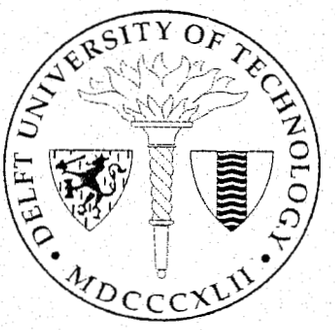
MODERN RAILWAY TRACK

Second Edition

Coenraad Esveld
Professor of Railway Engineering
Delft University of Technology



2001
MRT-Productions

Delft University of Technology

Editing: Dior Zwarthoed-van Nieuwenhuizen
Layout: Jan van 't Zand, TU Delft
Drawings: TU Delft
Production: Koninklijke van de Garde BV

ISBN 90-800324-3-3
SISO 696.3 UDC 625.1

This book can be ordered from:
MRT-Productions . P.O. Box 331 . NL-5300 AH Zaltbommel . The Netherlands
Tel.: +31 418 516369 . Fax: +31 418 516372 . Email: mrt@esveld.com
Internet: www.esveld.com

Major contributions to this Second Edition were made by the following persons:

Norbert Frank, Voest Alpine Schienen

Stanislav Jovanovic, TU Delft

Anton Kok, TU Delft

Amy de Man, TU Delft

Valeri Markine, TU Delft

Rainer Oswald, VAE

Peter Scheepmaker, TU Delft

Rainer Wenty, Plasser & Theurer

Gerard van der Werf, TU Delft and Arcadis

Jan van 't Zand, TU Delft

Arjen Zoeteman, TU Delft

Jan Zwarthoed, TU Delft and Volker Stevin Rail & Traffic

To my grandchildren
Thomas
Maud
Fieke
and
Douwe

Publication of this book has been made possible thanks to the sponsoring of the following companies:

Main Sponsors
Delft University of Technology,The Netherlands
BAM NBM Rail bv, The Netherlands
German Track Systems Projektgesellschaft mbH, Germany

Sponsors
Plasser & Theurer, Austria
Vossloh Rail Fastening Systems, Germany

Advertisers
Elektro-Thermit, Germany
Speno International, Switzerland

Front Cover Sponsor
Vossloh Rail Fastening Systems, Germany

Spine Sponsor
Pandrol, England

Acknowledgement

With the preparation of this Second Edition many experts have assisted to provide and check existing material and to write additional sections. In the first place I would like to thank my staff of the railway engineering group of Delft University of Technology: Jan van 't Zand, Peter Scheepmaker, Gerard van der Werf, Anton Kok, Valeri Markine, Ivan Shevtsov, Pedja Joksimovic and the secretaries Jacqueline Barnhoorn and Sonja van den Bos. I am most indebted to my Ph.D. students: Akke Suiker, Amy de Man, Arjen Zoeteman, Søren Rasmussen, Stanislav Jovanovic and Jan Zwarthoed for their invaluable ideas, suggestions and contributions. Those who have drafted significant parts have been mentioned explicitly in the outset of the book. From TU Delft I would like to mention in particular Jan van 't Zand who made the entire layout of the book in Framemaker.

I would also like to express my gratitude to my colleagues of the management team of the Section for Road and Railway Engineering at the Civil Engineering Department of TU Delft: André Molenaar, Peter Scheepmaker, Lambert Houben, Martin van der Ven and Abdol Miradi for their support in producing this Second Edition.

For the high-speed section I would like to thank the Korean High Speed Rail Corporation for contributing information of the high-speed project between Seoul and Pusan. In this respect I would also like to refer to the many interesting discussions in the Special International Track Advisory Committee (SITAC), comprised of Dr. Kee-Dong Kang, Dr. Yoshihiko Sato, Mr. Serge Montagné, Prof. Klaus Riessberger, Mr. Gerhard Kaess and myself, with the active assistance of Mr. Arne Svensoy, Mr. Bertold Pfeifer and Mr. Ki-Jun Son.

Valuable information was received from my Japanese colleagues Dr. Yoshihiko Sato from the Railway Track System Institute, Dr. Katsutoshi Ando and Mr. Noritsugu Abe from the Railway Technical Research Institute (RTRI) and Mr. Tetsuhisa Kobayashi from the Japan Railway Construction Public Corporation (JRCPC) for which I would like to express my gratitude.

I very much appreciated the indirect contributions by the companies and members participating in the Coordinating Committee for Railway Engineering of the Information and Technology Centre for Transport and Infrastructure (CROW) in The Netherlands.

I also owe much gratitude to Mr. Rainer Wenty from Plasser and Theurer for revising the section on track maintenance and renewal, and providing information on various other related subjects.

I highly appreciated the input on stone blowing from Mr. Peter McMichael of Railtrack and Mr. David Hill-Smith of AMEC Rail.

The section on rail grinding was checked by Mr. Wolfgang Schöch for which I would like to express my thanks.

For the section on rails I am very grateful for the contribution of Dr. Norbert Frank from Voest Alpine Schienen, who revised large parts of the original text.

I very much appreciated the assistance of Mr. Paul Godart of NMBS/SNCB for providing the information on the work of CEN and UIC concerning new rail standards.

I would like to express my gratitude to Mr. Hugo Goossens of TUC Rail for the many interesting discussions on track maintenance.

I owe much gratitude to Mr. Rainer Oswald from VAE, for his suggestions on revising the section on switches and crossings.

I would like to thank Dr. Frank Kusters of Elektro-Thermit for checking the section on ET welding.

Thanks to Dipl.-Ing. Hans Bachmann, Dipl.-Ing. Jens Kleeberg and Dipl.-Ing. (FH) Martin Kowalski of Pfleiderer I was able to incorporate the latest information of the Rheda system in the chapter on slab track.

Furthermore essential information on track components were provided by the suppliers, for which I would like to thank in particular Mr. Gerrien van der Houwen of Edilon, Mr. Dirk Vorderbrück from Vossloh, Mr. Chris Eckebus from Phoenix Benelux, Mr. David Rhodes from Pandrol, Mr. Patrick Carels of CDM and Mr. Olaf Unbehaun of Cronau.

For the parts on inspection systems we received many contributions from the industry. I would like to express my thanks in particular to Mr. Anton Weel and Mr. Han Wendt of Eurailscout, Mr. Jaap Roos and Mr. Erwin Giling of TNO-TPD, Mr. Aad van der Linden and Mr. Jan van der Schee from Koninklijke BAM NBM, Mr. Wido de Witte from Erdmann Softwaregesellschaft, Mr. Kevin Kesler of Ensco, Mrs. Danuše Marusicová of Czech Railways (CD), Prof. Willem Ebersohn of Amtrack, Mr. Charles Penny of Balfour Beatty, Mr. Paolo Redi of S.E.I. Sistemi Energetici Integrati and Mr. Ted Slump of NS Rail Infrabeheer.

Finally I would like to thank Dior van Nieuwenhuizen for her magnificent work to check and correct the English text.

I would like to conclude with expressing the hope that this Second Edition will once again prove to be a useful contribution to the training of students and railway engineers.

Coenraad Esveld

PREFACE

After the success of Modern Railway Track this Second Edition is an extension and complete revision of the original book, in which the developments of the last ten years have been incorporated. The research projects carried out at the Railway Engineering Group of Delft University of Technology have played a central role. The theory of railway track and vehicle track interaction has been substantially enhanced and much more attention has been given to dynamics. Undoubtedly one of the most important extensions was the part on slab track structures. But also track management systems have been given much more attention. Numerical optimization and testing, as well as acceptance are new chapters.

When revising the lecture notes for the railway course at the Civil Engineering Department of TU Delft in the period 1994 - 2000 the first edition of this book was taken as a starting point. The first edition and the TU Delft lecture notes, together with various publications and research reports, mainly of the railway engineering group of TU Delft, were then forming the base for the second edition.

The staff of the railway engineering group at TU Delft has made a great contribution to the composition and revision of the various chapters. Also the industry provided some important contributions, specifically on the chapters dealing with rail manufacturing, track components, maintenance and renewal, as well as inspection systems.

The first seven chapters are dealing with the basic theory of the wheel rail interface and track design. In the design attention is given to both static and dynamic aspects, whereby a number of examples is given of results obtained from computer models like RAIL, GEOTRACK and ANSYS. In the part on stability and longitudinal forces the CWERRI program is extensively discussed.

The discussion of track structures has been split up into a chapter on ballasted track and one on slab track. The first one is dealing with the conventional structures and modern ballasted designs, whereas the slab track chapter focuses on developments of the last decades. Both continuous slabs and prefabricated solutions are addressed in combination with discretely supported and continuously supported rails.

The chapter on rails has been brought to the state of the art, with introducing the new EN standards and discussing the latest inspection systems. Also the latest information on bainitic rail steels has been incorporated.

For switches and crossings high-speed turnouts are discussed, together with the geometrical design criteria, and also modern inspection systems for controlling switch maintenance.

In railway engineering practice track maintenance and renewal forms a key factor. The latest track maintenance methods and the associated machines are presented, being a major extension compared to the first edition of this book. The part on track deterioration has now been incorporated in this chapter.

Optimization was one of the issues very much underestimated in railway engineering. Such techniques are not only applicable to components and structures, but also to decision support systems and resource optimization. A separate chapter has been added called numerical optimization with the main emphasis on structural components.

From the outset railway engineering has always had a strong component in experimental work. Therefore a new section has been added on testing and acceptance, in which also the issue of acceptance criteria for new railway components is addressed.

The chapter on noise and vibration is describing the fundamentals and has been taken over from the first edition with only a few modifications.

The chapter on inspection and detection methods has been completely revised. The original chapter was primarily based on NS experience. Now the state of the art inspection systems have been introduced. However the fundamental parts of the first edition have been left in tact.

The chapter on high-speed tracks contains some applications of high-speed projects and some dedicated issues such as pressure waves in tunnels. Also a section is devoted to magnetic levitation.

In track maintenance management systems various issues on track maintenance and renewal decision support are described, as well as monitoring of phenomena relevant to the various maintenance processes. Special attention is given to the ECOTRACK system, developed under the auspices of UIC and maintained and supported by TU Delft.

Railway assets involve a large capital and need to be managed carefully. The chapter on this issue deals with the general principles of asset management and the way in which such systems can be set up.

The final chapter is dealing with life cycle cost analysis. After describing the general principles a number of case studies are discussed.

Zaltbommel, Summer 2001

Coenraad Esveld

TABLE OF CONTENTS

1 INTRODUCTION

2 WHEEL-RAIL INTERFACE

3 CURVES AND GRADIENTS

4 TRACK LOADS

5 STATIC TRACK DESIGN

7 TRACK STABILITY AND LONGITUDINAL FORCES

8 BALLASTED TRACK

9 SLAB TRACK

10 THE RAIL

11 SWITCHES AND CROSSINGS

12 TRACK MAINTENANCE AND RENEWAL

17 HIGH-SPEED TRACKS

1 INTRODUCTION

1.1 Historic development

The rail as supporting and guiding element was first utilised in the sixteenth century. In those times the mines in England used wooden roadways to reduce the resistance of the mining vehicles. The running surface was provided with an uprising edge in order to keep the vehicles on the track.
During a crises as a result of overproduction in the iron industry in England in 1760, the wooden rails were covered with cast iron plates which caused the running resistance to diminish to such an extent that the application of such plates soon proliferated. About 1800 the first free bearing rails were applied (Outtram), which were supported at the ends by cast iron sockets on wooden sleepers. Flanged iron wheels took care of the guiding, as we still practice now. In the beginning the vehicles were moved forward by manpower or by horses.

The invention of the steam engine led to the first steam locomotive, constructed in 1804 by the Englishman Trevithick. George Stephenson built the first steam locomotive with tubular boiler in 1814. In 1825 the first railway for passengers was opened between Stockton and Darlington. On the mainland of Europe Belgium was the first country to open a railway (Mechelen - Brussels). Belgium was quick to create a connection with the German hinterland bypassing the Dutch waterways. The first railway in The Netherlands (Amsterdam - Haarlem) came into existence much later: only in 1839. Here the railway was regarded as a big rival of the inland waterways.

The railways formed a brand new means of transportation with up till then unknown capacity, speed, and reliability. Large areas were opened which could not be developed earlier because of the primitive road and water connections. The railways formed an enormous stimulus to the political, economical and social development in the nineteenth century. Countries like the United States and Canada were opened thanks to the railways and became a political unity. In countries like Russia and China the railway still plays a crucial roll.

The trade unions originated when the railways were a major employer (railway strikes in England in 1900 and 1911 and in The Netherlands in 1903). The railway companies were also the first line of business which developed careful planning, organisation and control systems to enable efficient management. Moreover, they gave the impulse to big developments in the area of civil engineering (railway track building, bridges, tunnels, station roofing).

1.2 Railways

While the railways found themselves in a monopoly position up to the twentieth-century, with the advent of the combustion engine and the jet engine they had to face strong competition in the form of buses, cars and aeroplanes.

Mass motorization after World War II expressed by the growing prosperity brought about many problems, especially in densely populated areas: lack of space, congestion, lack of safety, emission of harmful substances and noise pollution. Exactly in these cases railways can be advantageous as they are characterized by the following:

- Limited use of space compared to large transport capacity;
- Reliability and safety;
- High degree of automation and management;
- Moderate environmental impact.

Moreover, railways have a comfort level comparable with automobiles and have the possibility of attaining high speeds which can compete with planes on the middle range distance. Regarding passenger transport, this potential should be translated into:

- High-quality commuter and urban transport;
- Fast intercity and high-speed daily connections up to 800 km;
- Comfortable intercity night connections up to 1500 km;
- Season charter transport (possibly with car).

Furthermore, in case of freight transport, high-grade connections exist on the medium-range and long-range distance.

In Table 1.1 an estimate is given of the length of global rail networks

Length 1000 km	Existing	Under Construction
Europe	530	10
Asia	250	110
Africa	80	10
North-America	420	3
Middle- and South-America	150	15
Australia	50	3
Total	**1500**	**150**

Table 1.1: Length of railway track network anno 1990

The railway systems are the proper means for massive passenger transport over short distances to and in within big conurbations. The quality of the railway system has been substantially increased in the last years by implementing large star-shaped networks around the big cities which are run frequently by quickly accelerating and decelerating stopping trains. If necessary, trains can enter the cities via special tunnel routes, which open up the city centres and enable connections to be made. Examples are commuter services like the S-Bahn (Munich, Hamburg) and the RER (Paris). A good integration with other means of pre- and post- transport (metro, tramway, bus, car, and bicycle) is very important.

Railway companies are unprofitable and governments have to support them financially to enable the companies to operate trains. This will be the case as long as the railways - contrary to road traffic and inland waterway shipping - have to carry the full costs of the infrastructure. Infrastructure is expensive. One kilometre of rail track costs about EUR 7 - 10 million; big structures not included.
In an increasing number of countries, however, the property and management of the railway infrastructure is taken over by the government while (private) railway companies pay for its use. This will also be the case in The Netherlands where the government demands the operating expenses to be fully covered by the profits.

1.3 Tramways and metro

The railways, developed as a fast interurban means of transport, are less suitable for local transport functions. They do not fit into the scale of the city (curve radii, clearances) while the capacity of a train is too big to suit the local traffic needs with a diffuse pattern of displacements. Therefore, in the second half of the nineteenth-century low-scale forms of rail traffic were developed which can also use the public road. At first horse power was used for traction and sometimes steam power; in the period 1890 - 1920 these traction forms were almost completely replaced by electric traction. In Table 1.2 an estimate is given of the number of global metro and tramways, while Table 1.3 and Table 1.4show some train and transport characteristics. Other characteristics like loads, tonnage, and speeds are dealt with later in the chapter 'Train Loads'.

	METRO	TRAM
Europe	35	225
Asia	15	65
Africa	0	5
North-America.	10	25
Middle- and South-America	5	5
Australia	0	2
total	**65**	**325**

Table 1.2: Metro- and tramway companies anno 1990

	LENGTH [m]	CAPACITY [persons]
Tram	30	175
Metro	100	1000
Suburb train	200 - 300	2000
Regional train	30 - 100	100 - 300
Intercity train	500	1000
Freight train	600	40.000 kN

Table 1.3: Average values for length and capacity of trains

	modal split	paskm/ inhabit	mio paskm/ km track	mio tonkm/ km track
The Netherlands	7%	650	3	1
Western Europe	8%	800	1	1.5
USA	1%	100	0.07	4.4
Japan	38%	2700	10	0.7

Table 1.4: Some transport characteristics

The tram is also used for minor suburban and rural transport, but here the bus has taken over for the greater part as is the case in little and medium large cities. Only in the big cities (above ca. 300.000 inhabitants) the tram has survived, thanks to its large transport capacity and the possibility to operate on closed track independent of road traffic. Subsequently, the tramway has more and more acquired the character of a low-scale railway, although interaction with road traffic by no means resembles the absolute priority the railway has. This demands special requirements with regards to the braking power of this vehicle and the layout of the track.

