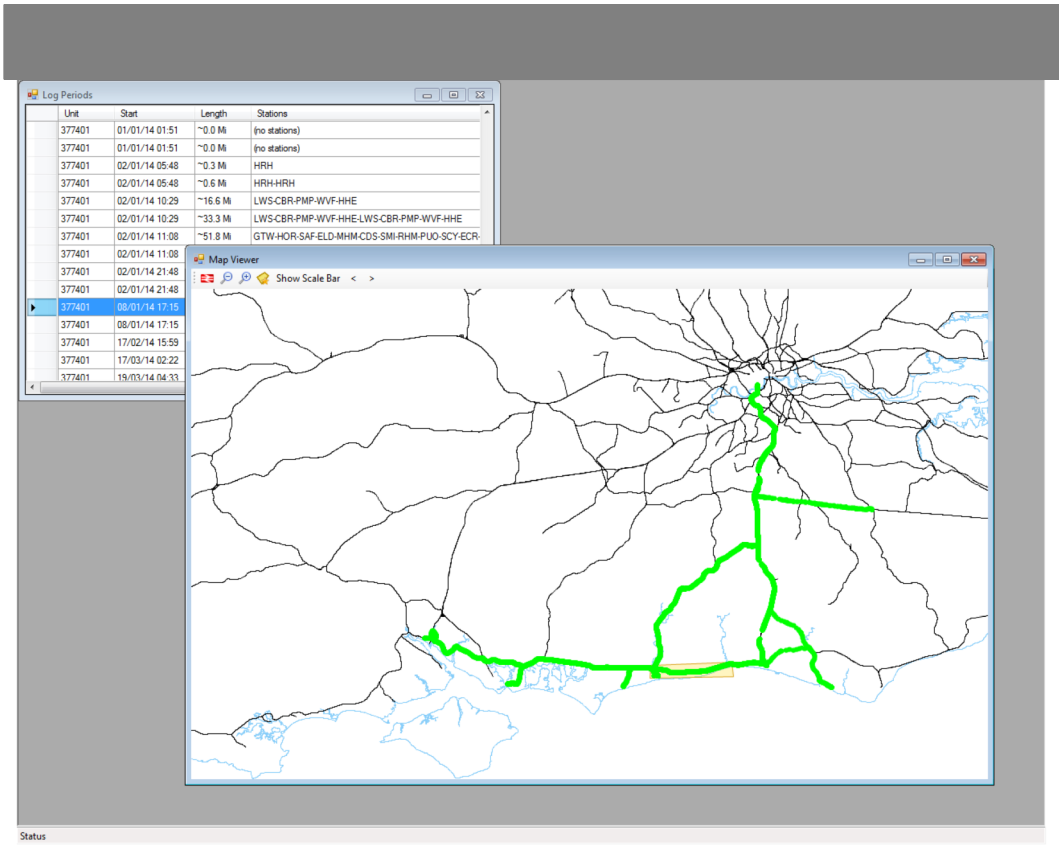


Figure 4.1: Screenshot of data management software

The foreground window shows GPS locations for a journey plotted on a map

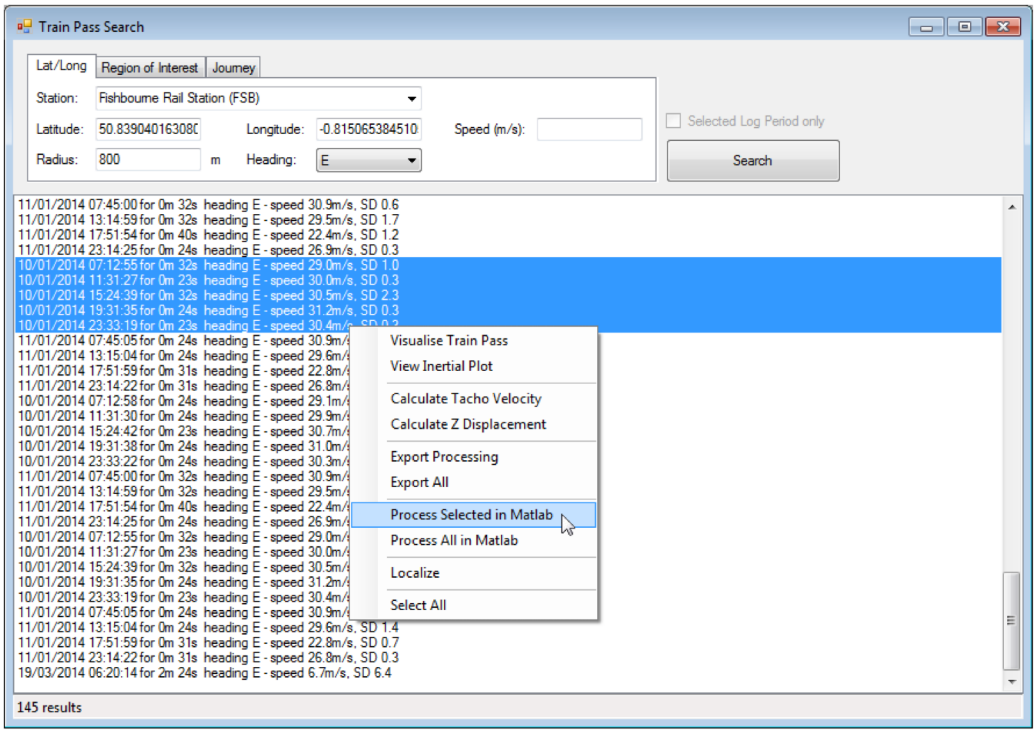


Figure 4.2: Screenshot of the train pass search window

4.2 Vertical Displacement Processing

4.2.1 Geometry Model

In subsequent work, the movements experienced by the IMU are considered in three dimensions. The six degree of freedom (6DoF) model is used (Figure 4.3), in which a body may move in 3 perpendicular axes, and may rotate about those axes. This geometry frame is oriented with the bogie of the IMU, so that the x axis is always parallel to the track. It is also oriented such that the positive z axis is downwards.

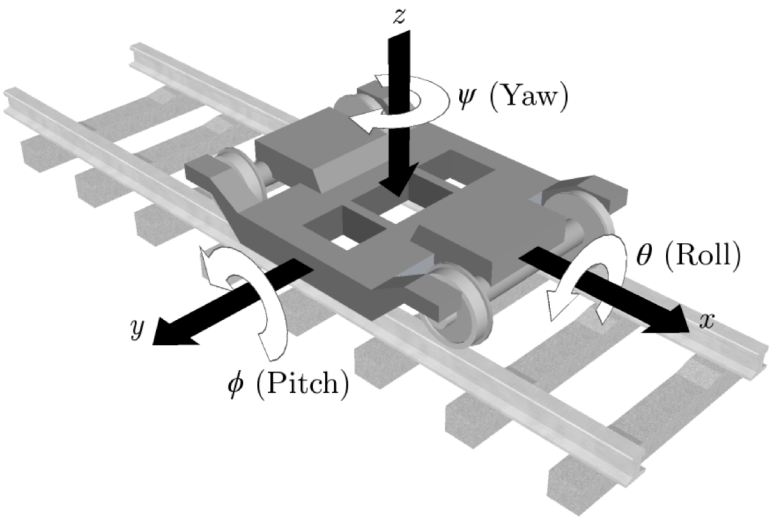


Figure 4.3: Six degree of freedom model applied to a vehicle bogie
Arrows point in the positive direction of each degree of freedom

The IMU is mounted on the bogie such that each of its faces are parallel to one of the three axes. Consequently, each of the three accelerometer channels directly measures acceleration in the x, y, and z axes, and each of the three gyroscopes directly measure the rate of change of θ, ϕ, and ψ. Some of the sensors require inversion of their output polarities, but this is done through calibration at the data import stage described in Section 4.1.1.

It is important to note that measured accelerations will vary depending on the placement

of the IMU on the bogie. For example, an IMU placed nearer to the right-hand side of the bogie will experience a greater component of the vertical displacement generated by variations in the right-hand rail's vertical geometry, than that of the left-hand rail. Whilst the accelerometer measurements will differ based on the IMU's positioning on the bogie, the same is not true of rotational velocities, which will be the same in each axis when measured at different locations on the bogie. For example, if the bogie is pitching at a certain angle, this angle will be the same when measured at any point on the bogie's surface. Ideally, the IMU should be mounted in the centre of the bogie as it will measure an average of accelerations in each direction. However, in practice this is impossible, as the bogie's centre position is where the carriage above is supported. For the IMU mounted on the Class 377, a position is provided with four tapped holes for the addition of extra equipment to the bogie. This position is not central laterally or longitudinally. The effect of the IMU's positioning on the bogie is addressed in Sections 5.5 and 5.6, in which processing is performed to compensate for the off-centre positioning.

4.2.2 Calculating Vertical Displacement

A simple method of estimating the vertical displacement, z, of the bogie uses the acceleration measured by the vertically-sensing accelerometer, double integrated with respect to time (4.1).

$$z = \iint \ddot{z} \, \mathrm{d}t \, \mathrm{d}t \tag{4.1}$$

Where $\ddot{z}$ is the acceleration measured by the vertically-sensing accelerometer, which includes some error. The result of this double integration will usually contain a large drift in either the positive or negative direction. This is due to the fact that there is usually an offset in the accelerometer's zero measurement (*i.e.* the accelerometer rarely gives a 'zero' output when acceleration is exactly zero), and often additional low frequency drift. Even a small offset is accumulated by the integration process, and the second integration

further exacerbates the problem, often making the result resemble a second order curve (*e.g.* $y = x^2$). In order to remove this effect, a high-pass filter is applied to minimise the low frequency offset and drift.

The accelerometer output will also contain some high-frequency random noise, and further noise is introduced by the sampling electronics. In this instance, the random noise is very small (within 2 LSB) owing to the high-quality accelerometer used and the low-noise signal conditioning electronics. The double integration acts as a 2^{nd} order low-pass filter, making the resulting high-frequency noise insignificant.

Vertical Curvature

Raw accelerometer samples provide samples of the rate of change of vertical speed with respect to time. It has been seen how a double integration with respect to time can provide an estimate of vertical displacement, but these data are in the time domain. In order to provide meaningful information about the track itself, it is useful to view the vertical displacement with respect to the position along the track. This conversion is necessary if comparisons are to be made between passes of the train over the same section of track at differing speeds, as comparisons will not be possible in the time domain. Due to factors such as signal aspects, temporary speed restrictions, and differing driver styles, it is unlikely that a train will always traverse a section of track at the same speed.

It is therefore necessary to convert the vertical acceleration with respect to time into vertical displacement in the spatial domain. This can be approximated by first finding the vertical curvature of the track measured by the accelerometer:

$$\kappa_z = \frac{\ddot{z}}{v_x^2} \tag{4.2}$$

Where v_x is the vehicle speed in the x direction.

Vertical curvature is equal to the second derivative of z with respect to x:

$$\kappa_z = \frac{\mathrm{d}^2 z}{\mathrm{d}x^2} \tag{4.3}$$

Therefore, double integrating κ_z with respect to the distance along the track, x, gives the vertical displacement in the spatial domain:

$$\hat{z} = \iint \kappa_z \, \mathrm{d}x \, \mathrm{d}x \tag{4.4}$$

As with $\hat{z}$ in the time domain, a high-pass filter must be applied to remove the effects of accelerometer drift. For most of the work in this thesis, a 2$^{\text{nd}}$-order high-pass Butterworth filter with a cutoff wavelength of 35 m is applied to the waveform once in a forwards direction and again in a backwards direction. This achieves a 4$^{\text{th}}$-order filter with a zero phase shift. The 35 m cutoff was chosen to match Network Rail's filtering; they calculate 'top' at both 35 and 70 m wavelengths [86]. The lower value was chosen here because there is less focus on very long wavelength geometry features. A 4$^{\text{th}}$-order filter was chosen to provide a sharp cutoff in frequency response.

Domain Conversion

Now $\ddot{z}$ and v_x are considered as vectors $\ddot{Z}$ and V_x containing discrete samples from the vertically-sensing accelerometer and calculated tacho speed. Samples are taken in the time domain, separated by δt. Since data from the Southern IMU installation is sampled at 250 Ss^{-1}, the sample separation, $\delta t = 4$ ms.

After conversion to $\hat{Z}$ in the spatial domain, samples are still separated by δt, not δx as would be necessary for samples to form a discrete sequence in the spatial domain. Resampling is performed using an algorithm, given in Appendix A.1. This converts samples from discrete temporal spacing to discrete spatial spacing.

The original IMU sample rate is 250 Ss^{-1}. This is converted to a spatial sample rate of 8 Sm^{-1}. This sample rate was chosen to give an acceptable resolution for the detection of most track defects. It also means that upsampling is only required at vehicle speeds above 31.25 ms^{-1} (approximately 70 mph). The vehicle's maximum in-service speed has been found to be approximately 40.25 ms^{-1} (90 mph), which would result in upsampling by a factor of ~1.3. Data could be sampled at a higher rate when travelling at higher speeds to eliminate the need for upsampling. This was not done, as the serial connection between the IMU and the PC was already operating close to its maximum bandwidth.

Comparing to Other Systems

In Chapter 2, several of the systems reviewed [42–44] use measurements from vertically-sensing accelerometers to calculate the vertical profile of the track. Accelerometer-based systems are capable of accurate results when the train is operating at a high speed, because the resultant rate at which the train moves vertically on the track is also high. The accuracy of accelerometer-based systems reduces as the vehicle speed decreases. This is because for a geometry feature of fixed vertical curvature, κ_z, it is known from (4.2) that:

$$\ddot{z} = \kappa_z \, v_x^2 \tag{4.5}$$

Where $\ddot{z}$ is the vertical acceleration, and v_x is the vehicle speed. Because the speed is squared, if v_x is low, $\ddot{z}$ becomes very small. This can cause it to fall below the noise threshold of the accelerometer, creating a very low signal-to-noise ratio and making the acceleration unmeasurable. The effects of low vehicle speed on vertical displacement derived from both accelerometer and gyro measurements is explored in Section 4.2.4.

Several of the systems reviewed in Chapter 2 also use axle-box accelerometers, which complicates the installation and maintenance of the system. In the case of the instrumentation of the Class 377 EMU, axle-box accelerometers are not required, as the system is

not required to detect very short wavelength defects such as rail squats and corrugation. The vehicle's primary suspension will only have an effect on the measurement of short wavelength features [66].

Low Speed Operation

The low-speed limitation of the accelerometer-based system may not be considered a problem when used on a dedicated measurement vehicle which may be operated continuously at line speed. However, it is a significant problem when using data recorded from an in-service vehicle which slows and stops regularly. When the train is moving at low speed, the resultant vertical acceleration is very small and can easily fall below the noise threshold of the accelerometer. An alternative method of estimation which suffers less from this problem uses the pitch-rate gyro instead of the vertical accelerometer. This system has been used by the University of Birmingham for many years [65,87], and takes advantage of the fact that the IMU is mounted on the vehicle bogie, which pitches as it traverses dips and bumps in the rail.

The curvature of the track can be calculated using the pitch rate of the bogie (4.6).

$$\kappa_\phi = \frac{\dot{\phi}}{v_x} \tag{4.6}$$

Here, κ_ϕ, which is the approximate vertical curvature of the track measured by the gyro, is found by dividing the pitch rate, $\dot{\phi}$, by the speed of the train along the track, v_x. The resulting curvature samples must be converted to the spatial domain using the algorithm listed in Appendix A.1, after which it is double integrated with respect to distance along the track, shown in (4.7). This is also high-pass filtered to remove the effects of low frequency offset and sensor drift. The result is an approximation of the vertical displacement in the spatial domain.

$$\hat{z} = \iint \kappa_\phi \, \mathrm{d}s \, \mathrm{d}s \tag{4.7}$$

As with the vertical displacement derived from the accelerometer, a 4^{th}-order high-pass Butterworth filter with zero phase shift and 35 m cutoff wavelength is used for high-pass filtering.

With the sensors used on board the Southern Trains installation, the pitch-rate gyro method of deriving vertical displacement gives acceptable results at speeds as low as 2 ms^{-1}, which are to within 0.2 mm of the vertical displacement measured at higher speeds (> 20 ms^{-1}).

When the pitch-rate gyro is being used for calculation of vertical displacement data, the direction of travel of the vehicle (in the case of the Class 377 EMU, whether the 'A' or 'B' cab is leading the train) must be accounted for and the sign of the speed, v_x, chosen accordingly. The direction can be identified in one of two ways. The data can be inspected to find the most recent time that the train stopped (using tacho pulses or GPS speed) and then observing the sign of the x (longitudinally-sensing) accelerometer measurement when the train next accelerates. This may be done automatically using an algorithm. Alternatively, the vertical displacement can be calculated without regard to direction, after which the resulting data and its negative version can be matched to an existing data set to identify the correct sign. This method is described in Section 6.2.4.

4.2.3 Calculating Vehicle Speed

Two sources are available for speed measurement. These are the speed measured by the GPS receiver, and the speed measured by the tachometer. Both of these have advantages and disadvantages resulting in varying levels of accuracy. The GPS speed is only updated once per second which means a high level of upsampling is required to obtain speed samples at the same rate as IMU samples. This also means that the speed measurements tend to be delayed, as the measurement is reported up to one second after the vehicle was travelling at that speed. Errors in the measured GPS positions (which are used by the GPS receiver to calculate the speed), mean that at low speeds the signal-to-noise ratio

of the reported speed can be very low, resulting in inaccurate speed samples. It is also found that whilst the GPS speed can become very accurate whilst the receiver moves in a straight line, it quickly becomes less accurate around a curve.

The tachometer on the Class 377 does not suffer from these problems. The IMU counts the number of pulses produced by the tachometer and reports this count with samples from the other IMU channels once every 4 ms. As stated in Section 3.2.2, the tachometer produces around 2000 pulses per second at 100 mph. Using the values from Section 3.2.2 it can be found that a speed of around 5.8 ms^{-1} will produce one pulse per IMU sample, so speed upsampling is only required at very low vehicle speeds. However, the tachometer is subject to inaccuracies due to wheel slip and slide, where the wheel rotates at a different speed to the true track speed. This occurs when an axle is driven or braked too hard, causing a loss of adhesion between the wheels and rails. The tachometer connected to the IMU is mounted on an axle which is not motored (driven), so a loss of adhesion can only occur during braking. The tachometer measurement also suffers from inaccuracies caused by wheel wear. Throughout the course of their lives, wheels reduce in diameter due to wear and due to turning; a maintenance procedure performed on the wheels to remove flat spots caused by wheel slide. Consequently, the circumference of the wheel reduces and the distance travelled per tacho pulse reduces.

The tacho is the most consistently accurate method of measuring the vehicle speed, but wheel wear and maintenance means that regular calibration will be required. This can be done using the GPS speed when the vehicle is travelling at high speed and the signal-to-noise ratio of the GPS speed is high. By choosing a calibration time when the speed is constant, any inaccuracies caused by GPS measurement delays or interpolations can be minimised.

4.2.4 Vertical Displacement Results

Figure 4.4 shows the vertical displacement calculated from both the accelerometer and gyro methods described in this section. The figure shows a pass of the IMU over a 100 m section of track at a speed of ~30 ms^{-1}. The figure also shows data from Network Rail's New Measurement Train (NMT), recorded within a few weeks of the IMU passes. This was used to verify data from the IMU against an industry-accepted system. The pre-processed NMT data had a 3rd-order high-pass Butterworth filter applied to it in the reverse direction to make its filtering a better match to the filter applied in the IMU data processing.

In this instance, because NMT data was available, the decision was made to use this as a reference, as it is accepted by the UK rail infrastructure operator as a standard measurement. Without access to such data, it would be possible to use a pre-calibrated vibration or motion device to introduce known movements to the IMU under laboratory conditions and verify that a valid output is given. Some basic tests were carried out before installation. The IMU was placed on a flat surface in different orientations such that each axis measured the acceleration due to gravity. Rotational tests were also carried out, where the IMU was rotated through exactly 90° on each axis, and the resulting rotational velocity was integrated to verify that a change of 90° had been observed by the IMU. These tests proved basic operation of the three accelerometer axes and the three gyros.

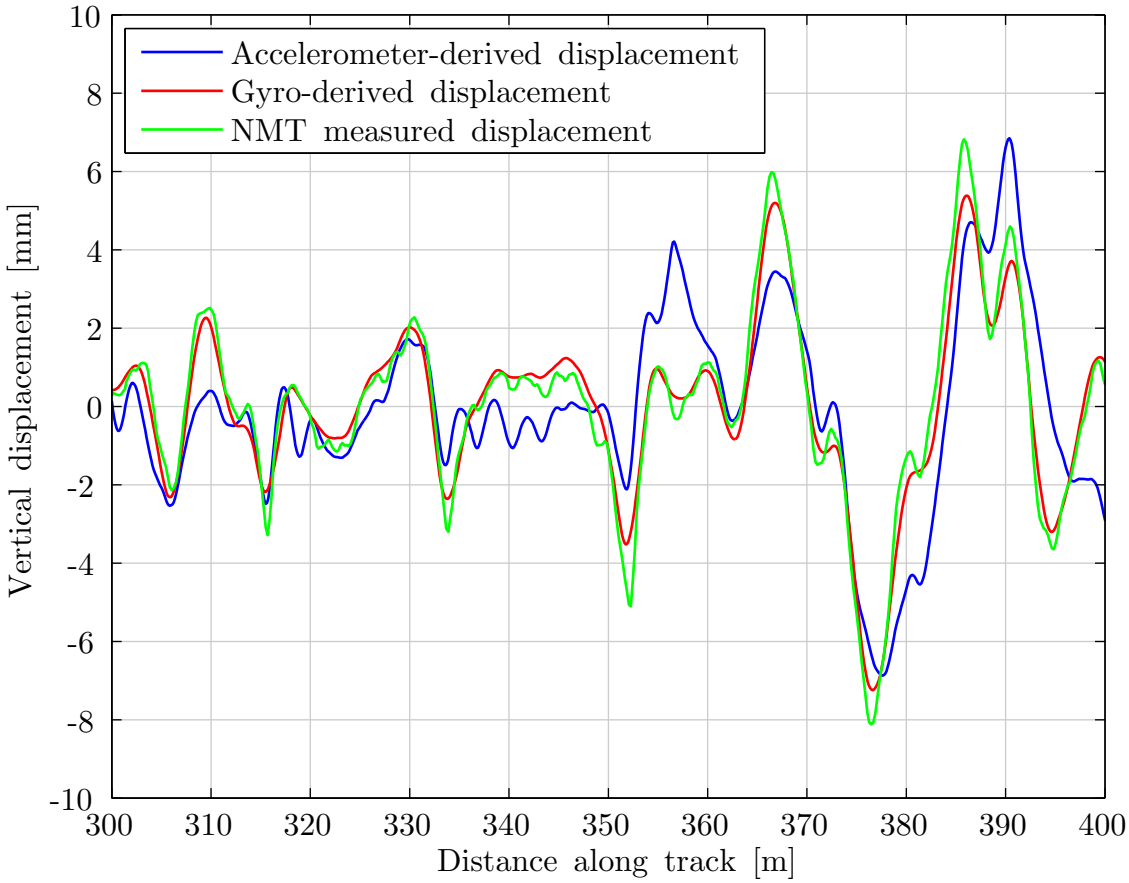


Figure 4.4: Calculated vertical displacement compared to NMT data

It can be seen from the accelerometer- and gyro-derived results that each of the two sensors are better at measuring different components of the vertical displacement. When comparing displacement derived from each sensor to the NMT reference data, it can be seen that the gyroscope is more accurate at representing the magnitude of the larger peaks and troughs, whereas the accelerometer is better at representing some of the high-frequency components, such as that in the range 338 m to 346 m. In this range it can be seen that the gyro-derived displacement is very flat compared to the NMT data, whereas the accelerometer-derived displacement shows a good quantity of the high-frequency information. Chapter 5 considers why these effects occur.

Figure 4.5 contains another vertical displacement plot, which shows both accelerometer- and gyro-derived vertical displacements. An aligned plot of the vehicle speed is shown immediately above it. Both displacement measurements are known to be within a normal

range throughout (around ±10 mm). It can be seen that just after 400 m, where the speed falls below ~12 ms^{-1}, the accelerometer measurement begins to deviate from the gyro measurement. Equations (4.2) and (4.4) dictate that as the vehicle speed reduces, a smaller acceleration is required to show the same displacement. The acceleration is very small at low speeds, and starts to fall below the level of the accelerometer drift. This causes the calculated displacement to show a large component of the accelerometer drift.

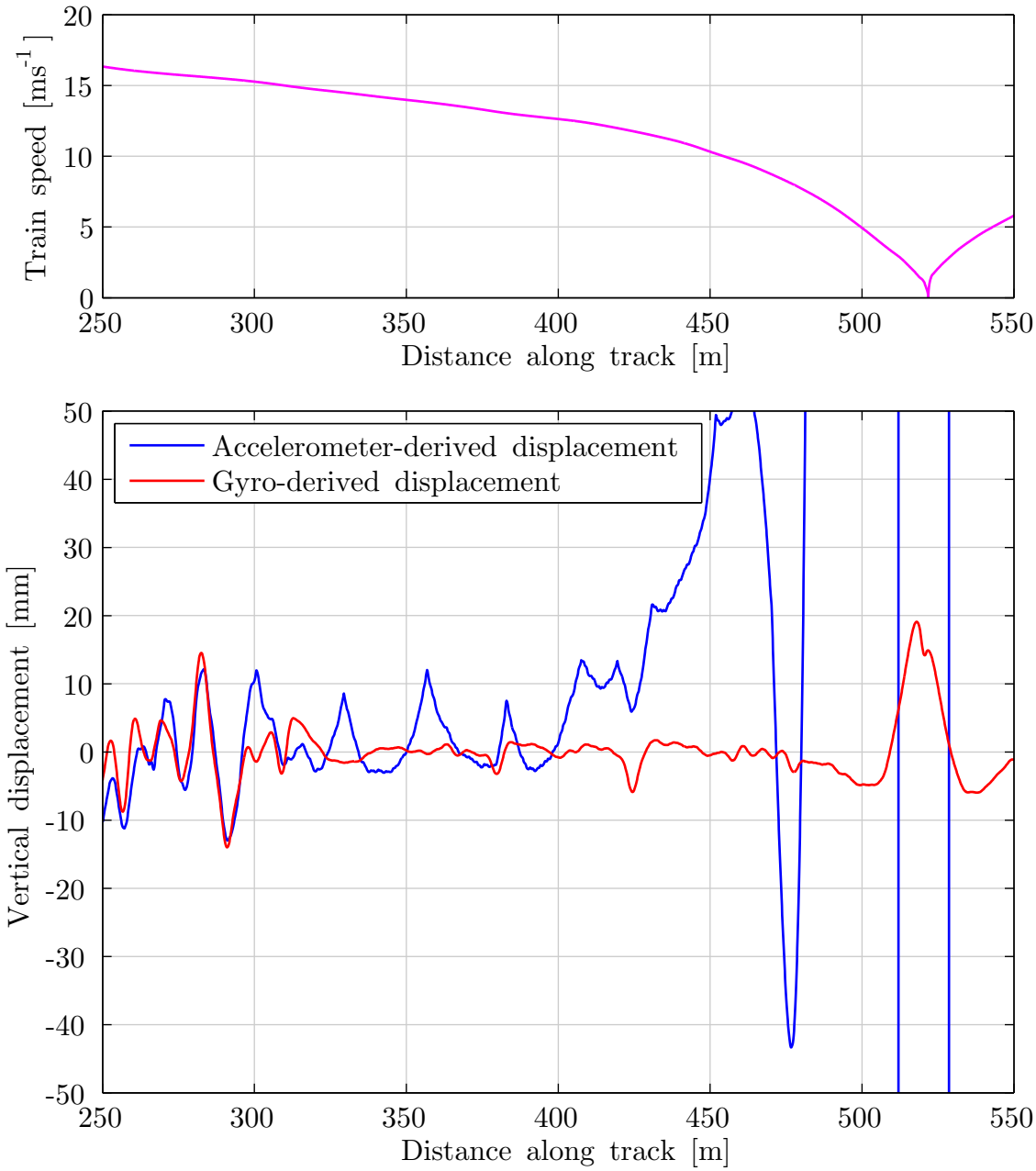


Figure 4.5: Effect of low vehicle speed on calculated vertical displacement

It can also be seen that the gyro continues to give acceptable results until the speed falls below ~2 ms^{-1} at around 510 m. The gyro suffers from a similar effect to the accelerometer, but this effect is lessened due to the lower drift of the accelerometer, and the fact that in the accelerometer-measured curvature equation (4.2) the speed is squared, whereas in the gyro-measured curvature equation (4.6) it is not.

It can be seen from both displacement figures (4.4 and 4.5) that there would be some advantage to combining data derived from the two sensor outputs to produce an optimal vertical displacement. The gyro works best at lower speed, whereas the accelerometer provides useful information about some of the higher-frequency features of the track displacement. The effective frequency responses of the two sensors are considered in Chapter 5, after which a method of combining processed outputs from the two sensors is devised.

CHAPTER 5

COMBINING ACCELEROMETER AND GYROSCOPE RESULTS

So far, consideration has been given to how either the vertically-sensing accelerometer or the pitch-rate gyro can be used to estimate the vertical displacement of the bogie. Each method has its own advantages and disadvantages. The accelerometer method is less effective at low speeds, whereas the gyro is unable to detect short wavelength geometry features.

In fact, each method of vertical displacement estimation has a different response to varying wavelengths of track geometry features. Gillespie provides a good explanation of the natural filtering due to the spacing of a car's front and rear axles, ignoring the suspension [88]. He explains the effect of different wavelength features in the road's surface as the car traverses it. Graphs of the response gain from both the vertical bounce and the pitch are given. Figure 5.1 illustrates this effect when applied to the bogie of a railway vehicle.

If the bogie travels over a geometry feature which is twice the wavelength of the bogie wheelbase, it will pitch, and the gyro will be able to measure the feature. However, the IMU will stay at the same vertical position and consequently the vertical-sensing accelerometer will not measure it. Conversely, a feature with a wavelength the same as the bogie wheelbase will result in no gyro rotation, as the bogie does not pitch, but the vertical-sensing accelerometer will detect the feature well.

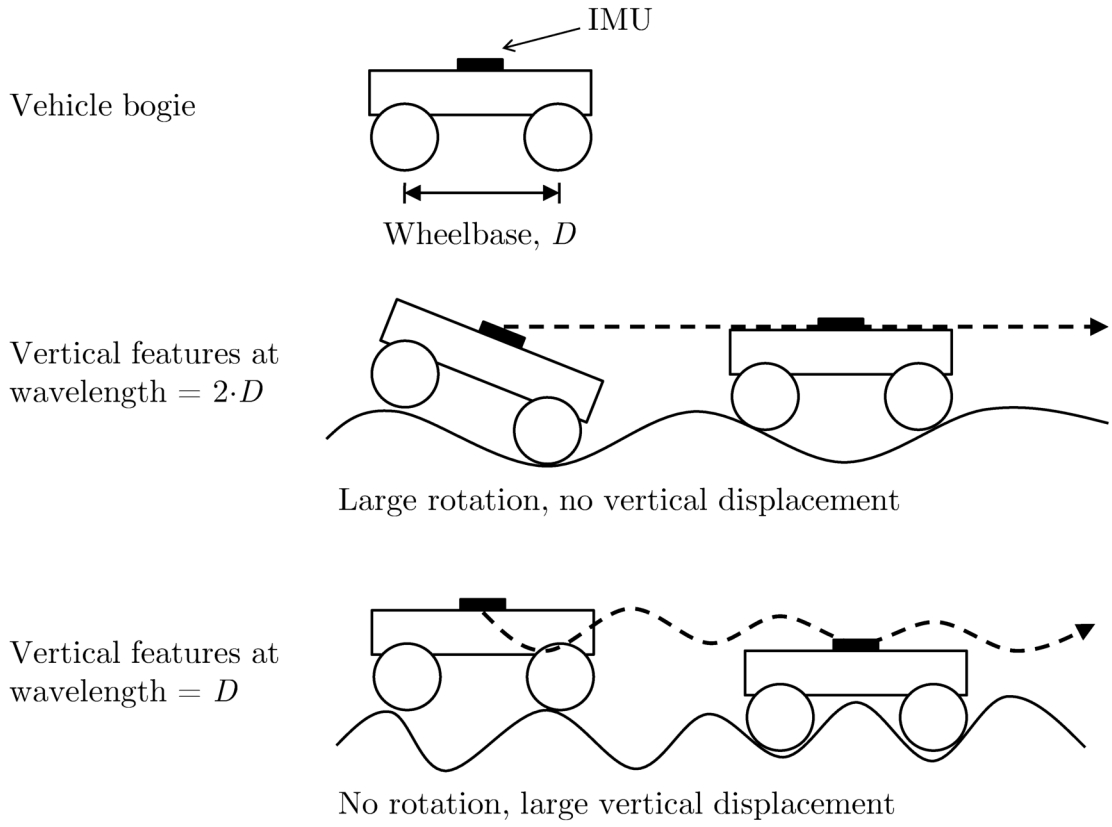


Figure 5.1: The filtering effect of a vehicle bogie

A method of combining the two sensor inputs to estimate the optimal vertical position of the bogie is required. A Kalman filter was used to do this.

5.1 The Kalman Filter

In a Kalman filter [89], multiple inputs are used to estimate the state of multiple outputs. A discrete Kalman filter works using an iterative process for each step of a changing signal. The outputs being predicted are known as state variables. In each iteration there are two steps.

1. Predict the new state and its uncertainty (error)

2. Correct the state based on the latest measurement

The initial prediction for the state, P, at step $k+1$, where k is the current step, is calculated in the following manner for a system containing n state variables:

$$P_{k+1} = AP_kA^T + Q \tag{5.1}$$

Where P_{k+1} is the new state prediction and P_k is the previous state prediction. Each of these are $n \times 1$ column vectors containing the state variables. Q is an $n \times n$ matrix containing the process noise covariance, and A is the state transition matrix which is an $n \times n$ matrix which specifies how each predicted state changes based on the previous state.