In cities where millions of people live, urban railway systems (metro's) have been developed with a complete infrastructure of their own which, by necessity, are built underground or on viaducts. The high expense of this infrastructure is justified by the heavy traffic, which is dealt with quickly and reliably using long metro trains, much longer than the tram (100 to 150 m against 30 to 50 m).

To achieve the same large advantages of the metro in cities with fewer inhabitants, one tries nowadays to realise in-between forms of metro and tram. This form, indicated as *light rapid transit*, is built partly at street level (as much as possible on closed track, but sometimes in the street with level crossings) and partly in tunnels and on viaducts. Examples of this development are for instance to be found in Rotterdam, Brussels, Cologne and Calgary (Canada).

1.4 Operational aspects

1.4.1 Functions of a railway company

A railway company no doubt belongs to the category of most complicated enterprises. Not only the product (the seat kilometre) cannot be delivered from stock, but it also must be produced on the very moment of acceptance. Moreover, a railway company must generally supply, administer, and maintain the means of production (infrastructure, safety equipment, rolling stock, and personnel). Finally, the connection between the different means of production is very firm, so all elements need to match each other very accurately.

In aid of the operation, meaning the use of the means in favour of the customer, a good preparation is necessary, not only for the daily processes, but also in the long term in order to make sure that the necessary production means will be available on time. These means, such as rolling stock and especially infrastructure, demand a long period of preparation: a new railway line will sometimes take up to 20 years.

1.4.2 Infrastructure

The main demands on rail infrastructure are:

- For the passenger: travel time as short as possible (by short distance and/or speed);
- For local residents: nuisance as minimal as possible (sufficient integration);
- For the railway company: travel time as short as possible, sufficient capacity, sound, and acceptable costs.

The layout determines the maximum speeds and hence the minimal possible travel times. The speed can be restricted by:

- Curves and gradients;
- Switches (when negotiated in diverging direction);
- Performance of stock (for instance power);
- Catenary voltage (decline of voltage may lead to speed loss);
- Signal system (location of signs should not cause speed loss).

A special traffic engineering aspect of layouts are the crossings with roadways and waterways. Level crossings with roads (level crossings) should be prevented as much as possible. Although the train has priority under all circumstances, collisions can hardly be prevented when a train approaches a vehicle on the crossing. The braking distance of the train is mostly too long.

In principle fly-overs are applied to motorways and to railways with an admissible speed of 160 km/h and more or with more than two tracks. The loss of time for the road traffic would be unacceptably high in these cases.

Crossings with waterways take place via tunnels or movable or unmovable bridges. Movable bridges mean a loss of capacity for the railway line. The bridge should be opened according to a fixed regime and at these moments trains cannot use the railway. Nevertheless, the advantage of a movable bridge is that the train has to overcome less difference in height.
This can be financially or operationally attractive, for it prevents a long gradient (saving of space and costs) or a steep gradient (an undisturbed passage of a freight train on the spot also leads to loss of capacity). Therefore, one may be forced to build more tracks (profit in alignment and loss in cross section).

1.4.3 Rolling stock

The rolling stock can be divided into:

- Passenger and freight stock;
- Hauled and powered stock;
- Electric and diesel stock (the latter diesel-electric or diesel-hydraulic);

Generally electric rolling stock can make a faster start and reach a higher speed. Some considerations related to the choice of hauled or powered rolling stock are:

- With trains of greater length the locomotive power is better used and the operation with hauled trains will be cheaper; with train-sets the number of motors is in proportion to the number of carriages;
- Simple combining and splitting;
- Simple change of direction (is also possible with so-called pull-and-push trains; these are pulled or pushed by a reversible train set at the other end of the train);
- Multiple use of rolling stock (one locomotive may pull passenger trains in the daytime and freight trains at night).

A relatively new development is the tilting coach train. This train will adjust itself regarding curves in such an angle with respect to the vertical axis that the centrifugal force is completely cancelled. This means that in tight curves with a maximum cant and limited speed, the tilting train can nevertheless pass with higher speed. In this way a fast train service can be operated without adapting the infrastructure (building of spacious curves).

1.4.4 Personnel

The personnel can be subdivided into the categories: management, execution, and maintenance. The operational department consists mainly of executive personnel subdivided into production and sales. The production personnel consist of drivers, ticket collectors, and traffic controllers.

Personnel constitute the most expensive part of the production (more than half of the operating expenses) and also require much attention in social respect. In planning operations, personnel should never be considered as a balancing item, on the contrary. The following items should be taken into account in due time:

- Desired numbers with respect to quality and kind of work;
- Desired and actual place of residence;
- Employability (set of tasks, road knowledge, duty and rest period);
- Permitted weekend- and night work;
- Participation and motivation;
- Job training and retraining;
- Possible use of third party personnel.

1.4.5 Electrification

The electrification of a railway line means investment in both infrastructure for power supply (catenary wires with suspension system) and expensive electric rolling stock. These investments entail a cheaper operation because:

- The stock has more power available per volume unit (compare a electric locomotive with 3000-7500 kW with a diesel locomotive with 2000-3500 kW);
- The stock requires less maintenance and the lifetime of motor units is longer;
- The stock has a higher acceleration.

The result of this is that for specific services less rolling stock needs to be purchased, as is the case with non-electric operation. Moreover, electric operation is more beneficial for the environment in terms of noise and air pollution.

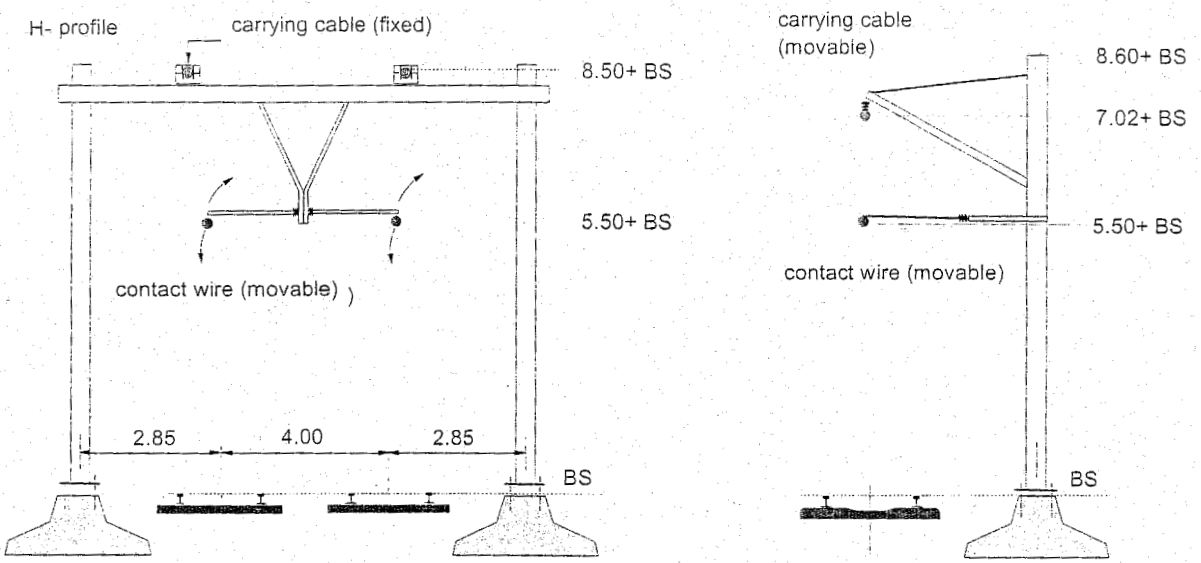
Generating electric energy is possible using different fuel materials. This makes the electric system flexible and less dependent on the price levels of these materials. When choosing electrification, the costs of energy also play a part. In countries with cheap electricity (Switzerland) many railway lines are electrified.

Electrified railway lines can especially be found in city areas with a large number of stops.

Two electric systems can be distinguished: direct current and alternating current. Direct current generally uses low voltages (tram and metro 600 to 750 V, train 1500 to 3000 V), alternating current has high voltages (15.000 V/$16^2/_3$ Hz to 25.000 V/50 Hz).

Because in the past the production of electric current with high voltage was more expensive and there was some doubt with regards to safety, direct current systems are found in a lot of countries (for instance The Netherlands). In urban areas high voltage is not applied because it may interfere with the electrical equipment in households and companies.
Nevertheless, alternating current with high voltage has many advantages:

- The cross section of the (copper) catenary wires can be smaller (hence lower costs);



Catenary system via portal structure: V ≤ 140 km/h

Catenary system via rotating lever: V > 140 km/h

Figure 1.2: Catenary systems: fixed and flexible suspension

Regarding this two catenary systems can be distinguished:

a. fixed suspension: for instance portal structures (Figure 1.2, left picture),

b. flexible suspension: for instance poles (Figure 1.2, right picture)

With a. the portals and carrying cables are fixed to each other. Because of this the carrying cables sag a little with warm weather and hence the contact wire as well. With b. the poles and carrying cables are fastened to each other flexibly. The cantilevers of the poles are movable and will change along with the carrying cables when temperature changes. The contact wire can now stay completely flat. The single pole structure is used with speeds of 140 km/h and higher. At lower speeds portal structures can be used.

Concerning metro systems the so-called third rail takes care of the supply of the current. This rail is installed next to the two rails carrying and guiding the train. Because of this no catenary system is necessary. This leads to a smaller clearance and therefore a smaller and cheaper tunnel construction is possible. When crossing over metro lines, the required height under the bridge can be lower. However, people should always be kept away from this third rail; the track should be inaccessible and without level crossings.

1.4.7 Road crossings

In practice the view of road crossings (or level crossings) is often insufficient. The nature and extent of road traffic and the speed and frequency of the train traffic necessitate in many cases the provision of road crossings with barriers, warning facilities or safety devices.
Road crossings can be provided with:

- Lifting gates;
- Semi-barrier in combination with flashing lights;
- Flashing lights;
- Gates which do not swing to the side of the railway.

For road crossings several standard constructions have been developed, amongst others some for very intensive road and train traffic.

1.4.8 Major rail infrastructure projects

In prosperous countries with substantially flourishing economies two problems can be identified which give rise to the building of high quality railway lines:

- Insufficient capacity of the existing rail (and road) infrastructure;
- Harmful effect on the environment due to road and air traffic.

In order to generate sufficient competition with respect to the use of cars and planes at distances of some hundreds of kilometres, fast passenger rail services are necessary. Trains should be moving at a maximum speed of about 300 km/h and an operational speed of 200-250 km/h. Freight traffic by rail may be competitive with road traffic at distances of more than 300 km if train services are offered with speeds in the range 120-160 km/h.

In Europe (France, Germany, Italy, Spain), Japan, and the United States high-speed rail links have been established during the last decades and new lines are under construction. Europe and Japan produce their own systems. The U.S. buys systems from other countries. France and Germany are leaders in building high-speed lines; moreover, they are mutual competitors on the world market. Both try to obtain a position and are involved in projects in the United States and South Korea.

A parallel development is taking place in the case of trains which are made to run faster making use of existing infrastructure: so-called tilting coaches. These tilting coaches produce an additional higher cant in curves compared to the track cant. Tilting coach trains are used, amongst others, in Italy and Sweden.

Railway lines for freight traffic (and more specifically for higher speeds) are an exception. The plans for the Dutch "Betuweroute" are an example of this. The higher axle load on this category of railways is more characteristic than higher speed. 'Heavy haul' lines can also be found in South Africa (for ore transport). There trains run with a weight of 200.000 kN. The world record is 700.000 kN.

1.4.9 Developing countries

The developing countries cannot be regarded as one with respect to quality. There are countries with operating systems which work well, although not according to our western standards. Especially in India and the Peoples Republic of China, a large network of railway lines is available operating substantial (overburdened) train services.

On the other hand, in many other developing countries the railway network is underdeveloped. Mostly there are remnants from a colonial past. Lack of maintenance has deteriorated the track condition which demands urgent renovation. Moreover, the routes should in many cases be adapted. Often heard wishes are increasing the permitted axle loads and speeds as well as improving the safety system. The curve radii should therefore be increased and gradients should be less steep.

Besides renovations there is an enormous need for newly built track. The bad state of the roads plays a part in this. Most of the new projects are being developed for freight transport, mainly of low value (ores and other raw material).
Especially the high oil prices have given a push to making new plans. Building rail connections between the mines in the interior part of the country and the harbours is most urgent. But also the passenger transport needs improving. The suburban traffic cannot cope with the rush of passengers and the long distance transport is very defective. It is illustrative that a railway journey through Africa, from Cairo to Cape Town, still takes four weeks when a part of the journey is made by boat.

The most important problem is of a financial nature. It may be true that modern, gigantic excavating machines offer all sorts of possibilities for building railway lines in hilly areas, but the construction is expensive. Because of this assistance from outside parties is essential.

Available technical aid is sufficient. Consultants from industrialised countries provide the necessary completed designs. The realisation of the construction and improvement of railway lines is mostly carried out by European and American firms, which mainly introduce the technical know-how and are concerned with the supply of materials. For the building activities, local labour is called in.

Especially India and the Peoples Republic of China are Third World countries which are active in the railway field. They are - by own experience - well informed about the important social and cultural problems. These deal primarily with the transfer of knowledge as well as instructions and attending local personnel.
The transport problem in the exploding cities of the Third World where millions of people live asks for rail solutions in terms of suburban rail, metro, and light rail lines. Here and there metro or light rail projects have been carried out (mostly South-America, the Middle East and Southeast Asia: Mexico City, Caracas, Cairo, Teheran, Singapore, Hong-Kong, Manila). Elsewhere plans are ready and waiting for financing (for instance by means of the World Bank) before they are able to be carried out (Bangkok, Jakarta).

1.5 Geometry of a railway line

1.5.1 Clearances

Above and next to the tracks a certain space should be reserved to ensure the unrestricted passage of vehicles. The dimensions of this *structure gauge* (or *clearance*) is also based on the internationally approved *vehicle gauge* of railway rolling stock and the *loading gauge*, within which the loading of the railway vehicles should be kept. In this clearance extra effects are dealt with:

- deviation of the correct track geometry;
- Swinging movement;
- Deviation due to wind loading;
- Tilting due to cant;
- Unequal loading of vehicles;
- Tolerance of vehicle dimensions.

Figure 1.3 shows the normal clearance; the left side applies to straight track, the right side to curves with a radius greater than 250 m. Height measures are measured from top of rail.

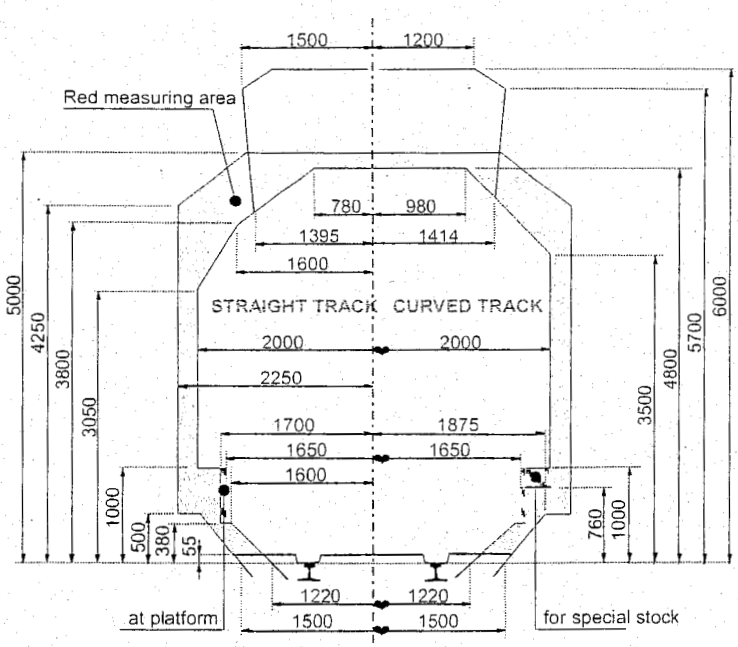


Figure 1.3: Normal clearance (structure gauge)

In a number of cases the clearance has other measures:

- Decrease of the width under the train (only bogie is present);
- Increase of the width at smaller curve radii;

- Rotation due to track cant;
- Widening with vertical curves;
- Widening in sections with frequent traffic outside gauge (Red measuring area).

Fixed objects located within the clearance are registered.

All present or future fixed objects located within the 'red measuring area' (about 20.000 objects) are also registered. If a transport is presented outside the clearance, a quick evaluation is possible to judge if and under which conditions the transport can take place.

To make a comparison between the loading gauge, the vehicle gauge and the normal clearance, Figure 1.4 has been drawn. The clearance with the 'red measurement area' was already pictured.

The distance between the centres of two tracks in double track amounts to 4.25 m. (in old situations a minimum of 3.60 m can still be found). At higher speeds (more than 160 km/h) a greater distance of up to 4.70 m is applied.
In curves the distance is increased as well.

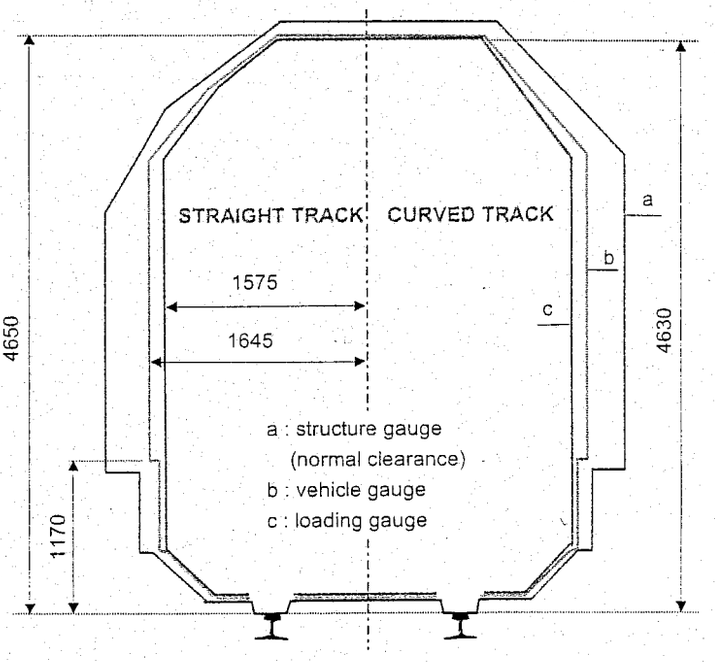


Figure 1.4: Loading gauge, vehicle gauge and clearance

The increase is higher as the curve radius is smaller or the cant difference becomes more unequal.

In multitrack sections an alternatively standard and a higher distance should be applied between the tracks (depending on the speed 6-8 m). It depends on the local circumstances if in case of a fourfold track the construction according to A or B is applied (both drawn in Figure 1.5).

With an expansion from two to four tracks, the choice for one- or two-sided expansion depends on:

- The available space;
- The soil (one should be careful with soil excavation next to a track which has become compact over the years in order to prevent shear);
- The intended use of the track.

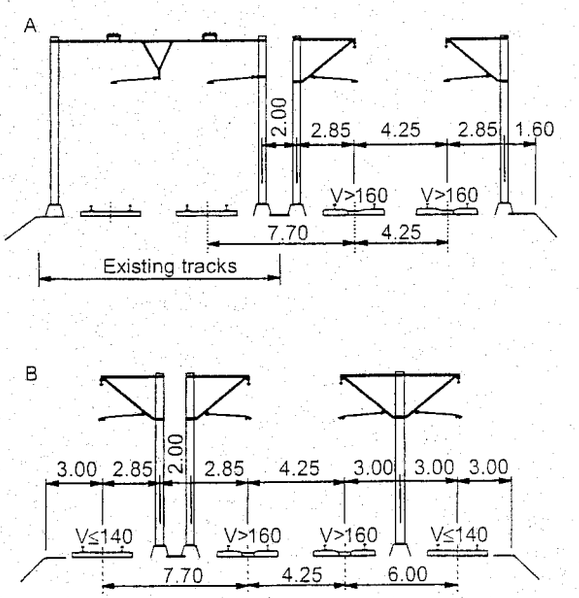


Figure 1.5: Two situations with fourfold track

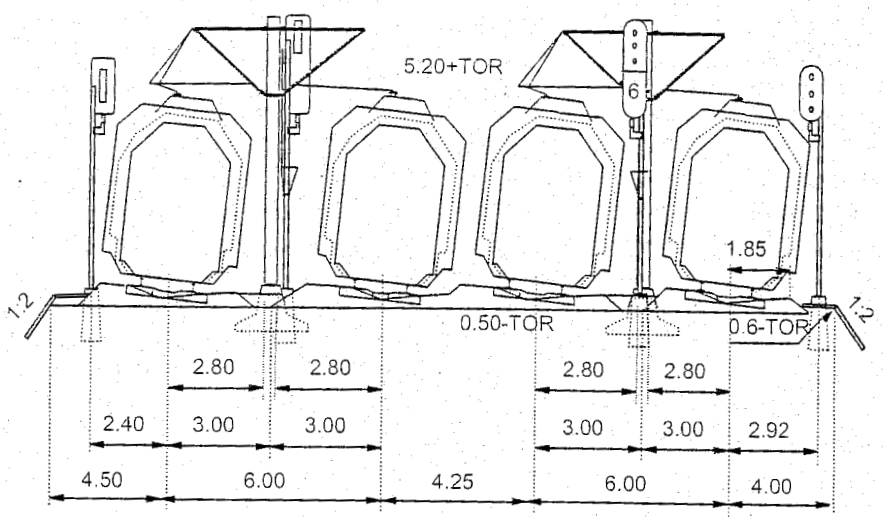


Figure 1.6: Cross section with fourfold track

In Figure 1.6 an example is given of the cross section of a fourfold track. Especially at the outer tracks provisions are made to allow a clearance with a 'red measuring area'.

The distance between newly built yard tracks is preferably 4.50 m. Moreover there should be, after each 4 to 5 tracks, a wider track distance of 5 m to enable personnel to walk safely over a 1 m path. If catenary portal structures are to be built, the track distance should be 6-8 m.

Clearance around the rail track

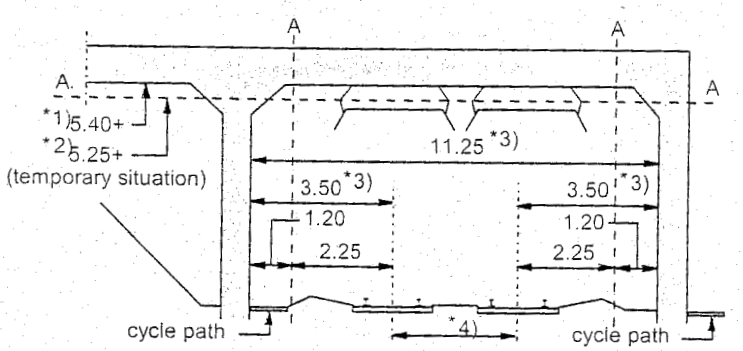
The railway law stipulates that no object may be erected that causes an obstruction of the view:

- Within a distance of 8 m in straight track and on the outside of a curve with a radius of more than 1000m;
- Within a distance of 20 m on the inside of a curve with a radius of less than 1000 m.

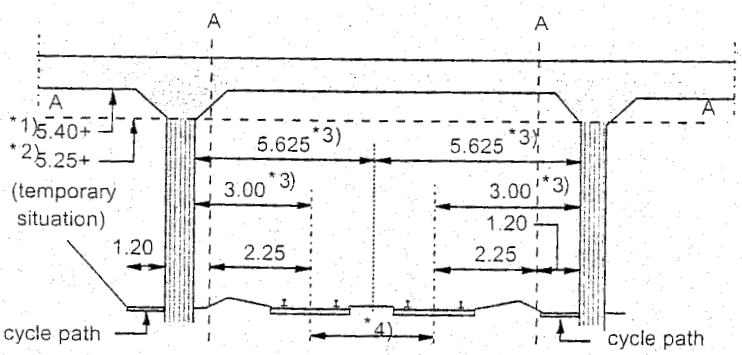
Figure 1.7 indicates which areas should be cleared of (the building of) structures (bridges, tunnels, etc.). It should be kept in mind that in due course the track distance may be increased to 4.25 m.

Between a track and a closed wall there should be at least 3.50 m room. Between a track centre and an open wall (columns) offering the possibility to run away, at least 3.00 m should be maintained.

The width of the cycle path is measured from the sidewall to the place where a ballastbed connects to the cycle path.



a. Overhead crossing with closed sidewalls



b. Overhead crossing with columns

*1) At V>140 km/h minimal height is 5.65 m.

*2) Lines A-A indicate extreme limits of the formwork, the suspension of the carrying cable should then be located within the formwork limit.

*3) In case of bad view or placement of fixed objects within the overbridge these measures should be adapted.

*4) As a structure has a long service life the possibility of expanding the track distance to 4.25 m should be kept in mind.

Figure 1.7: Clearances with structures

Noise barriers may not be installed closer than 4.50 m from the track in order to ensure the safety of personnel and allow the necessary room for cycle paths and space to put away mechanical manual equipment for maintenance purposes. In curves this distance is 4.80 m.

The height of the barriers should be limited and installing barriers between tracks is not allowed. Moreover, there should be a safety door in the barrier after every 100 m. If the view of the track is less than 1500 m due to a barrier, the barriers should be made less high or warning lights should be installed.

1.5.2 Alignment

The alignment of a railway line exists of gradients (the steepness is expressed in a permillage) and vertical rounding off curves. With gradients of 5‰ no difficulties will arise on the open track. All rolling stock in The Netherlands will be able to move off from standstill.
Steeper gradients can be applied; maximum values cannot be given as they depend on:

- The length of the gradient;
- The possibility to develop a starting speed;
- The characteristic of the applied pulling force and train loading.

It should be kept in mind that electric locomotives, when climbing gradients, are not allowed to apply the maximum force at low speed for a longer period because the series resistances may burn.

If gradients and curves coincide, the gradient should be decreased a little on the spot to keep the total resistance constant. Descending gradients extend the braking distance.

1.6 General track considerations

1.6.1 Track requirements

The term railway track or "permanent way" entails tracks, switches, crossings, and ballast beds. The track is used by locomotives, coaches, and wagons which in Europe normally have maximum axle loads of 22.5 t and which, on NS, run at speeds of up to 140 km/h. The fact that the purpose of the track is to transport passengers and freight and that operation is required to be as economical as possible, gives rise to a number of requirements to be met by the track. These are formulated as follows:

- Bearing in mind permissible speeds and axle loads, the rails and switches must be safe for vehicles to run on. To ensure this the track components, such as the rails, must be of such dimensions that they do not fail under the traffic load. Moreover, the correct geometry must be maintained whether the track is under load or not.
- Tracks and switches must allow comfortable passage at all times. Even if safety is not jeopardised, the locomotives and coaches may experience such vibrations and oscillations during the journey that passage becomes unpleasant for the passengers. An unfortunate combination of switches, curves and reverse curves may, even if the track is very well constructed and has perfect geometry, cause such strong movements in a vehicle that the passengers experience most unpleasant and sometimes even frightening sensations.
- Track must be electrically insulated so that the track circuits required for signalling continue to function even under the least favourable weather conditions. It should also be electrically insulated to ensure that on electrified sections the return current does not return to earth as stray current.

- Track must be constructed in such a way that the trains running on it do not cause excessive environmental pollution in the form of noise and ground vibrations.
- Costs of the total service life of the track must be as low as possible.
- Maintenance should be low and as inexpensive as possible.

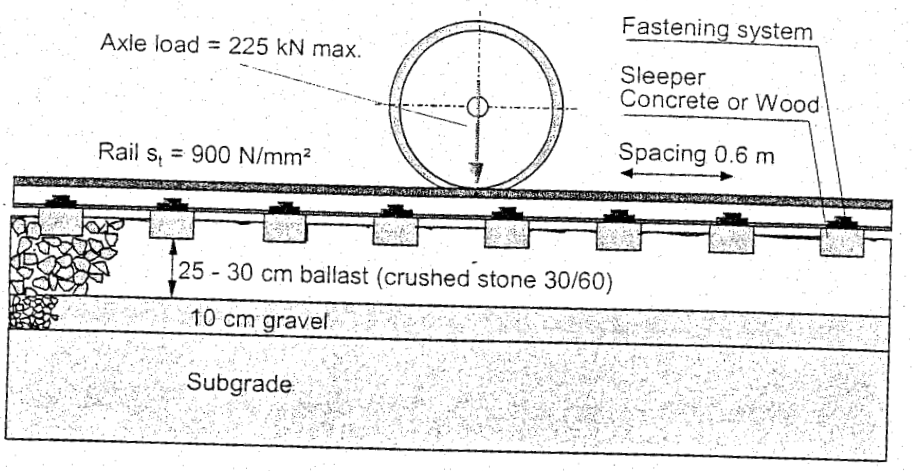


Figure 1.8: Conventional track structure

Tracks and switches are assets which will last for quite some years. The choice of a particular track system and the decision to use this system on certain lines, therefore, generally involves a decision which will hold good for 20 to 50 years. Consequently, such decisions must be taken with the future in mind, however difficult it may be to make a valid prediction. The only sure factor is that a certain degree of objectivity must be maintained vis-à-vis the present day situation, and not too much emphasis placed on random everyday events.

When choosing a track system, the above-mentioned requirements must all be given due consideration and it is clearly necessary to form some idea of the axle loads and maximum speeds to be expected in the decades to come. After this the situation regarding the various track components, such as rails, sleepers, fastenings, switches, and ballast should be examined so that the optimum track design is obtained.

1.6.2 Load-bearing function of the track

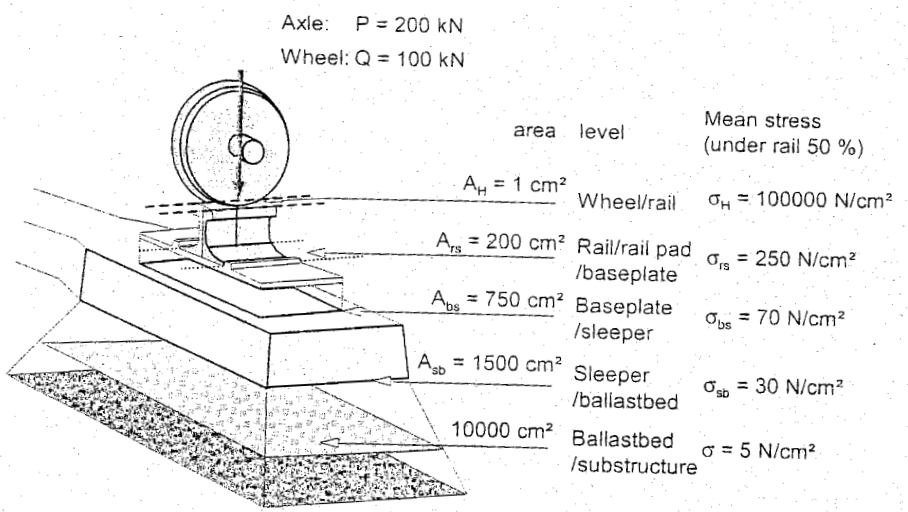


Figure 1.9: Principle of load transfer

The purpose of track is to transfer train loads to the formation. Conventional track still in use consists of a discrete system made up of rails, sleepers, and ballastbed. Figure 1.8 shows a principle sketch with the main dimensions.

Load transfer works on the principle of stress reduction, which means layer by layer, as depicted schematically in Figure 1.9. The greatest stress occurs between wheel and rail and is in the order of 30 kN/cm^2 (= 300 MPa). Even higher values may occur (see chapter 2). Between rail and sleeper the stress is two orders smaller and diminishes between sleeper and ballast bed down to about 30 N/cm^2. Finally the stress on the formation is only about 5 N/cm^2.

1.6.3 Indication of rail forces and displacements

In rail track literature the wheel-rail forces (Figure 1.10) are usually indicated with Q for the vertical force (z-direction), Y for the lateral force (y-direction), and T for the longitudinal force (x-direction).

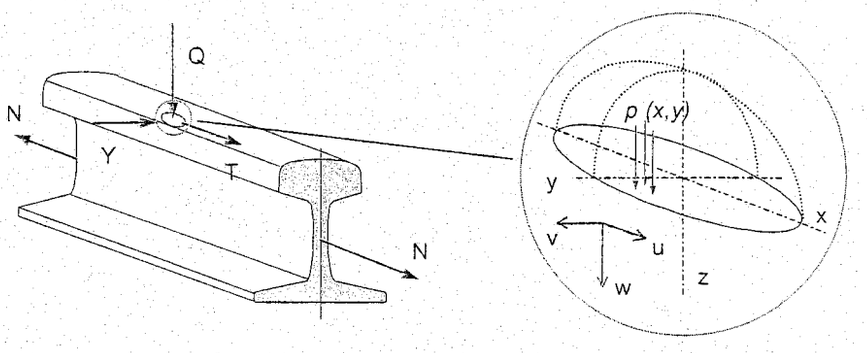


Figure 1.10: Rail forces and displacements

Apart from these external forces, internal longitudinal temperature forces may be present and are indicated with the symbol N.

The Hertz contact area and contact stress distribution, shown in the detail view, will be discussed in a later chapter.

In the given xyz-coordinate system, rail or track displacements are usually indicated as u, v, and w.

1.6.4 Track geometry components

An important aspect of construction is track geometry. The primary geometry components are measured in the cross-section of the track as drawn in Figure 1.11.

Each rail has 2 degrees of freedom. These 4 degrees of freedom are normally replaced by an equivalent system consisting of cant, level, alignment, and gauge.