The prediction is then corrected based on the input measurements using the following method:

$$\hat{\boldsymbol{x}}_{k+1} = A\hat{\boldsymbol{x}}_k + K\left(\boldsymbol{y}_k - CA\hat{\boldsymbol{x}}_k\right) \tag{5.2}$$

Where $\hat{\boldsymbol{x}}_{k+1}$ is an $n \times 1$ column vector containing the new state estimates, and $\hat{\boldsymbol{x}}_k$ contains the current state estimates. $\boldsymbol{y}_k$ is an $m \times 1$ column vector containing the current measurement values from m inputs, and C is the $n \times m$ observation matrix which defines how the observations are expected to be affected by the state variables.

K is the Kalman filter gain, found using:

$$K = P_{k+1}C^T\left[CP_{k+1}C^T + R\right]^{-1} \tag{5.3}$$

Here, R is an $m \times m$ matrix containing the observation noise covariance. The observation noise is the possible error incurred by factors such as sensor error, analogue-to-digital conversion errors, and errors arising from the sampling resolution (bit depth) of the digital version of the waveform.

5.2 Designing a Kalman Filter for the IMU

In the IMU scenario, the outputs (the curvatures estimated from sensor measurements) are already known but the input (the true curvature of the track) is unknown.

The vertical acceleration and pitch-rate of the bogie is measured. In a simple model, the vertical displacement derived from each of these sensors is an average of the vertical displacement of the track under each of the bogie wheels.

For this reason, the Kalman filter design estimates a state (containing the vertical displacement) for every location along the track length between the two wheels of the bogie. The bogie wheelbase is 2.6 m and the spatial sample rate is 8 Sm^{-1}. This means that 21 state variables are required in the system, which represent locations on the track δx apart. Here, $\delta x = 0.125$ m. However, the necessary sensor differentiations are carried out implicitly in the observation matrix, so the filter is widened to 23 state variables. This will be explained fully later. Figure 5.2 illustrates this arrangement.

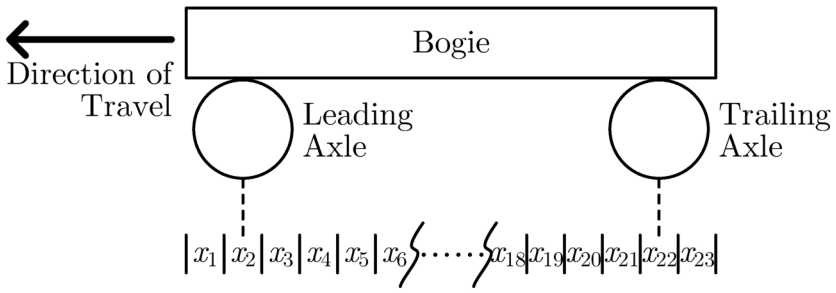


Figure 5.2: State variables representing discrete track locations beneath a bogie

The 23×23 state transition matrix, which defines how each prediction changes based on the previous prediction is as follows:

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \tag{5.4}$$

With each step, k, the bogie is moving forwards by δx, so rows 2 onwards of A cause the state predictions P_{k+1} to be equal to the previous predictions offset by δx. The first row of A predicts that the track δx ahead of the prediction P_k will be the same height as the previous sample. Process noise, w, is added to this state variable, as this is the only variable which is unknown.

The 23×23 matrix of process noise covariance is therefore:

$$Q = \begin{bmatrix} w & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{5.5}$$

In the next stage of the Kalman filter, the prediction is corrected based on the latest sensor measurements.

The vector of observations, y_k, is formed of the k^{th} samples from two inputs; one from the vertically-sensing accelerometer, and one from the pitch-rate gyro. Because the filter works inherently in the spatial domain, the sensor samples must first be converted to translate them into the spatial domain.

5.2.1 Using Accelerometer Samples in the Kalman Filter

In Section 4.2.2, the vertical curvature measured by the vertically-sensing accelerometer, κ_z, was found in (4.2) and converted to the spatial domain.

This value is one of the two outputs of the Kalman filter. The other will be the curvature calculated from the pitch-rate gyro. A method of deriving the output, κ_z, from the inputs (the discrete heights of the track) is now required.

Figure 5.3 illustrates three discrete positions along the length of the track. The vertical displacements of these points shall be referred to as z_1, z_2 and z_3. The angles ϕ_1 and ϕ_2 are the angles between points, measured from the horizontal.

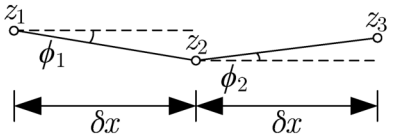


Figure 5.3: Angles between discrete points on a track

The curvature is equivalent to the rate of change of angle with respect to distance along the track, x.

$$\kappa = \frac{\Delta\phi}{\Delta x} \tag{5.6}$$

This can be put in terms of the example in Figure 5.3. $\Delta\phi$ is the difference between the two angles ϕ_1 and ϕ_2, and Δx is the distance between discrete steps, δx. So:

$$\kappa = \frac{\phi_2 - \phi_1}{\delta x} \tag{5.7}$$

It can be proven (Appendix A.2) that the angle, ϕ, turned in distance travelled, x, is approximately equal to the rate of change of z displacement:

$$\phi \simeq \frac{\mathrm{d}z}{\mathrm{d}x} \tag{5.8}$$

Using this, it can be said that:

$$\kappa \simeq \frac{(z_3 - z_2) - (z_2 - z_1)}{\delta x^2} \tag{5.9}$$

Which can be simplified to:

$$\kappa \simeq \frac{z_1 - 2z_2 + z_3}{\delta x^2} \tag{5.10}$$

This gives an approximation of how the vertical track displacements at discrete track locations, z_k affect the output, κ. This equation only defines how the discrete track heights under one axle affect the curvature. It is assumed at this stage that the curvature measured by the accelerometer on the bogie, κ_z, is the average of the curvatures experienced by the leading (front) and trailing (rear) axles, κ_F and κ_T respectively.

$$\kappa_z = \frac{\kappa_F + \kappa_T}{2} \tag{5.11}$$

From Figure 5.2, it can be seen that the three discrete values nearest the leading wheel are $k = 21, 22, 23$ and the three nearest the trailing wheel are $k = 1, 2, 3$.

So by combining (5.10) and (5.11):

$$\kappa_z \simeq \frac{z_1 - 2z_2 + z_3 + z_{21} - 2z_{22} + z_{23}}{2\delta x^2} \tag{5.12}$$

This defines how the system inputs affect the output. In this case the inputs are the unknown vertical heights of the track at discrete locations, and the output is the known

curvature calculated from the vertically-sensing accelerometer samples, measured on the bogie.

The top row of the observation matrix, C, is the impulse response for the accelerometer curvature input. This row shall be referred to as C_z. The equation (5.12) is used to determine the value in each column:

$$C_z = \left[\frac{1}{2\delta x^2} \quad \frac{-1}{\delta x^2} \quad \frac{1}{2\delta x^2} \quad 0 \quad 0 \quad \cdots \quad 0 \quad 0 \quad \frac{1}{2\delta x^2} \quad \frac{-1}{\delta x^2} \quad \frac{1}{2\delta x^2} \right] \tag{5.13}$$

The second row of the observation matrix defines how the curvature calculated from the gyro samples is affected by the track.

5.2.2 Using Gyroscope Samples in the Kalman Filter

In Section 4.2.2, the curvature experienced by the gyro, κ_ϕ, was calculated in (4.6) and converted to the spatial domain. In order to use this as an output to the Kalman filter, a method of deriving the output κ_ϕ from the discrete track heights is now required.

For symmetry, the same three-element groups are used for the translation from track height to pitch-rate gyro-derived curvature. The gyro measures the rotational velocity of the bogie. The angle of the bogie itself is derived from the difference in height between the two axles (Figure 5.4).

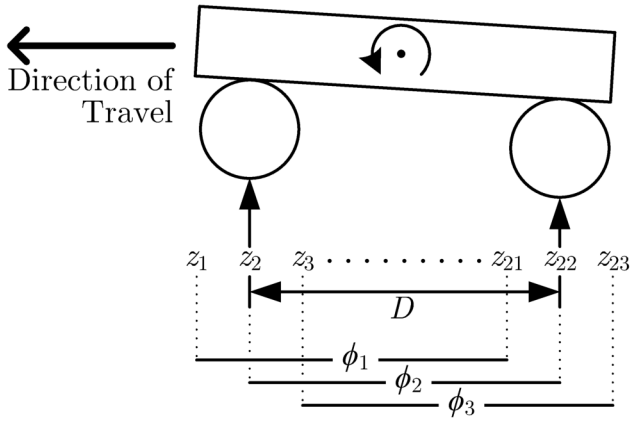


Figure 5.4: Bogie rotation derived from the two axle heights

The angles of each of the three pairs of discrete track locations can be found, using approximation (5.8). The distance between axles is referred to as D which, in this case, is equal to $20\delta x$.

$$\phi_1 \simeq \frac{z_{21} - z_1}{D} \tag{5.14}$$

$$\phi_2 \simeq \frac{z_{22} - z_2}{D} \tag{5.15}$$

$$\phi_3 \simeq \frac{z_{23} - z_3}{D} \tag{5.16}$$

Using (5.7), the curvature between each adjacent pair can be found. An average curvature is found between the ϕ_1 and ϕ_2, and ϕ_2 and ϕ_3. This is the curvature experienced by the gyro, κ_ϕ.

$$\kappa_\phi = \frac{\left(\frac{\phi_2 - \phi_1}{\delta x}\right) + \left(\frac{\phi_3 - \phi_2}{\delta x}\right)}{2} \tag{5.17}$$

This can be simplified, and it is found that angle ϕ_2 cancels out:

$$\kappa_\phi = \frac{\phi_3 - \phi_1}{2\delta x} \tag{5.18}$$

Finally, (5.14) and (5.16) are used to put this in terms of the vertical displacements of the track:

$$\kappa_\phi \simeq \frac{z_{23} - z_3 - z_{21} + z_1}{2\delta x \cdot D} \tag{5.19}$$

As $D = 20\delta x$, the denominator can be simplified:

$$\kappa_\phi \simeq \frac{z_{23} - z_3 - z_{21} + z_1}{40\delta x^2} \tag{5.20}$$

This is now in a form which can be transferred to the bottom row of the observation matrix. This row is referred to as C_ϕ. Equation (5.20) is now used to determine the value in each column:

$$C_\phi = \begin{bmatrix} \frac{1}{40\delta x^2} & 0 & \frac{-1}{40\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{-1}{40\delta x^2} & 0 & \frac{1}{40\delta x^2} \end{bmatrix} \tag{5.21}$$

As with C_z there are 17 zeros between the first three and last 3 elements, giving a total width of 23 elements.

5.2.3 Observation Matrix

Rows C_z and C_ϕ of the observation matrix have now been found, which define how the system inputs affect both curvature outputs. (5.22) shows the finished 2×23 observation matrix.

$$C = \begin{bmatrix} \frac{1}{2\delta x^2} & \frac{-1}{\delta x^2} & \frac{1}{2\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{1}{2\delta x^2} & \frac{-1}{\delta x^2} & \frac{1}{2\delta x^2} \\ \frac{1}{40\delta x^2} & 0 & \frac{-1}{40\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{-1}{40\delta x^2} & 0 & \frac{1}{40\delta x^2} \end{bmatrix} \tag{5.22}$$

5.2.4 Observation Noise

The final component required for implementation of the Kalman filter is the Observation Noise Covariance matrix, R. This defines the noise incurred by sensor error, conversion errors, resolution errors, and any other noise introduced between the value experienced by the sensor and the Kalman filtering stage.

From [89, pp. 119], it can be found that the covariance, R, at step k is:

$$R(k) = E\left[\boldsymbol{v}(k) \cdot \boldsymbol{v}(k)^T\right] \tag{5.23}$$

$\boldsymbol{v}(k)$ is a row vector containing the noise present on each measurement input. For each element, $v_i(k)$, in $\boldsymbol{v}(k)$, the value is the difference between the measured value of input i and its true value. The noise covariance is the expected value of the multiplication of $\boldsymbol{v}_k$ with its transpose. As there are two measurement inputs in the IMU system, R is a 2×2 matrix. By expanding 5.23 it is found that:

$$R(k) = \begin{bmatrix} E\left[v_1^2(k)\right] & E\left[v_1(k)v_2(k)\right] \\ E\left[v_1(k)v_2(k)\right] & E\left[v_2^2(k)\right] \end{bmatrix} \tag{5.24}$$

Because there is no cross-correlation between the noise of each sensor, elements $(1,2)$ and $(2,1)$ become zero.

At this point, consideration must be given to how error fits into the translation from sensor measurements to vertical track curvature. Curvature was calculated from the two sensors previously in (4.6) and (4.2). In fact each of these measurements should include a noise value. The existing equations can be updated to contain the error values:

$$\kappa_z = \frac{\ddot{z} + e_z}{V_x^2} \tag{5.25}$$

$$\kappa_\phi = \frac{\dot{\phi} + e_\phi}{V_x} \tag{5.26}$$

Where e_z and e_ϕ are the accelerometer and gyro errors, respectively. From this, the error covariance can be deduced:

$$R(k) = \begin{bmatrix} \left(\frac{e_z(k)}{V_x^2(k)}\right)^2 & 0 \\ 0 & \left(\frac{e_\phi(k)}{V_x(k)}\right)^2 \end{bmatrix} \tag{5.27}$$

This defines how the noise of each sensor is affected by the vehicle speed. Values for e_a and e_g were selected through experimentation.

5.2.5 Results from the Standard Kalman Filter

Previously in Figure 4.4, vertical displacements derived from single-sensors were plotted with NMT data as a reference. Figure 5.5 shows the same data, with the output of the Kalman filter added to the graph. The beneath below it shows the difference between Kalman filter output and NMT measured displacement, for clarity.

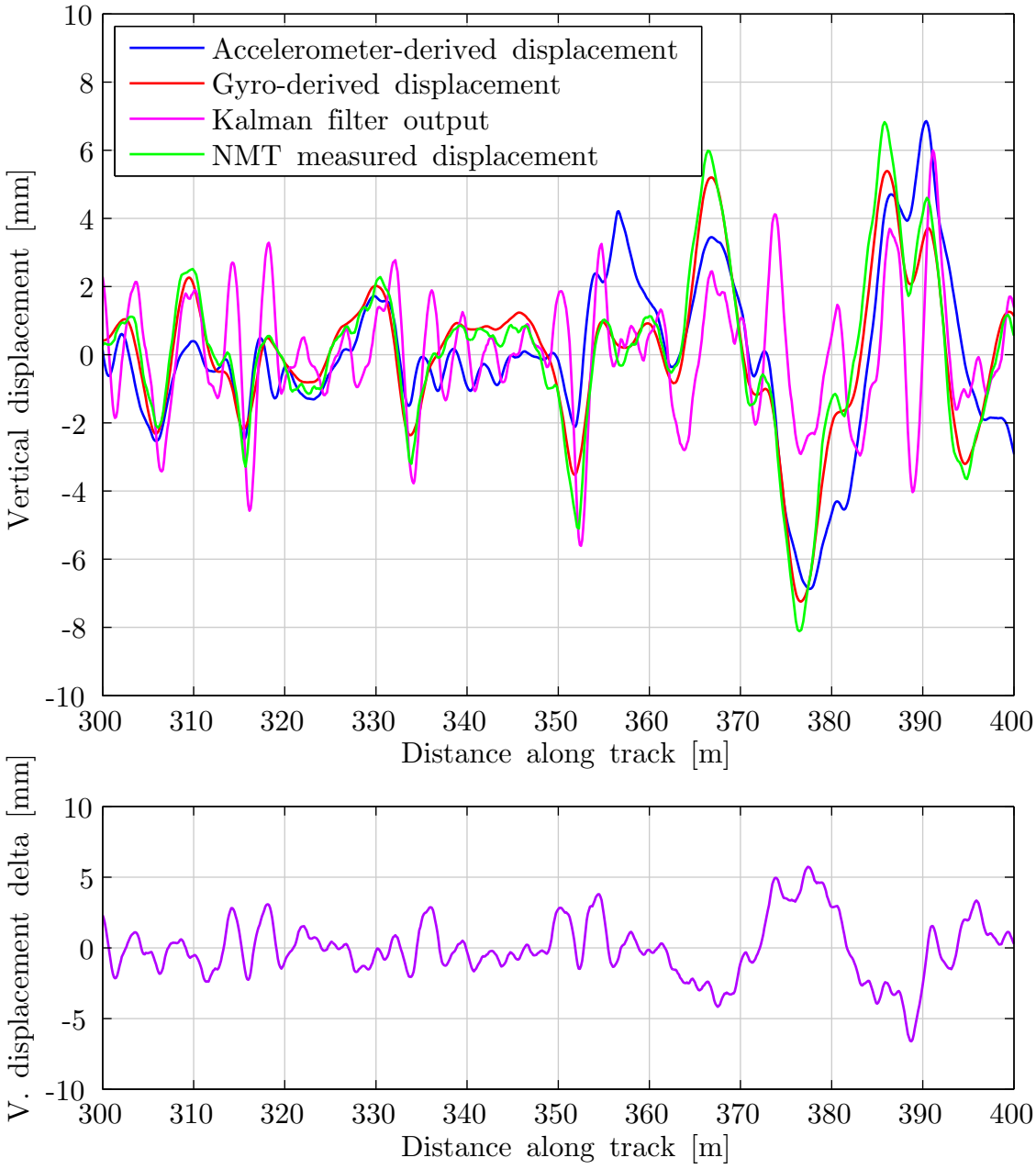


Figure 5.5: Comparison of Kalman filter output
showing accelerometer- and gyro-only displacements (top),
and the delta between Kalman filter and NMT displacements (bottom)

Upon initial inspection, it appears that the Kalman filter output is not a good fit to the NMT reference data. However, it can be seen that the Kalman filter is much better at indicating the magnitude of large dips in the track than either the accelerometer or gyro alone. It can also be seen, however, that generally the output is greater in magnitude than the NMT reference data. This is due to the fact that a standard Kalman filter is a causal system. For example, the presence of a large positive peak at 374 m is started by the sensors indicating a bogie movement in the upward direction between 371 m and 373 m. This upward movement is assumed to continue until the filter is confident that the displacement value is decreasing, which does not happen until 374 m. The causal nature of the standard Kalman filter means that it cannot "look ahead" in order to learn that the best estimate of the track profile is one which begins to go negative earlier on.

5.3 Modified Bryson-Frazier Filter

An improved variation to the Kalman filter is a modified Bryson-Frazier (mBF) filter [90, 91] which provides an optimally smoothed output. The mBF filter uses two passes; one in the forward direction, and one in reverse which operates on the saved output from the forward pass. This variation does not suffer from the same causality problems as the standard Kalman filter.

The impulse responses for the mBF filter are identical to those used in the standard Kalman filter; (5.22). The implementation of the filter is different to the standard Kalman filter, which used the three equations (5.1), (5.2), and (5.3).

The forward pass of the filter starts in a similar fashion to the standard Kalman filter:

$$\hat{\boldsymbol{x}}_k = A\hat{\boldsymbol{x}}_{k-1} \tag{5.28}$$

$$P_k = AP_{k-1}A^T + Q \tag{5.29}$$

The following update sequence is then followed:

$$N_k = CP_kC^T + R \tag{5.30}$$

$$K_k = P_{k-1}C^T N_k^{-1} \tag{5.31}$$

$$B_k = I - K_kC_k \tag{5.32}$$

$$P_{k+1} = B_kP_kB_k^T + K_kRK_k^T \tag{5.33}$$

$$\boldsymbol{z}_{k+1} = \boldsymbol{y}_k - C\boldsymbol{x}_{k-1} \tag{5.34}$$

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} + K_k\boldsymbol{z}_k \tag{5.35}$$

The results of (5.30), and (5.32) to (5.35) are stored for every step, k, so that they can be operated on by the reverse pass. Before the reverse pass of the filter, the following initialisations are performed:

$$\hat{\Lambda}_N = 0 \tag{5.36}$$

$$\hat{\lambda}_N = 0 \tag{5.37}$$

The following sequence is then performed with k descending, using values stored in the forward pass:

$$\widetilde{\Lambda}_k = C^T N_k^{-1} C + B_k^T \hat{\Lambda}_k B_k \tag{5.38}$$

$$\widetilde{\lambda}_k = -C^T N_k^{-1} \boldsymbol{z}_k + B_k^T \hat{\lambda}_k \tag{5.39}$$

$$\hat{\Lambda}_k = A^T \widetilde{\Lambda}_k A \tag{5.40}$$

$$\hat{\lambda}_k = A^T \widetilde{\lambda}_k \tag{5.41}$$

The smoothed state and the covariance are combined in each step:

$$P_N = P_k - P_k \hat{\Lambda}_k P_k \tag{5.42}$$

$$\hat{\boldsymbol{x}}_N = \hat{\boldsymbol{x}}_k - P_k \hat{\lambda}_k \tag{5.43}$$

The contents of $\hat{\boldsymbol{x}}_N$, which is the optimally smoothed input to the filter, is stored for every step, k. In the case of the state variables defined for the IMU system in Section 5.2, only the middle value of the 23 state variables is needed. This is the estimate of the vertical displacement of the rail directly underneath the middle of the bogie.

5.3.1 Results from the mBF Filter

Figure 5.6 shows the output of the mBF filter compared against the original output from the standard Kalman filter. The NMT data is once again shown for reference. The plot beneath shows both the difference between standard Kalman and NMT data, and the difference between mBF and NMT data.

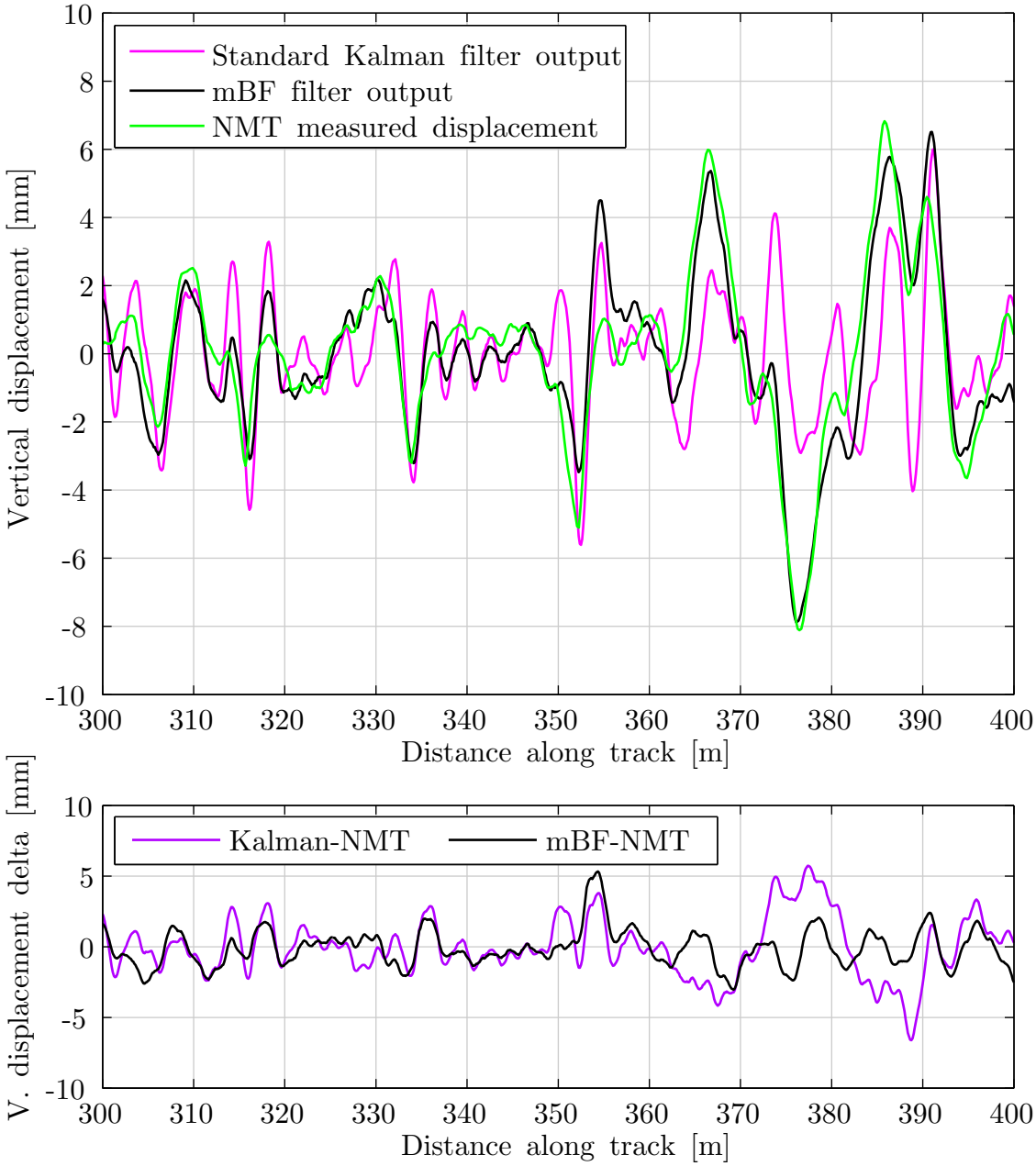


Figure 5.6: Comparison of outputs from standard Kalman and mBF filters shown with NMT reference data (top), and the deltas between Kalman filter/mBF filter and NMT displacements (bottom)

It can be seen that the mBF output is closer to the NMT reference data than the standard Kalman filter output. The use of the mBF filter has removed problems introduced by the causal nature of the standard Kalman filter, such as the large peak at 374 m which was studied in Section 5.2.5.

5.4 Variations Caused by Vehicle Orientation

It was found when comparing data produced from multiple passes of the train over the same section of track, that processed displacement data differed noticeably depending on the vehicle orientation (*i.e.* 'A' or 'B' cab leading the train). This difference is shown in Figure 5.7, which shows displacement data from 12 passes of the train over a 100 m section of track. These measurements were taken over the course of one month. The orientation of the train in each pass is indicated by the colour of the line. It should be noted that the direction of travel along the track is the same in every pass, only the orientation of the train changes.

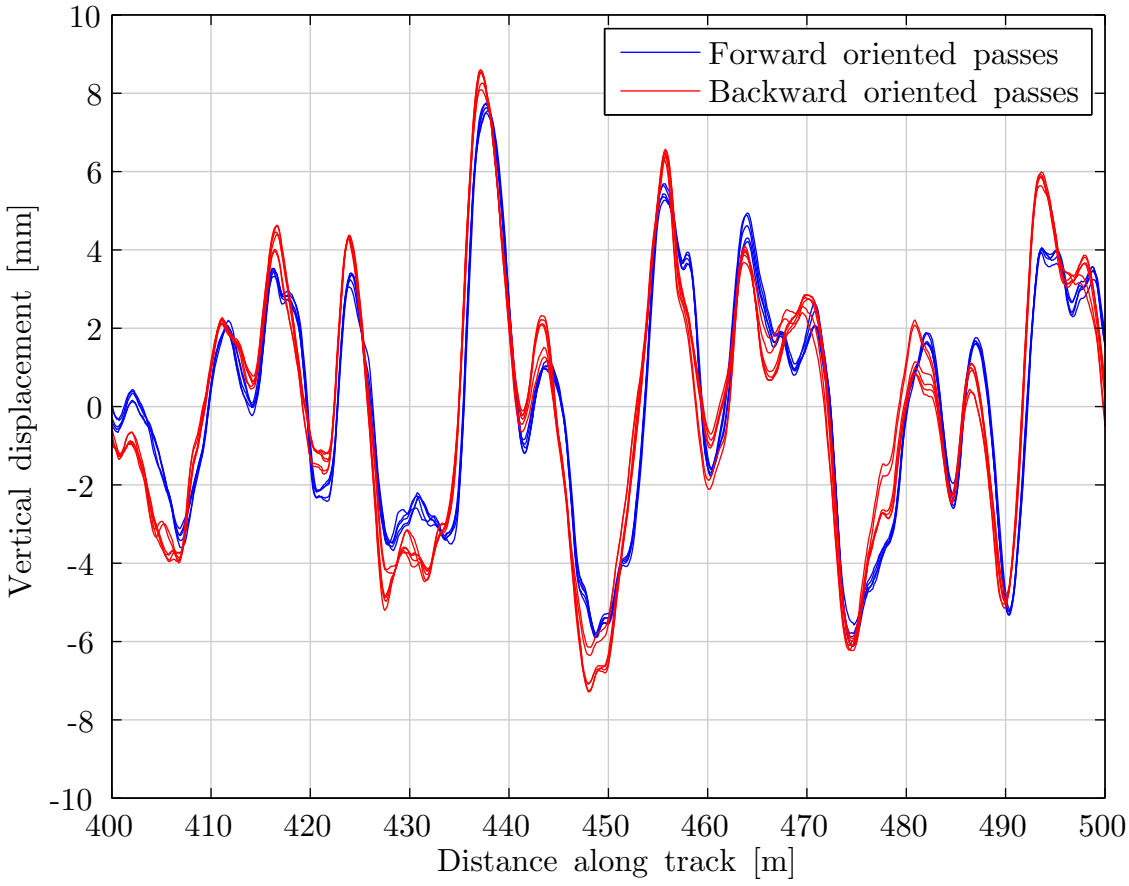


Figure 5.7: Vertical displacements for multiple train passes in differing orientations

It can be seen that all of the forward-oriented passes show displacements very similar to one another, and that reverse-oriented passes are also similar. However, the variation between orientations is noticeable.

Song *et al.* consider the effects of wheelbase filtering [92] and note that whilst the angular motion of the body in motion is constant across it, the vertical acceleration can differ depending on position. In the paper, vertical acceleration filtering is calculated at various points of the body and differing response gains can be seen. This means that the response of the IMU's vertically sensing accelerometer will be affected by its position on the bogie.

Based on this information, there are a number of possible factors which could influence the difference due to vehicle orientation:

- *The longitudinal position of the IMU on the bogie*

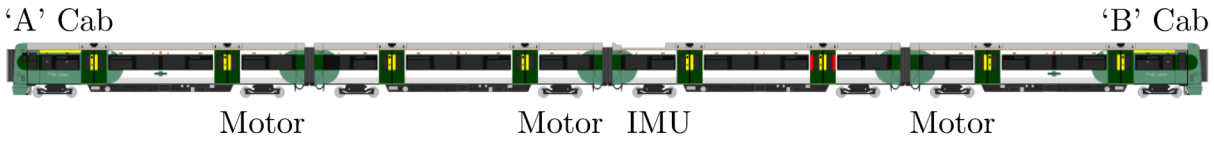
  The non-central longitudinal position of the IMU on the bogie causes one axle to have more influence on the movement experienced by the IMU. Processing so far has assumed that the IMU is centred between the two axles.

- *The lateral position of the IMU on the bogie*

  The non-central lateral position of the IMU on the bogie causes one rail to have more influence on the movement experienced by the IMU than the other.

- *Dynamic effects of the carriage and other coupled carriages*

  The IMU is mounted on a bogie at one end of a carriage near to the middle of the train. Not all bogies of the train are motored. Figure 5.8 shows the positioning of the IMU on the train. This could cause differing forces to be exerted on the IMU bogie depending on the train orientation.



Original image by 'Unisouth'
Reproduced and modified under the Creative Commons Attribution 3.0 Unported license

Figure 5.8: The locations of the 3 motored bogies and the sensorised bogie on the 4-car Electrostar EMU

The first two of these factors can be accounted for in the Kalman filtering stage.

It should be noted that other factors can affect the dynamics of the bogie, which may have an impact the measured geometry. Instrumenting the leading or trailing bogie will have an effect on the dynamics [93], with the leading bogie typically being subject to larger forces. Instrumentation of a motored bogie will also mean that a more heavily loaded track will likely be measured, as the motor equipment on the bogie adds mass. These factors are not considered in this thesis, as they lie outside its scope, however these are only important factors to consider when comparing between IMUs installed on different bogies. All data in this thesis is taken from an IMU in the same location on the same bogie, so this is not an issue.

5.5 Compensating for Longitudinal IMU Position

In Section 5.2, the Kalman filter outputs were defined to be the curvatures experienced by the vertically-sensing accelerometer, κ_z, and the pitch-rate gyro, κ_ϕ. The longitudinal placement of the IMU on the bogie cannot affect the curvature κ_ϕ because at any given time, the pitch is the same at all points on the bogie. Therefore only the output κ_z needs to be adjusted for IMU position.

In equation (5.11) it was assumed that the curvature measured by the accelerometer, κ_z, was the average of the curvatures experienced by the leading and trailing axles, κ_F and κ_T. This equation can be adjusted to take into account the longitudinal position of the IMU.

$$\kappa_z = \alpha\kappa_F + (1-\alpha)\kappa_T \tag{5.44}$$

In this equation, α, is a parameter with a range from 0 to 1 which specifies the position of the IMU between the two axles. A value of 0.5 would represent a central IMU location, whereas $\alpha = 0$ represents a position directly above the trailing axle, and $\alpha = 1$ directly above the leading axle. This parameter must be adjusted according to the orientation of

the train. α_A is used when the 'A' cab is leading the train and α_B is used when the 'B' cab is leading.

$$\alpha_A = \frac{x_B}{D} \tag{5.45}$$

$$\alpha_B = \frac{x_A}{D} \tag{5.46}$$

Where x_A is the distance between the IMU and the axle on the same bogie nearest the 'A' cab of the train. x_B is the distance from the IMU to the axle at the other end of the bogie. D is the distance between the two axles. These distances are indicated in Figure 5.9.