An additional important parameter is twist, which is defined as the difference in cant over a given length. All quantities are functions of the length coordinate of the track (x-axis).

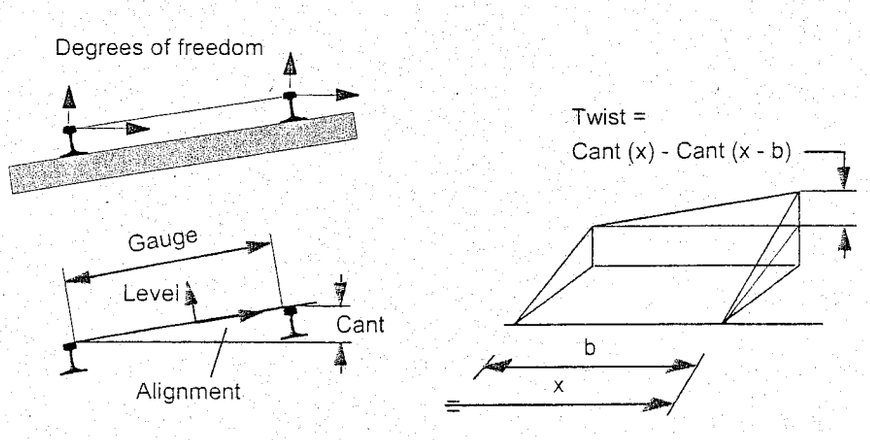


Figure 1.11: Geometric track components

2 WHEEL-RAIL INTERFACE

2.1 Wheel-rail guidance

A rail vehicle basically consists of a body supported by secondary suspension on bogies in which the wheelsets are mounted and damped by means of primary suspension. Track guidance of the wheel is achieved in principle by making the following two provisions:

- The tires are conical instead of cylindrical which means that in straight track a centering force is exerted on the wheelset if there is slight lateral displacement. The centering effect promotes a better radial adjustment of the wheelset in curves. This leads to more rolling, less slipping and hence less wear.

- The tires have flanges on the inside of the track to prevent derailment. In case of more considerable lateral displacement both in curves and on switches, the lateral clearance between wheelset and track is often no longer sufficient to restrict lateral displacements adequately by means of the restoring mechanism previously discussed. Should the wheel flange touch the rail head face, this can result in high lateral forces and wear.

2.2 Wheelset and track dimensions

Generally the track gauge is used as a distance measured between the two rails, more specifically the distance between the inside of the railheads measured 14 mm below the surface of the rail. By choosing 14 mm the measurement is less influenced by lipping or lateral wear on the rail head and by the radius r = 13 mm of the rail head face. On normal track the gauge is 1435^{+10}_{-3} mm with a maximum gradient of 1:300. For new track, however, NS apply the following standards:

- Mean gauge per 200 m: 1435^{+3}_{-1} mm

- Standard deviation within a 200 m section less than 1 mm.

Figure 2.1 illustrates the definition of the track gauge as well as some other commonly used dimensions such as:

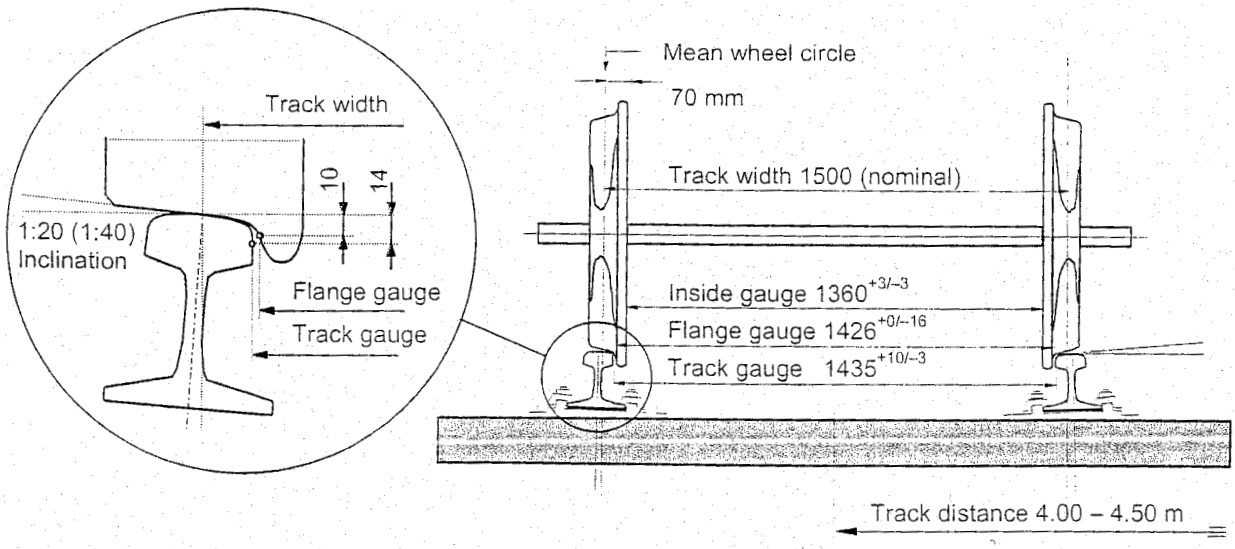


Figure 2.1. Wheelset and track dimensions for straight normal gauge track

- *Track width*: distance between the points of contact of the mean wheel circles with the rails, having a nominal value of 1500 mm. This dimension is important for calculations and should not be confused with gauge nor track distance.

- *Track distance*: distance between lines of adjacent tracks.

For the wheelset the following dimensions are used:

- *Flange gauge*: distance across the wheel flanges, measured 10 mm below the rail surface (wheelset in the centred position) on standard track is 1426^{+0}_{-16} mm.

- *Inside gauge*: distance between the insides of the wheels, on standard track is 1360^{+3}_{-3} mm.

- *Flangeway clearance*: clearance between wheelset and track, i.e. the distance the wheelset can be displaced laterally. This is not the same as the difference between track gauge and flange gauge.

It should be mentioned that a specific method of design process applies to switches and crossings.

The following summary gives some values for narrow gauge, standard gauge, and broad gauge.

- Narrow gauge:
  - 750 mm: parts of Indonesia
  - 1000 mm: parts of Switzerland, tram lines etc.
  - 1067 mm: (3½') (Cape gauge), South Africa, Japan, Indonesia, etc.
- Standard gauge:
  - 1435 mm: (4' 8½"). Gauge used by George Stephenson in 1825 based on existing mail coaches. Most commonly used nowadays.
- Broad gauge:
  - 1524 mm: (5'). Russia, Finland
  - 1665 mm: Portugal
  - 1667 mm: Spain

2.3 Conicity

Originally conical tire profiles with an inclination of 1:20 were used. Since a centrally applied load on the railhead is desired, a rail inclination of 1:20, as shown in Figure 2.1, was also selected; this for instance still applies to NS profile NP 46. UIC 54 rail usually has an inclination of 1:40. This inclination matches the S 1002 worn wheel profile which is in general use in Europe. During manufacturing the tires are given a profile which matches the average shape caused by wear. In contrast to the straight conical profile this has a hollow form.

2.4 Lateral movement of a wheelset on straight track

2.4.1 Theory according to Klingel

If a wheelset with conical tire profiles is laterally displaced from the central position, this displacement is counteracted due to different rolling radii of the wheels. This results in a periodical movement of the wheelset which was described theoretically by Klingel in 1883 and is therefore often referred to as the Klingel movement. To analyse this case, the wheelset is modelled as a biconus travelling on an ideally straight track as shown in Figure 2.2.

The following parameters are used in the mathematical formulation:

γ = conicity of the wheel tread (inclination)
r = wheel radius in central position wheelset
R = curve radius of the Klingel movement path y(x)
s = track width
y = lateral displacement of the Klingel path
v = speed
x = distance co-ordinate

In a perfect rolling movement, the wheelset displaces laterally over a distance y with respect to the central position. A difference in rolling radius of $2\gamma y$ will occur and the lateral displacement path will show a corresponding radius R.

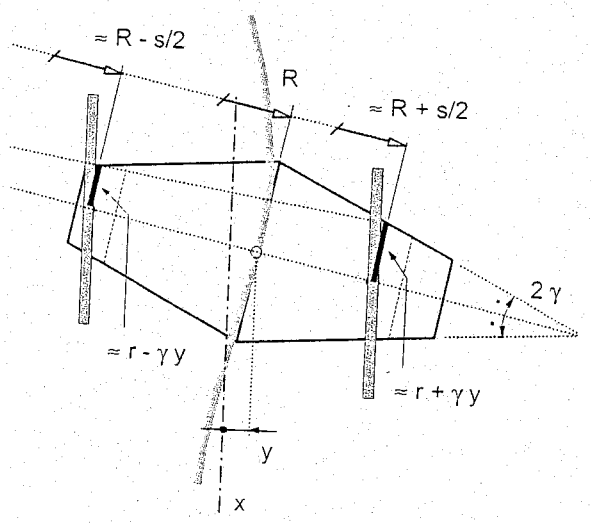


Figure 2.2: Wheelset biconus in general position

Figure 2.2 shows that this leads to the following geometrical condition:

$$\frac{r+\gamma y}{r-\gamma y} = \frac{R+\frac{1}{2}s}{R-\frac{1}{2}s} \tag{2.1}$$

Furthermore, the curvature may be approximated in the usual way by:

$$\frac{1}{R} = \frac{d^2y}{dx^2} \tag{2.2}$$

From these two expressions the differential equation follows:

$$\frac{d^2y}{dx^2} + \frac{2\gamma}{rs}y = 0 \tag{2.3}$$

If y(0) = 0 the solution of this differential equation reads:

$$y = y_0 \sin 2\pi\frac{x}{L} \tag{2.4}$$

in which y_0 and L are the amplitude and the wavelength of the lateral displacement. The wavelength is dependent on r, s and γ according to:

$$L_K = 2\pi\sqrt{\frac{rs}{2\gamma}} \tag{2.5}$$

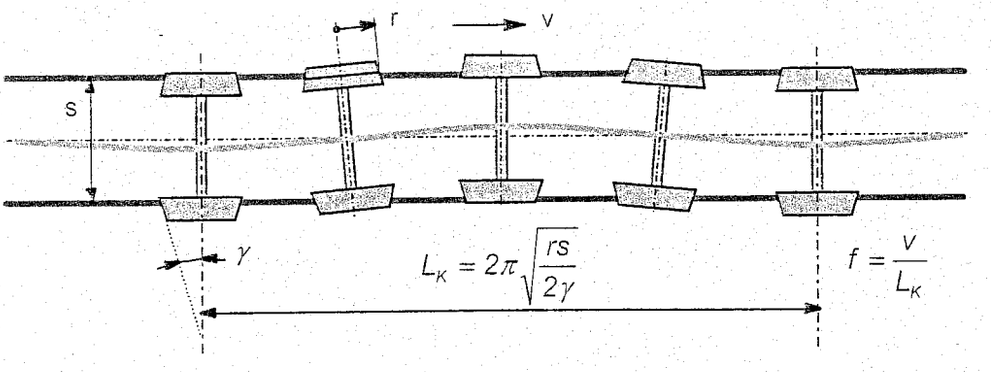


Figure 2.3: Klingel movement

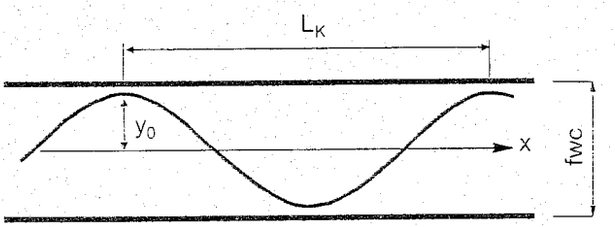
The Klingel movement is therefore purely a kinematic movement in which forces play no part in the derivation (Figure 2.3). The lateral displacement y is a harmonic, undamped function of the distance co-ordinate x as long as the amplitude moves within the flangeway clearance fwc. This is illustrated in Figure 2.4. Thus, if for instance r = 0.45 m, s = 1.5 m and γ= 1/20 then L_K = 16 m;

Introducing the speed, the time domain frequency of the Klingel movement is:

$$f = \frac{v}{L_K} \qquad (2.6)$$

and hence the maximum lateral acceleration can be calculated as:

$$\ddot{y}_{max} = 4\pi^2 y_0 \frac{v^2}{L_K^2} \qquad (2.7)$$



fwc = flangeway clearance

Figure 2.4: Undisturbed lateral movement of a wheelset

If the frequency f coincides with one of the natural frequencies of the rolling stock, the vehicle ride becomes unstable. The lateral acceleration, which is a measure of the forces, shows the adverse effect of high speed and/or small wavelength. A conicity, for example, of 1:40 in comparison with 1:20 therefore gives a greater wavelength and a lower lateral acceleration at the same speed. The progressively increasing conicity in the case of worn profiles due to increasing lateral axle movement, therefore, has an adverse effect in this respect.

2.4.2 Hunting movement

It should be noted that the Klingel theory is simple and instructive but does not include the effect of coupled axes, mass forces, and adhesion forces. In reality, the amplitude y_0 of the Klingel movement is dependent on alignment, dynamic vehicle behaviour, and the speed of the rolling stock.
Generally speaking, y_0 due to slip will increase with speed until it is equal to half the flangeway clearance. Flanging then occurs as a result of which the axle will rebound.

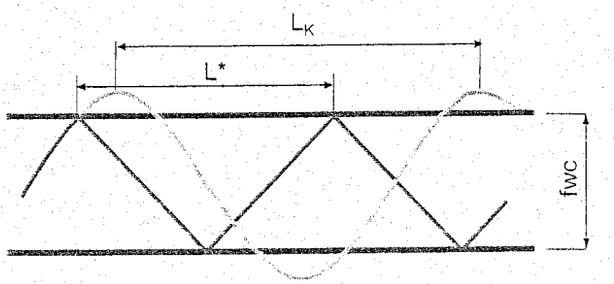


Figure 2.5: Influence of flanging on lateral wheelset movement

This means that the lateral movement takes on a completely different behaviour which is known as hunting. As shown in the drawing in Figure 2.5 the movement changes from a harmonic to a zig-zag shape. The wavelength becomes shorter and the frequency increases quickly until it is in the critical range for the rolling stock and resonance occurs.

This phenomenon is shown in Figure 2.6. The bogie design, as far as conicity and flangeway clearance are concerned, must be such that stable running is always guaranteed for the speed range in which the vehicle is to be used.

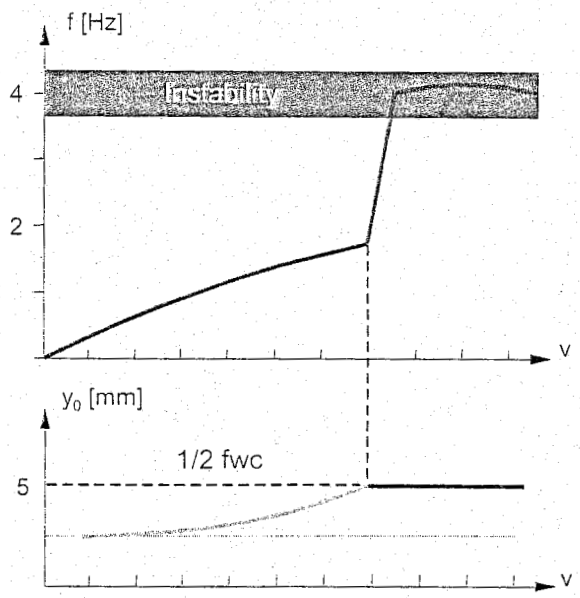


Figure 2.6: Increase in amplitude and frequency with speed and the development of instability

2.5 Equivalent conicity

It is clear that regarding a worn profile the conicity depends on the actual shape of the rail head and tire, including any wear, track gauge, and rail inclination. Likewise, elastic deformation of the wheelset and rail fastenings plays a role.

Generally, the effective or equivalent conicity is defined as:

$$\gamma_e = \tfrac{1}{2}\frac{\Delta r}{y} = \tfrac{1}{2}\frac{r_1 - r_2}{y} \qquad (2.8)$$

Here r_1 - r_2 is the instantaneous difference in rolling radius of the wheel treads; generally speaking this is a non-linear function of the lateral displacement y of the wheelset with respect to the central position. The difference between conical and worn profiles is given in Figure 2.7. To enable numerical comparisons γ_e is determined at a certain lateral displacement $y = \bar{y}$.

With a conical profile the conicity is constant and (2.8) becomes (see also):

$$\gamma_e = \tfrac{1}{2}\frac{\Delta r}{y} = \tfrac{1}{2}\frac{(r + \gamma y) - (r - \gamma y)}{y} = \gamma \qquad (2.9)$$

In the next paragraph the effects resulting from progressive non-linear behaviour of the effective conicity and its influence on the running stability of vehicles and rail wear are dealt with in greater detail.