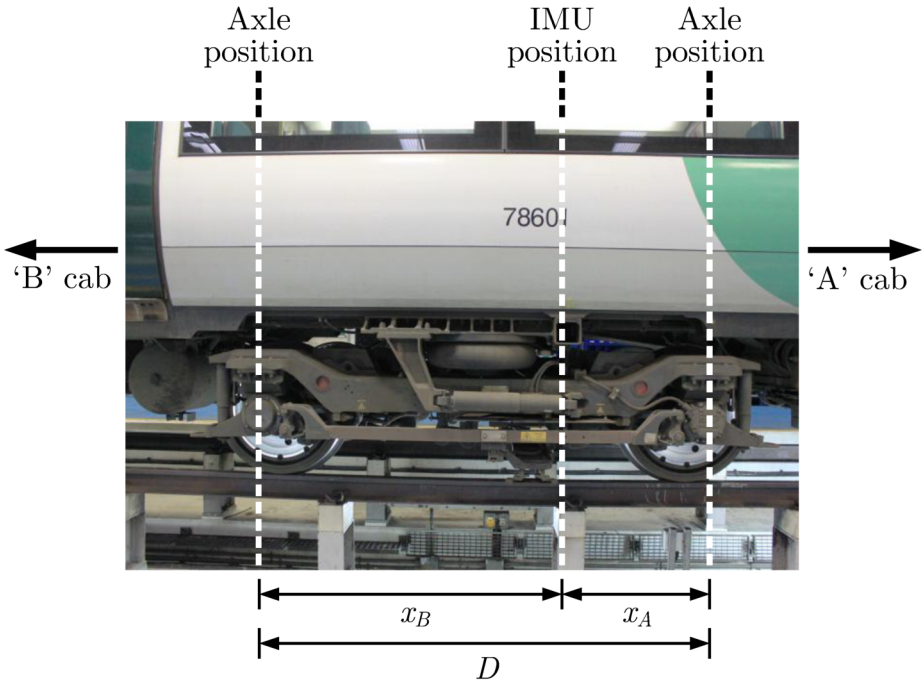


Figure 5.9: Longitudinal distances on the instrumented bogie

Since $x_B = D - x_A$, it can also be said that:

$$\alpha_B = 1 - \alpha_A \tag{5.47}$$

Using the modification made in (5.44), the approximation of curvature, κ_z can be updated to include α:

$$\kappa_z \simeq \frac{\alpha(z_1 - 2z_2 + z_3) + (1-\alpha)(z_{21} - 2z_{22} + z_{23})}{\delta x^2} \tag{5.48}$$

From this, the observation matrix, C, can also be updated:

$$C = \begin{bmatrix} \frac{\alpha}{\delta x^2} & \frac{-2\alpha}{\delta x^2} & \frac{\alpha}{\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{1-\alpha}{\delta x^2} & \frac{2\alpha-2}{\delta x^2} & \frac{1-\alpha}{\delta x^2} \\ \frac{1}{40\delta x^2} & 0 & \frac{-1}{40\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{-1}{40\delta x^2} & 0 & \frac{1}{40\delta x^2} \end{bmatrix} \tag{5.49}$$

The following values were obtained through measurement of the bogie:

$D = 2.60$ m

$x_A = 0.87$ m

Using these measurements values can be found for α_A and α_B:

$\alpha_A = 0.67$

$\alpha_B = 0.33$

5.5.1 Results After Longitudinal Position Compensation

The data used in Figure 5.7 were reprocessed using the updated observation matrix and the calculated values of α_A and α_B. Figure 5.10 shows the results.

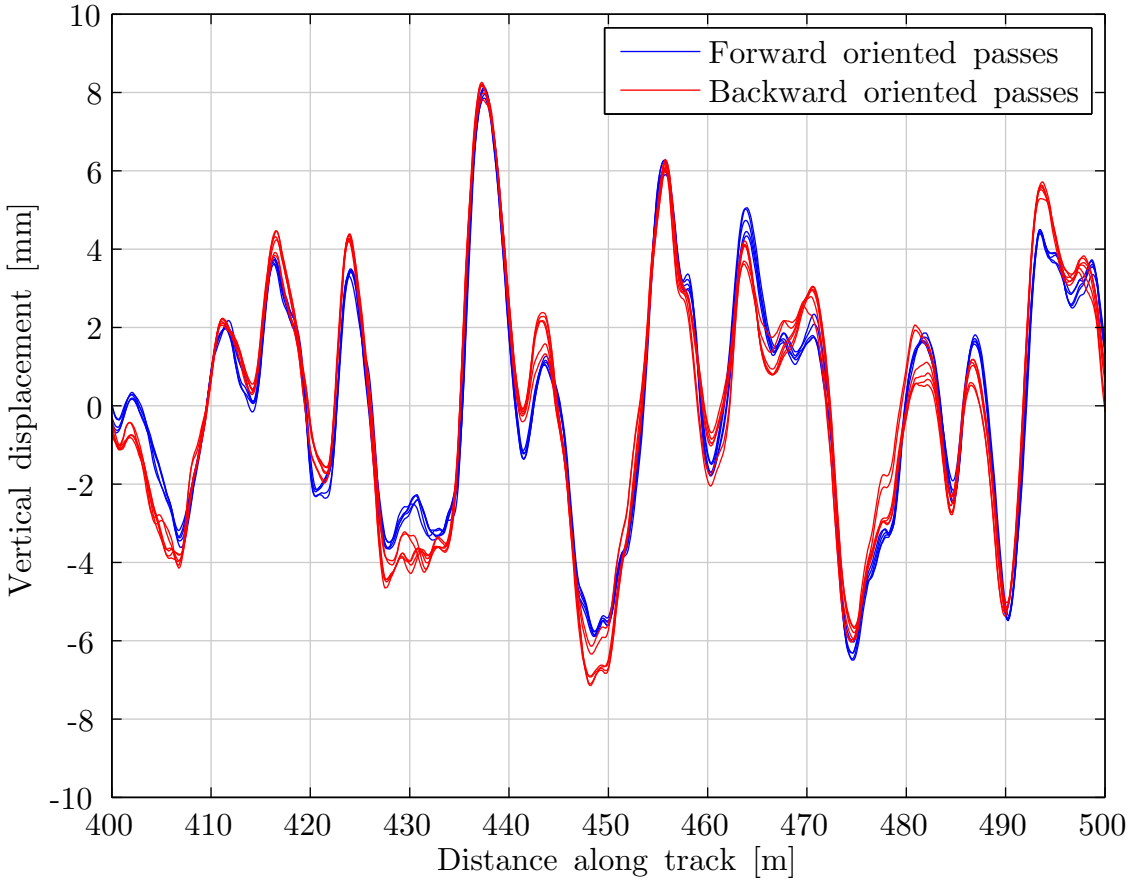


Figure 5.10: Vertical displacements for multiple train passes after compensation for longitudinal IMU position

It can be seen from Figure 5.10 that the differences between differently oriented passes have been reduced when compared to the differences in Figure 5.7. The differences around 437 m and 456 m are now negligible, however, many of the other peaks and troughs still contain noticeable differences of up to 1 mm.

5.6 Compensating for Lateral IMU Position

The second factor identified in Section 5.4 which can affect the difference between forward- and backward-oriented passes, was the lateral position of the IMU on the bogie. A non-central lateral IMU position causes the IMU to measure the influence of one of the running rails more than the other. If the orientation of the vehicle is reversed, the IMU will be influenced more by the opposite running rail.

This can be compensated for by further expansion of the Kalman filter. Previously, 23 state variables were used to represent the vertical displacements of the discrete locations beneath the bogie. These displacements were assumed to be the average height of the two running rails. By expanding the filter to 46 state variables, the displacements of both rails can be tracked at all 23 discrete locations.

There are two sensors which experience the difference in height between running rails. The first of these is the roll-rate gyro. This measures the rate of change of angle across the bogie, $\dot{\theta}$. This can be incorporated into the Kalman filter by adding a third output to the filter. Secondly, the vertically-sensing accelerometer is influenced by one rail more than the other depending on the orientation of the train. The first row of the observation matrix, C_z, can be adjusted to take the difference in rail heights into account.

5.6.1 Adding State Variables

As described previously, the filter is expanded to use 46 state variables; 23 for each rail. The state variables are referred to as x_{L1} to x_{L23} for the left running rail, and x_{R1} to x_{R23} for the right running rail. This configuration is shown in Figure 5.11. In order to

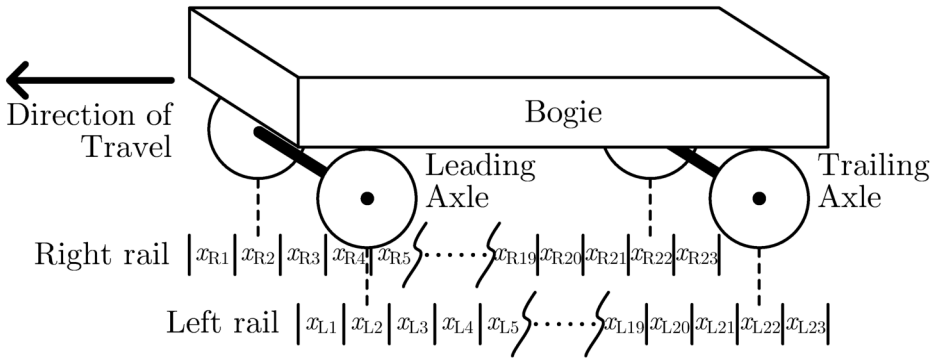


Figure 5.11: State variables representing discrete track locations on two rails beneath a bogie

implement the left and right rails in the filter, the existing 23×1 column vector, $\hat{\boldsymbol{x}}_k$, which contains the current state estimate is widened to 46 elements.

$$\hat{\boldsymbol{x}}_k = \left[\begin{array}{ccccc||ccccc} \hat{x}_{k|L1} & \hat{x}_{k|L2} & \cdots & \hat{x}_{k|L22} & \hat{x}_{k|L23} & \hat{x}_{k|R1} & \hat{x}_{k|R2} & \cdots & \hat{x}_{k|R22} & \hat{x}_{k|R23} \end{array} \right] \tag{5.50}$$

The double vertical line is used here to represent the middle of the vector.

The state transition matrix, A, must be updated to take into account the widening of $\hat{\boldsymbol{x}}_k$. The new 46×46 state transition matrix is now as follows:

$$A = \left[\begin{array}{cccccc||cccccc} 1 & 0 & 0 & \cdots & 0 & 0 & & & & & & \\ 1 & 0 & 0 & \cdots & 0 & 0 & & & & & & \\ 0 & 1 & 0 & \cdots & 0 & 0 & & & & & & \\ 0 & 0 & 1 & \cdots & 0 & 0 & & & 0 & & & \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & 1 & 0 & & & & & & \\ \hline\hline & & & & & & 1 & 0 & 0 & \cdots & 0 & 0 \\ & & & & & & 1 & 0 & 0 & \cdots & 0 & 0 \\ & & & & & & 0 & 1 & 0 & \cdots & 0 & 0 \\ & & 0 & & & & 0 & 0 & 1 & \cdots & 0 & 0 \\ & & & & & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & & & & & & 0 & 0 & 0 & \cdots & 1 & 0 \end{array}\right] \tag{5.51}$$

Again, the double lines here represent the horizontal and vertical center lines of the matrix. The upper-left and lower-right quadrants of the matrix are identical to the previous version of A from (5.4), which specifies that on each transition of k the bogie moves forward by one position. The prediction for the new position ahead of the bogie is the same as the value found for the previous front position. The upper-left quadrant specifies that this is the case for the left rail, and the lower-right for the right rail. The lower-left and upper-right quadrants specify how the left rail states affect the right rail states, and vice-versa. In this model there is no interaction between the two rails. It could be said that the front (unknown) positions of the two rails have some affect on each other, as the rail is unlikely to have a large difference between rail heights (cross-level). This effect is ignored, as high-pass filtering after the Kalman filter stage removes any drift between the rails which may represent an unlikely or impossible cross-level.

The process noise covariance, Q, is updated from (5.5) in a similar fashion so that it measures 46×46 elements:

$$Q = \left[\begin{array}{cccc||cccc} w & 0 & \cdots & 0 & & & & \\ 0 & 0 & \cdots & 0 & & & & \\ \vdots & \vdots & \ddots & \vdots & & & 0 & \\ 0 & 0 & \cdots & 0 & & & & \\ \hline\hline & & & & w & 0 & \cdots & 0 \\ & & & & 0 & 0 & \cdots & 0 \\ & & 0 & & \vdots & \vdots & \ddots & \vdots \\ & & & & 0 & 0 & \cdots & 0 \end{array}\right] \tag{5.52}$$

Q specifies that the front positions of each rail are the only unknown values in the system, which are therefore given process noise, w.

Finally the observation matrix, C, must be expanded. With the previous (single-input) filter, the sensor measurements were affected by the vertical rail displacement, which was assumed to be the average height of the two running rails. With the two-rail filter, both rails affect the measurements of the sensors. The roll-rate gyro observations are also to be added to the filter. The resultant observation matrix must therefore be 3×46. Changes to each row of the observation matrix are now considered.

5.6.2 Modifying the Accelerometer Response

As the IMU is not mounted centrally to the bogie in either the longitudinal or lateral position, measurements from the vertically-sensing accelerometer are affected by changes in both the pitch and roll of the bogie.

Equation (5.10), shows how the curvature experienced by one axle of the bogie is affected by the vertical displacement of the track. After compensation for the longitudinal position of the IMU on the bogie, equation (5.44) was introduced which defines how much the

displacement of each axle affects the vertical curvature experienced by the vertically-sensing accelerometer. A similar equation can be formulated for the influence of the lateral position on the experienced curvature from the vertical curvature of the left and right running rails, κ_L and κ_R respectively:

$$\kappa_{z\theta} = \beta\kappa_R + (1-\beta)\kappa_L \tag{5.53}$$

Here, the parameter β is introduced, which specifies the lateral position of the IMU on the bogie. β has a range from 0 to 1, with $\beta = 0$ representing an IMU position directly above the left rail (in the direction of travel), and $\beta = 1$ representing a position above the right rail. As before, two values of β are defined, one for when the 'A' cab is leading the train, and the other for the 'B' cab:

$$\beta_A = \frac{y_{LA}}{G} \tag{5.54}$$

$$\beta_B = \frac{y_{LB}}{G} \tag{5.55}$$

Here, y_{LA} is the lateral distance from the left wheels of the bogie when the 'A' cab is leading the train, and y_{LB} is the distance from the left wheels when the 'B' cab is leading the train. G is the distance between the contact points of the two wheels of an axle on each rail. This is a value slightly larger than the gauge of the track. In the case of the British rail network, the gauge is 1.435 m, so the value used is $G = 1.48$ m, which includes an approximate value for the distance from the gauge face of the rail to the contact area for both rails. Through measurement, y_{LA} was found to be 0.28 m.

In a similar manner to longitudinal displacements, it is known that $y_{LB} = G - y_{LA}$, so it can also be said that:

$$\beta_B = 1 - \beta_A \tag{5.56}$$

Using the known values of y_{LA} and G, the values of β_A and β_B are found to be:

$$\beta_A = 0.19$$

$$\beta_B = 0.81$$

Equations (5.44) and (5.53) can be combined to take into account both the longitudinal and lateral IMU positions. Here, four curvature inputs are used; one from each wheel. These are leading-left, κ_{FL}, leading-right, κ_{FR}, trailing-left, κ_{TL}, and trailing-right, κ_{TR}.

$$\kappa_z = \alpha\beta\kappa_{FR} + (1-\alpha)\beta\kappa_{TR} + \alpha(1-\beta)\kappa_{FL} + (1-\alpha)(1-\beta)\kappa_{TL} \tag{5.57}$$

By combining this with the approximation of vertical curvature (5.10), the following can be deduced:

$$\kappa_z \simeq \frac{1}{\delta x^2}\Big[\alpha\beta(z_{R1} - 2z_{R2} + z_{R3}) + (1-\alpha)\beta(z_{R21} - 2z_{R22} + z_{R23}) + \alpha(1-\beta)(z_{L1} - 2z_{L2} + z_{L3}) + (1-\alpha)(1-\beta)(z_{L21} - 2z_{L22} + z_{L23})\Big] \tag{5.58}$$

This can be used to update the first row of the observation matrix, C_z, which has 1×46 elements:

$$C_z = \Bigg[\begin{array}{ccccccc} \frac{\alpha(1-\beta)}{\delta x^2} & \frac{-2\alpha(1-\beta)}{\delta x^2} & \frac{\alpha(1-\beta)}{\delta x^2} & 0 & 0 & \cdots & \\ \cdots & 0 & 0 & \frac{(1-\alpha)(1-\beta)}{\delta x^2} & \frac{-2(1-\alpha)(1-\beta)}{\delta x^2} & \frac{(1-\alpha)(1-\beta)}{\delta x^2} & \Big\| \\ \frac{\alpha\beta}{\delta x^2} & \frac{-2\alpha\beta}{\delta x^2} & \frac{\alpha\beta}{\delta x^2} & 0 & 0 & \cdots & \\ \cdots & 0 & 0 & \frac{(1-\alpha)\beta}{\delta x^2} & \frac{-2(1-\alpha)\beta}{\delta x^2} & \frac{(1-\alpha)\beta}{\delta x^2} & \end{array}\Bigg] \tag{5.59}$$

This matrix row has been split onto four lines on the page. The double vertical line represents the center of the row, with dots representing an in-fill of zeros.

Consideration is now given to the remaining two rows of the observation matrix.

5.6.3 Modifying the Pitch-Rate Gyroscope Response

The pitch-rate gyroscope does not experience any changes due to roll. Consequently its impulse response, from (5.20), remains unchanged. Each discrete element, z_n, can be considered to be the average of the discrete heights of the two running rails:

$$z_n = \frac{z_{Ln} + z_{Rn}}{2} \tag{5.60}$$

The resultant row of the observation matrix is:

$$C_\phi = \left[\begin{array}{cccccccccc|cccccccccc} \frac{1}{80\delta x^2} & 0 & \frac{-1}{80\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{-1}{80\delta x^2} & 0 & \frac{1}{80\delta x^2} & \| & & & & & & & & \\ & & & & & & & & & & & \| & \frac{1}{80\delta x^2} & 0 & \frac{-1}{80\delta x^2} & 0 & 0 & \cdots & 0 & 0 & \frac{-1}{80\delta x^2} & 0 & \frac{1}{80\delta x^2} \end{array} \right] \tag{5.61}$$

Again, the double vertical line is used to indicate the middle of the vector. This matrix is essentially two copies of (5.20); one for the left rail and one for the right. Each element has been divided by 2 to achieve the average described by (5.60).

5.6.4 Adding the Roll-Rate Gyroscope

The roll-rate gyro measures the rate of change of angle around the x axis, $\dot{\theta}$. As the Kalman filter operates in the spatial domain, the roll-rate is converted to the spatial domain, using the same method used for the pitch rate gyro in (4.6).

$$\kappa_\theta = \frac{\dot{\theta}}{v_x} \tag{5.62}$$

κ_θ is the curvature around the x axis, travelling along the x axis. As the axes are the same, the curvature can be thought of as a twist motion, rather than curves as with vertical or horizontal track curvature. Fundamentally, κ_θ represents the change of θ with respect to distance travelled in the x axis; $\frac{\mathrm{d}\theta}{\mathrm{d}x}$.

The approximation in Appendix A.2 can again be used to show how the angle, θ, is related to the difference in vertical displacements between the rails.

$$\theta \simeq \frac{\Delta z}{G} \tag{5.63}$$

Where Δz is the difference between the vertical displacements of the two rails.

A similar system of three angles used when considering the the pitch-rate gyro (Figure 5.4) can be applied to the roll-rate. The difference is that the angles are calculated between parallel pairs of displacements. For symmetry, these are averaged between leading and trailing axle pairs

$$\theta_1 \simeq \frac{1}{2}\left(\frac{z_{R1} - z_{L1}}{G} + \frac{z_{R21} - z_{L21}}{G}\right) \tag{5.64}$$

$$\theta_2 \simeq \frac{1}{2}\left(\frac{z_{R2} - z_{L2}}{G} + \frac{z_{R22} - z_{L22}}{G}\right) \tag{5.65}$$

$$\theta_3 \simeq \frac{1}{2}\left(\frac{z_{R3} - z_{L3}}{G} + \frac{z_{R23} - z_{L23}}{G}\right) \tag{5.66}$$

And using (5.7), it can be deduced that:

$$\kappa_\theta = \frac{\theta_3 - \theta_1}{2\delta x} \tag{5.67}$$

And therefore:

$$\kappa_\theta \simeq \frac{z_{R3} - z_{L3} + z_{R23} - z_{L23} - z_{R1} + z_{L1} - z_{R21} + z_{L21}}{4G\delta x} \tag{5.68}$$

This equation can now be used to create a new 1×46 row for the observation matrix, C_θ, which describes how the vertical displacements of each rail affect the curvature calculated from the roll-rate gyro.

$$C_\theta = \left[\begin{array}{cccccccccc|cccccccccc} \frac{1}{4G\delta x} & 0 & \frac{-1}{4G\delta x} & 0 & 0 & \cdots & 0 & 0 & \frac{1}{4G\delta x} & 0 & \frac{-1}{4G\delta x} \\ \frac{-1}{4G\delta x} & 0 & \frac{1}{4G\delta x} & 0 & 0 & \cdots & 0 & 0 & \frac{-1}{4G\delta x} & 0 & \frac{1}{4G\delta x} \end{array} \right] \quad (5.69)$$

Again, the double vertical line is used here to indicate the middle of the vector.

5.6.5 Completing the New Kalman Filter

The three new rows of the observation matrix are combined to form the finished observation matrix, in a similar manner to that used for (5.22). The complete matrix is not shown here for reasons of size, but the matrix is formed of the three rows identified in Sections 5.6.2 to 5.6.4 as follows:

$$C = \begin{bmatrix} C_z \\ C_\phi \\ C_\theta \end{bmatrix} \quad (5.70)$$

The final component to consider is the observation noise covariance, R. The elements for the vertically-sensing accelerometer and pitch-rate gyro remain the same as found previously in (5.27), but an element must be added for the observation noise of the roll-rate gyro. When error is considered in the calculation of roll curvature, it is found that:

$$\kappa_\theta = \frac{\dot{\theta} + e_\theta}{V_x} \quad (5.71)$$

Where e_θ is the roll-rate gyro error. The observation noise covariance matrix can now be expanded to include this element:

$$R(k) = \begin{bmatrix} \left(\frac{e_z(k)}{V_x^2(k)}\right)^2 & 0 & 0 \\ 0 & \left(\frac{e_\phi(k)}{V_x(k)}\right)^2 & 0 \\ 0 & 0 & \left(\frac{e_\theta(k)}{V_x(k)}\right)^2 \end{bmatrix} \tag{5.72}$$

5.6.6 Results After Two-Dimensional Position Compensation

The data used in Figures 5.7 and 5.10 were again reprocessed using the updated filter. Figure 5.12 shows the results. It can be seen that there is a large overall reduction in the differences caused by the vehicle orientation. These differences are mostly within an error margin of 0.5 mm from pass to pass in a single orientation.

There are some parts of the graph where some backward oriented passes differ from others, for example at ~449 m and ~478 m. After inspection it was found that these differences were introduced through differing curvatures calculated from the vertically-sensing accelerometer. As this is the sensor most likely to be affected by the speed of the train, the vehicle speed for each pass was plotted in Figure 5.13. In this figure, the two passes identified to be outliers in Figure 5.12 are highlighted. It can be seen that the speed of these passes are both around 19 ms^{-1}, whereas the speeds of all other passes fall between 21.5 ms^{-1} and 24.5 ms^{-1}.

When vertical acceleration, $\ddot{z}$, is converted to vertical curvature, κ_z, as in (4.2), the resulting curvature is theoretically unaffected by changes in the vehicle speed, V_x. In practice, the primary suspension of the bogie causes the bogie to be excited more by certain frequencies of vertical displacement. Traversal of a specific track feature at one speed will result in a different vertical displacement frequency to traversal of the same feature at another speed. Consequently some features may cause more bogie excitation at certain speeds. This is the effect which was seen in the outlying passes of Figures 5.12 and 5.13.

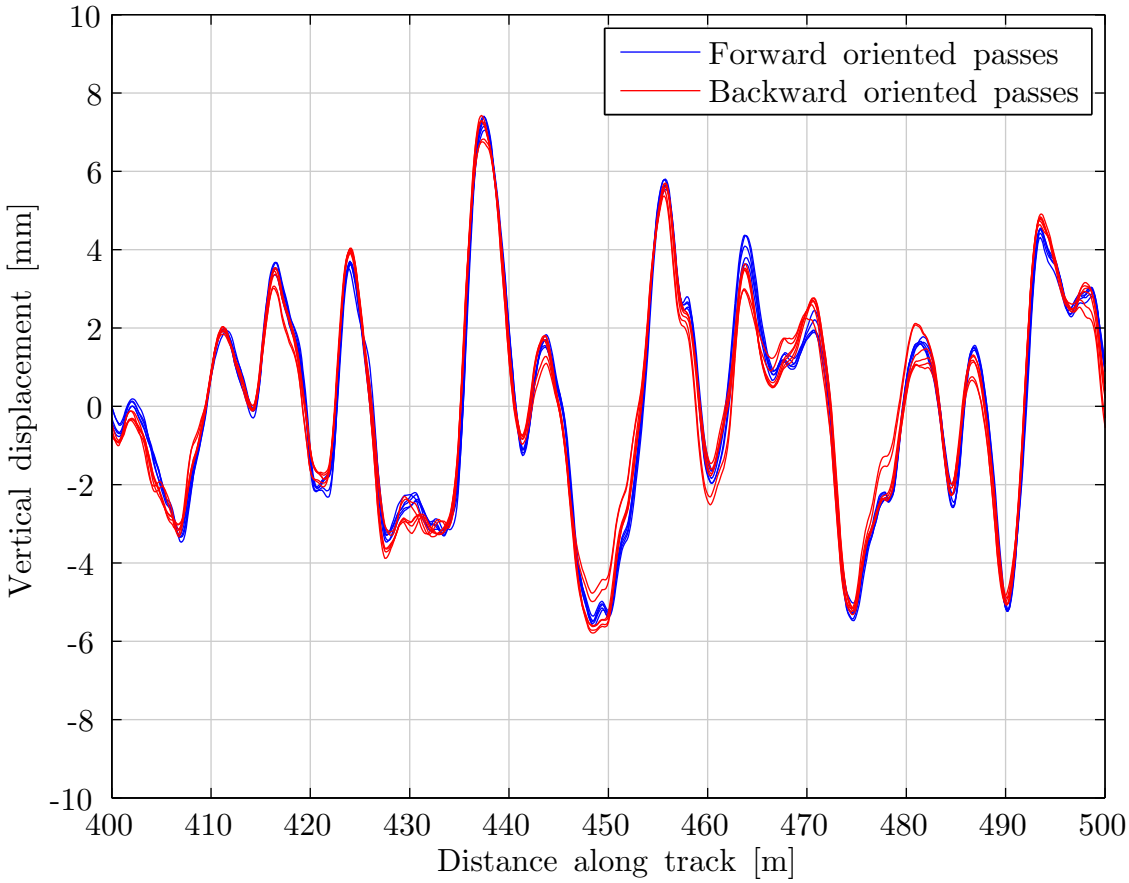


Figure 5.12: Vertical displacements for multiple train passes after compensation for 2-dimensional IMU position

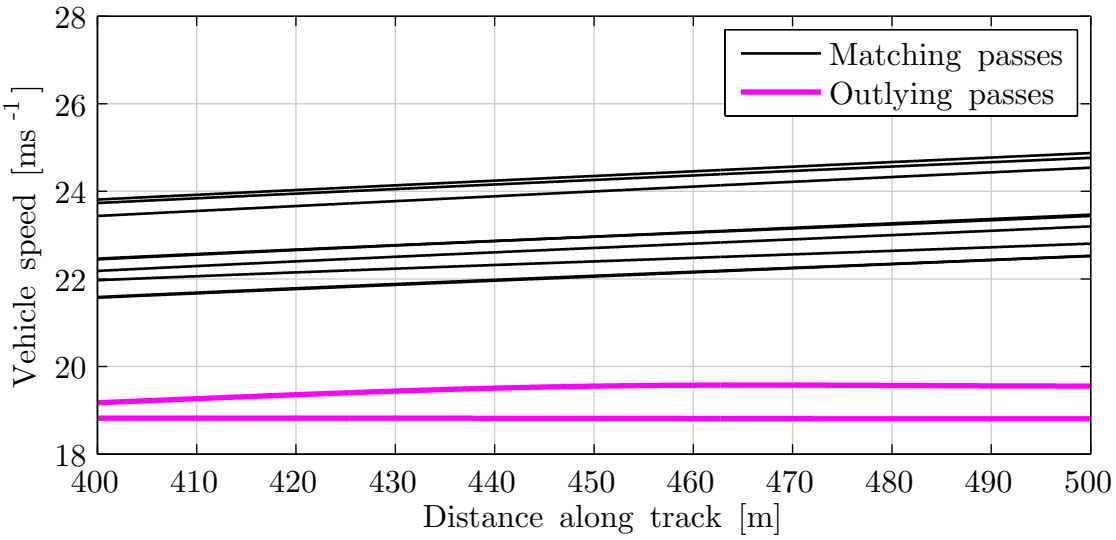


Figure 5.13: Vehicle speed for multiple passes, showing two outlying passes

5.7 Further Work for Orientation Compensation

In Section 5.4, a third factor was identified which may contribute to the differences in vertical displacement introduced by differing vehicle orientations.

It was decided that the results after compensation for longitudinal and lateral IMU position were accurate enough for the purposes of further work in this thesis. It would be possible to model the dynamics of the train, and compensate for this factor. This is a broad area of work which could form the basis of further research in this area.

It could also be possible to remove variations caused by vehicle speed, as found in Section 5.6.6, by modelling the dynamics of the primary suspension of the bogie. Once again it was decided not to pursue this, in favour of completing further work in other areas.

It should be noted that there is a possibility that the true geometry of the track changed throughout the time over which the repeated passes were recorded. All passes shown in Figures 5.7 to 5.13 were recorded within a 1 month time period, and no known maintenance took place during that time. This means that the variation should be minimal. Throughout the time period, there are a mixture of forward and backward oriented passes, meaning that these differences are highly unlikely to be due to major changes in the true track geometry.

CHAPTER 6

ALIGNING AND COMPARING PROCESSED DATA

Now that vertical displacement data can be calculated in a repeatable form (to within around 0.2 mm if passes of the same orientation are chosen), data from multiple passes of the same section of track can be compared to provide information about changes in the state of the track over time.

Before comparisons can be made, data must be mutually aligned in the spatial domain. Usually data extracted from the data management software is aligned within around ±20 m based on the GPS position of the vehicle. This is inadequate for numeric comparisons between data sets which would require data to be aligned to the nearest sample; in this case the nearest 0.125 m.

6.1 Improving the Accuracy of Train Positioning

A great deal of work is ongoing into improving the accuracy of location data collected from railway vehicles. A pair of journal papers by Saab [94,95] published in 2000 describe an algorithm developed to use a yaw-rate gyro and wheel tachometer to keep track of the train's position on a digital map of the railway network. In his work, Saab recognises the importance of good decision making when choosing 'candidate' track sections from the map (*i.e.* track sections where the train is most likely to be located).

More recent research at the German Aerospace Center (DLR) [96] has extended this approach to use a Bayesian filter, allowing a probabilistic approach to the selection of track sections, as well as positioning along them. This allows erroneous decisions to be corrected later. For example if the train passes a set of points, and the algorithm chooses the incorrect route, the Bayesian filter will continue to track both possibilities until one of them becomes improbable, for example a curve is experienced which only exists on one route.

Further research at the DLR has used tight integration of the GPS, inertial sensors, and digital map to generate more accurate positional information [97]. A normal 'loosely coupled' system would pre-calculate a GPS location, usually within the GPS receiver itself, and then use inertial sensors to provide corrections to the location. The 'tightly coupled' system used by Crespillo *et al.* integrates inertial sensor measurements and track sections from the map, along with the raw GPS satellite range data. This allows a much better tracking of the vehicle's position, and has the advantage of being able to use GPS information even if there are too few 'visible' satellites to calculate a location in the normal way.

The tight integration method also has an advantage that the Kalman filter parameters used to calculate GPS position can be re-defined to suit a railway vehicle. A typical off-the-shelf GPS receiver will be suited to a road vehicle, which has a turning circle of a few metres. Adjustments could be made so that tight turns and steep sloped are not a possibility when considering a train-based GPS system.

Whilst this research clearly has an application within the alignment of data collected from multiple passes of track sections within a railway network, the focus of this thesis is on the data collected itself. It was decided not to develop a complex positioning system, rather to focus on mutually aligning data sets known to be within around ±20 m of the correct position, and travelling on the same track where there are multiple parallel tracks. Usually the track sections selected for study are double-track sections, where there is

one track in each direction of travel. This makes the selection of track easy, as only one direction of travel (*e.g.* westbound) is selected. The data management software described in Section 4.1.1 can be used to find multiple passes of the train through a location of interest, and in a specific direction.

6.2 Aligning Multiple Vehicle Passes

6.2.1 Manual Alignment

For the mutual alignment of a pair of data sets over a short length of track (≤ 500 m), alignment can be performed by hand, adjusting the x position of one of the data sets until both sets are visibly aligned. The data sets can then be trimmed so that only the portion of track where data is available in both data sets remains.

Clearly this is impractical and time consuming for aligning many data sets. Another issue arising when aligning data sets over large distances is that they may drift in and out of alignment due to errors in position and speed measurements. Some differences in alignment can be corrected by a single factor throughout, for example where the wheel diameter has reduced by wheel turning, causing a difference in wheel circumference between passes. This affects the distance calculation derived from the tacho.

A graphical user interface (GUI) was developed (Figure 6.1) to make manual alignment of data sets faster. This allows one data set to be used as a reference, and other data sets to be dragged into alignment over the top. Markers can be placed on the waveform, allowing parts of it to be stretched or shortened. This allows any areas of the waveform drifting out of alignment to be pulled back into alignment.