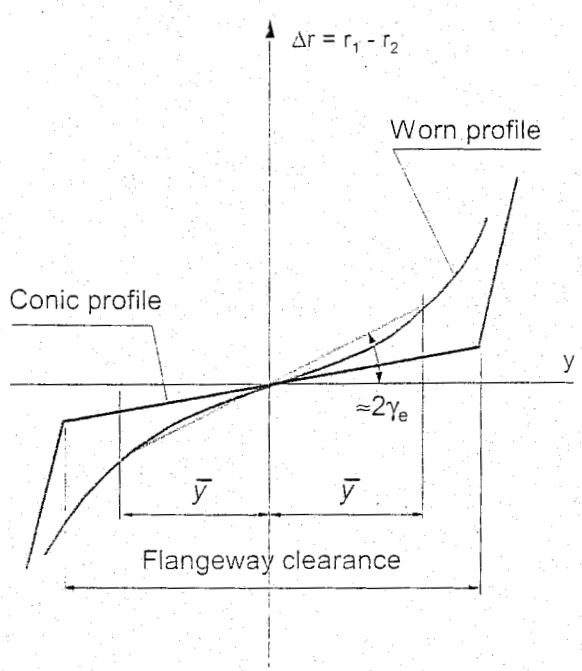


Figure 2.7: y-Δr curves. Difference between conical and worn wheel profiles

2.6 Worn wheel profiles

A perfectly conical wheel profile is unstable as far as its shape is concerned, but will take on a shape that is stable as the effect of wear.
In addition, conical profiles have the disadvantage that a substantial lateral movement will, because of the two-point contact, lead directly to an impact. If the profile of the rail and wheel tire at the point of contact is assumed to be circular, it can be deduced from Figure 2.8 that in the case of lateral displacement y of the wheelset with respect to the track, the contact point on the rail will over a distance translate to:

$$\Delta s = \frac{\rho_r}{\rho_w - \rho_r} y \qquad (2.10)$$

Moreover, if the value of y is small compared to the radii, the following relationship holds true:

$$\Delta r = \tan\phi \frac{\rho_w}{\rho_w - \rho_r} 2\gamma = \gamma_e 2y \qquad (2.11)$$

In the case of conical profiles the radius is infinite and therefore the contact point on the rail does not move. This means that rail wear is very much concentrated on one point.

An interesting case arises if the radii of the wheel tire and rail are almost equal. If there is slight lateral displacement of the axle, the contact point jumps which results in less passenger comfort. In the y-Δr graph this manifests itself as a jump.

Confirmation of this is given in the y-Δr curve in Figure 2.9, which is derived from the theoretical S 1002 wheel profile in combination with UIC 54 rail and track gauge 1434 mm.

Practical research has shown that over a period of time wheel profiles stabilise with wear at an equivalent conicity of 0.2 to 0.3. With regards to running stability, the equivalent conicity must remain below 0.4 and to ensure the centering effect it must be greater than 0.1.

Figure 2.10 gives a summary of equivalent conicity values for the S 1002 profile in combination with UIC 60 rail for different track gauges and rail inclinations. This information shows that, apart from a rail inclination of 1:20, the conicity increases as track gauge decreases and rail inclination becomes steeper. With a rail inclination of 1:20 the conicity is very small and is independent of the track gauge.

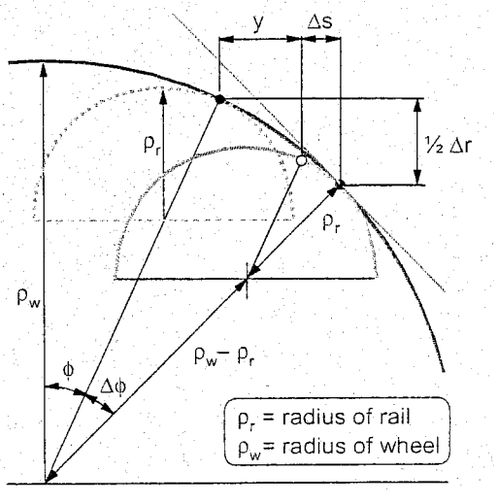


Figure 2.8: Displacement of contact point on rail due to wheelset displacement

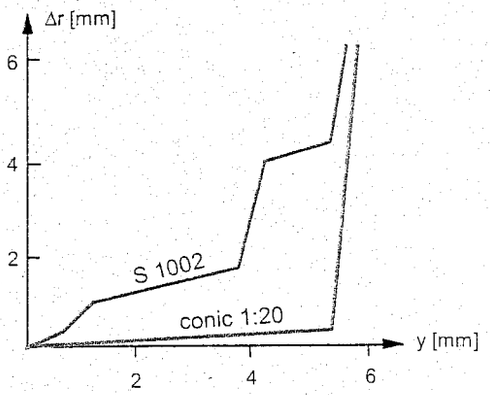


Figure 2.9: y-Δr curve for a theoretical S1002 profile in combination with a UIC 54 rail 1:40. gauge 1434 m

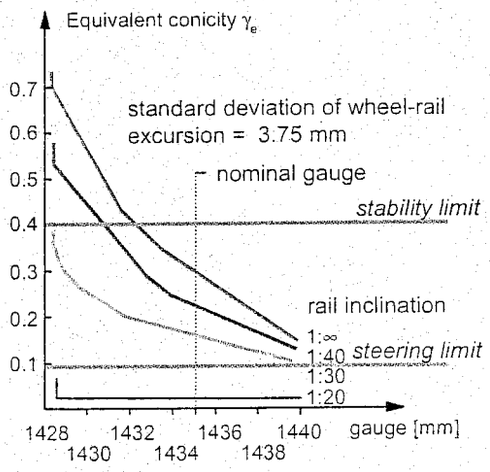


Figure 2.10: Equivalent conicity for S1002 on UIC 60.

Figure 2.11 shows a measurement result of both new and worn wheel profiles in which the wear is quantified in zones as specified [235].

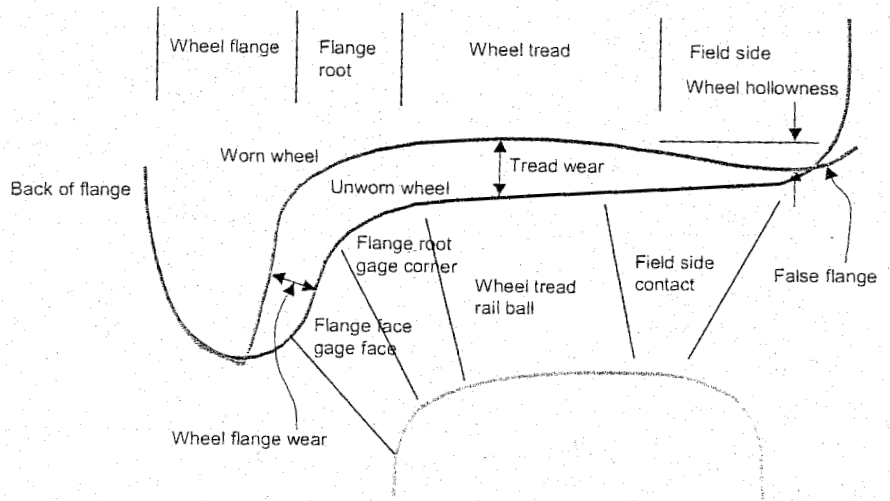


Figure 2.11: Wheel wear zones

Experience over the past few years with worn wheel profiles has resulted in a number of variants of the original S 1002 profile. It is still possible to further optimize the profile based on the measured data gathered over a long period of time for a specific type of rolling stock on a given section of the network.

2.7 Wheel-rail contact stresses

2.7.1 Hertz theory

The combination of rail vehicle and track should be regarded as one system. This applies to the function as a transport system, but also with respect to the technical point of view. A strong interaction exists between infrastructure and vehicle. The separation between both subsystems and the place where the interaction manifests itself is the contact between wheel and rail. Of course, the two main aspects are dominant here, viz. the bearing and guidance of the vehicles.

The Hertz theory (1887) explains what happens: the elastic deformation of the steel of the wheel and the rail creates an elliptic contact area. The dimensions of the contact ellipse are determined by the normal force on the contact area, while the ratio of the ellipse axes a and b depends on the main curvatures of the wheel and rail profiles.

Figure 2.12 gives an illustration of how the shape, expressed by the ratio b/a, of the contact ellipse relates to the location of the wheel-rail contact area across the railhead.

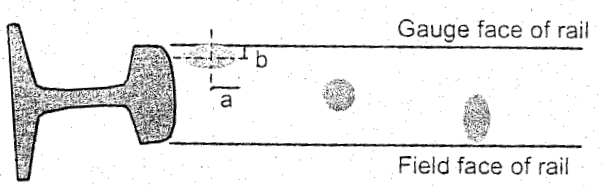


Figure 2.12: Examples of contact ellipses on the railhead

Inside the contact area a pressure distribution develops which in a cross section is shaped in the form of a semi-ellipse with the highest contact pressure occurring at the centre given by:

$$\sigma_{max} = \frac{3}{2}\frac{Q}{\pi a b} \qquad (2.12)$$

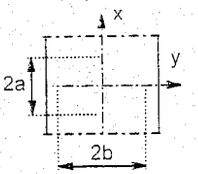
R_{wheel} [mm]	$R_{wheelprof}$ [mm]	$R_{railprof}$ [mm]	a [mm]	b [mm]	σ_N [N/mm²]	
460	∞	300	6.1	4.7	1012	
460	-330	300	3.9	14.6	502	
460	-330	80	7.1	2.7	1520	
150	-330	80	4.2	3.3	2103	

Table 2.1: Contact ellipse for different combinations of wheel and rail curvatures. Q = 60 kN.

As can be seen in Table 2.1the compressive stresses due to a wheel load of 60 kN can sometimes be higher than 1000 MPa and for small wheels may even exceed 2000 MPa.

2.7.2 Hertz spring constant

In dynamic wheel-rail models the so-called Hertz spring is of importance. This stiffness can also be derived from Hertz's theory and results in general in a rather complicated equation. Assuming a circular contact area, $R_{wheelprof} = \infty$, and adopting the geometric mean of the other radii, this relatively simple formulae can be established:

$$k_H = \sqrt[3]{\frac{3E^2 Q \sqrt{R_{wheel} R_{railprof}}}{2(1-\nu^2)^2}} \qquad (2.13)$$

where:

E = modulus of elasticity
ν = Poisson's ratio
Q = vertical wheel load
R_{wheel} = radius wheel
$R_{railprof}$ = radius railhead

Note the non-linear relationship between the Hertzian spring constant and the load Q.

Example: using the data of the first case in Table 2.1, with Q = 60 kN and ν = 0.3, the value of the Hertz spring constant then is: k_H = 1.2 GN/m².

The application of the Hertzian spring in dynamic models is discussed in a later chapter.

2.7.3 Single and two-point contact between wheel and rail

In the case of single-point contact, according to Figure 2.13, wheel load and lateral force act on the same point. This situation occurs when using worn wheel profiles. In the case of two-point contact, shown in Figure 2.14, the application points do not coincide.

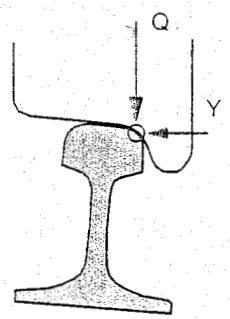


Figure 2.13: Single contact point

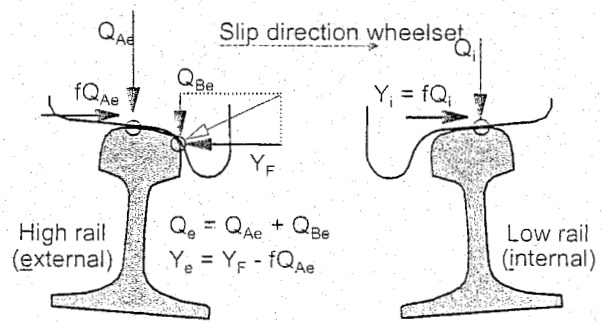


Figure 2.14: Double contact point. Forces on rails in case of lateral slip in curves

This two-point contact situation arises in curves where the wheelset lacks the freedom to position itself radially. The first wheelset of the vehicle that enters the curve makes an angle of attack producing a flange force on the high rail. This flange force forces both wheels to slip in the direction of the inside of the curve causing friction forces on the rails as indicated in Figure 2.14 and Figure 2.17.

According to Section 2.7.1 the contact force is the resultant of a stress distribution which acts on the contact area. Figure 2.15 and Figure 2.16 show pictures of both contact situations which were obtained by means of photo-elastic measurements. The calculation of the stress distribution within the rail head will be discussed in Section 5.6.4.

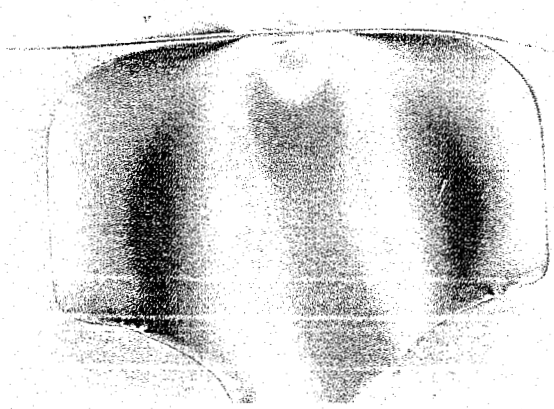


Figure 2.15: Internal cross section stress distribution in the rail head due to one-point contact

Figure 2.16: Internal cross section stress distribution in the rail head due to two-point contact

2.7.4 Spreading forces

In order to satisfy equilibrium, the horizontal component of the flange force equals the sum of the horizontal friction forces. It can be concluded that under these circumstances both rail heads are pushed apart.

Figure 2.18 gives another explanation of this phenomenon based on the rolling and creep movements to be discussed in the following sections. This figure, which by the way is viewed from the top (not a perspective view), shows the situation of a moving wheelset making lateral contact with the high rail (a).

The rotation vector of the wheelset can be resolved in a lateral component (rolling) and a longitudinal component (creep). These movements cause forces to act on the wheelset as drawn (b) which also includes the lateral flange force. From this, the resultant forces can be derived that act on the rail (c). Apparently, the lateral forces on the railheads are identified as spreading forces.

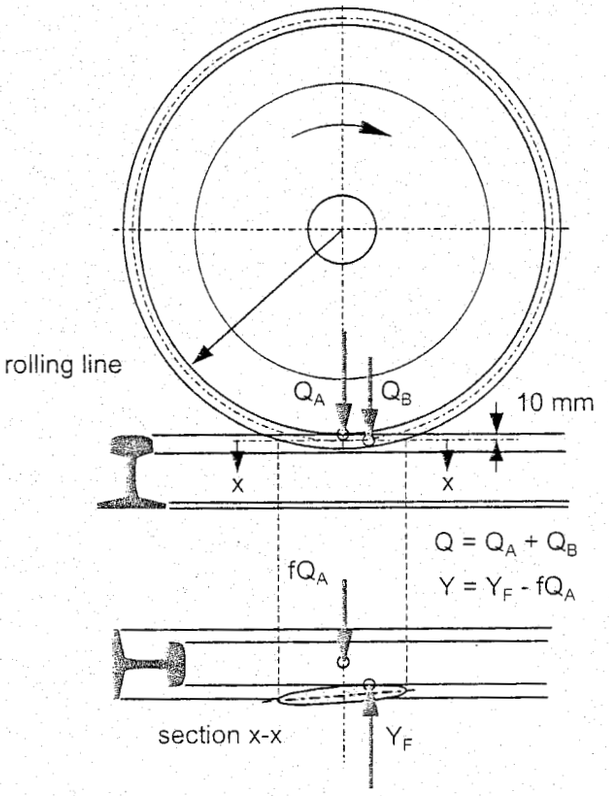


Figure 2.17: Contact between wheel flange and railhead

This observation leads to some important effects. First the track components are loaded by this mechanism. It should be noted that we deal here with internal forces.

If there are also external forces, for instance due to non-compensated accelerations, the resulting flange force can be considerably higher.

Secondly, as a result of the striking angle the flange contact point is located at some distance in front of the contact point on the running surface. This induces a sliding velocity in the vertical direction between the flange and the railhead. In combination with the normal force, this effect creates a friction force on the wheel which pushes the wheel upward. If this force is high enough to surmount the vertical force on the wheel, the wheel may climb up the railhead followed by a derailment. A simple model for derailment risk is discussed in the next section.