Whilst the use of a GUI significantly reduces the time taken to align data sets, an automated method would be preferable and would ultimately enable full automation of the comparison process. This would be required in an automatic fault detection and monitoring system.

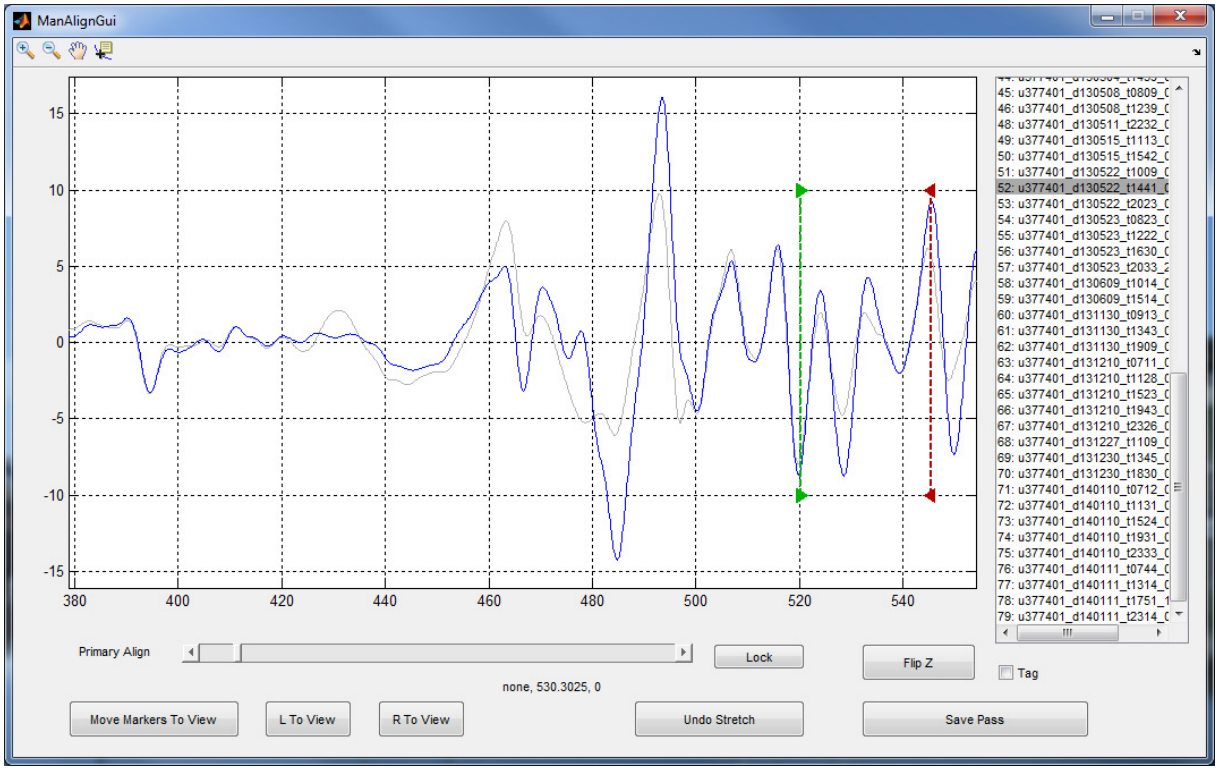


Figure 6.1: Manual Alignment GUI

6.2.2 Fixed-Shift Automated Alignment

Discrete Comparison Method

Initially, a basic method of aligning two passes of the same track was developed. This is based on a discrete cross-correlation of the two data sets. One data set is shifted to a range of positions in the x direction, and for each shift value, ℓ, the difference between each set is taken. An operation is performed on the result to give the a score at that position, q_ℓ. The value of ℓ giving the highest value of q_ℓ is selected as the best shift, $\ell_{\max}$.

The score for each shift value, ℓ, in the alignment range is obtained using the following equation:

$$\frac{1}{q_\ell} = \frac{1}{N} \sum_{i=K_1}^{K_N} |\boldsymbol{z}_1(i) - \boldsymbol{z}_2(i + \ell/\delta x)| \tag{6.1}$$

Here, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ are the two data sets, i is the sample number, ℓ is the applied offset in

metres, and δx is the sample spacing in metres. K is an N-long set of sample indices which would be valid in both $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$.

An example of the scores given to each shift value is shown in Figure 6.2. In this figure an alignment range of ±350 m is used. It should be noted that the y-axis units are arbitrary; the output is used only to determine the shift value with the maximum score.

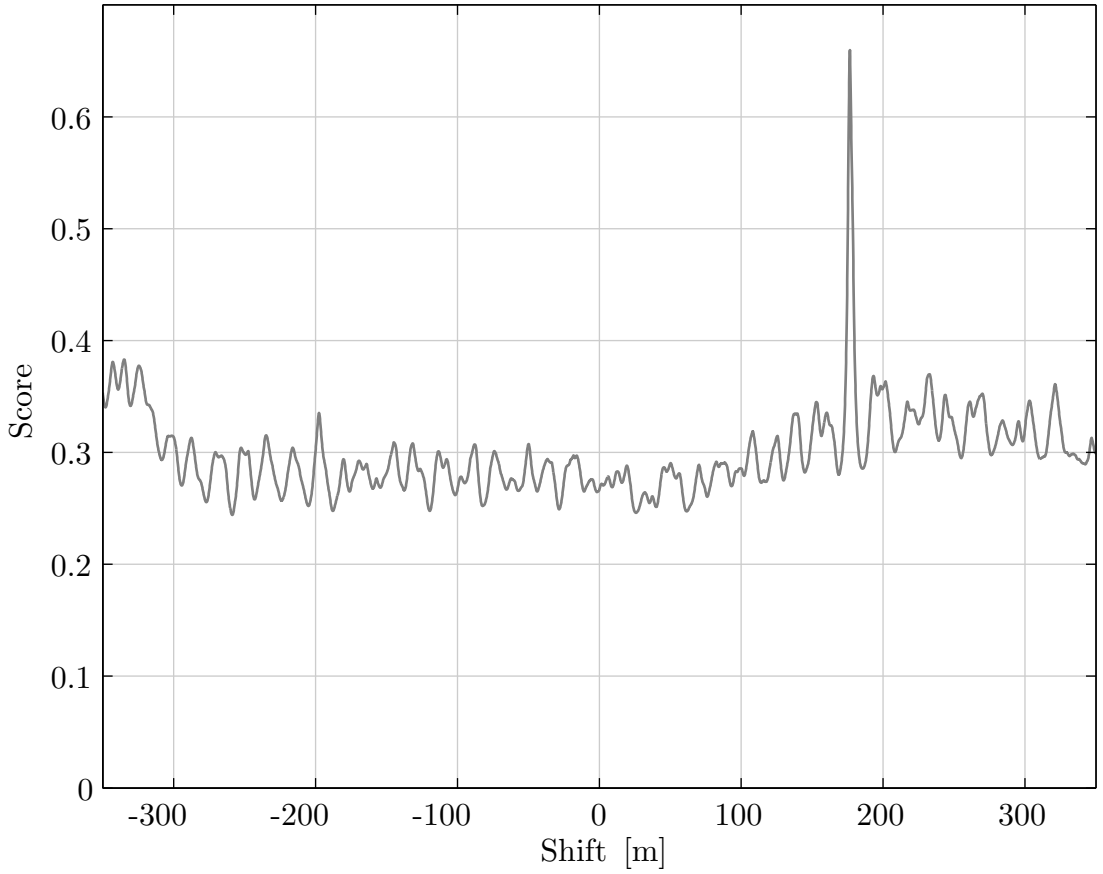


Figure 6.2: Example scores for a discrete comparison alignment between two data sets

In this figure, each value of ℓ is shown on the x axis, with q_ℓ on the y axis. It can be seen that in this example, the highest score, $\ell_{\max}$ occurs at 177 m. This is the offset value which must be applied to the second data set to best align it with the first.

$$\boldsymbol{x}_{2\ \text{aligned}} = \boldsymbol{x}_2 + \ell_{\max} \tag{6.2}$$

Where $\boldsymbol{x}_2$ is a vector containing the x values in the second data set.

With the fixed-shift auto-alignment algorithm it is not necessary to resample the data after the shift is applied, as the shift is constant throughout the data set, and is aligned to the nearest sample of the first data set. This means that in this scenario, samples will remain aligned to locations which are multiples of 0.125 m, which is the delta used throughout the processing. As stated in the requirements, this was chosen to give an accurate representation of geometry features as small as 1 m in length.

FFT Cross-Correlation Method

The previous method operated through comparison of every sample at every shift value in the chosen range. This makes it a relatively slow method of comparison. A faster method is to use a classical cross-correlation method. Cross-correlation may be implemented using Fast Fourier-Transforms (FFTs) [98]. The cross-correlation of two vectors can be found by calculating the product of their respective Fourier transforms, with one of the transforms conjugated.

$$\mathcal{F}\{\boldsymbol{z}_1 \star \boldsymbol{z}_2\} = \mathcal{F}\{\boldsymbol{z}_1\}.(\mathcal{F}\{\boldsymbol{z}_2\})^* \tag{6.3}$$

Where $\mathcal{F}$ denotes a Fourier transform, and * denotes the complex conjugate. The five-point star ($\star$) operator indicates a cross-correlation between two vectors.

Implementation of the cross-correlation method using FFTs proved to be significantly faster than the discrete comparison method explored previously. On the computer used for development, the discrete comparison method executed 30 times in 28.89 seconds aligning within limits of ±350 m of track. The FFT method executed 30 times in 0.16 seconds to align within all possible shifts; approximately 180 times faster than the discrete comparison method.

Figure 6.3 shows the result of the cross-correlation using the same data sets as used in Figure 6.2. Once again the y-axis units are arbitrary.

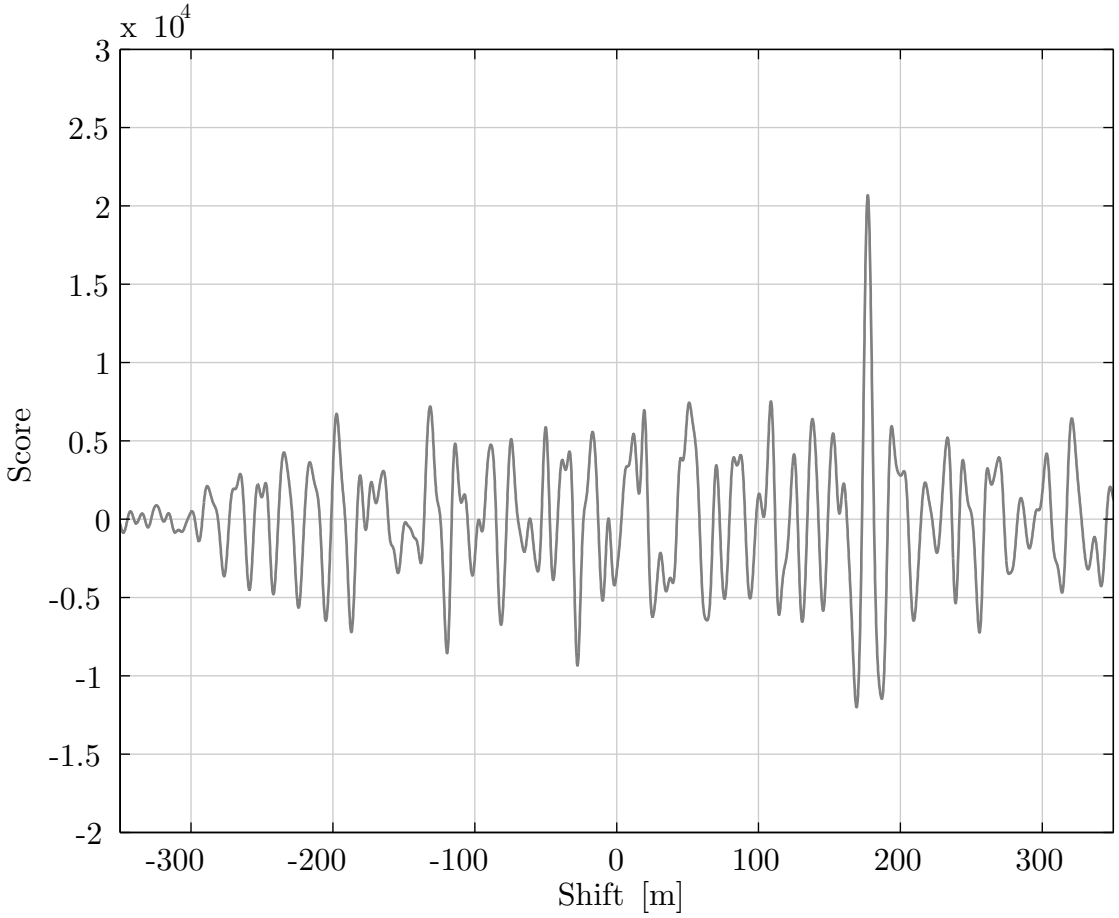


Figure 6.3: Example scores for an FFT cross-correlation between two data sets

It can be seen that the cross-correlation method also identifies the aligned shift value as 177 m. Upon comparison of Figures 6.2 and 6.3, it can be seen that the discrete comparison method gives a much more distinctive peak at the aligned shift value. This means that it is more likely to be able to produce a correct alignment when two signals vary more greatly. However, the significant speed increase afforded by the FFT method outweighs this benefit, particularly as it is unlikely that different passes of the same track will produce significantly varied results throughout their entire length.

6.2.3 Dynamic-Shift Automated Alignment

Both fixed-shift alignment methods investigated in the previous section are equivalent to manual alignment through modification of the x position of one data set. As with the

manual method, the fixed-shift alignment is acceptable for a short section of track, but positional errors may cause longer sections to drift in and out of alignment. As well as this, even disagreements correctable by a single factor can confuse the cross-correlation process, giving poor suggestions for $\ell_{\max}$. Figure 6.4 shows two vertical displacement data sets recorded on different days, which have been manually aligned. It can be seen that the data sets begin well-aligned, but at around 300 m the first set drifts behind the second. However, by 380 m the first set is ahead of the second. A method is required to correct alignment drifts such as these.

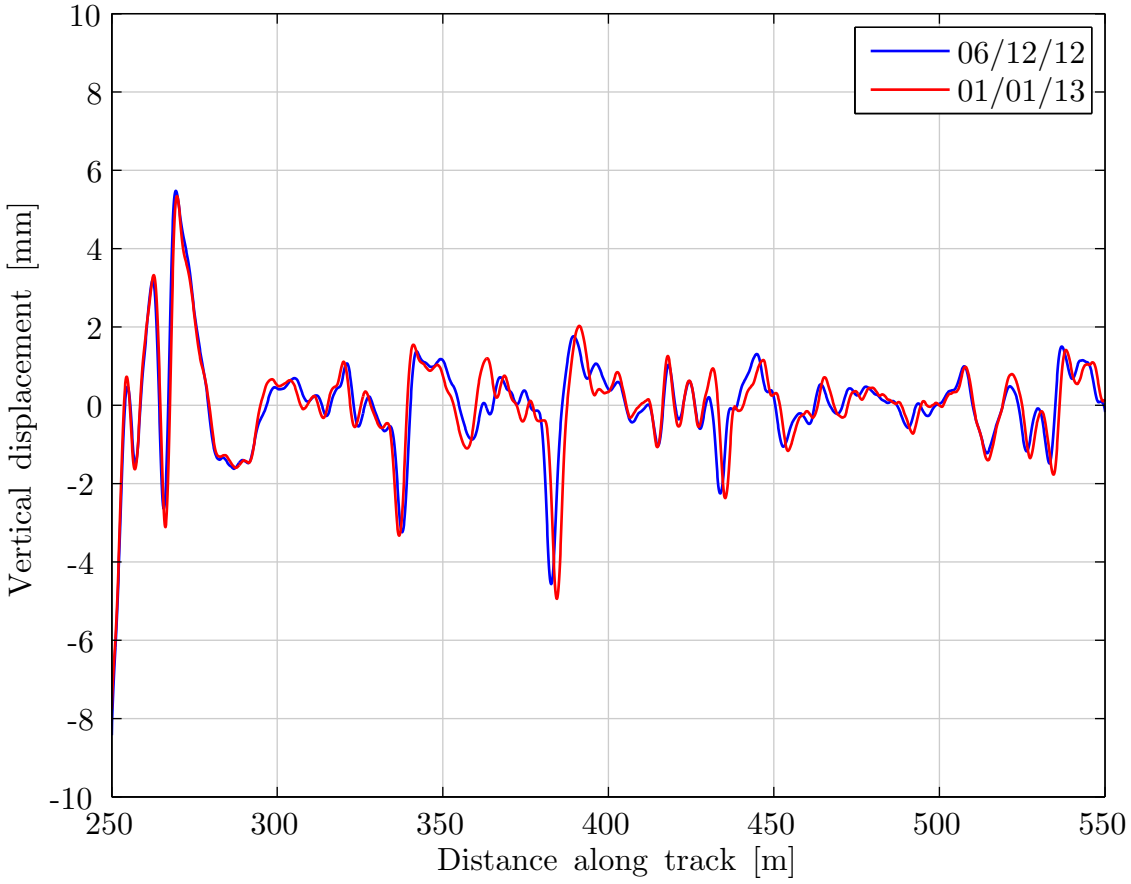


Figure 6.4: Two vertical displacement data sets, showing alignment drift

It should be noted that normally, drift tends to occur over several kilometres, rather than the tens of metres shown in this example. This is an extremely small-scale example, used to better illustrate the effects of dynamic alignment.

A dynamic auto-alignment method splits one waveform into windows, the length of which can be adjusted according to application, but for these purposes were chosen to be a length of 50 m. Each window, w, is aligned to the other waveform in order to determine a shift value, $\ell_{\max w}$, for that window. Stitching these offset windows back together with differing values of $\ell_{\max w}$ applied to each, could easily result in sharp transitions between windows, which are not representative of the true track displacement. For this reason, each sample of the entire dataset is assigned with an ℓ parameter which is calculated by drawing a smoothed line through the values of $\ell_{\max w}$ for every window. For the example in Figure 6.4, the smoothed offsets are shown in Figure 6.5.

Finally, the assigned ℓ values are applied to the second waveform's x values, and resampling is performed so that samples are once again evenly spaced at 8 Sm^{-1}. Figure 6.6 shows the result of alignment after the dynamic-shift alignment algorithm is performed.

The dynamic-shift alignment method has a clear advantage over fixed-shift methods, in that it allows alignment of track over many kilometres of track, even when longitudinal track position is not consistently accurate. The downside to this method is that it requires a larger number of alignments due to the windowing method. However, if the FFT cross-correlation method is used to align each window, rather than the discrete comparison method, the alignment will still execute within an acceptable length of time.

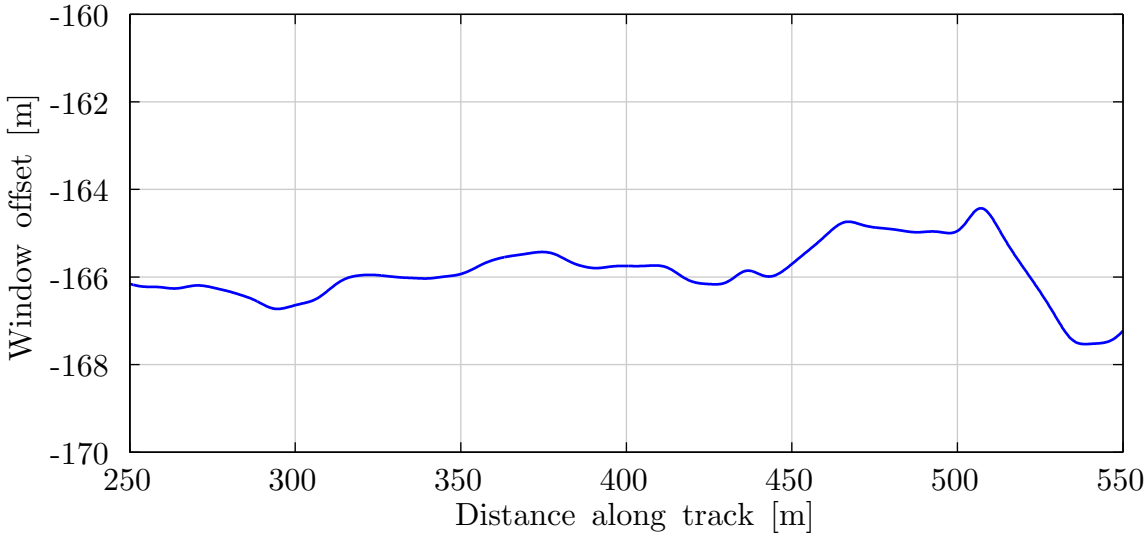


Figure 6.5: Smoothed line showing offsets (ℓ values) to be applied to each sample

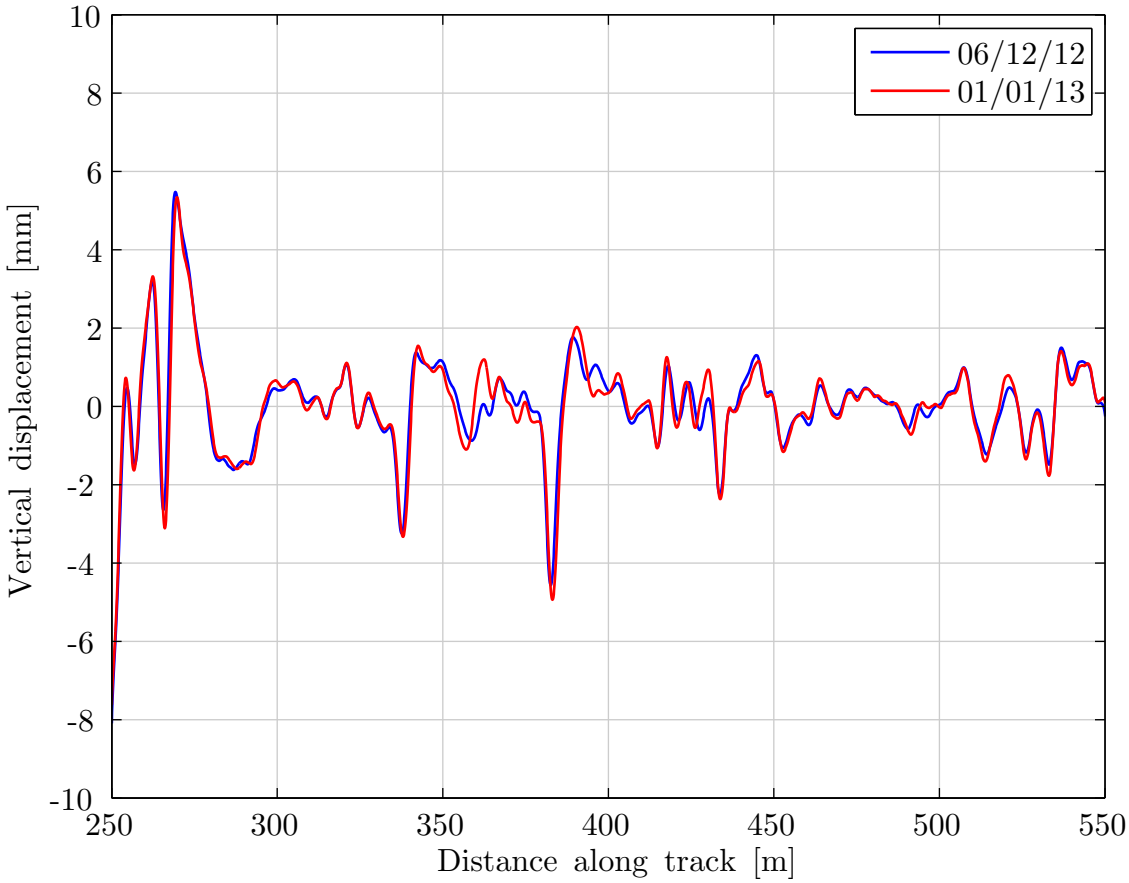


Figure 6.6: Two vertical displacement data sets, after dynamic-shift alignment

6.2.4 Using Auto-Alignment to Determine Train Orientation

It was determined in Section 4.2.2 that the polarity of vertical displacement data derived from the pitch-rate gyroscope is dependent upon the orientation of the train (*i.e.* whether the 'A' or 'B' cab of the train is leading). The same is also true of the roll-rate gyro. One method of determining the orientation of the train is to initially process the pitch-rate gyro without regard to the direction of travel of the vehicle, after which auto-alignment can be used to determine the orientation.

Vertical displacement is first calculated using the gyro-only method described in Section 4.2.2 (equations (4.6) and (4.7)). Next, the following two comparisons are performed to find their best-shift values, $\ell_{\max}$, and their highest score values $q_{\ell\max}$.

$$\{\ell_{\max A}, q_{\ell\max A}\} = \text{align}\,(\boldsymbol{z}_1, \boldsymbol{z}_2) \tag{6.4}$$

$$\{\ell_{\max B}, q_{\ell\max B}\} = \text{align}\,(\boldsymbol{z}_1, -\boldsymbol{z}_2) \tag{6.5}$$

Here, comparisons are made between displacement data $\boldsymbol{z}_1$ which must be known to have the correct polarity, and an unknown data set, $\boldsymbol{z}_2$. The comparison is performed again with negative $\boldsymbol{z}_2$. After this, the highest value of $q_{\ell\max A}$ and $q_{\ell\max B}$ is selected. If the B value is chosen, the data set $\boldsymbol{z}_2$ is permanently inverted. Next, the chosen shift value $\ell_{\max A}$ or $\ell_{\max B}$ is applied to the $\boldsymbol{x}_2$ vector, and resampling is performed.

If sensor measurements are to be combined, as in Chapter 5, the curvatures κ_ϕ and κ_θ must be inverted accordingly as well. Vectors κ_ϕ, κ_z and κ_θ must also be resampled.

This method of determining the vehicle orientation can be faster than searching data for the last stop, and finding the polarity of x acceleration as the train next accelerates (as described in Section 4.2.2). This is because in some cases the last stop may be long ago in time, meaning that a lot of data will need to be searched to find the stop. However, when alignment is performed over a large section of track, finding orientation by using the

double-pass alignment method can be slower than the search method. This is due to the fact that twice the number of alignments must be performed in the alignment process.

The selection of which method is to be used should be based on three factors:

1. Which variation of the auto-alignment algorithm is being used – The dynamic-shift alignment method requires more time to execute than the fixed-shift method, particularly when using the discrete comparison method.

2. The required range of alignment – If discrete comparison is being used, and the initial alignment from GPS is very poor (typically more than ±50 m), the range of alignment must be large to allow the correct alignment to be achieved. This results in more comparisons being made.

3. The amount of data available in the data sets – Occasionally a delay to the on-board PC's boot sequence, or a temporary loss of power to the logging during operation of the train will result in a set of log files which begin with the train already in motion. In this scenario, data for a previous stop is unavailable.

Generally, it is found that the double-pass alignment method is the fastest way of determining the train orientation. One technique which may be used to significantly shorten the process when using dynamic alignment, is to determine direction using only the first alignment window. After this the direction of travel is known, and the other window alignments can be made only once using the known polarity.

6.3 Comparing Vertical Displacement Data

Once samples from multiple vehicle passes are aligned, comparisons can be drawn between them. Comparing data sets over a number of days can allow the following information to be derived with a relatively fine granularity of time:

- The rate of degradation of the track.
- Improvements to the track after maintenance.
- The effectiveness of maintenance performed.
- How long improvements caused by maintenance last before maintenance is required again.
- Prediction of future degradation rates allowing creation of more efficient maintenance schedules.

In the remainder of this chapter, three comparison methods are shown, which can be used to evaluate the degradation of the track. Other information from the list above can also be found using these comparison methods. In a case study in Section 8.2, improvements are found following maintenance, and the effectiveness of maintenance is discussed.

6.3.1 Comparison Methods

The most simple form of comparison is to find the absolute difference between an initial data set, Z_A and a data set Z_B recorded later on, at each location x:

$$D_z = |Z_B - Z_A| \tag{6.6}$$

In order to represent the change in a section of track over time, it is useful to compare each pass of the track to a reference pass at the start of the time period. Data from the i^{th} pass is referred to as Z_i, and the first (reference) pass as Z_0.

$$D_{zi} = |Z_i - Z_0| \tag{6.7}$$

Figure 6.7 shows a comparison between multiple passes of the same section of track. In this figure, each pass of the train along the track is represented by one row. The colour of each segment represents the difference from the first pass, D_{zi}. The first pass is shown on the bottom row, with later passes (higher values of i) stacked above. It should be noted that because the first pass is compared to itself, it is entirely green, representing all zeros.

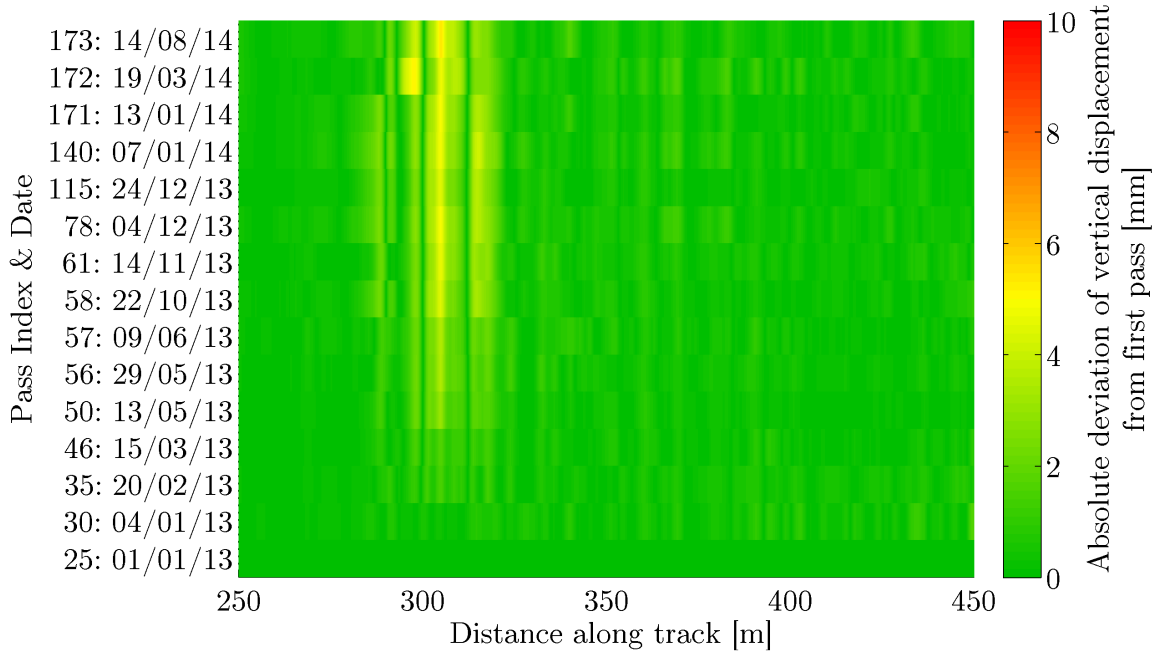


Figure 6.7: Absolute differences between measured vertical displacements for multiple passes of the same track

It can be seen from this figure that as time progresses, a change occurs in the region of track between 280 m and 320 m. It can also be seen that in later months, the size of the change does not increase, with the largest areas of change remaining at around

5 mm. The problem with this representation is that the yellow and red areas representing larger changes do not necessarily represent a degradation of the track geometry. In some cases, maintenance may be performed between passes causing the vertical displacement to return to values closer to zero. Because equations (6.6) and (6.7) take the absolute value of any change, be it positive or negative, an improvement in vertical geometry cannot be distinguished from a degradation.

A better comparison takes the difference between absolute values of the displacements:

$$E_{zi} = |Z_i| - |Z_0| \tag{6.8}$$

This method of comparison can produce positive values representing a change away from a zero displacement, or negative values representing a change towards zero. The same data used in Figure 6.7 was used to calculate values for E_{zi} which are shown in Figure 6.8.

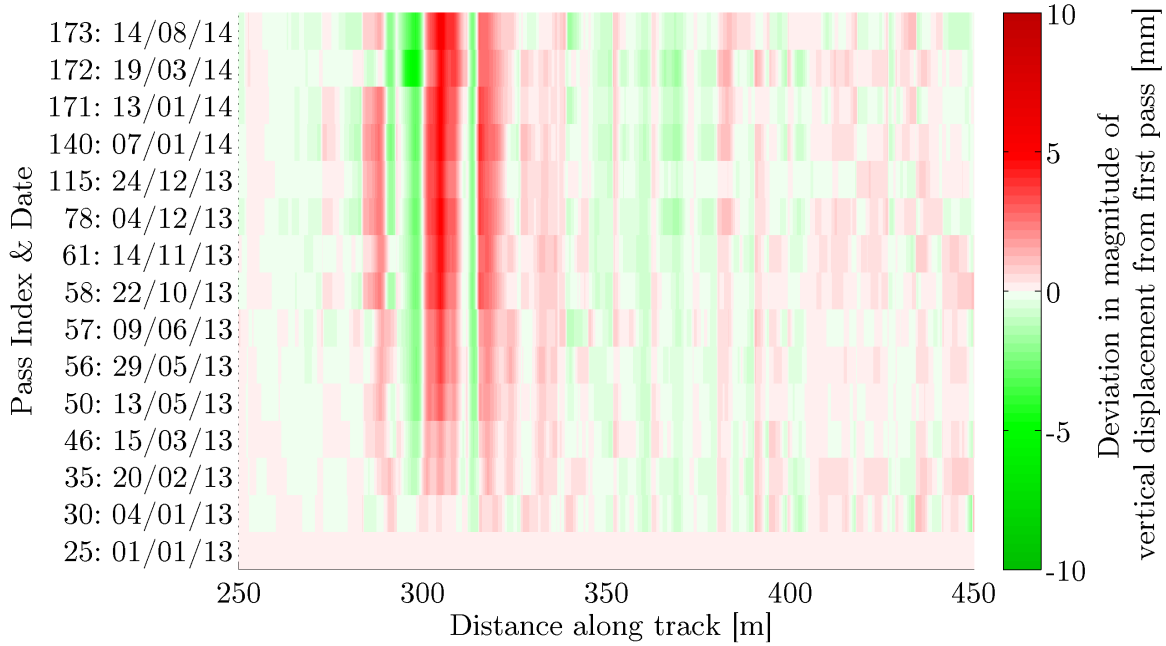


Figure 6.8: Differences in magnitude of measured vertical displacements for multiple passes of the same track

This figure uses red to represent a deviation away from zero (a degradation), and a green to represent a change towards zero (an improvement). The same change can be seen as was visible in Figure 6.7, but the red areas indicate a change away from zero, suggesting a degradation in vertical track geometry. It can also be seen that there is a green area at around 298 m where the track geometry actually changes towards zero. This kind of improvement is often seen in areas where the loaded geometry has degraded, due to factors such as ballast movement.