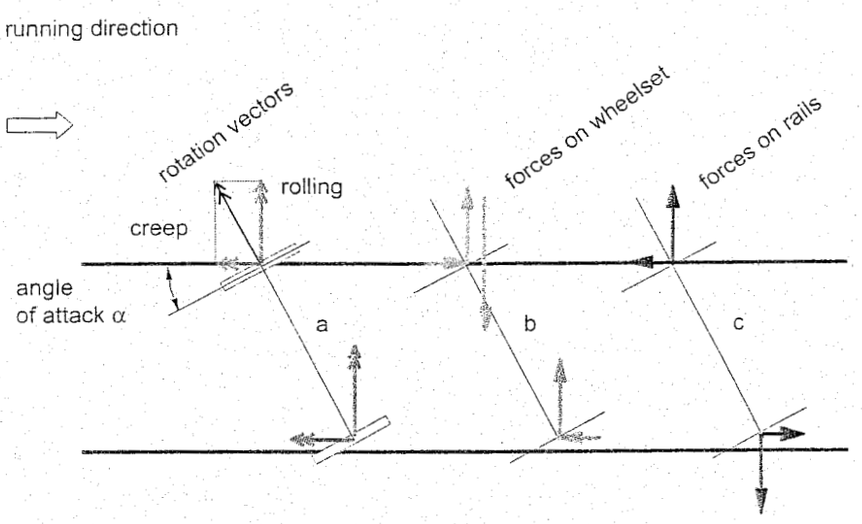


Figure 2.18: Origin of spreading forces

2.7.5 Wheel-rail creep

According to the Hertz theory the contact area is elliptic and the inside pressure distribution is also elliptic. In the case of perfectly smooth surfaces the highest pressure occurs at the centre. In reality, the surfaces show a certain roughness resulting in an irregular pressure distribution with high local peak values. At these spots, the high pressure may lead to plastic deformation of the material causing loose particles and thus wear.
Also, the pressure distribution will be changed in such a way that the maximum is moved a bit in the forward direction. Consider for instance the pulled wheel situation in Figure 2.19 (see other possibilities in Figure 2.25).

Because the resultant of the pressure is also moved forward, a moment will be created in combination with the wheel load. This moment should be surmounted during the movement and is called rolling resistance. The small asymmetry of the pressure distribution, ca. 0.5 mm, is in practice not taken into consideration.

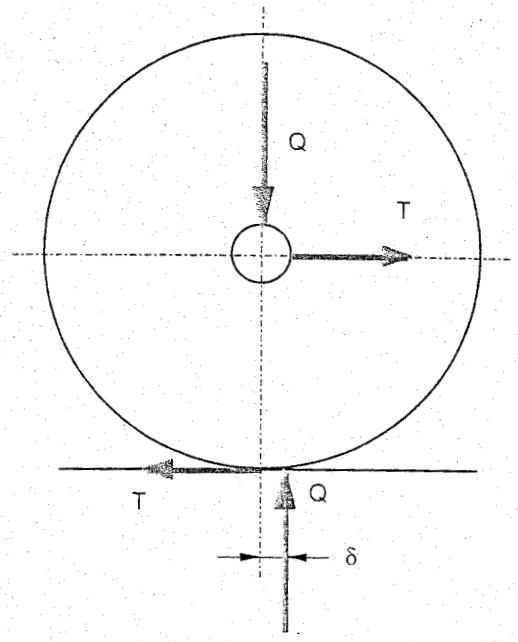


Figure 2.19: Asymmetrical pressure distribution

The tangential force, which must be transmitted in the contact area, is based on friction and can be quite substantial, e.g. with driving wheels or a braking operation.

It seems logical that for each particle of the contact area the maximum shear force to be transmitted is equal to the product of the friction coefficient and the local current pressure. This assumes that all points are about to glide at the same time. However, this is not the case because the shear forces create elastic deformations of the material of both wheel and rail. Therefore, gliding will sooner occur at the rear of the contact ellipse than at the front. (Figure 2.20).

For small values of the transmitted friction force, the elastic effects are dominant and the force will increase with the magnitude of the apparent slip. This apparent slip or 'creep' arises if part of the surfaces of wheel and rail are not gliding but are deformed elastically. As the transmitted force becomes higher, gliding will take place over a constantly increasing part of the contact area until the whole contact area takes part in gliding. The total transmitted force is then obviously equal to the friction coefficient μ times the total transmitted normal force.
The slip is defined as relative quantity. For practical purposes a simplified form is used:

$$\varepsilon_x = \frac{v - \omega r}{v} \qquad (2.14)$$

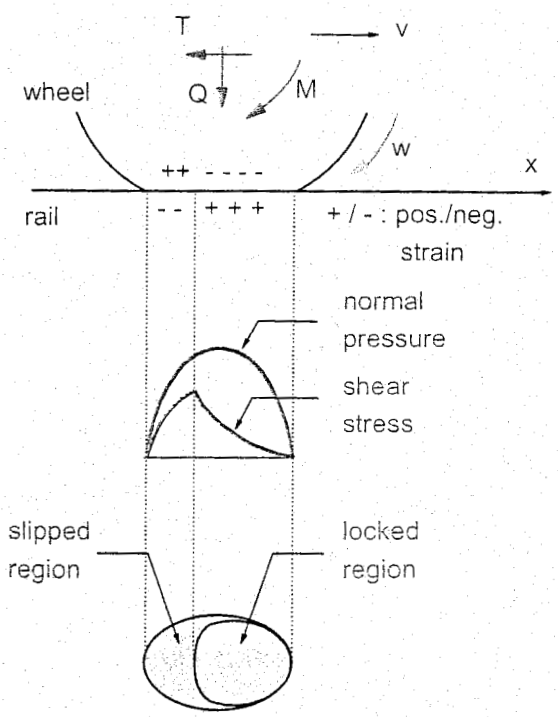


Figure 2.20: Explanation of creep phenomenon wheel/rail

The theoretical path of the friction force as a function of the creep or slip is drawn in Figure 2.21. For low values of the creep the behaviour is dominated by elastic deformations and there is an almost linear relation. For larger values, the slip zone expands and the curve becomes more and more flattened until finally it runs horizontal when slip occurs in the whole contact area.

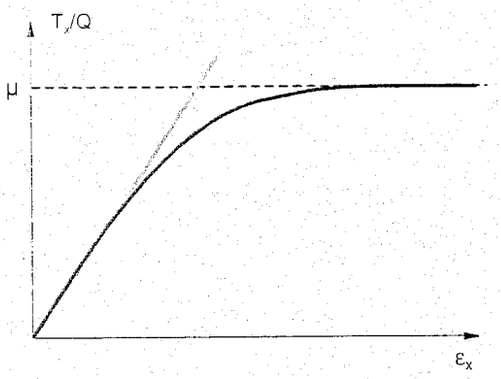


Figure 2.21: Relative longitudinal friction force versus slip

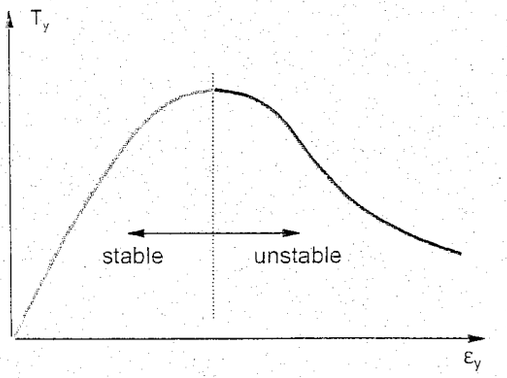


Figure 2.22: Actual lateral friction force versus slip

In the preceding consideration it was more or less assumed that the friction force and slip act in the longitudinal direction. A complete analogue mechanism acts in the lateral direction with the same functional relationship between the force T_y and the slip, the latter defined by:

$$\varepsilon_y = \frac{V_{lateral\ velocity}}{V} \tag{2.15}$$

The displayed function of the creeping force is theoretical. In reality, once full gliding occurs, the force will decrease in accordance with the increasing slip (Figure 2.22). The descending part of the curve, combined with the elasticity of the material of wheel and rail and one of the own value mode shapes of the wheel body, produces a phenomenon called stick-slip. This mechanism is the cause of the well-known squealing of wheels in curves.
Because of the presence of the attack angle of the wheel which causes the flange to push against the inside of the railhead, a lateral slip in the running surface exists and consequently a lateral slip force will be developed. The negative relation between this force and the slip will not give a stable equilibrium, but will instead induce a highly frequent jumping to and fro between two points on the curve.

2.7.6 Spin

Apart from slip in the longitudinal and lateral direction a third quantity also exists: the so-called spin which also participates in transmitting the friction force. Spin or rotational slip arises if the small contact area between wheel and rail is not parallel to the rotation axis of the wheelset. The rotation vector of the wheelset can then be decomposed in a component parallel to the contact area (this is pure rolling), and a component perpendicular to it, which is the rotational slip or spin as indicated in Figure 2.23.

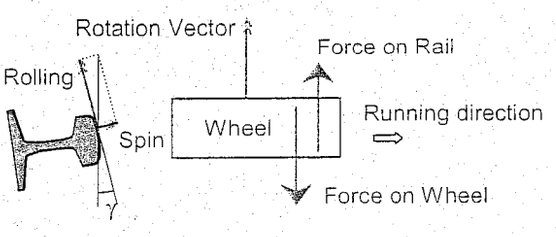


Figure 2.23: Effect of spin

The spin is defined as:

$$\phi = \frac{\omega \sin\gamma}{V} = \frac{\sin\gamma}{r} \tag{2.16}$$

It should be noted that the spin ϕ has a dimension [1/m].

When spin takes place, the relative movements between wheel and rail will also partly be taken up by elastic distortion and partly by slip. The result will be that in the contact area forces are generated with varying magnitude and direction, the resultant of which produce a force in the lateral direction. This can be clarified by means of Figure 2.24.

In this picture, which was used in numerical considerations about contact mechanics, the contact ellipse is divided into a grid of small elements. Each element shows the magnitude and direction of the slip regarding that element.

The resulting lateral force that acts on the wheel is directed to the 'high' side. Here too, small values of the force increase linearly with the magnitude of the spin. However, for a certain value of the spin a maximum is reached which decreases with the subsequent increasing of spin values.

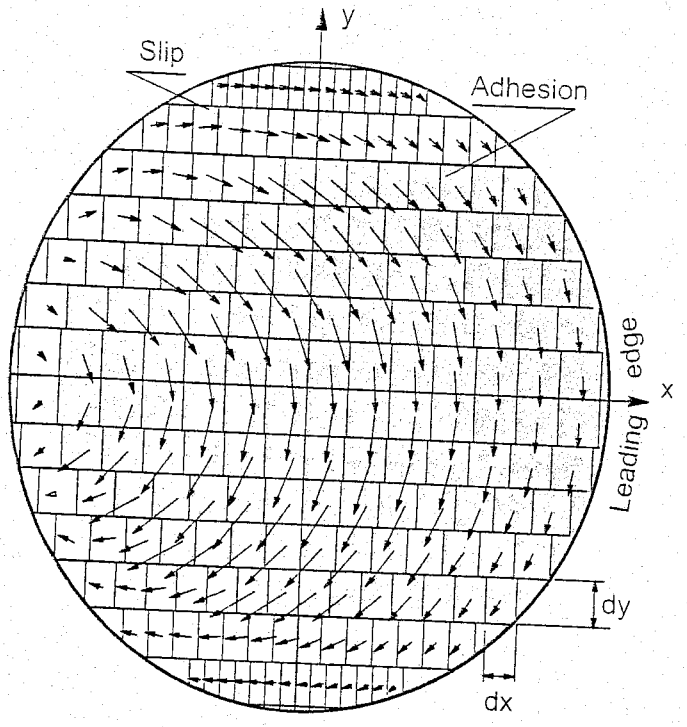


Figure 2.24: Origination of pure spin

A typical effect of this spin force is the almost complete compensation of the lateral component of the normal force for low values of the conicity. Hence the expected spreading of both rails by the lateral components of the normal force will be much less.

With greater values of the creep and slip these relations will not only be non-linear, but mutual influences also exist which make the connections much more complicated.

2.7.7 Creepage coefficients

According to Kalker [147], for both creep and spin it may be assumed that for small values the relations between these quantities and the resulting generated forces are about linear and can be expressed by:

$$T_x = Gc^2 C_{11} \varepsilon_x \tag{2.17}$$

$$T_y = Gc^2 (C_{22} \varepsilon_y + C_{23} c\phi) \tag{2.18}$$

$$M_z = Gc^3 (C_{23} \varepsilon_y + C_{33} c\phi) \tag{2.19}$$

with:

G = shear modulus
c = geometric mean of semi-axes of contact ellipse: $c = \sqrt{ab}$
C_{ij} = the so-called Kalker coefficients: constants determined by the ratio between the semi-axes a and b and the normal force N on the contact area.

In Table 2.2 the Kalker coefficients are given for the friction forces T_x and T_y (the moment M_z is disregarded).

		C_{11}			C_{22}			C_{23}		
g		ν = 0	1/4	1/2	ν = 0	1/4	1/2	ν = 0	1/4	1/2
	↓0	$\pi^2/\{4(1-\upsilon)\}$			$\pi^2/4$			$\frac{\pi\sqrt{g}}{3(1-\upsilon)}\left\{1+\upsilon\left(\frac{1}{2}\Lambda+ln4-5\right)\right\}$		
	0.1	2.51	3.31	4.85	2.51	2.52	2.53	0.334	0.473	0.731
	0.2	2.59	3.37	4.81	2.59	2.63	2.66	0.483	0.603	0.809
	0.3	2.68	3.44	4.80	2.68	2.75	2.81	0.607	0.715	0.889
	0.4	2.78	3.53	4.82	2.78	2.88	2.98	0.720	0.823	0.977
b>a	0.5	2.88	3.62	4.83	2.88	3.01	3.14	0.827	0.929	1.07
	0.6	2.98	3.72	4.91	2.98	3.14	3.31	0.930	1.03	1.18
	0.7	3.09	3.81	4.97	3.09	3.28	3.48	1.03	1.14	1.29
	0.8	3.19	3.91	5.05	3.19	3.41	3.65	1.13	1.15	1.40
	0.9	3.29	4.01	5.12	3.29	3.54	3.82	1.23	1.36	1.51
	1.0	3.40	4.12	5.20	3.40	3.67	3.98	1.33	1.47	1.63
	0.9	3.51	4.22	5.30	3.51	3.81	4.16	1.44	1.59	1.77
	0.8	3.65	4.36	5.42	3.65	3.99	4.39	1.58	1.75	1.94
	0.7	3.82	4.54	5.58	3.82	4.21	4.67	1.76	1.95	2.18
	0.6	4.06	4.78	5.80	4.06	4.50	5.04	2.01	2.23	2.50
a>b	0.5	4.37	5.10	6.11	4.37	4.90	5.56	2.35	2.62	2.96
	0.4	4.84	5.57	6.57	4.84	5.48	6.31	2.88	3.24	3.70
	0.3	5.57	6.34	7.34	5.57	6.40	7.51	3.79	4.32	5.01
	0.2	6.96	7.78	8.82	6.96	8.14	9.79	5.72	6.63	7.89
	0.1	10.7	11.7	12.9	10.7	12.8	16.0	12.2	14.6	18.0
	↓0	$\frac{2\pi}{(\Lambda-2\upsilon)g}\left\{1+\frac{3-ln4}{\Lambda-2\upsilon}\right\}$			$\frac{2\pi}{[(1-\upsilon)\Lambda+2\upsilon]g}\left\{1+\frac{(1-\upsilon)(3-ln4)}{(1-\upsilon)\Lambda+2\upsilon}\right\}$			$\frac{2\pi}{3g\sqrt{g}}\frac{1}{(1-\upsilon)\Lambda-2+4\upsilon}$		

$\Lambda = ln(16g^2)$; $g = min(a/b, b/a)$; $ln4 = 1.386$; υ: Poisson's ratio

The analytical forms involving Λ are accurate for very small g only.

Table 2.2: Creepage and spin coefficients.

2.8 Train resistances

2.8.1 Types of resistances

A train must overcome several train resistances to start moving or to continue moving. Instead of the train resistance itself, denoted by W, mostly the relative value is given with respect to the train weight as indicated by the symbol w. The relative value is independent of the train weight and its unit is N/kN or, more customary, ‰. The most important train resistances are:

- *Running resistance*, consisting primarily of the rolling resistance between wheel and rail and the journal friction (see Figure 2.25).The order of magnitude of the running resistance is 1.5 to 2 ‰, which is much lower than in the case of road vehicles (10 to 30 ‰);

- *Air resistance*, which depends on the dimensions and the cross section shape of the rolling stock. The air resistance is proportional to the square of the speed while a possible headwind should also be taken into account;

- *Curve resistance*, to be discussed in Chapter 3;
- *Gradient resistance*, to be discussed in Chapter 3;

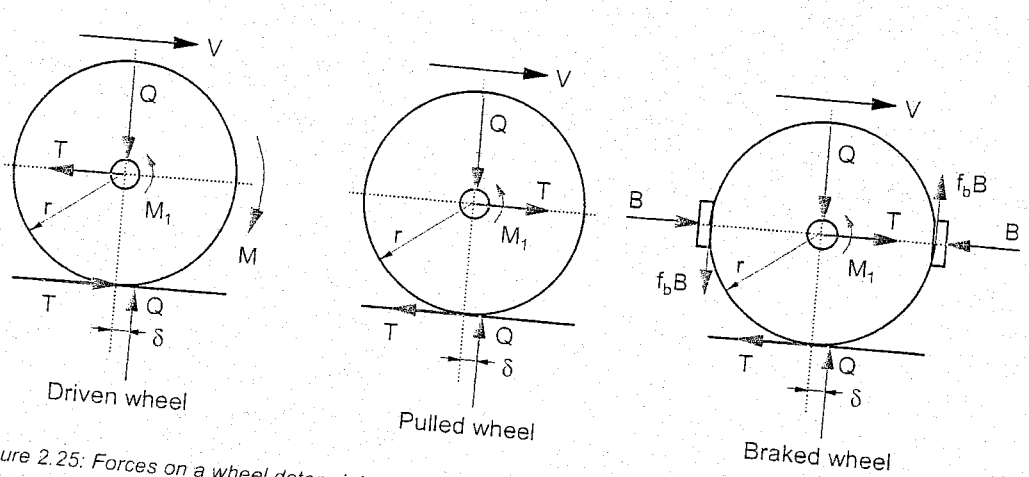


Figure 2.25: Forces on a wheel determining running resistances

Furthermore, the following resistances may be of interest:

- *Switch resistance*, in diverging direction;
- *Tunnel resistance*, (metro's, high speed track in tunnel);
- *Pulling up resistance*, originating from pulling up stock after a long standstill.

The total resistance is equal to the sum of all separate resistances which are applicable in the given situation:

$$R_{tot} = \sum r_i W_i \qquad (2.20)$$

where:

r_i = the i^{th} relative resistance in ‰, which arises on the track section considered;
W_i = the train weight, which experiences the resistance concerned. This is not always the total train weight (for instance, long train in short curve).

2.8.2 Required pulling force

Of importance is the pulling force, which is necessary to move a particular train formation with a certain speed on a railway line and with given resistances.
The pulling force is dependent on:

- The available power of the propelling installation;
- The desired speed on the section considered;
- The available adhesion force of the driven wheels.

If the power is assumed to be constant, the pulling force is simply:

$$F = \frac{P}{v} \tag{2.21}$$

where:

P = traction power [kW]
v = speed [m/s]

This relation can be drawn in a diagram (Figure 2.26) which is known as the force-speed diagram. In the same diagram the connection between the total resistance and the speed is drawn for two cases.

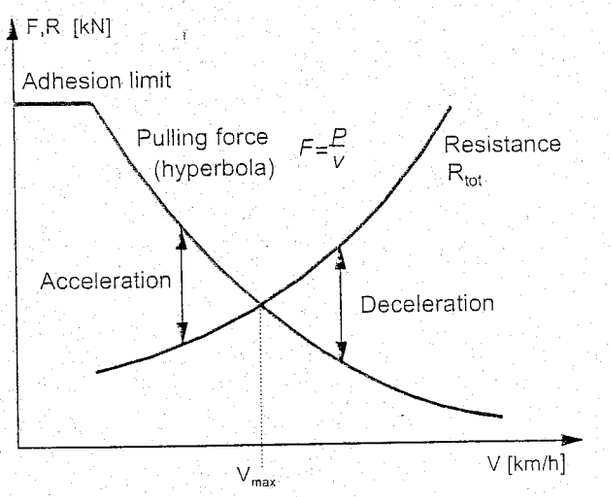


Figure 2.26: Force-speed diagram

If the pulling force F is greater than the total resistance R_{tot}, then the train mass experiences an acceleration (leading to a speed rise) determined by:

$$F - R_{tot} = \gamma M_{tot} a \tag{2.22}$$

where:

a = acceleration
M_{tot} = total train mass
γ = multiplication factor for the mass, which takes account of the inertia of the rotating masses. (γ = 1.06 - 1.10).

From (2.22) it follows that (with γ = 1 and g ≈ 10 m/s²) the acceleration is 0.01 m/s² per 1 ‰ pulling force excess.

The theoretical maximum speed in Figure 2.26 will only be reached after a very long time. In practice the maximum speed will therefore be lower.

If at a certain speed the resistance becomes higher than the available pulling force, the train mass will undergo a decreasing of the speed.

2.8.3 Adhesion force

The adhesion force T between a driven wheel and the rail is limited as given by Coulombs law:

$$T \leq fQ_{adh} \tag{2.23}$$

where:

f = friction coefficient (adhesion coefficient) between steel wheel and rail. Global values are:

- f = 0.25 for dry rails;
- f = 0.1 for wet rails.

Q_{adh} = wheel load of the driven wheel.

As the pulling force is the sum of all adhesion forces, the following result can be obtained by using (2.23):

$$F = \sum T \leq f \sum Q_{adh} \leq f W_{adh} \tag{2.24}$$

where:

W_{adh} is the total adhesion weight of the train. In case of pulled vehicles, the adhesion weight is the weight of the locomotive.

From (2.22) and (2.24) it follows with $R_{tot} = 0$, $M_{adh} = M_{tot}$, $\gamma = 1$, for the acceleration the physical upper limit:

$$a_{max} = fg \tag{2.25}$$

The order of magnitude of acceleration for maximum pulling is about 0.2 - 1.2 m/s². The maximum braking deceleration can be substantially higher than the maximum starting acceleration because all axles are braked. However, in order to protect passengers or loading a maximum value of 0.5 to 1.3 m/s² is maintained. Tramway cars are also equipped with an electro-magnetic rail brake, with which in case of emergency (suddenly obstructing street cars) a substantially higher deceleration can be obtained.

3 CURVES AND GRADIENTS

3.1 General considerations

Longitudinal track alignment consists of the following geometrical elements:

- In a straight direction: straight flat track sections. These are the easiest to make and have the lowest resistance;
- In the horizontal direction: curves, cant, transition curves, and transition gradients;
- In the vertical direction: gradients and vertical curves.

The physical appearance of these elements is determined by the characteristics of vehicle behaviour, translated into simple instructions for use and formulae based on safety, comfort, and cost-effectiveness criteria.

From a commercial point of view, it is desirable to increase speeds as much as possible. It is essential in this respect that the load on the track and the accelerations on the vehicle body, which are the determining factors for passenger comfort, do not exceed specific limit values and that safety against derailment and lateral displacement of the track are ensured at all times. In the case of new lines, such operating requirements can be worked in at the design stage. On existing lines the possibilities for increasing speed are usually restricted.

The standards presented in this chapter are taken from UIC leaflet 703 R [277] which was revised in 1988 and which takes recent experience on the European railways into account. The values recommended in these regulations relate to modern ballasted track and modern passenger stock which meet UIC standards. In addition to desirable values for accelerations, maximum values are also given, as are values which are only permissible in exceptional cases. Some examples are also given from NS regulations for the design of new lines [192].

3.2 Curvature and superelevation in horizontal curves

3.2.1 Curve radius/curvature

The convention in Europe is to express horizontal curve geometry as the curve radius R or as curvature 1/R. Elsewhere curve radius is often given as the angle α, in degrees, of a 100 ft-long (30.48m) curve. The relationship between α and R [m] is:

$$\alpha = \frac{1746}{R} \qquad (3.1)$$

This relationship is shown in Figure 3.1.

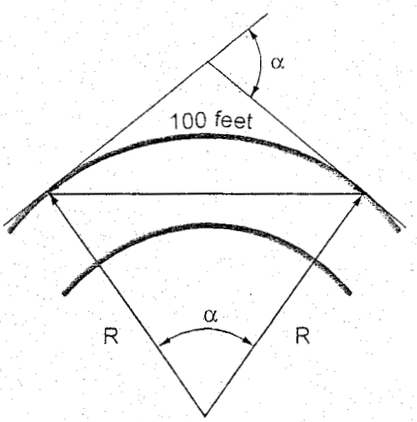


Figure 3.1: Relationship between radius and curve angle

3.2.2 Curve effects

It is a known fact that a vehicle running at a speed v in a curve with a radius R experiences a centrifugal lateral acceleration of $a = v^2/R$ which results in a number of undesirable effects:

- Possible passenger discomfort;
- Possible displacement of wagon loads;

- Risk of vehicles overturning;
- Risk of derailment caused by the wheels mounting the outer rail or by loosening of rail fastenings;
- High lateral forces on the track, which increase:
  - curve resistance;
  - wear of rails and wheel flanges;
  - risk of rail tilting;
  - risk of lateral displacement of the whole track;
  - noise nuisance.
- High lateral forces on track structure and substructure.

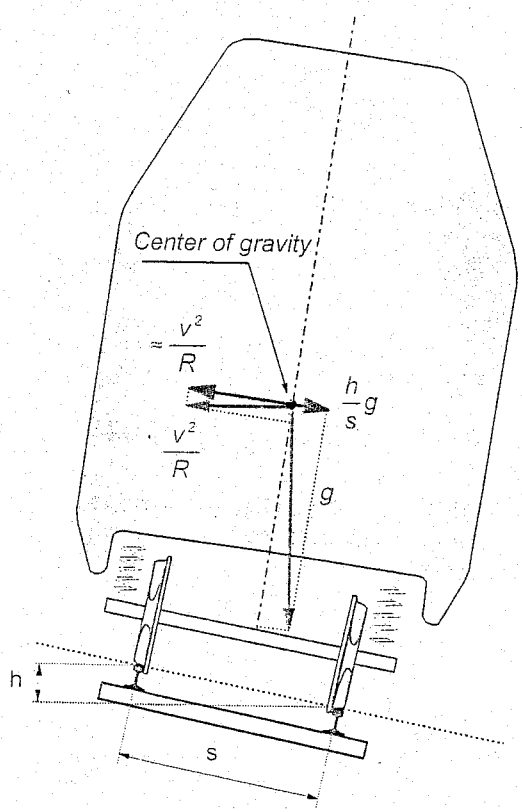


Figure 3.2: Accelerations in curves

To limit, if not prevent, these phenomena the following measures can be implemented:

- Use of maximum possible curve radius R, preferably so that no superelevation is necessary;
- Use of cant in curves so that lateral acceleration is entirely or partly compensated by the gravity component;
- Speed restrictions. This is not an attractive option because of the consequent increase in transit time and the loss in capacity.

3.3 Superelevation

3.3.1 General considerations

If it is not possible to make a suitably large curve radius, the curve with a smaller radius must have cant in order to reduce or eliminate the influence of the centrifugal acceleration. This can improve the situation as regards passenger comfort and horizontal loading of the track. This provision is only made in curves where V > 40 km/h.

Superelevation is achieved on NS by raising the outer rail in the curve, whilst keeping the inner rail at the same level.

If we now consider the general situation depicted in Figure 3.2, in which a vehicle is running at a speed of v in a curve with radius R and cant h, the resultant non-compensated lateral acceleration on the vehicle is:

$$a_d = \frac{v^2}{R} - \frac{gh}{s} \qquad (3.2)$$

The ideal cant appears as $a_d = 0$. In this case the resultant of $a = v^2/R$ and g is perpendicular to the track and equals:

$$h_{id} = \frac{sv^2}{gR} = \frac{11.8V^2}{R} \tag{3.3}$$

in which:

v : running speed [m/s];
V : running speed [km/h];
R : curve radius [m]
g : acceleration due to gravity (= 9.81 m/s²)
h : cant [mm]
s : track width (= 1500 mm)

For practical reasons the thus calculated cant is rounded up to the nearest 5 mm. If the calculated cant is less than 20 mm it can be disregarded.

3.3.2 Cant deficiency

Ideal cant applies to just one speed and therefore in principle can only be used by railways with uniform traffic. Generally speaking, however, passenger and freight trains run on the same track at different speeds, which means that ideal cant for the top speed would result in considerable excess cant for the slow-running traffic. This would in turn produce excess wear on the low rail. A compromise is, therefore, to accept a certain degree of cant deficiency for the fast trains, producing flanging on the high rail and thus lateral wear of the rail head. This, however, outweighs the disadvantages of speed limits. Cant deficiency h_d is the difference between ideal cant and actual cant and must satisfy the condition:

$$h_d = \frac{sv^2}{gR} - h = 11.8\frac{V_{max}^2}{R} - h < \bar{h}_d \tag{3.4}$$

in which V_{max} is the maximum speed in km/h. Non-compensated acceleration and cant deficiency are related as follows:

$$a_d = \frac{g}{s} h_d \tag{3.5}$$

Non-compensated acceleration must satisfy the following condition:

$$a_d = \frac{V_{max}^2}{12.96} - \frac{h}{153} < \bar{a}_d \tag{3.6}$$

The non-compensated acceleration causes a quasi-static lateral track load of the magnitude of the mass per axle times the acceleration. With a non-compensated acceleration of 1 m/s^2, which must be considered as the absolute limit at international level, and a mass per axle of 22.5 t, the lateral load on the track is 22.5 kN per axle. This load must be supplemented by the dynamic load, which to a large extent depends on the quality of the track geometry. The same reasoning applies to the vehicle body accelerations which determine passenger comfort. Both are shown schematically in Figure 3.3.

It should be noted that another limit is set on the time derivative of the acceleration da_d/dt in transition curves (see also 3.5.2). This 'yerk' should generally not exceed 0.2 m/s^3.

The cant deficiency allowed in practice is determined by the following factors:

- Track construction;
- State of track components;
- Track alignment;
- Type of bogie;
- Axle loads and unsprung masses;
- State of maintenance of the rolling stock.

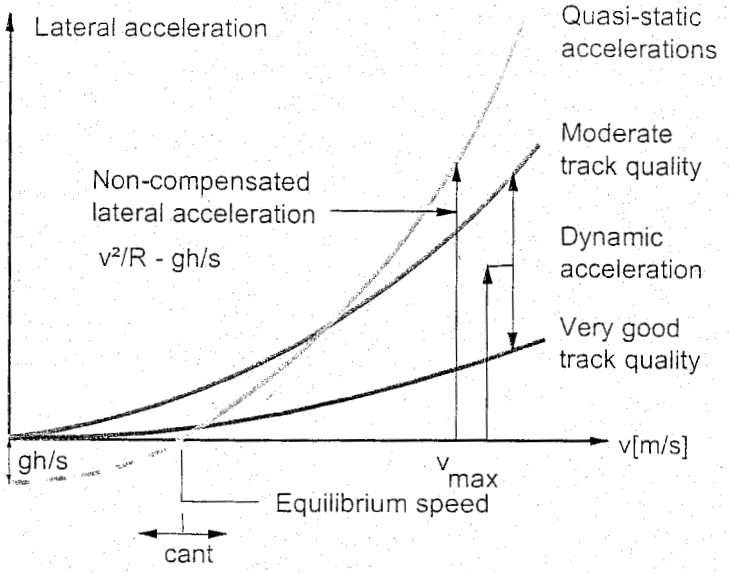


Figure 3.3: Quasi-static and dynamic accelerations in curves versus speed

If high values are allowed for non-compensated acceleration, the track components must be designed accordingly and there must be no risk of exceeding the lateral track resistance immediately after tamping. A large cant deficiency causes a large increase in wear on the face of the high rail head. Moreover, an increase in cant deficiency leads to faster deterioration of the track geometry.

3.3.3 Effect of suspension on lateral acceleration

With a non-compensated lateral acceleration, the centrifugal force exerted at the centre of gravity will cause an overturning moment which means that the spring system of the vehicle body is compressed unevenly and the body tilts somewhat towards the outside. Since this cancels out some of the cant, the vehicle body and hence the passengers and load are subjected to a higher lateral acceleration of:

$$a_{ds} = (1 + \varepsilon) a_d \qquad (3.7)$$

The acceleration a_i must in all cases remain below 1.5 m/s², and preferably below 1 m/s². The cant deficiency must be adapted to the suspension factor ε. For modern rolling stock this value is of the order of 0.4 and can be reduced by special measures to 0.2.

3.3.4 Effect of body tilt coaches on cant deficiency

To be able, nevertheless, to run on heavily curved sections at high speed, the coaching stock is sometimes designed to tilt artificially. By means of such tilting mechanisms the cant deficiency in curves at high speed can be compensated by h_c, as a result of which speed can be increased by up to 50% in comparison with coaches without body tilt. The remaining lateral acceleration on the passengers and load is then:

$$a_{dt} = a_d - \frac{g}{s} h_{tilt} = \frac{g}{s}(h_d - h_{tilt}) \qquad (3.8)$$

The only improvement is increased passenger comfort and reduced risk of overturning; the track, on the other hand, is more heavily stressed by the higher lateral forces.

3.3.5 Switches and other constraints

Regarding switch curves, basically the same conditions apply to non-compensated accelerations as for plain line curves. The dynamic effects due to impacts are, however, much greater in switches. As already shown in Figure 3.3, the dynamic and thus the overall acceleration and force level increases considerably because of the relatively poor track geometry. For this reason cant deficiency in switches is kept lower than in plain line curves.