Figure 6.9 shows the vertical displacements of the first and last passes shown in Figures 6.7 and 6.8.

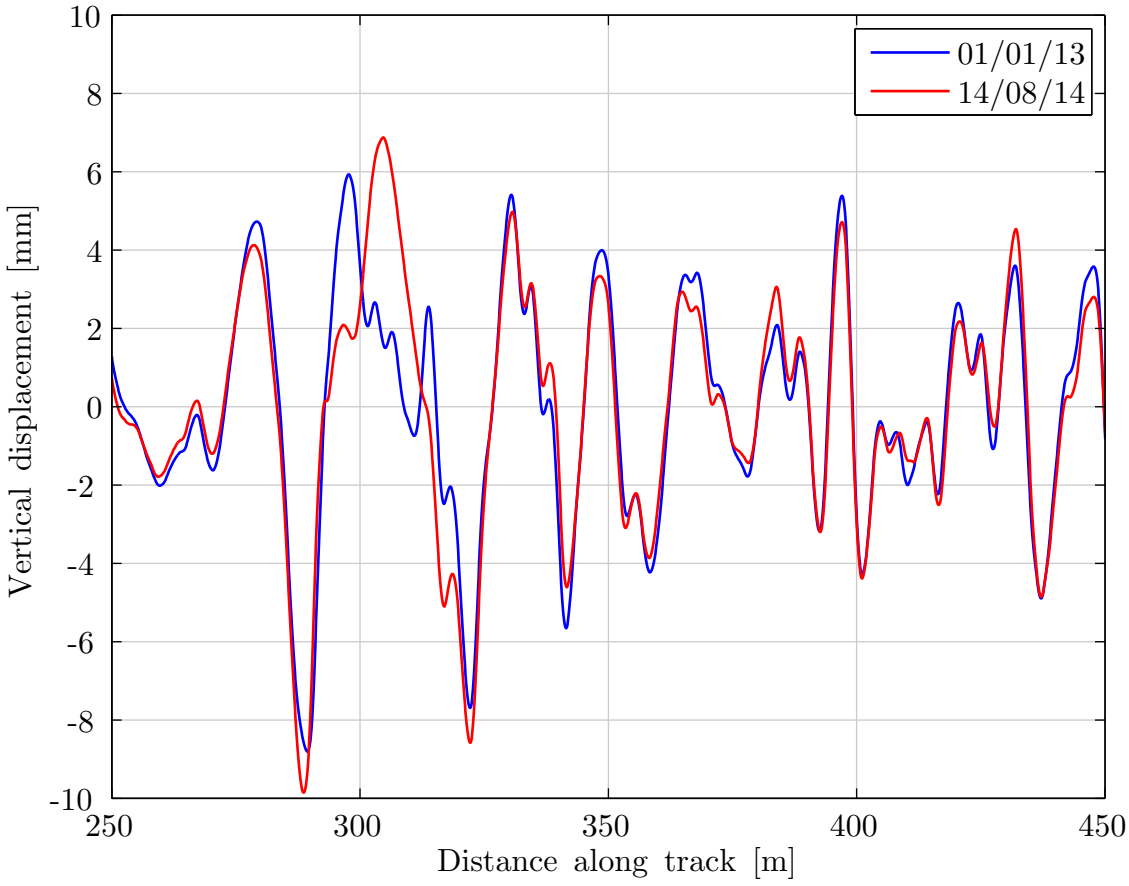


Figure 6.9: Vertical displacements of first and last passes in comparison set

It can be seen from the figure that there is indeed an improvement at the region between 290 m and 300 m, after which a significant degradation has occurred from 300 m to 320 m.

The data used in these figures are from a transition from ballasted track onto a steel bridge. At transition zones such as this, the repeated passage of trains can cause the ballast supporting the track to settle unevenly [17, 99, 100]. This 'differential settlement' could be a factor causing this change.

The magnitude comparison method of indicating degradations and improvements (as in Figure 6.8) does have some limitations, the first of which is the fact that a displacement of zero is assumed to be the ideal state for the track. Whilst this may largely be true, it may not always be the case, particularly preceding or following other large displacements. Another limitation is that if the displacement changes its sign (*i.e.* a positive displacement in the reference data set changes to a negative displacement in the i^{th} data set), a false change value could be indicated. Such a change could occur in areas of track with a small magnitude of displacement, or after maintenance has occurred.

CHAPTER 7

TRACK STIFFNESS

Part (c) of the hypothesis stated that repeated data from multiple passes of the same track could be used to estimate track stiffness. In Section 1.3.1, an approximate calculation was done which estimated that in order to estimate track stiffness, measurements repeatable to within 0.2 mm would be required. This is possible using the developed inertial measurement system, providing that data from passes with the same vehicle orientation are used, otherwise the value increases to 0.5 mm. In order to be able to estimate stiffness, accurate alignment (to within 0.125 m) between data sets is also required. This has also been achieved in Chapter 6.

Now that the inertial measurement system which has been developed satisfied these requirements, work was started on using the acquired data to estimate track stiffness from the bogie of the instrumented in-service vehicle. Firstly, further research into track stiffness and the problems it can cause was carried out. A recent guide produced by the Cross Industry Track Stiffness Working Group [21] provides a useful reference to track stiffness and its measurement, and to the effects of a train on a track in terms of deflection. The passage of a railway vehicle over a section of track causes it to deflect downwards due to the gravitational force of the vehicle acting on the track. An opposing upward force is provided by a combination of the bending stiffness of the rails, and the supporting force of the ballast and ground below the sleepers. The measure of resistance against the deflection of the track is called track stiffness.

7.1 Degradation Caused by Track Stiffness

The Guide to Track Stiffness [21] also provides a table of problems attributable to track stiffness factors, along with their symptoms. To summarise, a low value of track stiffness means that the track is poorly supported and the amount of downward deflection will be large. Excessive and repeated deflection of the track can cause a rapid loss of track geometry, weakening of the rails, and track clips becoming unfastened. Ultimately, these symptoms could result in rail breakage and/or derailment.

Track stiffness which is too high can also cause problems. Puzavac *et al.* [101] identifies that very stiff track causes large forces to be experienced by the rails and sleepers, which can rapidly cause damage to the track components. Puzavac *et al.* also recognise the existence of an 'optimum stiffness' value, which minimises the rate of deterioration of the track.

A report [12] states that abrupt changes in stiffness along the length of the track can also cause problems. The locations of such changes are known as transition zones, and can generally be found at places such as transitions onto and off of bridges, into and out of tunnels and over level crossings. Because of the resulting stresses on the rails, transition zones often require maintenance to be performed more frequently than on evenly supported plain track. A journal paper by Varandas *et al.* [102] provides a lot of information about the physical structure of transition zones, using a concrete culvert as an example. Work is also done to produce a train-track interaction model around the transition zone.

Another stiffness-related feature sometimes found on the railway is the 'void'. A void in railway terms refers to a poorly supported short section of track, typically occurring under a small number of adjacent sleepers. Poor support is usually due to a lack of adequate ballast underneath the sleepers. Figure 7.1 illustrates a void beneath two sleepers.

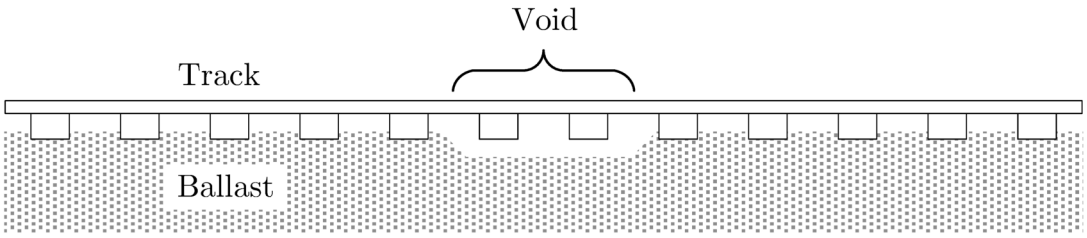


Figure 7.1: Illustration of voiding beneath two sleepers

Voids can cause a significant increase in the stresses experienced by the rail, and other track components such as rail pads and ballast [103]. This will cause a reduction in the life of the track components.

In Chapter 2, various systems were reviewed which are capable of measuring the stiffness of a track on a continuous basis, but all of them required the use of dedicated measurement vehicles. A method of deriving stiffness using instrumentation fitted to an in-service vehicle, such as the IMU presented in this thesis, would be a significant benefit and would reduce the overheads required to operate and maintain a dedicated vehicle.

7.2 Defining Stiffness

In order to measure stiffness, its definition must be known. An Innotrack report [12] defines a basic measure of track stiffness as:

$$k_t = \frac{F}{z} \tag{7.1}$$

Where F is the force being applied to the track, and z is the vertical deflection. This fundamental equation is known as Hooke's Law, and may be used to calculate the extension or compression of a spring with linear stiffness when a known force is applied. Based on this law, it can also be said that:

$$k_t = \left|\frac{\delta F}{\delta z}\right| \tag{7.2}$$

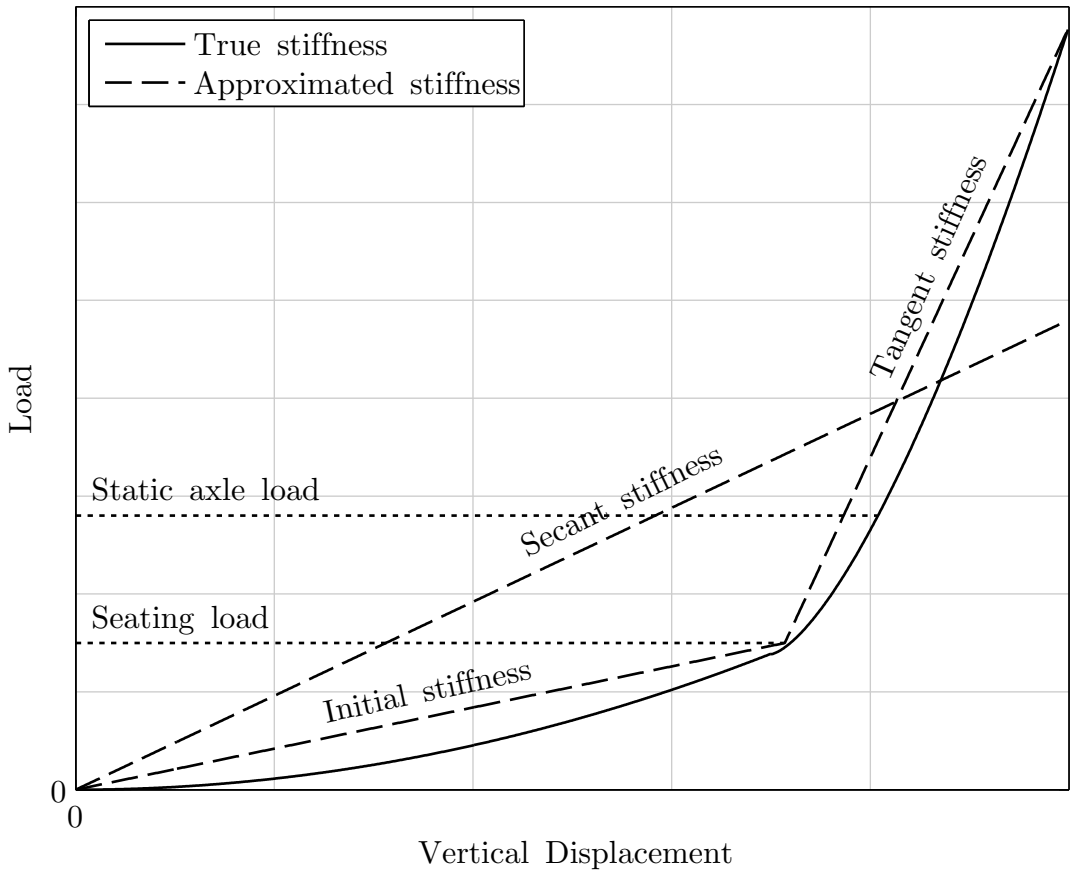
This shows that stiffness is the change in force, δF, over the change in vertical displacement, δz. This means that in order to determine a measure of track stiffness where linear stiffness is assumed, the deflection need not be measured from a completely unloaded state. This is the principle used by several of the stiffness measurement vehicles reviewed in Chapter 2 which measure the vertical displacement experienced by both heavy and lightweight wagons, which have a known difference in force, and can calculate the difference in experienced displacement.

Tangent and Secant Stiffness

Sometimes, particularly where voids have formed beneath the track, a small load will cause a large displacement, as the unsupported track deflects. Once the unsupported sleeper comes into contact with the ballast, further increases in load result in a lower rate of further displacement, as the track is now supported by the ballast. These are often approximated as two linear stiffness values. A typical displacement-load graph for a voided track is shown in Figure 7.2.

An initial stiffness is experienced at low loads, usually whilst the sleeper is still unsupported from below. Once the sleeper comes into contact with the ballast, the stiffness at higher loads is referred to as the tangent stiffness. The load force at which the 'knee' occurs between the initial and tangent stiffnesses, is called the seating load. An alternative method of stiffness measurement is known as the secant stiffness, which approximates the stiffness as a single linear equation.

Figure 7.2 also indicates a representation of the static axle load of a train. As a bogie-mounted IMU will only ever measure loaded geometry, it would seem unlikely that it will be able to measure the initial stiffness before the sleeper becomes seated on the ballast. However, there is a possibility that due to the natural damping in the track system, the track may take some time to fully seat, before which the initial stiffness may be seen for a short time. It was not possible to determine whether this was the case from the work done here, but further research may prove or disprove this possibility.



Redrawn from [12]

Figure 7.2: Example of track stiffness at a void

7.2.1 Measuring Stiffness from an In-Service Vehicle

Measurement of track stiffness using inertial sensors fitted to an in-service vehicle is not easy to achieve. Any such system will only ever be able to measure the loaded profile of the rail. It is not possible to measure the unloaded profile of the rail from an in-service vehicle without fitting measurement devices such as lasers to the underside of a carriage of the train. These would allow the height of the rail relative to the vehicle body to be measured away from the location of any wheelsets, and consequently away from the locations of loaded track.

If a downward displacement is measured using a bogie-mounted IMU, this can indicate that the track displaced downwards as the train passed over it. However, it could also simply indicate that the unloaded profile of the rail had a downward gradient at this location. From a single IMU measurement, it is difficult to determine which of these two scenarios is the case.

The hypothesis declared in Section 1.4 stated that repeated passes of the IMU over the track could be used to determine stiffness. Sometimes the train can exert different forces on the rail between one pass of the track and another due to factors such as vehicle speed which affects the dynamic axle load. Estimation of the forces, and calculation of the difference in displacements could be used to determine the track stiffness.

Some work has been ongoing in the area of detecting track stiffness and sleeper voids. Work towards a previous doctorate [104] concentrated on a finite element analysis (FEA) model of a track and railway vehicle, which was constructed in the software package ABAQUS. The model, which uses thousands of finite elements in three dimensions, is very computationally intensive to analyse, and can take several hours to compute the stiffness even in a static location. It is clear that such a model would not be suitable for continuous analysis of stiffness along a stretch of track several hundred metres long.

Bezin *et al.* [103] provide a useful background on multi-body system (MBS) models and vehicle-track interaction (VTI) models. Ultimately, a flexible track system model (FTSM) is developed which allows detailed evaluation of the forces on the track using a known stiffness input. Unfortunately, the fact that the track stiffness is the input to the model means that whilst some of the principles may be used, it is of limited use when estimating track stiffness from inertial measurement inputs.

7.3 Vehicle and Track Dynamics

Before attempting to derive stiffness from experienced forces and displacements, it is useful to understand the dynamics of the track and components of the railway vehicle. This will aid in reverse-engineering these models to derive a method of stiffness estimation.

7.3.1 Track Model

When considering the vertical displacements of a railway track structure, the system is often modelled in two dimensions as a simple beam on elastic foundation (BOEF) [93, pp. 159-160], where a single rail is represented as a beam supported at regular intervals by springs, as shown in Figure 7.3(a). The force, F, which must be applied to the rail to achieve a vertical deflection, w, may be calculated using (7.3).

$$F(x) = EI\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} + kw \tag{7.3}$$

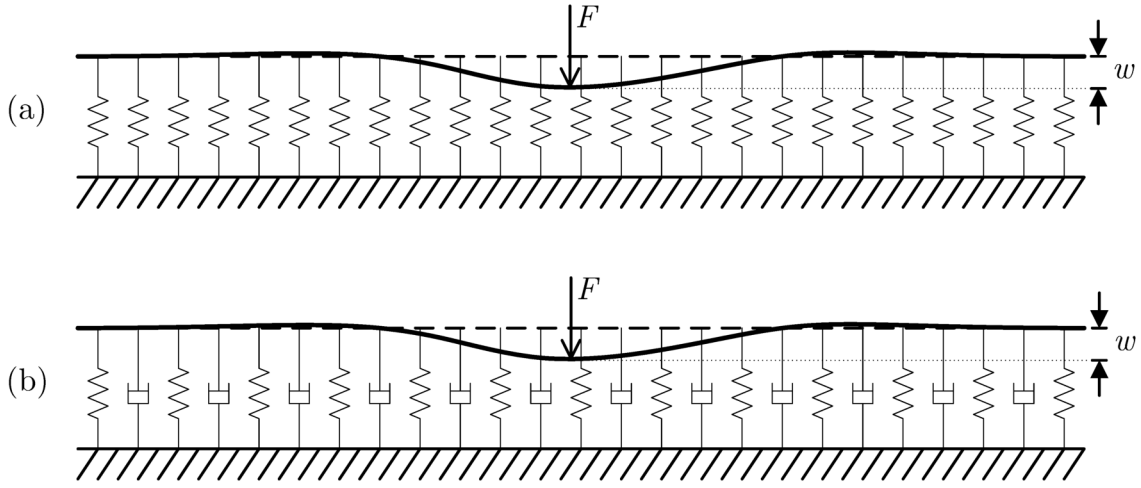
Where x is the horizontal position along the track. k is the stiffness of the track support per meter, E is the elastic modulus (or Young's modulus) of the track, and I is the area moment of inertia. The product EI is referred to as the bending stiffness of the rail.

An improvement on this model is the Beam on Pasternak foundation (BOPF) [105], which adds dampers to the rail support and takes into account the mass of the rail itself (Figure 7.3(b)). The force required to achieve deflection, w, is shown in (7.4).

$$F(x,t) = EI\frac{\mathrm{d}^4 w}{\mathrm{d}x^4} + \rho\frac{\mathrm{d}^2 w}{\mathrm{d}t^2} + c\frac{\mathrm{d}w}{\mathrm{d}t} - k_1\frac{\mathrm{d}^2 w}{\mathrm{d}x^2} + kw \tag{7.4}$$

Where t is time, ρ is the mass per unit length of the rail, c is the damping coefficient per meter of track, and k_1 is the shear parameter of the rail.

The deflection of the rail can be calculated directly from the static force from the mass of the wheelset, bogie and carriage resting on the rail. However, the train is moving, and



Redrawn from [93, 105]

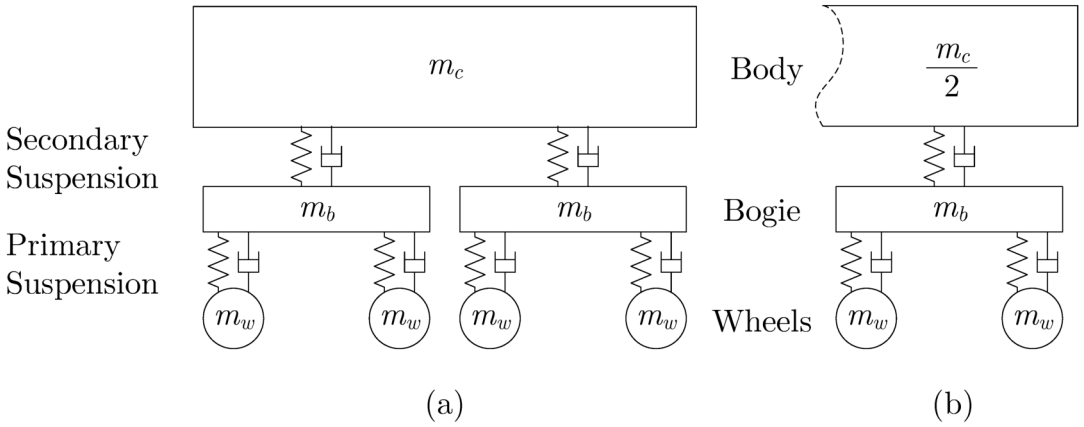
Figure 7.3: Deflected rail represented in two different models
(a) as a Beam on Elastic Foundation, and (b) as a Beam on Pasternak Foundation

the suspension systems of the train have an effect on the dynamic forces being applied to the track.

7.3.2 Bogie Suspension Model

A common one dimensional model of the suspension of a bogie models the primary suspension (between the wheelsets and the bogie) and the secondary suspension (between the bogie and the body) as parallel springs and dampers [106].

The model can be arranged in a number of ways. Banimahd *et al.* [106] demonstrates models for a single wheelset (a quarter coach) and for all four wheelsets (a whole coach). For this work, measurements are being taken from on a single bogie, so it is prudent to model at least the two wheelsets of a bogie. A whole coach could be modelled, but as the coaches in a Class 377 use a relatively rigid coupling system between coaches, the next coach would need to be modelled as well to accurately model the forces acting on the track. For simplicity, a single-bogie model was opted for, shown in Figure 7.4(b). Figure 7.4(a) shows a whole coach model alongside for comparison.



Adapted from [106]

Figure 7.4: One dimensional bogie models
(a) for a whole coach, and (b) for a half coach

The masses of each component are indicated in the figure. The half coach model shows that the effective mass of the body is considered to be half the whole coach body mass, m_c. It is presumed here that half of the mass is supported on the bogie at the other end of the coach. In reality, there is some dynamic variation in the mass supported by the bogie. For example, as the body pitches, the balance of mass across its two bogies changes. Also some mass will be supported by coupled coaches, and some of the mass of coupled coaches will be supported by the coach being measured. However, these effects are beyond the scope of this work and are disregarded.

7.4 Methods of Track Stiffness Estimation

Ideally a model which combines the Beam on Pasternak Foundation with the one-dimensional bogie model from Figure 7.4(b) would be used to estimate track stiffness. Initially, however, it was decided to opt for a simpler model to determine the feasibility of the principles involved.

7.4.1 Evaluating a Simplified Suspension and Track Model

Figure 7.5 shows a simplified version of the model which has three masses with a spring and damper between each. This is a commonly used configuration in classical mechanics. Track stiffness is represented as a single spring-damper pair and does not take into account complex factors such as the bending stiffness of the rail. It was seen in Figure 7.2 that track stiffness can be approximated with a linear (secant) stiffness, which is effectively what this model represents.

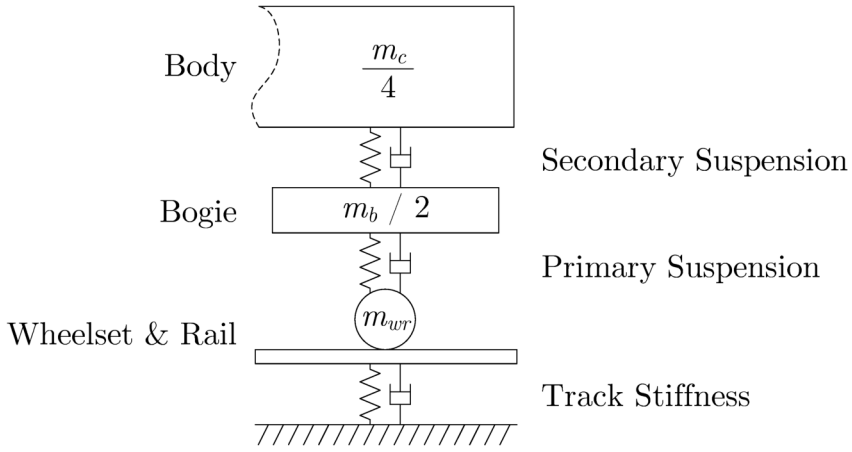


Figure 7.5: Simplified model of motion

This model assumes that there is a fixed trackbed beneath the track structure which does not displace due to forces exerted upon it from the rail and train above. However, the height of this bed, z_t, does vary as the train moves along the track. Changes in z_t represent variations in the height of the ground supporting the track and its substructure, over the length of the track.

Figure 7.6 shows the same model with each of the components labelled, allowing its equations of motion to be considered.

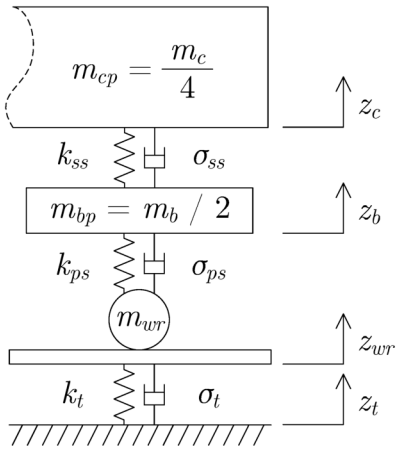


Figure 7.6: Simplified model of motion showing vertical displacement of components

The model in Figure 7.6 is similar to that of a standard 'quarter car' suspension model [107] used in automotive engineering. Using the same methods used in [107], the equations of motion can be found for each of the components in the model.

$$m_{cp}\ddot{z}_c = -k_{ss}(z_c - z_b) - \sigma_{ss}(\dot{z}_c - \dot{z}_b) \tag{7.5}$$

$$m_{bp}\ddot{z}_b = k_{ss}(z_c - z_b) + \sigma_{ss}(\dot{z}_c - \dot{z}_b) - k_{ps}(z_b - z_{wr}) - \sigma_{ps}(\dot{z}_b - \dot{z}_{wr}) \tag{7.6}$$

$$m_{wr}\ddot{z}_{wr} = k_{ps}(z_b - z_{wr}) + \sigma_{ps}(\dot{z}_b - \dot{z}_{wr}) - k_t(z_{wr} - z_t) - \sigma_t(\dot{z}_{wr} - \dot{z}_t) \tag{7.7}$$

The IMU system can determine values for parameters z_b, $\dot{z}_b$, and $\ddot{z}_b$. It can be assumed here that the masses remain constant while the train is in motion, and that the stiffness and damping coefficients of the primary and secondary suspensions remain constant also. z_t is the input to the system which has an unknown value. k_t and σ_t are also unknown for any given location on the track. As these factors are directly influencing the motion of the rest of the system, these are considered inputs as well.

As well as these inputs, at any given time, the motion of the coach (z_c, $\dot{z}_c$, and $\ddot{z}_c$) is unknown as well as the motion of the wheelset and rail (z_{wr}, $\dot{z}_{wr}$, and $\ddot{z}_{wr}$). It can be seen from equation (7.6) that the motion of the bogie measured by the IMU, is dependent upon the motions of the coach, and the wheelset and rail.

The last equation of motion (7.7) can be rearranged so that the system inputs are all on the left-hand side.

$$k_t z_t + \sigma_t \dot{(z)}_t = m_{wr}\ddot{z}_{wr} + k_t z_{wr} + \sigma_t \dot{z}_{wr} - k_{ps}(z_b - z_{wr}) - \sigma_{ps}(\dot{z}_b - \dot{z}_{wr}) \tag{7.8}$$

It can immediately be seen that the inputs, whose values need to be calculated, are factors of one another. This makes the equation impossible to solve by conventional methods, even if the motion of the wheel and rail were directly known.

Extended Kalman Filter

An Extended Kalman Filter (EKF) could be put to use to estimate the input values, as it is a non-linear filter. A similar system to the one used previously in Section 5.2 could be used, where the input is unknown but is predicted using the previous value with process noise added. Sensor measurements are then used to update the prediction.

A potential problem with this system is that the EKF is not an optimal estimator like the standard Kalman filter. This means that the filter output could generate a result which diverges far from the true values, rather than converging on an optimal estimate. If the EKF were designed such that it used sensor measurements from two passes of the train at differing speeds, greater confidence could be obtained when estimating the input values, as both passes of the system should 'agree' on the same values for z_t, k_t, and σ_t.

Particle Filter

An alternative system could be used to estimate the trackbed height, and track stiffness and damping. A particle filter could be designed, which would create multiple estimates (particles) for combinations of z_t, k_t, and σ_t which would produce a bogie motion equal to that measured by the IMU. This initial distribution (scattering) process is usually created using a normal distribution of values around likely values of one or more parameters. For example, a normal distribution for values of k_t could be created with its mean around

the known average track stiffness on the rail network. The variance of the distribution is adjusted so that all possible stiffness values are covered.

Using values from one particle estimate, along with the motion of the bogie measured by the IMU, the motions of the coach, and wheel and rail can be calculated for the particle. The probability of these motions having occurred can then be determined based on the laws of physical motion. This is done for every particle in the system.

The particle with the highest probability is the 'best-guess' based on the state of the system inputs. At each discrete step, every particle is updated. Eventually the probability of some particles will drop below a pre-determined threshold. When this happens, each particle below the threshold can be deleted, and new particles created in their place. These are generated using the original scatter process, or using a scatter process centred around the current best estimate.

If measurements from two passes of the train at different speeds were used within the filter, the probability of each particle could also be determined based on how well the particle fits both passes of the train. As z_t, k_t, and σ_t are factors of one another, several different combinations of values could describe the measured outputs of a single pass well. However, by matching these estimates to two passes at different speeds, the number of combinations of high-probability values should be reduced.

Time Limitations

Unfortunately, time limitations did not allow the EKF or particle filter methods to be developed. However, a simpler system was created which allows verification of the principle of using two train passes at differing speeds to estimate track stiffness.

7.4.2 Peak-Force Estimation Method

The methods considered in Section 7.4.1 rely on differences between the vertical displacements experienced during two passes of the train travelling at differing speeds. The work in this section aims to test the feasibility of this concept.

By creating an approximation of the downward force being exerted by the bogie on the track during the two train passes, and by measuring the displacements experienced during each pass, a value for track stiffness should be obtainable.

Li *et al.* [108] approximate dynamic wheel load using a formula used by the American Railway Engineering Association. This formula effectively adds a proportion of the train speed to the static wheel load.

$$P_{di} = \left(1 + \frac{0.0052V}{D}\right) P_{si} \tag{7.9}$$

Where P_{si} is the static wheel load, V is the train speed in $\mathrm{km\,h^{-1}}$, and D is the wheel diameter in metres. P_{di} is the resulting approximation for dynamic wheel load. Clearly this is a crude approximation of the peak axle load at a given speed, as many other factors affect the forces exerted by the wheelset on the track. However, it is a useful approximation here in proving the feasibility of the stiffness estimation method.

The static wheel load of the train does not change (unless the train is stationary and passengers are boarding or disembarking). As the only variable input parameter in the equation is speed, the resulting dynamic wheel load plotted against time or displacement, tends to be very smooth. Also, since the factor of speed is very small (0.0052), the dynamic load only fluctuates by a small amount (*e.g.* a change in speed of $10\ \mathrm{km\,h^{-1}}$ will give an increase in load of approximately 6% with a wheel diameter of 0.8 m). Clearly in a real-life scenario, the axle load changes constantly as the train moves over vertical irregularities in the track. The values calculated using (7.9) represent peak forces only,

so it will not be possible to estimate stiffness for every discrete location along the track; only where the axles experience peak force. This means that these locations need to be identified.

Identifying Locations of Peak Force

Initially, basic principles of motion were considered to attempt to identify peak force locations. Newton's second law of motion, $F = ma$, indicates the fact that with a constant mass, peak acceleration gives the peak force. In the case of the model being considered, a peak upward acceleration indicates peak force being exerted on the rail.

7.5 Verification of Peak-Force Estimation Method

In order to verify the stiffness estimation method, data was collected in a field trial from an IMU mounted on board a DMU-hauled coach, as well as from instrumented sleepers at the trackside. The aim of this trial was to gain deflection data from sleepers instrumented with geophones, which can then be compared to deflections estimated from the IMU to verify the effectiveness of stiffness estimation methods.

7.5.1 Field Trial Overview

The field trial was conducted at a railway facility at Long Marston in Warwickshire, UK. The site has a loop of railway track, including a relatively recently laid straight section, and various older sections.

The newer track section has modern concrete sleepers laid on large ballast stones (around 50 mm in length), whereas the older section has an old style of concrete sleeper laid on a much finer gravel and ash foundation (gravel stones around 5 mm in length). Figure 7.7 shows the older ash foundation in the foreground with the newer ballasted section in the distance.

Figure 7.7: Two ballast types at the Long Marston site

There is a visible difference in track deflection between the older track and the newly laid track. Upon visual inspection of the track as a train passes, a few sections of the older track appear to displace by up to 20 mm as the train passes, whereas most of the newer track appears to displace by 1 mm or less.

It was decided to instrument four locations along a straight section of track which includes both new and old track. Figure 7.8 shows an overview of this stretch of track, and Figure 7.9 shows a plan-view diagram of the locations of the four instrumented sites. In the photograph, Site 1 and the concreted level crossing can be seen in the foreground.

Trackside Instrumentation

The trackside instrumentation consisted of three geophones, whose outputs were captured by an off-the-shelf data acquisition unit from National Instruments connected to a laptop running NI's LabVIEW software. The equipment was powered from a 12 V car battery and a mains inverter. This allowed the setup to be moved throughout the course of the day between the four sites indicated in Figure 7.9. At each site, three adjacent sleepers were instrumented. Tapped baseplates were glued to the sleeper ends in advance, which

Figure 7.8: An overview photograph of the instrumented track

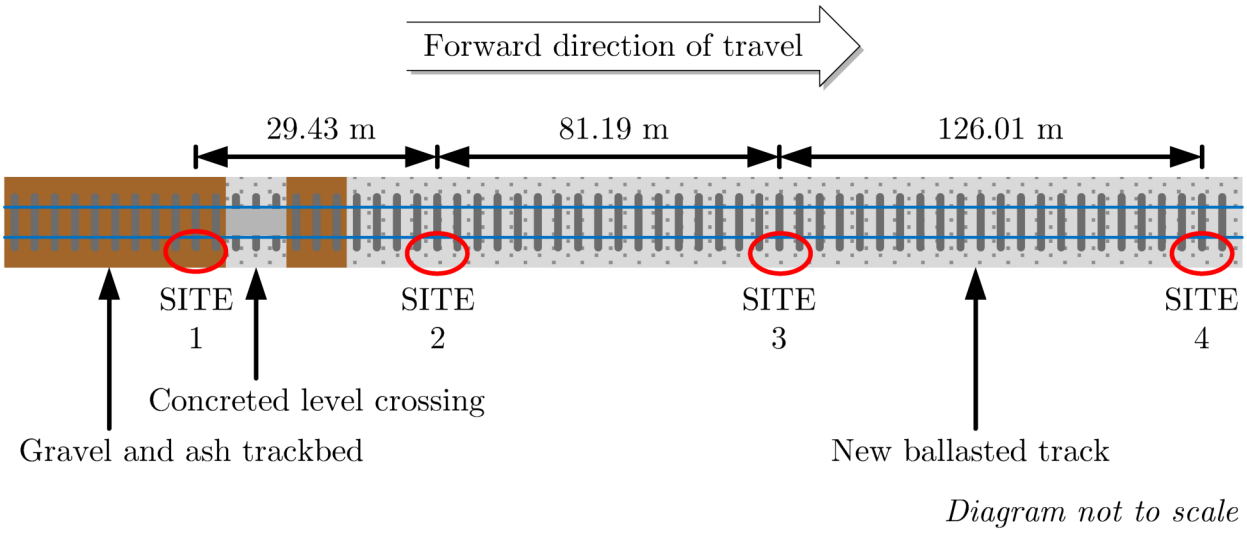


Figure 7.9: A plan-view diagram showing the instrumented locations

allowed the geophones to be screwed on in an upright orientation. Figure 7.10 shows this configuration.

On-board Instrumentation

The on-board instrumentation was attached to a bogie of the University of Birmingham's test coach, hauled behind a Class 117 DMU power car. The design of the IMU is similar to the one used on the Southern system, and includes a connection to a GPS antenna, and a connection to a wheelset tacho. The location of the instrumentation on the bogie of the coach is shown in Figure 7.11. This includes a 'sleeper detector' which is used for aligning on-board data with trackside data.

7.5.2 Aligning On-Board and Trackside Data

It was necessary to accurately align the data collected from on-board the train, with data recorded at the trackside. In order to do this, the IMU was modified to accept an input from a reflective light sensor. This was mounted on the outside edge of the bogie so that it was able to detect reflective objects at the trackside as the vehicle passes them. A white panel attached to a pole was driven into the ground at each site as it was instrumented. This can be seen on the right-hand side of Figure 7.10, opposite the geophones. As the instrumented bogie passes by the reflective board, a marker is recorded in the inertial data, indicating the location of the site of interest. Around the rest of the track were some other reflective objects at the trackside, so primary alignment was done by manually recording the time at which the train passed the site and finding this time in the recorded data. Both time sources were synchronised to GPS time. The optical system (called the 'sleeper detector') is accurate to within 4 ms, allowing fine alignment of the on-board data with the instrumented sleepers. The reflective board was always positioned directly opposite the middle sleeper of the three instrumented sleepers at each site.

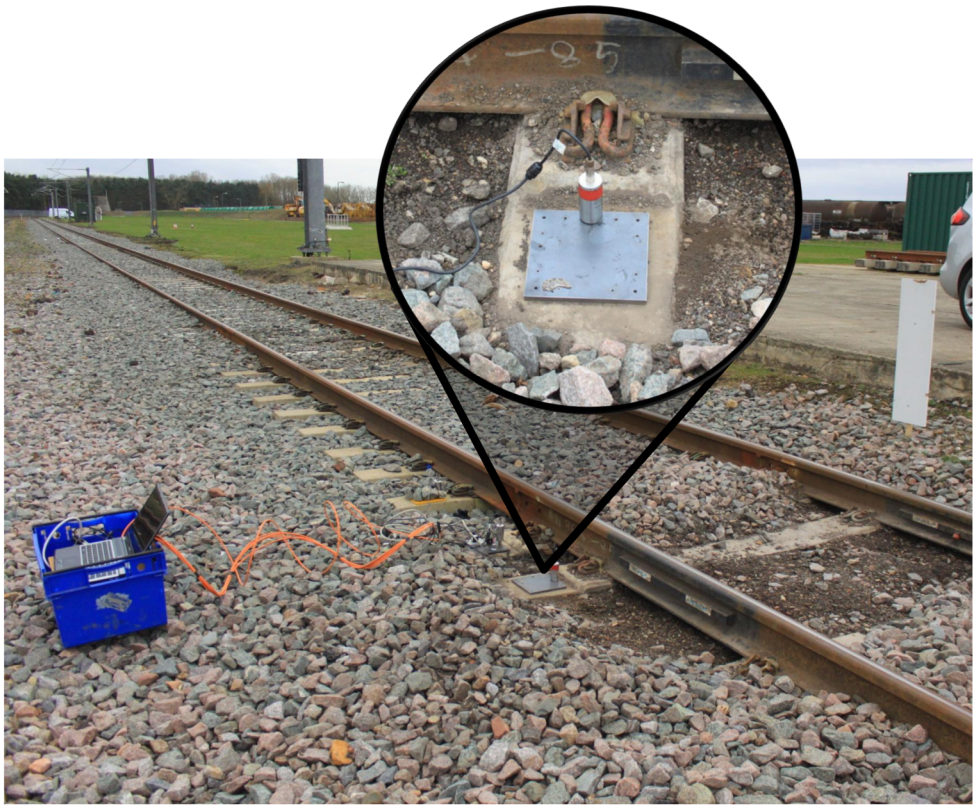


Figure 7.10: Geophone setup at Site 4

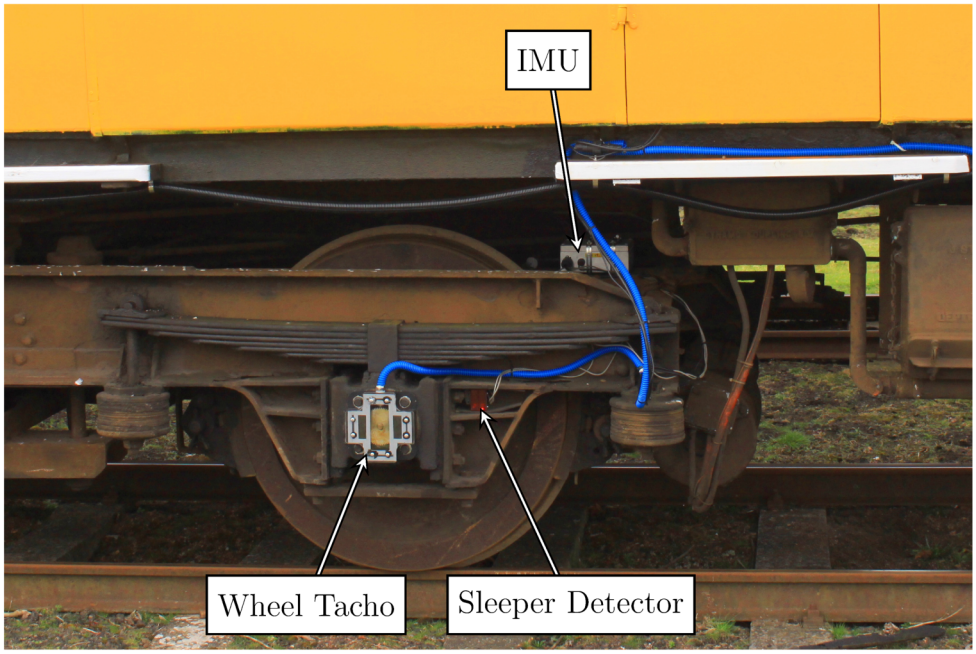


Figure 7.11: External instrumentation installed on University of Birmingham test coach

7.5.3 Processing Acquired Data

Vertical displacement was calculated from the data acquired from the IMU using the methods described in Chapters 4 and 5. Relative stiffness was then estimated using the method described in Section 7.4.2.

Geophone data was processed using similar methods to those used in systems of the same nature [8, 11, 68]. The geophones output a signal proportional to their vertical velocity. The signals are calibrated and integrated to derive the vertical displacement of each sleeper. A high-pass filter is applied to remove the effects of drift. A 4^{th}-order Butterworth filter with a cutoff wavelength of 1.5 s was chosen. In order to remove noise, a low-pass filter of the same order with a cutoff of 15 Hz was also applied to the signal.

It was found that the output signal had a lower amplitude than expected. This was due to the filtering effect of the geophones used, which naturally filter out frequencies below 1 Hz. This, coupled with the relatively low speed of the train, and consequently the low sleeper velocity, caused a low amplitude signal. A 'bass shelf' filter was designed to counteract the natural filtering effect of the geophone at low speeds. This filter corrects both the amplitude and phase of the velocity signal before it is integrated. Figure 7.12 shows the filtering effect of the geophone itself, as well as the designed shelving filter.

It can be seen from the figure that in the frequency range of interest (0.67–15 Hz, between the two cutoff filters), the resultant frequency magnitude response is almost completely flat at 0 dB, and the phase shift remains very close to 0°.

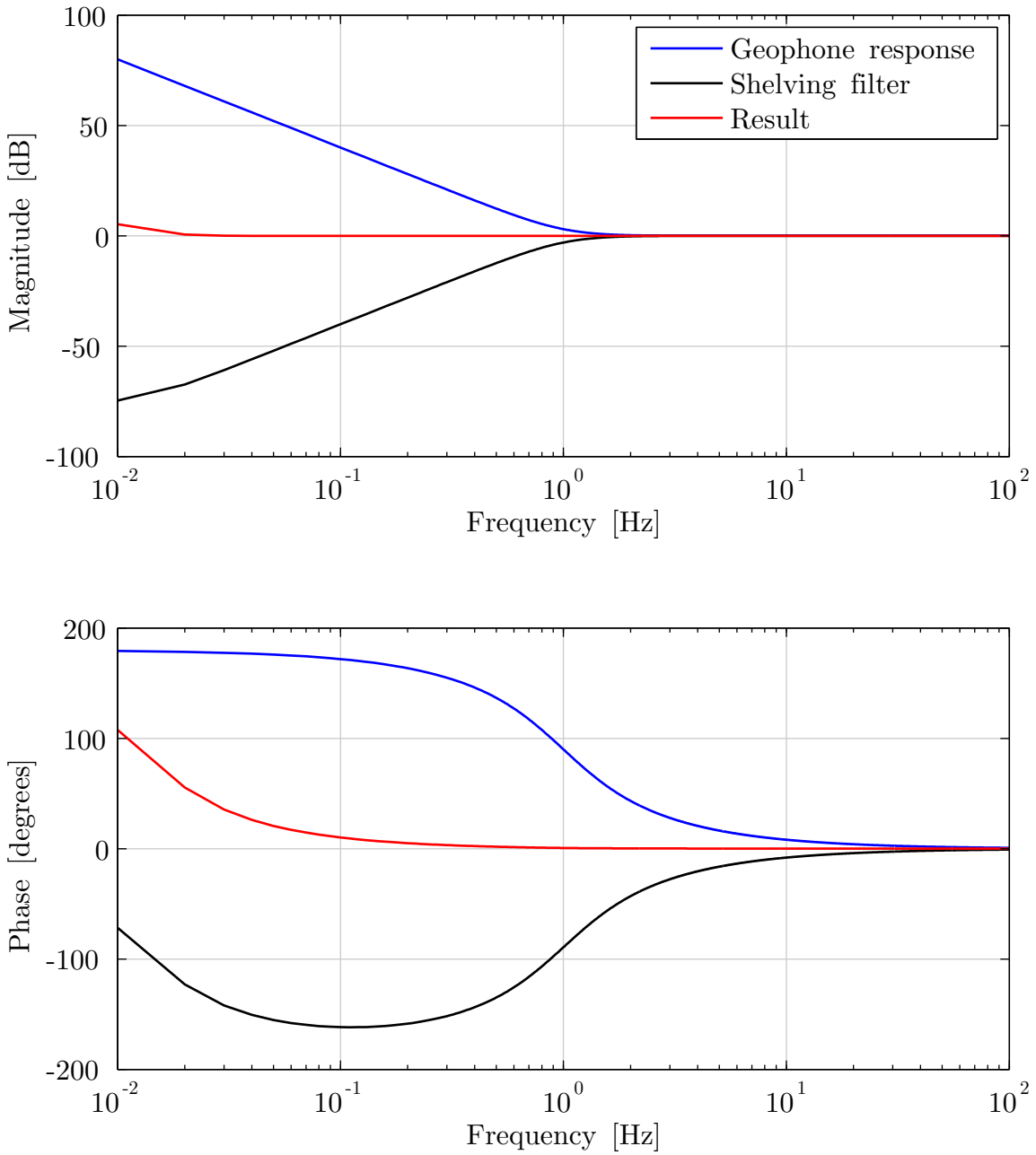


Figure 7.12: Filtering effects of geophone and shelving filter

7.5.4 Field Trial Results

The results of both the on-board and trackside stiffness measurements are plotted together against longitudinal distance in Figure 7.13. As the geophones each measure one specific point on the track, these are plotted as black crosses on the graph at locations found from measurements taken along the track (shown previously in Figure 7.9). The geophone stiffness is calculated by dividing the static axle force of the coach by the peak displacement measured as the instrumented axle passed each sleeper. It should be noted that this is not a high-accuracy method of determining stiffness because the force varies as the train passes vertical irregularities in the track.

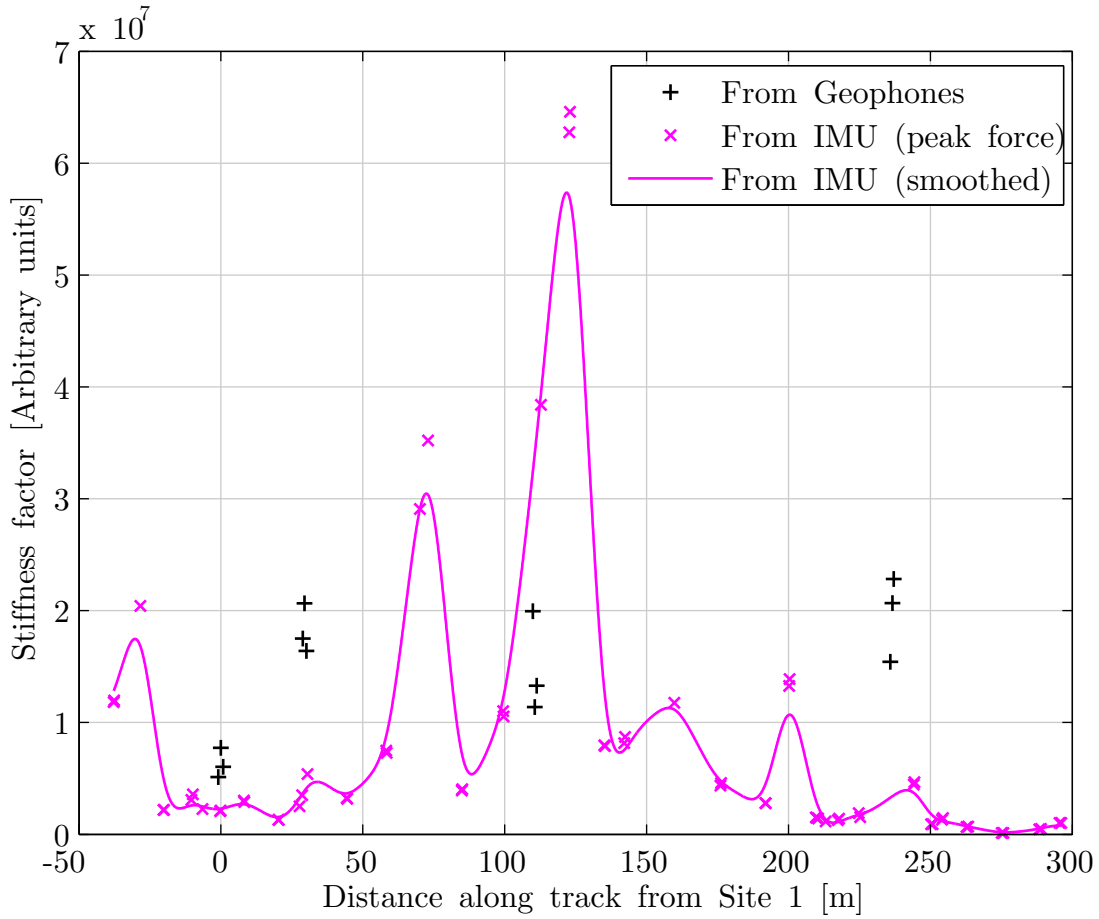


Figure 7.13: Estimated stiffness using simplified estimation method

The qualitative visual inspections performed before the experiments began identified that the track displaced by a large amount (~20 mm) at the area leading up to the level crossing, around Site 1. At the other three sites, the track appeared to displace by much smaller amounts (≤1 mm). This appears to correspond to the results obtained from the geophone measurements, which show a significantly lower stiffness at Site 1 than at the other three sites.

It is, however, immediately apparent that the IMU-derived stiffness estimation is a poor match to the geophone-derived stiffnesses found along the track. It suggests a low stiffness area around Site 1, which is congruent with the geophone measurements and visual inspection. However, whilst the stiffness does appear to increase after this, it varies by large amounts, and begins to decrease again after ~150 m. The geophone measurements suggest that Site 4 is the stiffest of all the sites, whereas the IMU estimation shows a very low stiffness value in this region.

Figure 7.15 shows the displacements measured by the IMU during fast and slow passes of the track. These have been calculated using the pitch-rate gyro only, as the slow train speed falls below the minimum speed required to use the accelerometer. It can be seen from the figure that from around -20 m to 50 m there are noticeable differences between displacements experienced at fast and slow speeds. Based on the underlying theory, this is to be expected for the track before Site 2, as the stiffness is low in this area. However, at Site 2 (around 30 m), the geophones indicate a higher stiffness. The IMU-derived stiffness does not rise until after 50 m. On the displacement figure, it can be seen between 50 m to 170 m there is little difference between the displacements of fast and slow passes, as the train traverses the newer, and stiffer, ballasted track. However, between 170 m and 260 m, there are very large magnitude differences between displacements at fast and slow speeds. The high-frequency components of the displacements appear to be similar, but there appears to be a low-frequency offset affecting the signal.

Low Stiffness Measurement at Site 2

Before traversing Site 2, the train passes over a 'compromise' joint, where two different types of rail are joined together. This is shown in Figure 7.14.

Figure 7.14: Compromise joint before Site 2

When traversing this joint at speed, a large amount of displacement was observed in the train's primary suspension, which continued as the train passed Site 2. The suspension displacement was significantly less at lower speeds. As the processing used does not consider the effects of the primary suspension, this could explain the low stiffness estimation derived from the IMU around Site 2 and its approach.

Low Stiffness Measurement at Site 4

Upon further inspection of the data, it was found that the speed of the train drops as it reaches Site 4 as it decelerates to stop at the end of the track. Figure 7.16 shows the speed profiles of both passes.

It was found at the end of Chapter 5 that vehicle speed had an impact on the measured track geometry. It is speculated that this effect could be due to the resonant frequency

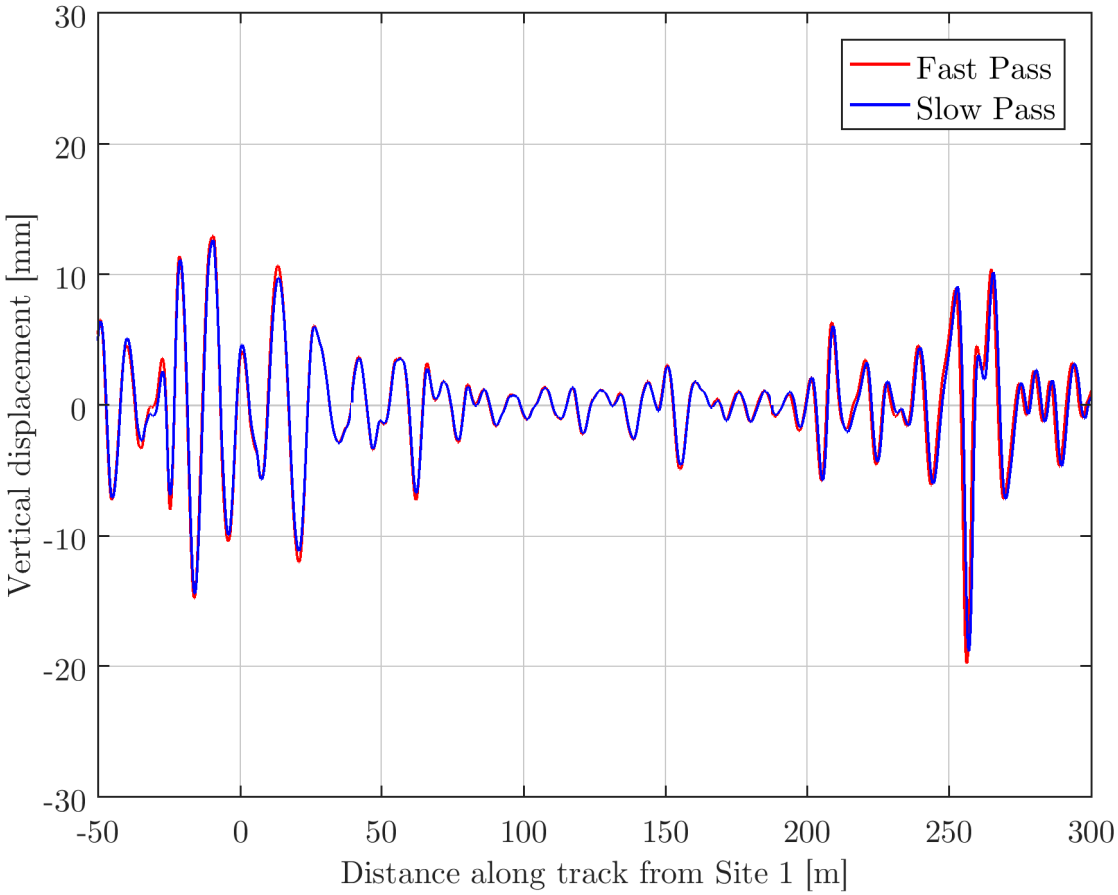


Figure 7.15: Displacements measured from the IMU on fast and slow passes

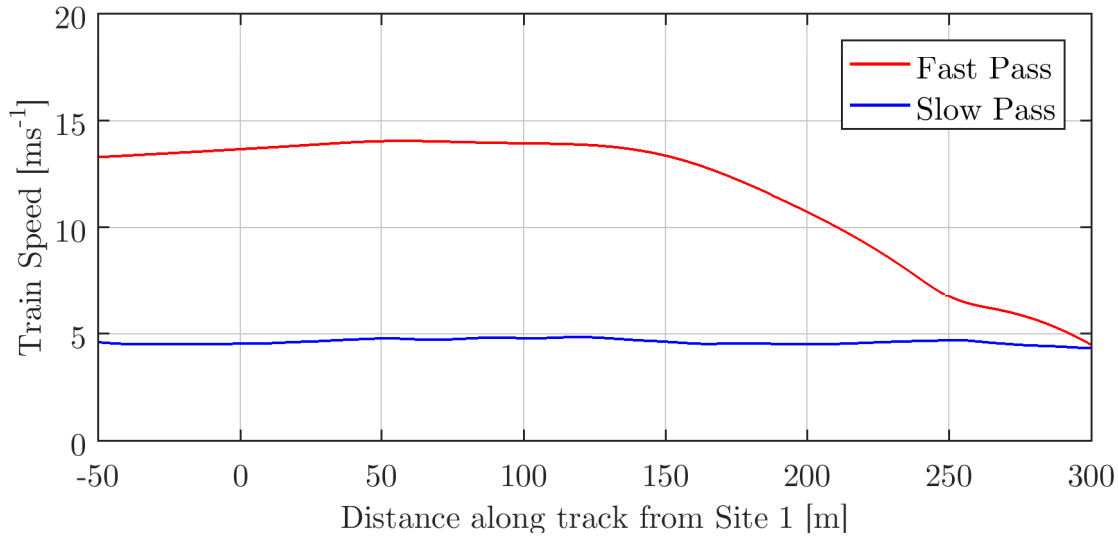


Figure 7.16: Train speeds of fast and slow passes

of the primary suspension. As the train travels at a constant speed, vertical geometry features with certain spatial wavelengths are amplified by the resonant frequency of the primary suspension. Features at other spatial wavelengths are attenuated. However, as the train changes speed, the time taken to traverse these features changes. The primary suspension, whose resonant frequency is constant in the time domain, now causes amplification or attenuation of features at different spatial wavelengths.

This effect is different on the University of Birmingham test coach compared to the Class 377. On the test coach, the primary suspension on each axle comprises of a pair leaf springs, resulting in a relatively soft primary suspension. This contrasts the Class 377 where the primary suspension is constructed from stiff rubber chevrons. The resonant frequency of a mass supported by a spring is dependent upon the stiffness of the spring, as shown in (7.10).

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{7.10}$$

Where k is the stiffness of the spring and m is the supported mass. This means that a higher spring stiffness will result in a higher resonant frequency.

The displacement of the primary suspension could be measured using displacement transducers as done in other research [49, 62]. Whilst this would be acceptable for a field trial at Long Marston, it would not be allowed within the specifications of the in-service system (Section 3.1.1).

7.5.5 Conclusions from Field Trial

The geophone measurements give results which appear to be consistent with visual inspection of the track. However, the peak-force method used to estimate track stiffness was not adequate to provide consistently accurate results.

The main problem with this method is that it fails to take into account the compressions and expansions of the vehicle suspension. Measurements observed at the level of the bogie are presumed to be the same as those present at track level. Another problem with the estimation system is that it interpolates values between points where peak forces are identified. Some stiffness values may differ on a per-sleeper basis. For example, at Site 3, one sleeper seems to be stiffer than the others. The peak-force method of estimation system makes heavy use of interpolation, and could easily miss variations at this level.

It is clear that the peak force method of estimation is inadequate for estimation of track stiffness. Whilst it may be possible to determine the stiffness while the train operates at two specific speeds, the variability of speed of a commuter train does not lend itself to this method.

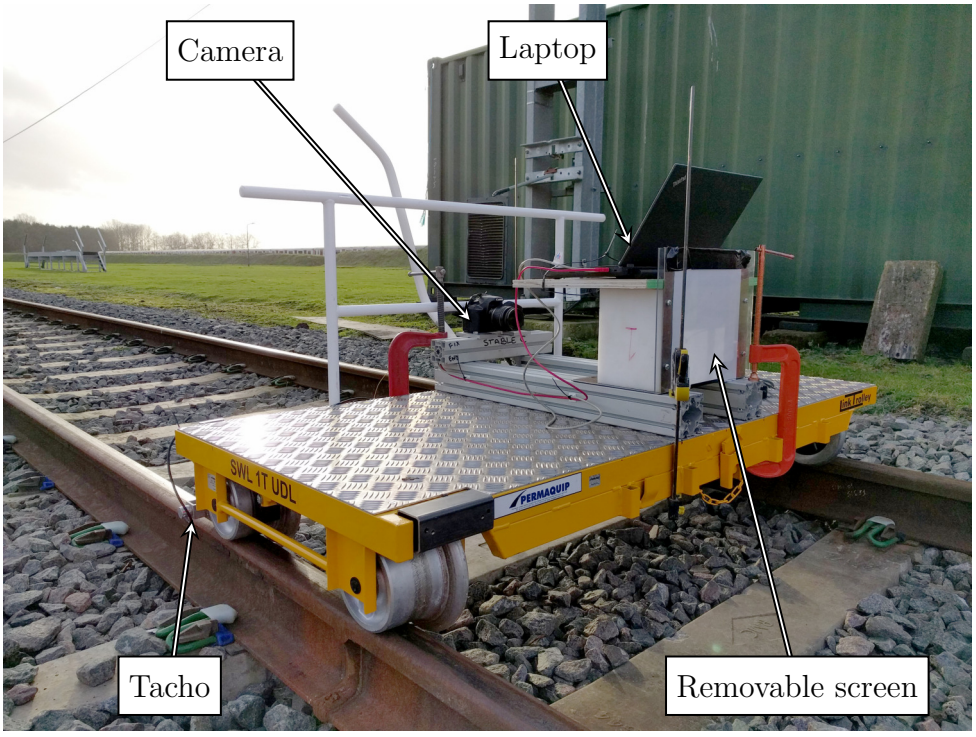
The field trials have proved that vertical displacement datasets acquired from from passes of the same track at two different speeds, contain differences which are measurable. It has also been determined that there does appear to be some correlation between the measured differences and the track stiffness. However, more work on other estimation methods, such as the EKF or particle filtering techniques, is needed to improve the model in order to prove or disprove this concept.

It is proposed that future trials include the use of a continuous stiffness measurement device, which can be used to measure stiffness values along the length of the track before IMU measurements take place. The geophone method only allowed the stiffness of a small number of sites to be measured along the length of the track. Consequently the variation of stiffness between sites was not known, and furthermore, it was not known whether the three measured sleepers of each site were typical of the other sleepers around them. A continuous stiffness measurement along the track would allow better assertion of any correlation between IMU-derived stiffness and the stiffness measured using a verified method.

7.5.6 Comparison to Unloaded Geometry

For a later experiment at the Long Marston site, a lightweight rail trolley was instrumented with a digital camera to measure vertical curvature of the track from images captured along the track [109]. The measurement trolley is intentionally very lightweight, so that the measurements taken can be considered to be on unloaded track.

This Vertical Profiling Trolley (VPT), shown in Figure 7.17, uses a tachometer to trigger the camera shutter at 0.125 m intervals along the track. The image sequence was then processed to produce vertical curvature for the track. This was done by determining the vertical position of a subject object in each image, and using trigonometry to determine the distance from it. After this the pitch angle of the trolley could be determined which can be differentiated by dx to give vertical curvature. Vertical displacement can then be found using the methods in Chapter 4. Exact details of the image processing methods are not given here, but can be found in the associated paper [109].



Note: Removable screen not fitted for this experiment

Figure 7.17: Vertical profiling trolley

Figure 7.18 shows a graph of the unloaded geometry measured by the VPT, with the loaded geometry measured by the IMU. In the figure, the mBF filtered method which combines gyro and accelerometer measurements is used to allow short wavelength features to be observed. Consequently only data from the fast pass is shown, as the slow pass falls below the minimum speed for use of the accelerometer.

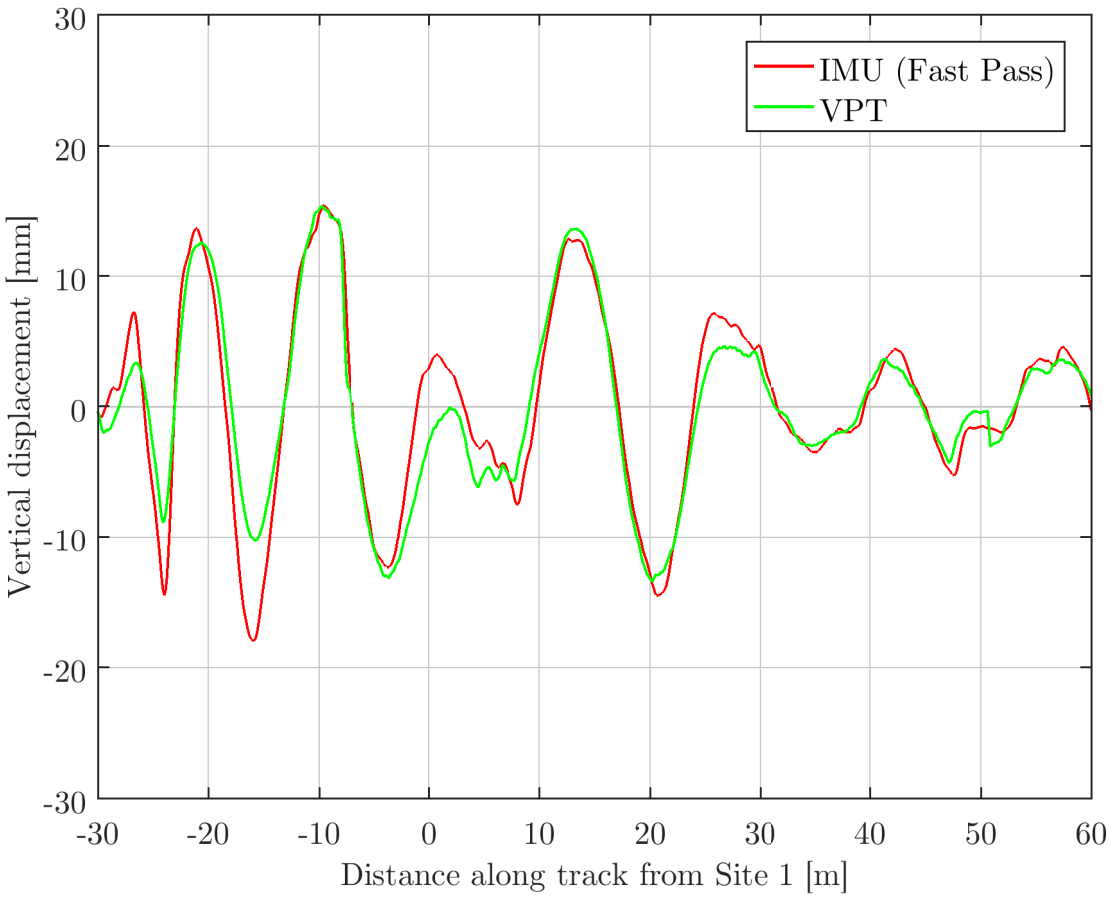


Figure 7.18: Measured vertical geometry from IMU and VPT

It can be seen from Figure 7.18 that the loaded geometry measurements show noticeably larger displacements compared to the unloaded geometry before Site 1, where the soft gravel and ash trackbed is. This is in contrast to the similar loaded and unloaded displacements on the stiffer ballasted section. The full effect of the primary suspension on the IMU measured geometry is still unknown. It could be said that the larger negative displacements at -24 m and -17 m are due to the suspension bounce after the preceding

large positive displacements, however it is encouraging that this does not happen at +21 m near the start of the stiffer track section, despite also having a large preceding positive displacement.

Unfortunately there was insufficient time to pursue the stiffness estimation work any further. However, the measurements taken so far suggest that it is worth continuing this work in the future to improve stiffness estimates.

7.5.7 Concurrent Work

Since the work on stiffness estimation started, some concurrent work at University College Dublin [1], has produced a method of vehicle-borne track stiffness measurement similar to the research work done towards this thesis. In Section 7.4.1 of this thesis, it was identified that an EKF or particle filter would be required to estimate the track stiffness parameter. In the paper [1] a cross-entropy optimisation method is used, which is indeed a form of particle filtering.

It is not known whether the algorithms demonstrated in this paper work when used with inertial data recorded from a train, as all of the processing is tested using only inertial data generated from a model. The authors do consider the uncertainty of vehicle parameters in their system, but evaluate this by adding gaussian noise to the parameters and simulated sensor readings. This seems to be an inadequate simulation of uncertainties for two reasons. Firstly, gaussian noise is statistically predictable, and is easily filtered out, leaving almost perfect data. Such filtering would inherently be achieved by the cross-entropy optimisation. Secondly, it is known from the work carried out towards this thesis that there are many unpredictable inputs which affect sensor readings, such as large jolts, producing sudden impulses and sometimes saturating sensors, and sensor drifts which affect the 'zero' measurements of the sensors.

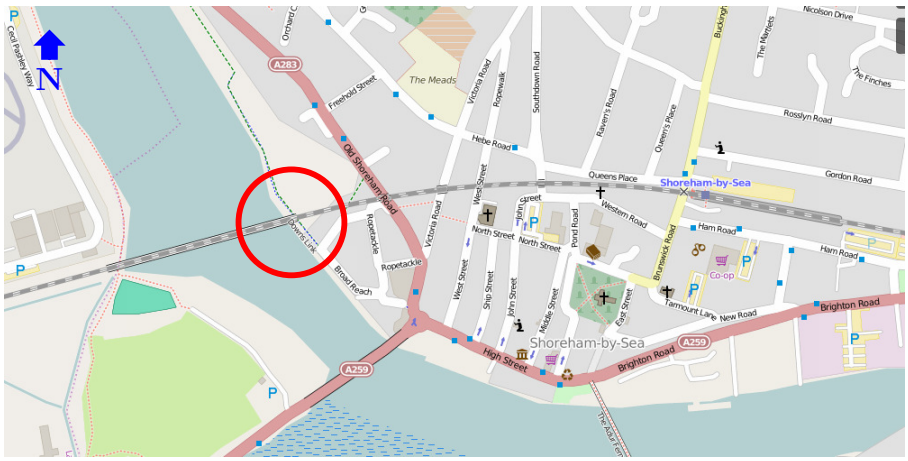
CHAPTER 8

CASE STUDIES

This chapter looks at two case studies at sites on the West Coastway Line in West Sussex, UK. These are both sites frequently traversed by the instrumented Southern Class 377 train.

8.1 Shoreham-by-Sea Bridge Transition

In the first case study, the transition onto a steel railway bridge at Shoreham-by-Sea (Figures 8.1 and 8.2) is examined. Specifically, the study focuses on the transition onto the bridge at its eastern end, on the westbound (down) line.



Coordinates: 50.8341, -0.2816 — © *OpenStreetMap contributors*

Figure 8.1: Map of Shoreham-by-Sea railway bridge and surrounding area

Photograph by 'Simon Carey', Source: geograph.org.uk
Reproduced under the Creative Commons Attribution-Share Alike 2.0 Generic license

Figure 8.2: The eastern end of the Adur Railway Bridge at Shoreham-by-Sea

The vertical track geometry was observed at the site over a 9 month period. Figure 8.3 shows the results plotted as a magnitude deviation chart.

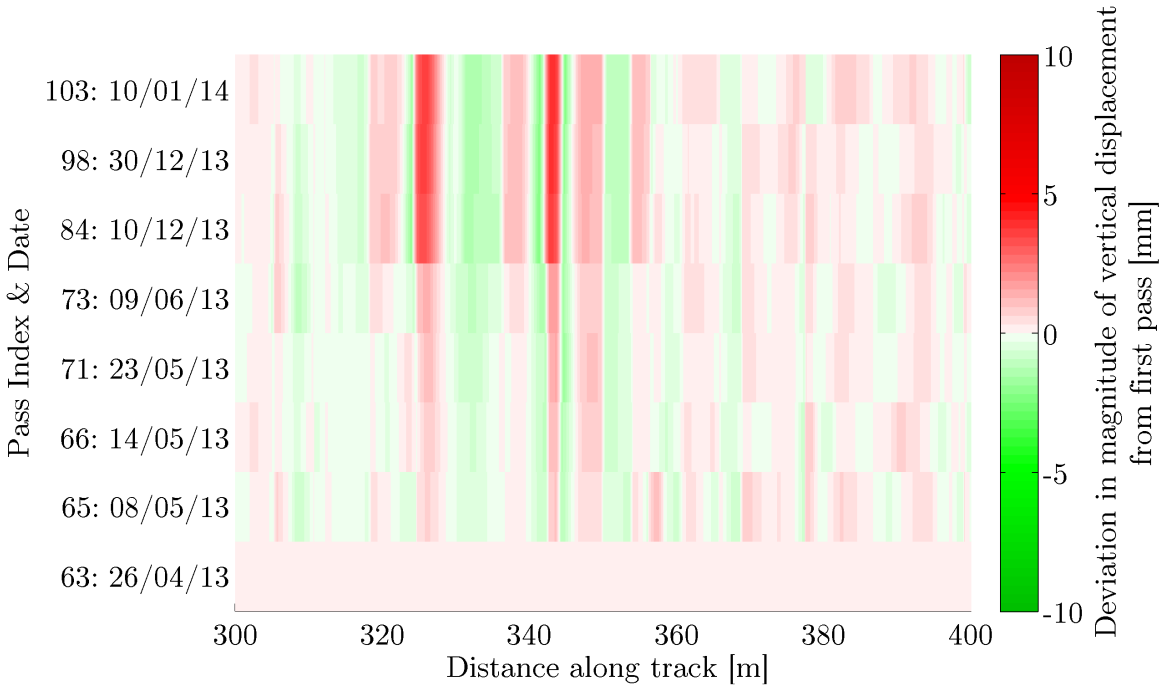


Figure 8.3: Differences between measured vertical displacements for multiple passes

It is evident from the chart that over time, faults emerge at around 326 m and 343 m, which is where the track makes its transition onto the bridge. Often, faults emerging at transition zones such as this can be attributed to changes in track stiffness between the ballasted track supported on soil, and the track supported on the much stiffer steel bridge structure.

It was decided to attempt to estimate the stiffness at the start and end of this 9 month period. For this, the peak-force track stiffness estimation method described in Section 7.4.2 was used. Figure 8.4 shows the results from the stiffness estimation using two passes at the start of the period (April 2013), and two passes from the end of the period (January 2014).

This figure shows an apparent increase in stiffness over the 9 month period in the region between 320 m and 375 m. Figure 8.5 shows the vertical displacements of the pairs of passes used to estimate the stiffness in Figure 8.5.

It can be seen that the displacement figure agrees with the stiffness estimation, as in the region from 325 m to 355 m, the April '13 passes differ by around 1 mm in places, despite the displacement magnitude being low, implying lower dynamic forces. Conversely, in this region, the January '14 passes are extremely close together despite a much larger displacement magnitude.

Whilst the stiffness estimation appears to show some change in the stiffness over time, the conclusions from the Long Marston field trial in Section 7.5 were that the method is not good enough to accurately represent track stiffness. It would be worthwhile generating further stiffness estimations using this data, once the estimation method has been improved in future work, potentially using the EKF or particle filter methods discussed in Section 7.4.

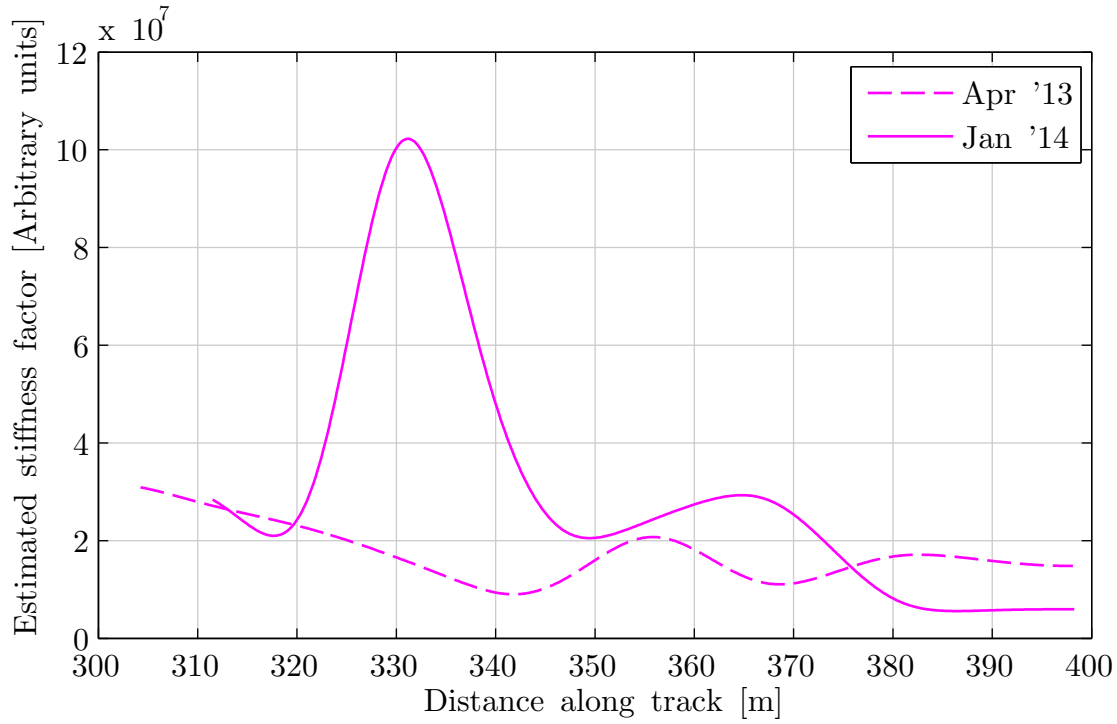


Figure 8.4: Estimated track stiffnesses in April '13 and January '14

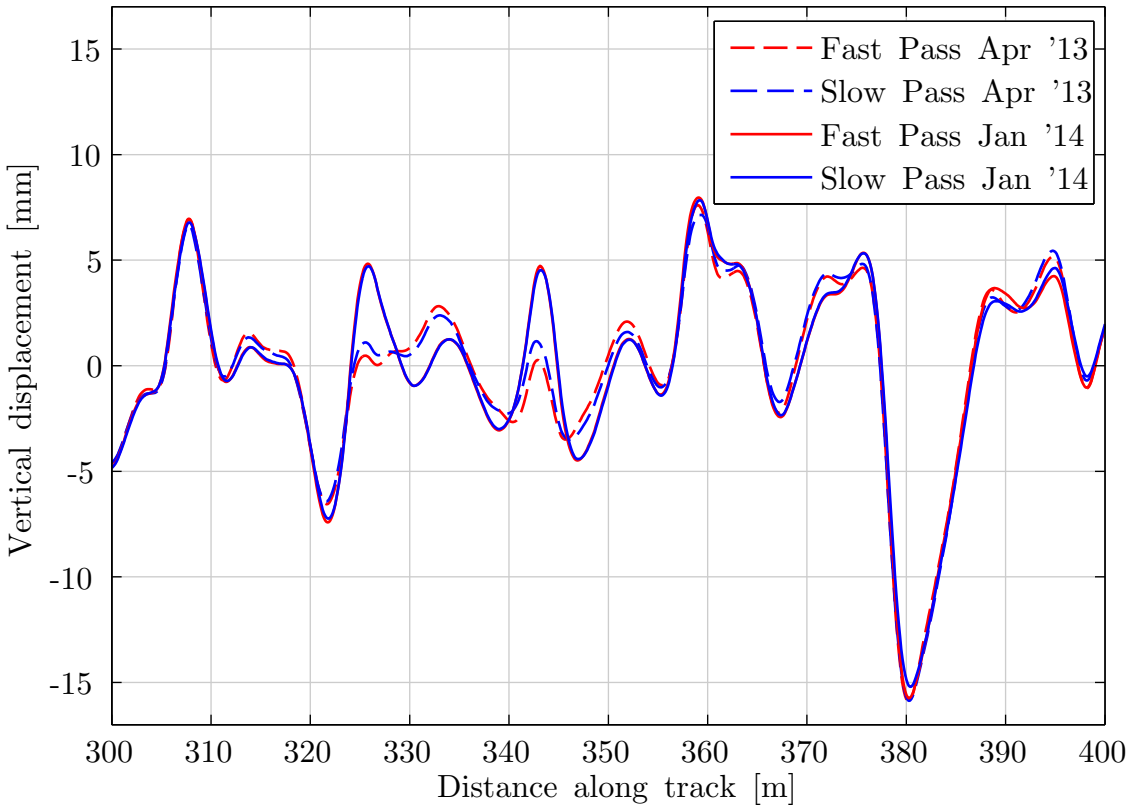
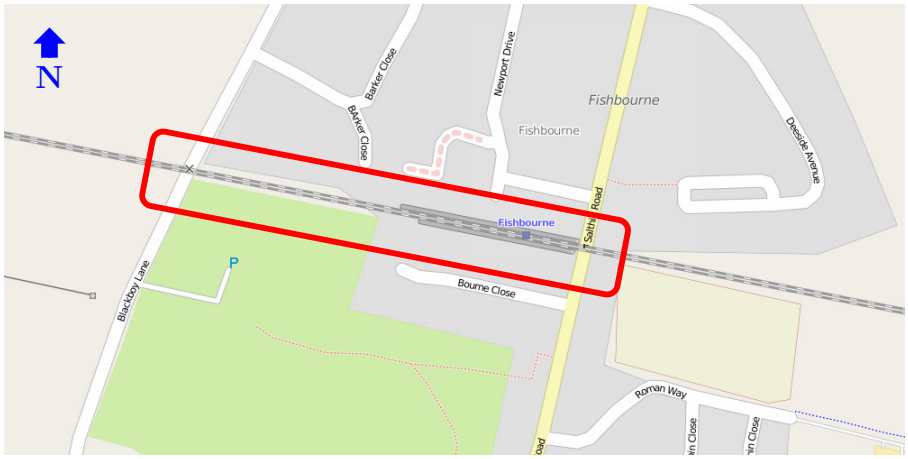


Figure 8.5: Vertical displacement of passes used to estimate stiffness

8.2 Fishbourne Station

The second case study looks at vertical displacement derived from the IMU, firstly after a track renewal through Fishbourne station (Figure 8.6), and secondly after tamping of the track has occurred. The study looks specifically at the eastbound (up) line.



Coordinates: 50.8391, -0.8154 — © *OpenStreetMap contributors*

Figure 8.6: Map of Fishbourne station and surrounding area

A section of the track between the two level crossings visible on the map had some particularly extreme vertical geometry features. A complete renewal of this section of track, including the ballast, took place on 24$^{\text{th}}$ October 2012. Figure 8.7 shows the vertical displacements measured before and after renewal. Installation of the Southern IMU had not been completed before the renewal, so NMT data is used.

Note that the distance to the 30 mile marker (on the x axis) decreases as the train is travelling. This is because the train is travelling in the 'up' direction, so mile marker values decrease in its direction of travel.

It can be seen from the figure that prior to renewal, there are some large peaks and troughs through the section, with two very large dips present. After renewal, the quality of the vertical geometry is much improved, with even the largest magnitude displacements falling within ±5 mm. It can be seen that almost none of the original track geometry is

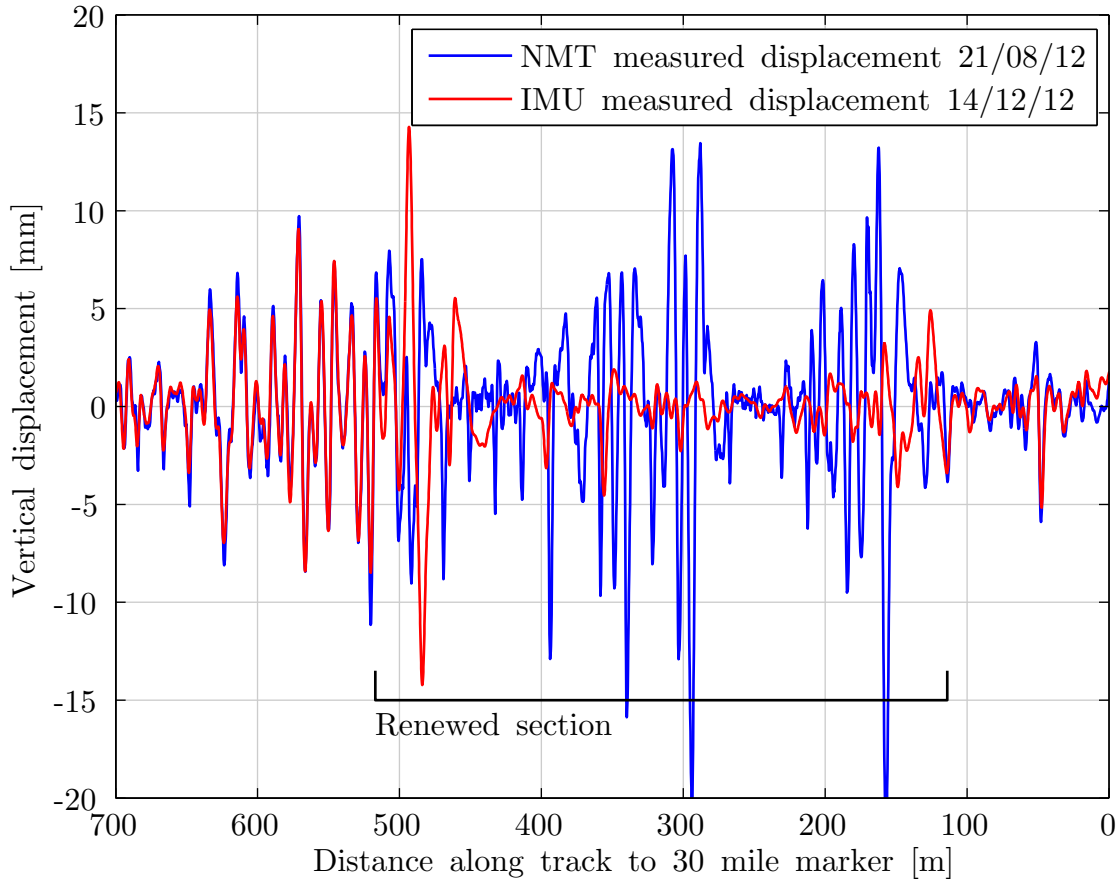


Figure 8.7: Vertical displacements measured before and after track renewal

evident in the geometry of the renewed track. This is because the renewal included the rails, sleepers, ballast, and trackbed.

Figure 8.8 shows a close up of the transition from the unrenewed track section at the start, to the renewed track section.

It can be seen from this figure that near the start of the renewal (493 m and 484 m) the the displacement magnitude of the renewed track is actually greater than that of the track pre-renewal. It is speculated that this is due to the transition between the renewed and original track. Usually when a full track and ballast renewal is performed, tamping is carried out prior to use of the track by in service vehicles. However, at the location of level crossings, track lifting cannot occur during tamping [110], because of implications this would have on the intersecting road. Whilst the crossing panels can be lifted to allow

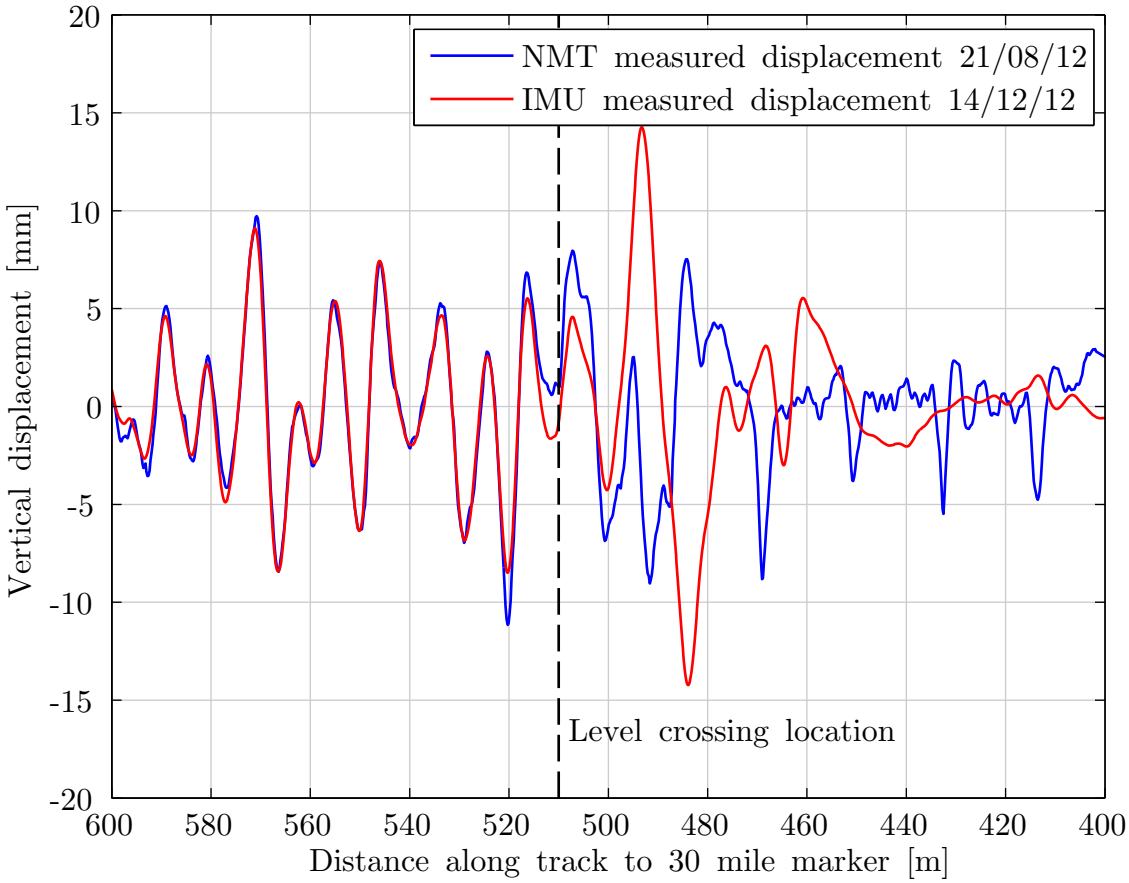


Figure 8.8: Vertical displacements measured before and after track renewal

tamping to be performed, sometimes this is not done, and instead tamping is stopped at the crossing and resumed on the other side. Handheld devices can be used to perform tamping closer to the crossing than is possible by a tamping machine, although this is not usually as effective. The result of this is that a low stiffness area may occur on the approach to the level crossing, where ballast has not been compacted to the same level as the plain track.

After the renewal was performed, the IMU continued to record measurements of the track for the following months. A magnitude deviation chart for this section of track is shown in Figure 8.9.

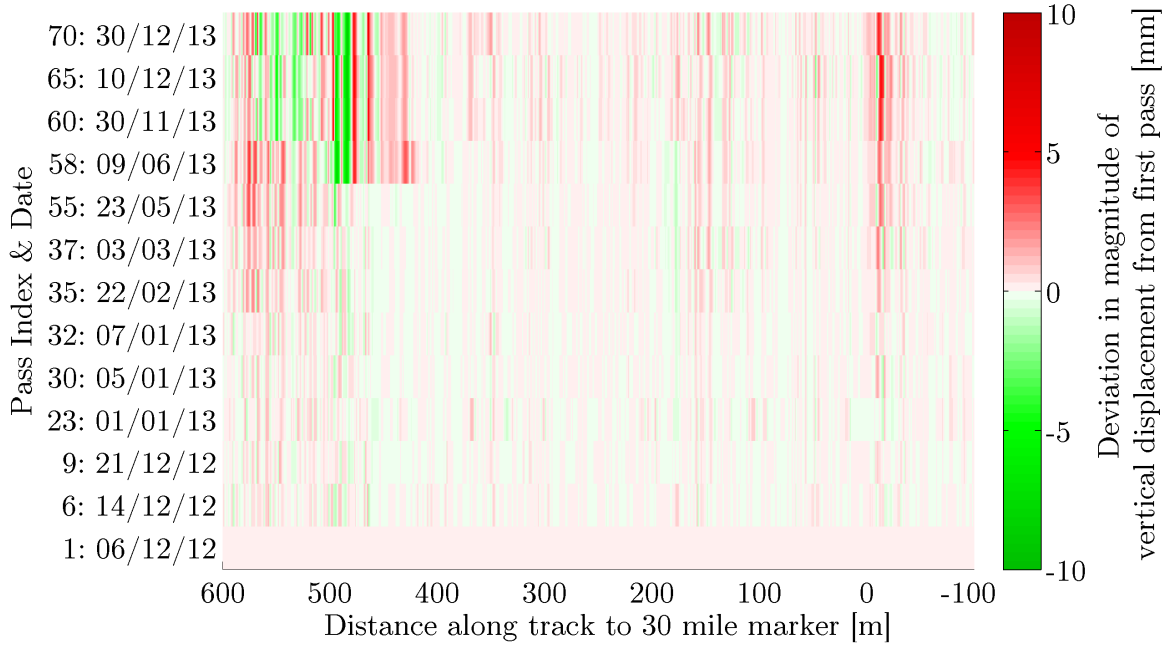


Figure 8.9: Differences in magnitude of measured vertical displacements for multiple passes of track through Fishbourne Station

From the figure, a degradation can be seen developing at around -20 m. This is the location of the level crossing to the west of Fishbourne station, where the railway crosses Salthill Road.

It can be seen from the figure that a significant change occurs between 600 m and 420 m on 9th June 2013. It was found that this was due to maintenance performed on this section of track on the 4th June 2013. Whilst some of this area of track seems to have a higher magnitude of displacement than previously, one part at ~480 m shows a significant improvement. Figure 8.10 shows the last pass of the IMU before maintenance, plotted with the first pass after maintenance, focused on the area between 600 m and 300 m.

It can be seen from the figure that before the maintenance, the large spikes at ~493 m and ~484 m which were introduced after the track renewal on 24th October 2012 (shown previously in Figure 8.8) are still present. In fact the positive spike has increased slightly in magnitude. The maintenance performed on 4th June 2013 has significantly reduced the severity of the displacements at the level crossing transition. It can also be seen that

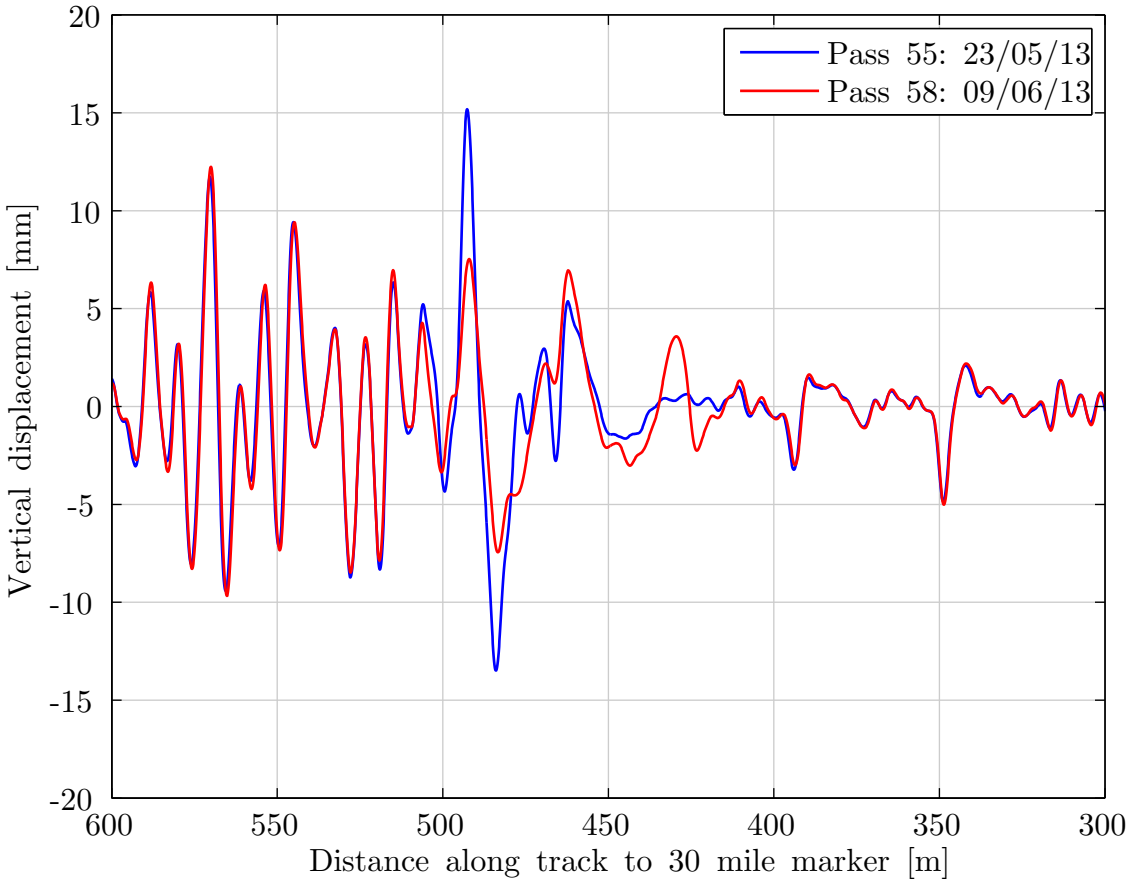


Figure 8.10: Vertical displacements measured before and after track maintenance on 04/06/13

the parts of the track which increased in displacement magnitude occur directly after the severe spike. This seems to be a symptom of the maintenance performed, but the resultant displacements are not severe in comparison to the spike which existed previously.

The approach to this level crossing was studied in depth as part of a collaborative study with the University of Southampton, also part of the Track 21 EPSRC programme grant. During the study, several aspects of the track condition were assessed, including track deflection, and noise and vibration. The results obtained using the IMU, as well as findings from track deflection measurements are presented in a journal paper [8].

CHAPTER 9

CONCLUSIONS

The work presented in this thesis has investigated whether a low-cost bogie-mounted IMU, fitted to an in-service railway vehicle can be used to effectively measure aspects of the track's geometry. This has been done with a view to using the information to better understand track degradation and to allow more effective maintenance scheduling.

9.1 Key Achievements

The following work has been carried out by the author, and has been presented in this thesis:

- A review of existing track measurement technologies and stiffness measurement methods. These have been used to identify advantages and disadvantages encountered by existing methods.

- An inertial measurement system has been designed, developed and installed on board an in-service train. Software has also been developed to manage data produced by the system. Data from this system has been gathered over a period of approximately two years.

- Methods of processing inertial data to obtain vertical displacement have been researched and implemented.

- Novel methods for fusion of sensor data to increase result accuracy have been developed. These include processing to account for sensor positioning on the bogie, and for the orientation of the train.

- Methods of mutually aligning data from multiple passes of a section of track have been developed through research.

- A novel dynamic alignment method has been developed using windowed data. This allows compensation for disagreement in longitudinal displacement estimations between multiple data sets.

- Methods have been developed to compare multiple passes of the train over the same track at different times.

- Theory behind stiffness calculation using two vehicle passes has been considered, and a simplified 'peak-force' method has been implemented.

- Developed processing methods have been applied to data collected in two case studies of sites traversed by the instrumented train. Conclusions about the condition of the track at these sites have been drawn from the data. Furthermore, information about the degradation and the effectiveness of maintenance at these sites has been found.

The thesis brings together research on the area of vehicle-based track measurement. Novel work has been done using a modified Bryson-Frazier filter to combine sensor data and produce very high accuracy vertical displacement estimations. A novel method of aligning data whilst allowing for longitudinal displacement disagreement has also been developed. Work has been done towards new methods of stiffness estimation, although at this stage their effectiveness has not been determined.

9.1.1 Answering the Hypothesis

The hypothesis defined at the beginning of this thesis made three statements which the work presented in this thesis has attempted to answer.

(a) *An unattended IMU mounted on the bogie of an in-service vehicle could be used to assess the condition of the railway track that the vehicle operates on.*

> This was answered during the literature review in Chapter 2, where it was proved through previous research that it is indeed possible to determine a measure of track condition by instrumentation of an in-service vehicle's bogie. Further work was carried out design, construct and install an IMU onto an in-service passenger train, which has showed that a bogie-mounted IMU is sufficient to assess the condition of the track to within < 1 mm of NMT measurements and with a self-consistent repeatability within 0.2 mm. This gave a good basis to answer the remaining two hypotheses.

(b) *Repeated data from multiple passes of the same track sections can be used to monitor the degradation of the vertical track geometry over time.*

> The research shows that it is possible to align and compare multiple passes of the same track, and a method of presenting comparison results to assess degradation and/or improvement of vertical track geometry has been defined. Case studies using this method have proven that it is possible to monitor and assess the degradation of track over time.

(c) *Repeated data from multiple passes of the same track sections can be used to estimate the stiffness of the track.*

> Methods of track stiffness estimation have been considered, although time restrictions have prevented these from being implemented. A simplified method was developed, and it has been shown that there is some correlation between the

estimated results and the trackside stiffness determination. However, it was not possible to determine whether the estimated results were accurate or not. Future work is required to fully answer this part of the hypothesis.

9.2 Potential for Further Research

Previous chapters have identified areas in which work could be extended in the future. These areas could form the bases of further research projects. The identified research areas were:

- Modelling of vehicle dynamics, for the purposes of minimising differences in measured geometry for passes with different vehicle speeds.

- Further work on stiffness estimation through in-depth modelling of suspension and track dynamics. Further research into work done to eliminate the effects of vehicle suspension, such as that presented by Tsunashima *et al.* [36–39] could be done. This work could then be used to improve stiffness estimation techniques. Proper verification of the improved techniques could be done using a proven stiffness measurement method capable of continuous measurements along a length of track.

- Improvement of vehicle location data, using research described in Section 6.1. Better positioning data would allow automated calculation of track degradation to be performed, as mutual alignment of multiple passes of track would be continuous along the train's route.

- Using the data collected during this PhD work to determine other track geometry features. The recorded information, which includes GPS, tacho and longitudinal acceleration data, also has potential for use in the assessment of the trackbed, and of the vehicle dynamics.

9.3 Publications

Three conference papers have been written relating to the work in this thesis. The author has also contributed work to two journal papers.

The first conference paper [111] was presented by the author at the IET Railway Condition Monitoring (RCM) 2014 conference in Birmingham, UK. It provides a background to the in-service monitoring system described in this thesis, along with an overview of the processing used to obtain vertical geometry information, and comparisons between multiple passes of the same track section. It also presents the differences found between forward- and backward-oriented passes (as described in Section 5.4). At the time the paper was written, the changes to the Kalman filter which were developed to minimise the differences due to orientation, had not yet been developed.

The second conference paper [85] was presented at the IMechE Stephenson Conference in April 2015 by the author. The work in this paper considers the future adaption of the system to detect track faults using processing on-board the vehicle, and to transmit fault reports using a cellular phone network connection.

A third conference paper [109] was presented as an e-poster presentation at the World Congress for Railway Research 2016 in Milan. It includes the work done towards measuring loaded and unloaded track geometry at the Long Marston railway facility, which was used in Chapter 7 of this thesis.

Work was contributed to a journal paper [8] published in Transportation Geotechnics, volume 1, issue 4. The paper describes a study of a level crossing at Fishbourne in Sussex, UK, carried out by the University of Southampton and the University of Birmingham. The site has also been studied in the work in this thesis (Section 8.2). Work in the journal paper describes the IMU system used on the instrumented Southern Class 377 train. Data collected from the IMU system is presented, and is used to verify trackside measurements

of sleeper displacements and velocities taken by the University of Southampton.

Some of the IMU results are presented in another journal paper [82] published in Vehicle System Dynamics. The paper also presents some of the mutual alignment and comparison techniques developed during the course of this thesis.

9.4 Concluding Remarks

The work carried out towards this thesis has shown that accurate estimates of the vertical track geometry can be obtained using a relatively compact and inexpensive inertial measurement unit mounted on the bogie of an in-service train. Information about the rate of degradation of the track can also be gained using the processing methods developed.

This has great potential in improving how the track condition of a railway network is monitored. Instrumentation fitted to multiple in-service vehicles can provide a rich database of track geometry data. Dedicated measurement vehicles would only be required to verify track geometry at sparse intervals, and to investigate faults at areas identified by in-service systems. This has potential for saving costs on operating and maintaining measurement vehicles, and for increasing track availability and ultimately capacity. It is not intended that an in-service inertial measurement system would replace dedicated measurement vehicles, which provide a greater number of parameter measurements, such as cross-level and gauge. It is also often the case that dedicated vehicles are frequently maintained and calibrated, so that data is extremely accurate.

The increased frequency at which a track can be monitored using an in-service system also means that faults can be detected in a more timely fashion. Using a dedicated measurement vehicle which traverses a track infrequently, a fault could go undetected for several weeks, whereas with an in-service system fitted to multiple trains, it is possible that a fault could be detected within an hour. Assessment of the degradation of a section of track, with frequent data from in-service vehicles would also enable accurate predictions

of failure times, meaning that faults are more likely to be predicted and corrected before they become failures.

The ability of to monitor track frequently also allows in-depth information about degradation rates and the effectiveness of maintenance to be gained. By inspecting measurements taken in the days following a maintenance operation, the effectiveness of the maintenance can be assessed. This enables determination of whether the kind of maintenance being performed at a specific location is the most effective possible.

A better understanding of degradation and maintenance allows more efficient scheduling of maintenance, greater availability of the railway, and reduced costs.

Appendix A

DATA PROCESSING

A.1 Domain Conversion Algorithm

An algorithm is used to convert samples in the temporal domain to the spatial domain. Input U is a vector of samples in the time domain, spaced δt apart. Corresponding elements of input vector V are the vehicle speed at the time of each sample. Output Y is a vector containing samples in the spatial domain, spaced δx apart. Output vector X contains the spatial positions for corresponding elements of Y. In the following pseudo-code, the ':=' operator is used to indicate assignment.

The following initialisations are made:

$t_{in} := \delta t$, $t_{out} := 0$, $t_{prev} := 0$, $x_{current} := 0$, $x_{target} := \delta x$, $y := 0$, $u := 0$, $j := 1$

After which the following is executed:

```
while i < length(U)-1 {
        if x_current + V_i(t_in - t_out) ≥ x_target {
                Δt := (x_target - x_current) × V_i
                Δu := U_{i+1} - U_i
                Y_j := (y + uΔt + [(Δt)²Δu / 2δt]) / (t_out + Δt - t_prev)
                X_j := x_target
```

$j := j + 1$

$y := 0$

$t_{out} := t_{out} + \Delta t$

$t_{prev} := t_{out}$

$x_{current} := x_{target}$

$x_{target} := x_{target} + \delta x$

$u := u + \dfrac{\Delta t \Delta u}{\delta t}$

}

else {

$y := y + \dfrac{(t_{in} - t_{out})(u + U_{i+1})}{2}$

$x_{current} := x_{current} + V_i(t_{in} - t_{out})$

$t_{out} := t_{in}$

$t_{in} := t_{in} + \delta t$

$u := U_{i+1}$

$i := i + 1$

}

}

A.2 Approximation of the Sine of an Angle

The situation in Figure A.1 is considered, where a shallow gradient exists, and δx is significantly larger than δy.

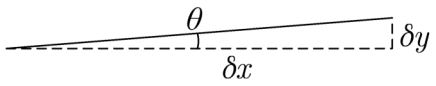


Figure A.1: Angle of a shallow gradient

For any values of δx and δy, basic trigonometry can be used to find the angle θ:

$$\frac{\delta y}{\delta x} = \sin(\theta) \tag{A.1}$$

However, for very small angles, expressed in radians, it can be said that:

$$\frac{\delta y}{\delta x} \simeq \theta \quad \text{where} \quad \delta x \gg \delta y \tag{A.2}$$

Figure A.2 shows the sine of an angle between -π and π radians, and also a magnified graph showing angles between -0.01 and 0.01 radians. It can be seen from the magnified graph that at small angles, the relationship is almost linear, with a factor of 1. This proves that (A.2) is true.

Trigonometric functions such as sine are relatively computationally intensive compared to basic operations such as multiplication and division. This means that the use of this approximation can give a significant speed advantage where a large number of angles must be calculated.

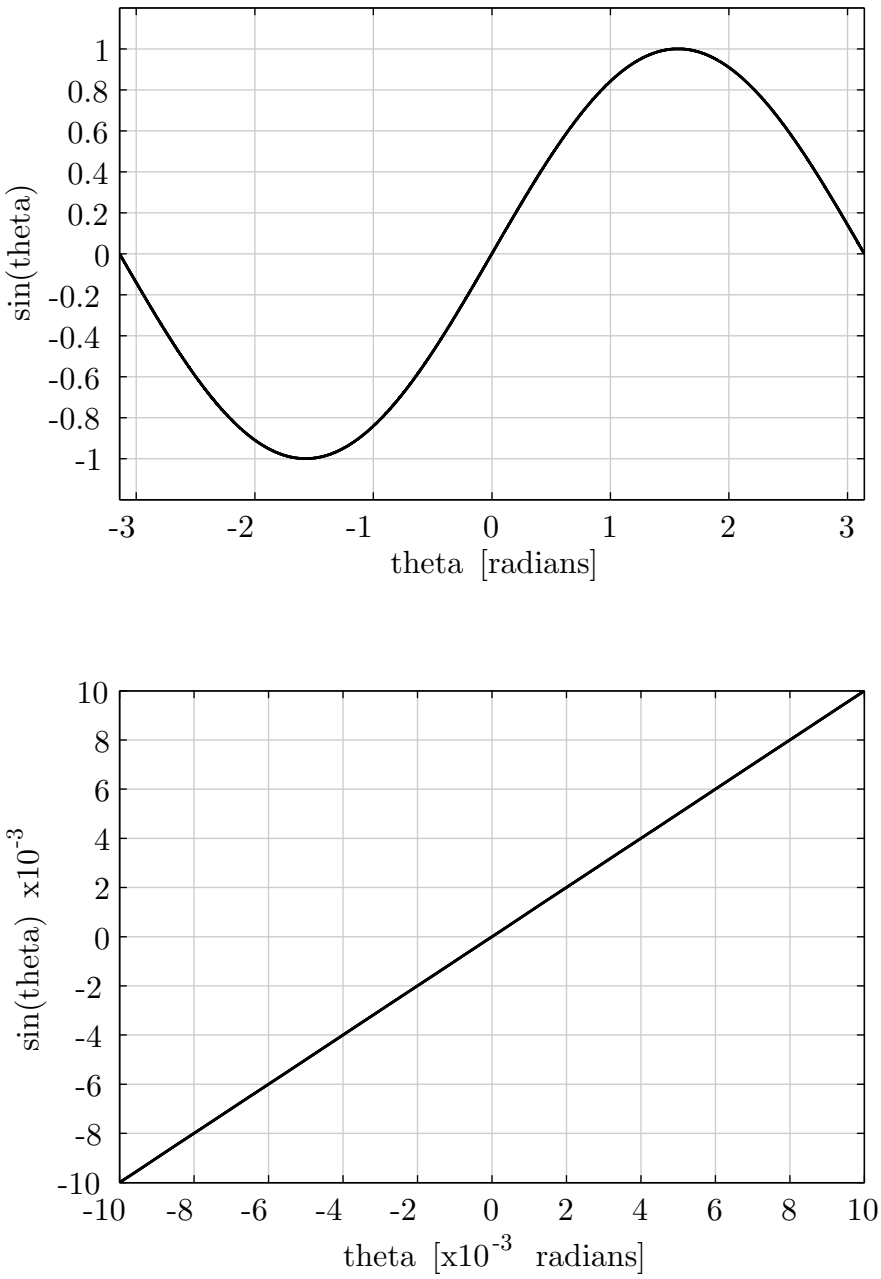


Figure A.2: The sine of an angle expressed in radians

LIST OF REFERENCES

[1] P. Quirke, D. Cantero, E. J. OBrien, and C. Bowe, "Drive-by detection of railway track stiffness variation using in-service vehicles", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, Mar. 2016.

[2] R. Lewis, *Track Geometry Recording and Usage*. Apr. 2011. Notes from a lecture to Network Rail, p. 6.

[3] British Railways Board, *Track Recording Handbook*. Group Standard GC/EH0038, Oct. 1993. appendix D.2.

[4] "BS EN 13848-1:2003 Railway applications. Track. Track geometry quality. Characterisation of track geometry", section 4.2, 2003.

[5] T. Dahlberg, "Railway track settlements - a literature review", technical report, Jan. 2004.

[6] W. Powrie and J. A. Priest, *Behaviour of ballasted track during high speed train passage*. Feb. 2011. Presentation given at the University of Southampton Railways Day.

[7] Network Rail, *www.networkrail.co.uk/timetables-and-travel/Delays-explained /Winter-weather/*. 2015.

[8] L. Le Pen, G. Watson, W. Powrie, G. J. Yeo, P. F. Weston, and C. Roberts, "The behaviour of railway level crossings: Insights through field monitoring", *Transportation Geotechnics*, vol. 1, pp. 201–213, Dec. 2014.

[9] P. Waterfall, B. Temple, C. Hardwick, J. Sharam, and A. Plumb, "Video measurement techniques for understanding wheel-induced lateral forces and track component deflections", in *Proceedings from the 2015 IMechE Stephenson Conference*, (London, UK), pp. 11–29, Institution of Mechanical Engineers, Apr. 2015.

[10] H. Kim, P. Weston, C. Roberts, and J. A. Priest, "Trackside measurement at railway critical zones using sensors and vehicle-borne instrumentation", *5th IET Conference on Railway Condition Monitoring and Non-Destructive Testing*, 2011.

[11] H. Kim, L. Saade, P. F. Weston, and C. Roberts, "Measuring the deflection of a sequence of sleepers at a transition zone", in *Proceedings of the 6th IET Railway Condition Monitoring Conference*, Sept. 2014.

[12] Innotrack, "Methods of track stiffness measurements", technical report, 2006.

[13] B. Coelho, J. A. Priest, P. Holscher, and W. Powrie, "Monitoring of transition zones in railways", June 2009.

[14] K. Tzanakakis, *The Railway Track and Its Long Term Behaviour*. Springer, 2013. p. 84.

[15] X. Lei, *High Speed Railway Track Dynamics: Models, Algorithms and Applications*. 2017. chapter 14.

[16] A. Paixão, E. Fortunato, and R. Calçada, "Design and construction of backfills for railway track transition zones", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 229, pp. 58–70, Jan. 2015.

[17] E. Tutumluer, T. D. Stark, D. Mishra, and J. P. Hyslip, "Investigation and mitigation of differential movement at railway transitions for US high speed passenger rail and joint passenger/freight corridors", in *Proceedings of the 2012 Joint Rail Conference*, (Philadelphia, PA), ASME, Apr. 2012.

[18] B. Coelho, P. Hölscher, J. Priest, W. Powrie, and F. Barends, "An Assessment of Transition Zone Performance", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 225, pp. 129–139, Mar. 2011.

[19] A. Ortega, S. Blainey, and J. Preston, "Reducing Maintenance Costs For Ballasted Track: Evidence From Britain", *International Journal of Transport Development and Integration*, vol. 1, no. 3, pp. 318–328, 2017.

[20] D. Briginshaw, "London Underground steps up Victoria Line frequency", *Article in International Railway Journal (IRJ)*, Feb. 2013.

[21] W. Powrie and L. Le Pen, *A Guide to Track Stiffness*. Cross Industry Track Stiffness Working Group, Aug. 2016.

[22] B. Solomon, *Railway Maintenance: The men and machines that keep the railroads running*. Voyageur Press, 2001.

[23] "Detecting Rail Defects", in *Railway Wonders of the World*, Part 29: Amalgamated Press, Aug. 1935.

[24] E. J. C. Stewart, *A Distributed Instrumentation System for the Acquisition of Rich, Multi-Dimensional Datasets from Railway Vehicles*. PhD thesis, University of Birmingham, School of Electronic, Electrical and Computer Engineering, 2012.

[25] R. H. Hronik, "Apparatus for high-speed measurement of track geometry", patent, 1970.

[26] Y. Jin, Y. Goto, Y. Nishimoto, H. Naito, Y. Fukumoto, and A. Iwake, "Advanced Dynamic Obstacle-detecting System for Railway Surroundings Using a Highly Accurate Laser-sectioning Method", 1990.

[27] M. Benfield, "The modern gauging process", *Journal and Report of Proceedings of the Permanent Way Institution*, vol. 127, no. 3, pp. 128–130, 2009.

[28] E. Yazawa and K. Takeshita, "Development of Measurement Device of Track Irregularity using Inertial Mid-chord Offset Method", Quarterly Report of RTRI 43(3), Jan. 2002.

[29] E. Yazawa, "Track inspection technologies", *Railway Technology Avalanche*, 2003.

[30] S. Ono, A. Numakura, and T. Odaka, "High-Speed Track Inspection Technologies", *JR EAST Technical Review*, vol. 2, 2003.

[31] E. D. Bokhman, A. M. Boronachin, Y. V. Filatov, D. Y. Larionov, L. N. Podgornaya, R. V. Shalymov, and G. N. Zuzev, "Optical-inertial system for railway track diagnostics", in *2014 DGON Inertial Sensors and Systems (ISS)*, pp. 1–17, Sept. 2014.

[32] A. M. Boronahin, Y. V. Filatov, D. Y. Larionov, L. N. Podgornaya, and R. V. Shalymov, "Fusion of heterogeneous sensor information for railway track diagnostics", in *2014 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, pp. 1–6, Oct. 2014.

[33] A. M. Boronahin, Y. V. Filatov, D. Y. Larionov, L. N. Podgornaya, and R. V. Shalymov, "Measurement system for railway track condition monitoring", in *2015 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EIConRusNW)*, pp. 155–158, Feb. 2015.

[34] A. M. Boronahin, A. S. Kukaev, D. Y. Larionov, L. N. Podgornaya, R. V. Shalymov, and E. D. Bokhman, "Application of regression analysis for data processing of inertial track monitoring system", in *2016 IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EIConRusNW)*, pp. 151–155, Feb. 2016.

[35] Y. Hayashi, T. Kojima, H. Tsunashima, and Y. Marumo, "Real Time Fault Detection of Railway Vehicles and Tracks", in *The Institution of Engineering and Technology International Conference on Railway Condition Monitoring, 2006*, pp. 20–25, Nov. 2006.

[36] H. Tsunashima, T. Kojima, Y. Marumo, A. Matsumoto, and T. Mizuma, "Condition monitoring of railway track and driver using in-service vehicle", in *2008 4th IET International Conference on Railway Condition Monitoring*, pp. 1–6, June 2008.

[37] H. Tsunashima, H. Mori, K. Yanagisawa, M. Ogino, and A. Asano, "Condition monitoring of railway tracks using compact size on-board monitoring device", in *Proceedings of the 6th IET Railway Condition Monitoring Conference*, Sept. 2014.

[38] H. Tsunashima, Y. Naganuma, A. Matsumoto, T. Mizuma, and H. Mori, "Japanese railway condition monitoring of tracks using in-service vehicle", in *5th IET Conference on Railway Condition Monitoring (RCM 2011)*, 2011.

[39] H. Tsunashima, Y. Naganuma, A. Matsumoto, T. Mizuma, and H. Mori, "Condition monitoring of railway track using in-service vehicle", in *Reliability and safety in railway*, ch. 12: InTech, Mar. 2012.

[40] H. Ishii, Y. Fujino, Y. Mizuno, and K. Kaito, "The study of train intelligent monitoring system using acceleration of ordinary trains", in *Proceedings from the Asia-Pacific Workshop on Structural Health Monitoring*, (Yokohama, Japan), 2006.

[41] O. Heirich, A. Lehner, P. Robertson, and T. Strang, "Measurement and analysis of train motion and railway track characteristics with inertial sensors", in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, pp. 1995–2000, Oct. 2011.

[42] J. Lee, S. Choi, S. Kim, Y. Kim, S. Kim, and C. Park, "Track condition monitoring by in-service trains: A comparison between axle-box and bogie accelerometers", in *5th IET Conference on Railway Condition Monitoring and Non-Destructive Testing (RCM 2011)*, pp. 1–6, Nov. 2011.

[43] J. S. Lee, S. Choi, S.-S. Kim, C. Park, and Y. G. Kim, "A Mixed Filtering Approach for Track Condition Monitoring Using Accelerometers on the Axle Box and Bogie", *IEEE Transactions on Instrumentation and Measurement*, vol. 61, pp. 749–758, Mar. 2012.

[44] C. Chellaswamy, V. Akila, A. Dinesh Babu, and N. Arasan, "Fuzzy logic based railway track condition monitoring system", in *2013 International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICE-CCN)*, pp. 250–255, Mar. 2013.

[45] M. Bocciolone, A. Caprioli, A. Cigada, and A. Collina, "A measurement system for quick rail inspection and effective track maintenance strategy", *Mechanical Systems and Signal Processing*, vol. 21, pp. 1242–1254, Apr. 2007.

[46] M. Molodova, Z. Li, A. Núñez, and R. Dollevoet, "Automatic Detection of Squats in Railway Infrastructure", *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, pp. 1980–1990, Oct. 2014.

[47] M. Molodova, M. Oregui, A. Núñez, Z. Li, J. Moraal, and R. Dollevoet, "Axle box acceleration for health monitoring of insulated joints: A case study in the Netherlands", in *2014 IEEE 17th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 822–827, Oct. 2014.

[48] P. Ackroyd, S. Angelo, B. Nejikovsky, and J. Stevens, "Remote ride quality monitoring of Acela train set performance", in *Proceedings of the 2002 ASME/IEEE Joint Rail Conference*, (Washington DC), Apr. 2002.

[49] S. King, "The UK's fastest Track Monitoring System as used on the Channel Tunnel Rail Link", in *IEEE seminar on Railway Condition Monitoring 2004*, 2004. AEA Technology Rail, UK.

[50] G. Presle, "The EM 250 high-speed track recording coach and EM-SAT 120 track survey car, as networked track geometry diagnosis and therapy systems", *Rail Engineering International Edition*, no. 3, 2000.

[51] Andian Technologies, *www.andian.com*. 2014.

[52] A. C. Bidaud, "Geometric track and track/vehicle analyzers and methods for controlling railroad systems", Patents US 7164975, US 6681160, CA 2489980, Jan. 2007.

[53] A. C. Bidaud, "Track analyzers and methods for controlling railroad systems", Patent CA 2474757, Apr. 2009.

[54] H. E. McAnaw, "The system that measures the system", *NDT&E International*, vol. 36, pp. 169–179, 2003.

[55] Donfabs & Consillia Ltd., *Vehicle Track Analysis Software (VTAS)*. Donfabs & Consillia Ltd., 2006. Software.

[56] "London Underground chose MERMEC Unattended Geometry Measuring Systems (UGMS) for line inspection", *MERMEC Group Press Release*, Sept. 2012.

[57] R. B. Lewis, "Track-recording techniques used on British Rail", *IEE Proceedings B - Electric Power Applications*, vol. 131, pp. 73–81, May 1984.

[58] "Intensive inspection at high speed", *Article in Railway Gazette*, July 2003.

[59] C. Baker, A. Quinn, H. Hemida, M. Sterling, M. Gallagher, and J. Morden, "A comparison of full scale and model scale measurements of train Aerodynamic characteristics", in *Proceedings from the 2015 IMechE Stephenson Conference*, (London, UK), pp. 637–648, Institution of Mechanical Engineers, Apr. 2015.

[60] T. Kolbe and R. Kratochwille, "ICE-S Vehicle reaction measurement and track geometry measurement on the same measuring train: Results of the comparison of the two different track inspection methods", in *Proceedings from the 2015 IMechE Stephenson Conference*, (London, UK), pp. 743–754, Institution of Mechanical Engineers, Apr. 2015.

[61] Automain, "Inspection of track from in-service freight trains", in *Augmented Usage of Track by Optimization of Maintenance, Allocation and Inspection of Railway Networks*, Deliverable report, Dec. 2013.

[62] C. Roberts, P. Weston, C. S. Ling, and C. Goodman, "New methods for track monitoring", in *IEEE seminar on Railway Condition Monitoring 2004*, 2004.

[63] P. F. Weston, P. LI, C. S. Ling, C. J. Goodman, R. M. Goodall, and C. Roberts, "Track and Vehicle Condition Monitoring during Normal Operation Using Reduced Sensor Sets", *HKIE Transactions*, vol. 13, pp. 47–54, Jan. 2006.

[64] P. F. Weston, C. S. Ling, C. J. Goodman, C. Roberts, P. Li, and R. M. Goodall, "Monitoring lateral track irregularity from in-service railway vehicles", *Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit*, vol. 221, pp. 89–100, Mar. 2007.

[65] P. F. Weston, C. S. Ling, C. Roberts, C. J. Goodman, P. Li, and R. M. Goodall, "Monitoring vertical track irregularity from in-service railway vehicles", *Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit*, vol. 221, pp. 75–88, Mar. 2007.

[66] C. P. Ward, P. F. Weston, E. J. C. Stewart, H. Li, R. M. Goodall, C. Roberts, T. X. Mei, G. Charles, and R. Dixon, "Condition monitoring opportunities using vehicle-based sensors", *Proceedings of the Institution of Mechanical Engineers Part F-Journal of Rail and Rapid Transit*, vol. 225, pp. 202–218, 2011.

[67] J. A. Priest and W. Powrie, "Determination of Dynamic Track Modulus from Measurement of Track Velocity during Train Passage", *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 135, no. 11, pp. 1732–1740, 2009.

[68] D. Bowness, A. C. Lock, W. Powrie, J. A. Priest, and D. J. Richards, "Monitoring the dynamic displacements of railway track", *Proceedings of the Institution of*

Mechanical Engineers Part F-Journal of Rail and Rapid Transit, vol. 221, pp. 13–22, Mar. 2007.

[69] C. Govan, *The Use of Falling-Weight Deflectometers in Determining Critical Velocity Problems*. Dec. 2013. Presentation given at the Railway Track Science & Engineering Workshop, UIC, Paris.

[70] M. P. N. Burrow, A. H. C. Chan, and A. Shein, "Deflectometer-based analysis of ballasted railway tracks", *Proceedings of the ICE - Geotechnical Engineering*, vol. 160, pp. 169–177, Jan. 2007.

[71] E. Berggren, *Railway Track Stiffness: Dynamic Measurements and Evaluation for Efficient Maintenance*. PhD thesis, KTH Royal Institute of Technology, 2009.

[72] E. Berggren and M. Berg, "Simulation, Development and Field Testing of a Track Stiffness Measurement Vehicle", in *Proceedings from the 8th International Heavy Haul Conference*, (Rio de Janeiro), 2005.

[73] M. Hosseingholian, M. Froumentin, and D. Levacher, "Continuous Method to Measure Track Stiffness - A New Tool for Inspection of Rail Infrastructure", *World Applied Sciences Journal*, vol. 6, no. 5, pp. 579–589, 2009.

[74] M. Hosseingholian, M. Froumentin, and A. Robinet, "Dynamic Track Modulus from Measurement of Track Acceleration by Portancemetre", in *Proceedings from the 9th World Congress on Railway Research*, (Lille, France), May 2011.

[75] G. Soldati, "Gleiseinsenkungen mit dem Einsenkungsmesswagen Achslast 20t", technical report, 2007.

[76] E. Berggren, "Beating stiff track", *International Railway Journal (IRJ)*, Oct. 2013.

[77] B. McVey, S. Farritor, C. Norman, N. Wood, R. Arnold, M. Fateh, and M. El-Sibaie, "Track modulus measurement from a moving railcar", in *Proceedings of the AREMA 2005 Annual Conferences*, (Chicago, IL), 2005.

[78] S. Lu, *Real-Time Vertical Track Deflection Measurement System*. PhD thesis, University of Nebraska - Lincoln, Feb. 2008.

[79] J. Amoore, P. Richards, A. Wright, P. Langdale, and J. Iaquinta, "Continuous assessment of the stiffness of tracks at line speed", in *5th IET Conference on Railway Condition Monitoring and Non-Destructive Testing (RCM 2011)*, pp. 1–6, Nov. 2011.

[80] E. Stewart, P. Weston, S. Hillmansen, and C. Roberts, "Using Bogie-Mounted Sensors to Understand the Dynamics of Third Rail Current Collection Systems",

Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, vol. 225, pp. 219–227, Mar. 2011.

[81] S. Green, D. Hickson, D. Ward, C. Roberts, P. Weston, and E. Stewart, "Monitoring the DC third rail interface using an in-service train", in *5th IET Conference on Railway Condition Monitoring and Non-Destructive Testing (RCM 2011)*, pp. 1–5, Nov. 2011.

[82] P. Weston, C. Roberts, G. Yeo, and E. Stewart, "Perspectives on railway track geometry condition monitoring from in-service railway vehicles", *Vehicle System Dynamics*, vol. 53, Apr. 2015.

[83] "BS EN 12299:2009 Railway applications - Ride comfort for passengers - Measurement and evaluation", 2009.

[84] Texas Instruments, *DS36276, DS36277, Application Note 1031 TIA/EIA-422-B Overview.* 2002.

[85] G. J. Yeo, P. F. Weston, and C. Roberts, "Detection of vertical track faults using an instrumented in-service vehicle: Considering efficient real-time fault detection", in *Proceedings from the 2015 IMechE Stephenson Conference*, (London, UK), pp. 755–763, Institution of Mechanical Engineers, Apr. 2015.

[86] Network Rail, "RfI: Data Collection Services - Data Collection Technology Capability", Feb. 2015.

[87] P. F. Weston, C. Roberts, C. J. Goodman, and C. S. Ling, "Condition Monitoring of Railway Track using In-Service Trains", in *2006 IET International Conference On Railway Condition Monitoring*, pp. 26–31, Nov. 2006.

[88] T. D. Gillespie, "Rigid Body Bounce/Pitch Motions", in *Fundamentals of Vehicle Dynamics*, pp. 168–171, ch. 5: SAE International, Feb. 1992.

[89] S. M. Bozic, *Digital and Kalman Filtering.* Edward Arnold, 2nd ed., 1994.

[90] G. J. Bierman, "Factorization Methods for Discrete Sequential Estimation", vol. 128 of *Mathematics in Science and Engineering*, pp. 223–225, 1977.

[91] P. F. Weston and J. P. Norton, "Detection and estimation of abrupt changes in input or state", *International Journal of Control*, vol. 67, no. 5, pp. 699–712, 1997.

[92] K. Song, X. Chen, and Y. Lin, "Wheelbase Filtering Effect on Vehicle Ride Dynamics", in *Proceedings of the FISITA 2012 World Automotive Congress*, vol. 8 of *Vehicle Design and Testing (II)*, pp. 1183–1195, 2013.

[93] S. Iwnicki, *Handbook of Railway Vehicle Dynamics.* CRC Press, May 2006.

[94] S. Saab, "A map matching approach for train positioning. I. Development and analysis", *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 467–475, Mar. 2000.

[95] S. Saab, "A map matching approach for train positioning. II. Application and experimentation", *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 476–484, Mar. 2000.

[96] O. Heirich, P. Robertson, A. C. García, and T. Strang, "Bayesian Train Localization Method Extended By 3d Geometric Railway Track Observations From Inertial Sensors", 2012.

[97] O. G. Crespillo, O. Heirich, and A. Lehner, "Bayesian GNSS/IMU tight integration for precise railway navigation on track map", in *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, pp. 999–1007, May 2014.

[98] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*. Cambridge, UK ; New York: Cambridge University Press, 3rd edition ed., Aug. 2007.

[99] M. Zaman, A. Gopalasingam, and J. G. Laguros, "Consolidation Settlement of Bridge Approach Foundation", *Journal of Geotechnical Engineering*, vol. 117, pp. 219–240, Feb. 1991.

[100] J. Long, S. Olson, T. Stark, and E. Samara, "Differential Movement at Embankment-Bridge Structure Interface in Illinois", *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1633, pp. 53–60, Jan. 1998.

[101] L. Puzavac, Z. Popović, and L. Lazarević, "Influence of Track Stiffness on Track Behaviour under Vertical Load", *PROMET - Traffic & Transportation*, vol. 24, pp. 405–412, Oct. 2012.

[102] J. N. Varandas, P. Hölscher, and M. A. G. Silva, "Dynamic behaviour of railway tracks on transitions zones", *Computers & Structures*, vol. 89, pp. 1468–1479, July 2011.

[103] Y. Bezin, S. D. Iwnicki, M. Cavalletti, E. de Vries, F. Shahzad, and G. Evans, "An investigation of sleeper voids using a flexible track model integrated with railway multi-body dynamics", *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 223, no. 6, pp. 597–607, 2009.

[104] M. Wehbi, *Developing a novel technique to extract track stiffness information from track geometry measurement*. PhD thesis, University of Birmingham, 2016.

[105] R. U. A. Uzzal, R. B. Bhat, and W. Ahmed, "Dynamic response of a beam subjected to moving load and moving mass supported by Pasternak foundation", *Shock and Vibration*, vol. 19, pp. 201–216, Jan. 2012.

[106] M. Banimahd, P. K. Woodward, J. Kennedy, and G. M. Medero, "Behaviour of train-track interaction in stiffness transitions", *Proceedings of the Institution of Civil Engineers-Transport*, vol. 165, pp. 205–214, Aug. 2012.

[107] R. N. Jazar, "Vehicle Dynamics: Theory and Application", pp. 931–975, 2008.

[108] D. Li and E. T. Sellig, "Method for Railroad Track Foundation Design. II: Applications", *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 124, no. 6, pp. 323–329, 1998.

[109] G. J. Yeo, P. F. Weston, and C. Roberts, "Assessing rail deflection data from inertial measurements collected from on board a railway vehicle", in *11th World Congress on Railway Research*, (Milan, Italy), May 2016.

[110] Network Rail, *Standard NR/L3/TRK/4041*. 2012.

[111] G. J. Yeo, P. F. Weston, and C. Roberts, "The utility of continual monitoring of track geometry from an in-service vehicle", in *Proceedings of the 6th IET Railway Condition Monitoring Conference*, (Birmingham, UK), Sept. 2014.