3.3.6 Cant excess

In curves with rather large radii there is generally a substantial difference between the maximum speed V_{max} of passenger trains and the lowest speed of freight trains V_{min}. At the lowest speed cant excess is of the order of:

$$h_e = h - 11.8\frac{V_{min}^2}{R} \tag{3.9}$$

A substantial cant excess produces a high load on the low rail in the case of slow-running freight trains. Figure 3.11 gives the maximum values used internationally.

3.3.7 Maximum cant

A maximum value is set for cant because of the following problems which arise if the train if forced to stop or run slowly in a curve:

- Passenger discomfort;
- Risk of derailment for freight trains due to uneven loading on the rails;
- Possible shifting of freight loads;
- Possible breakaway of freight trains from standstill because of the high level of friction of the wheels against the inner rail.

These problems can be aggravated by high cross winds. For these reasons the maximum cant in curves is:

$$\begin{aligned} &h = 120 \text{ mm track on gravel ballast} \\ &h = 150 \text{ mm track on crushed stone ballast} \end{aligned} \tag{3.10}$$

With just one type of traffic cant values up to 180 mm are used.

Large cant can also result in lateral displacement towards the centre of the curve during tamping at low temperatures. A smaller cant than in (3.10) is therefore preferable if possible. The maximum values used internationally are shown in Table 3.1.

3.4 Transition curves

3.4.1 General remarks

Basically, transition curves are used between straight track and curves or between two adjacent curves to allow gradual change in lateral acceleration. The centre line of the transition curve has the same tangent at the connection points as the adjacent part, whereas the curvature changes gradually from the value of the one connection point to the value of the other. Transition curves also introduce cant by means of transition gradients and, if necessary, gauge widening.

Transition curves are not used if:

- The curve radius is > 3000 m [192];
- A calculation shows that no cant is necessary;
- Between two adjacent curves in the same direction the discontinuity in acceleration remains limited to 0.2-0.3 m/s².

In the following, only transition curves between straight track and curve are examined. The track is presented in a simplified form by a single straight or curved line.

3.4.2 Clothoid

Between straight track and a curve, the curvature of the transition curve 1/r must gradually increase from zero in the straight track to the value 1/R in the curve. This can be achieved by starting from the linear relationship:

$$\frac{1}{r} = ks \tag{3.11}$$

Here 1/r is the curvature over a curve length s from the origin, as shown in Figure 3.4. At the origin the curvature is zero and therefore this is selected as the point of connection with the straight track. The factor k is a proportionality factor. The equation of the transition curve y is now to be found as a function of x. The exact form reads:

$$\frac{d\phi}{ds} = \frac{1}{r} \tag{3.12}$$

Using (3.11) ϕ can be solved:

$$\phi = \int \frac{1}{r} ds = \tfrac{1}{2} ks^2 \tag{3.13}$$

The Cartesian coordinates are thus determined and read:

$$x = \int \cos\phi \, ds = \int \cos\frac{ks^2}{2} ds \tag{3.14}$$

$$y = \int \sin\phi \, ds = \int \sin\frac{ks^2}{2} ds \tag{3.15}$$

The curve can therefore only be described in parameter form, expressed as Fresnel integrals. The curve is known as a clothoid or Cornu spiral [2]. Owing to the mathematically difficult formulation and the need to use a computer, a somewhat simpler transition curve, known as a cubic parabola, is usually used for manual calculations.

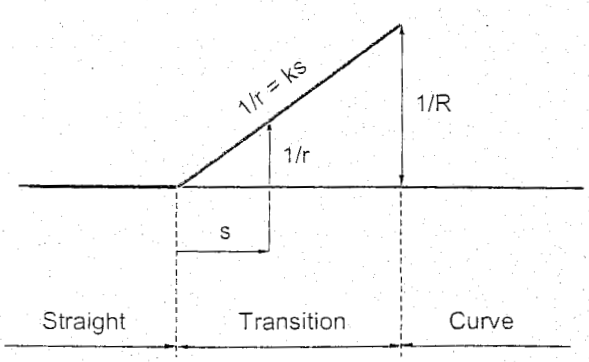


Figure 3.4: Curvature path of clothoid

3.4.3 Cubic parabola

To obtain this function for the transition curve two simplifications are introduced with respect to the clothoid. The curve is approximated by:

$$\frac{1}{r} = y'' \qquad (3.16)$$

The curve length s is approximated by the coordinate x. This applies only for small L/R ratios.

If linear behaviour is again chosen for the curve, beginning at zero at the point of connection with the straight section, the curvature is:

$$\frac{1}{r} = y'' = kx \qquad (3.17)$$

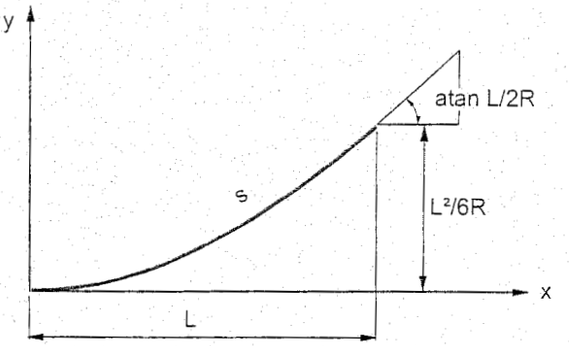


Figure 3.5: Cubic parabola

in which k is a proportionality factor. After twice integrating and then introducing the boundary conditions, the formula for the cubic parabola is found:

$$y = \frac{x^3}{6RL} \qquad (3.18)$$

in which L represents the length of the transition curve projected on the x-axis, as shown in Figure 3.5. This expression agrees with the first term of the series expansion for the clothoid.

DB use a fourth order parabola in which the curvature is quadratic (geschwungene Rampe).

3.4.4 Curve displacement

In case of the juxtaposition of a transition curve, the original curve is displaced with respect to the straight section as shown in Figure 3.6. According to (3.18):

$$tg\alpha = y'(L) = \frac{L}{2R} \qquad (3.19)$$

The curve displacement is therefore:

$$p = \frac{L}{2} - R\sin\alpha \cong \frac{L^3}{16R^2} \qquad (3.20)$$

$$q = \frac{L^2}{6R} + R\cos\alpha - R = \frac{L^2}{24R} \qquad (3.21)$$

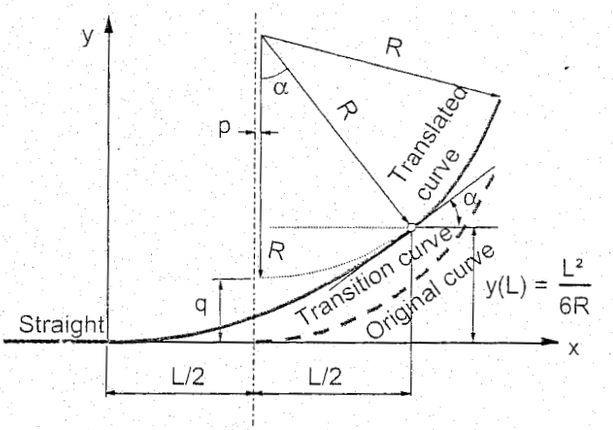


Figure 3.6: Curve transition

Displacement p is very small and can usually be disregarded. Displacement q = 0.25 y(L) and is therefore significant. Length L is obtained from the formulae given in section 3.5.

3.5 Cross level transitions

3.5.1 Relation with the transition curve

To allow transition from straight track to curved track with cant, use is made of a transition gradient. The cant is here changed from zero in the straight track to the value h in the curve by gradually raising the outer rail. Since the non-compensated lateral acceleration has to increase gradually from zero to the value a_d in the curve, it follows that:

- The transition gradient and transition curve must coincide;
- The increase in cant and the increase in curvature must follow the same function.

The cant h_x in the transition gradient is therefore as shown in Figure 3.7:

$$h_x = \frac{h}{L}x \qquad (3.22)$$

in which:
h = cant in the curve
L = length of the transition curve/gradient.

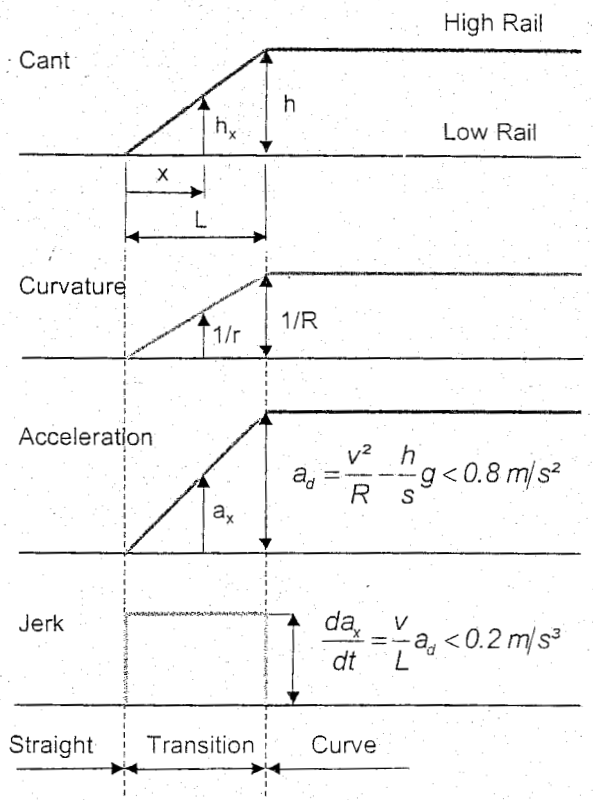


Figure 3.7: Lateral acceleration and jerk

Non-compensated lateral acceleration in the transition curve varies linearly according to:

$$a_x = \frac{a_d x}{L} \qquad (3.23)$$

whereas the change in acceleration:

$$\frac{da_x}{dt} = v\frac{da_x}{dx} = \frac{a_d v}{L} \qquad (3.24)$$

is constant in the transition curve but is discontinuous at the beginning and the end, as shown in Figure 3.7. During transition from straight track to a curve with cant, or vice versa, the rotation of the vehicle must be gradual in order to maintain comfort which is associated with lateral acceleration and its change per unit of time, and also for reasons of safety. This means that at maximum speed the following condition must be met:

$$\frac{da_d}{dt} = \frac{a_d V_{max}}{3.6L} < \frac{d\bar{a}_d}{dt} \qquad (3.25)$$

The permissible values are given in Table 3.1. It should be pointed out that safety is related to the difference in cant over the distance between two axles of one bogie. As a result of this distortion the four wheels are not in one plane and the wheel load is thus unevenly distributed. Combined with a horizontal force this can result in climbing of the wheel on the rail and derailment.

Most design standards for cant variations in curves specify 1 : 10V (V [km/h]) as the desirable value. On NS steeper transition gradients of up to 1 : 8V are allowed with a minimum of 1 : 600 at speeds lower than 80 km/h. The requirements concerning non-compensated acceleration with, for example, a_d = 0.8 m/s² and a_d = 0.2 m/s³ lead to the condition:

$$L \geq 1.11V \qquad \text{with L [m], V [km/h]} \tag{3.26}$$

3.5.2 Length of normal transition curve

The length of a transition curve is normally obtained from the ideal cant according to (3.3) and the maximum permissible change in acceleration according to (3.25). If a_d = 0.2 m/s³ this gives:

$$L = 0.118\frac{V^3}{R} \qquad \text{with: L [m], V [km/h], R [m]} \tag{3.27}$$

3.5.3 Adjacent curves

Generally speaking there should be a minimum straight section of 30 m, but preferably of V/2 (V [km/h]) between the transition curve because of the length of the vehicle as shown in Figure 3.8a. If this is not possible the following applies:

- For two adjacent curves in the same direction the cant must stay the same over at least 30 m between the transition gradients. Between curves, one transition curve or adjacent transition curves in opposite directions should be used in accordance with Figure 3.8b and Figure 3.8c.
- For two adjacent curves in opposite directions, known as curve and reverse curve, with only one point where the cant is zero, the cant should vary continuously from the one direction to the other. The transition curves should be extended so that they join. This situation is depicted in Figure 3.9.

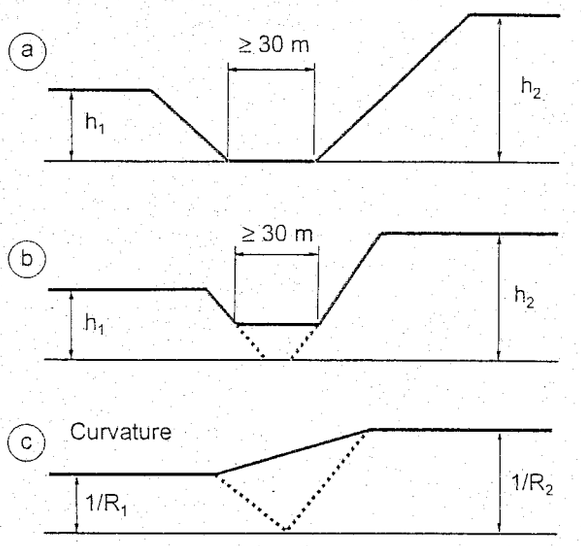


Figure 3.8: Situation with adjacent curves having the same curvature sign

3.6 Curve resistance.

A moving vehicle must overcome a series of resistances, including curve resistance, caused by slipping of the wheels and flanging as a result of non-radial adjustment of the wheelsets.

The curve resistance, which is defined as the total resistance force divided by the weight of the part of the train in the curve, is dependent on the curve radius, track gauge, condition of the rails, and the characteristics of the rolling stock. To reduce curve resistance rail lubrication or flange lubrication can be used. The primary objective of these measures is, however, to reduce rail and flange wear.

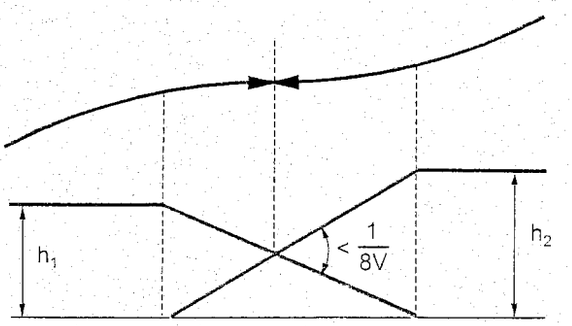


Figure 3.9: Situation with adjacent curves having an opposite curvature sign

For curve resistance, expressed as N/kN or as ‰, NS use the following formulae:

$$W_R = \frac{650}{R - 55} \text{ ‰} \quad \text{for } 400\ m < R$$
$$W_R = \frac{750}{R} \text{ ‰} \quad \text{for } 150\ m < R < 400\ m \qquad (3.28)$$
$$W_R = 5 \text{ ‰} \quad \text{for } R < 150\ m$$

3.7 Gradients

3.7.1 Gradient resistance

If differences in track level cannot be avoided, gradients are used. For a given slope, the minimum hauling force required to overcome the difference in level alone, according to Figure 3.10, is at least:

$$W_H = G \sin\delta \cong G\delta \quad (if\ \delta << 1) \qquad (3.29)$$

Similar to other resistances which trains must overcome, a gradient resistance can be defined as:

$$W_H = \delta \qquad (3.30)$$

where δ can be expressed in [N/kN], [mm/m], [m/km], or [‰].

The gradients cannot be too steep in connection with the maximum available adhesion force between the driven wheels and the rails. Nevertheless, the gradient resistance is an important part of the resistance as a whole.

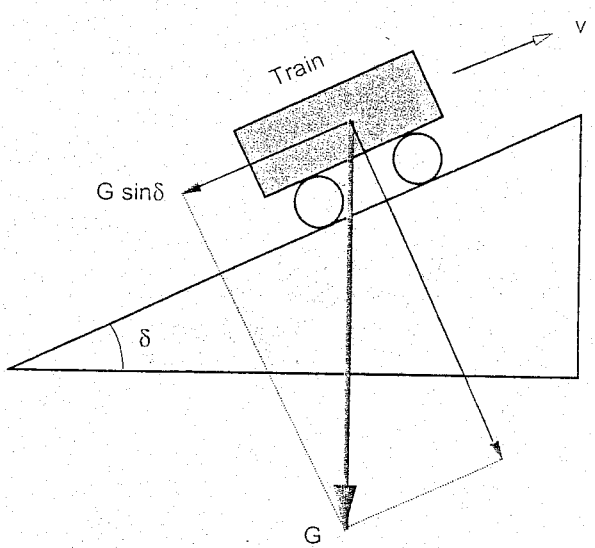


Figure 3.10: Gradient resistance

3.7.2 Magnitude of gradient

On plain line NS allow gradients of 5 ‰ (1:200). On such gradients any train can move away from standstill. If steeper gradients prove necessary, there may be restrictions on the types of train used, the tonnages and speeds, and this would result in an increase in operating and maintenance costs.

In exceptional cases NS allow gradients of up to 16.6 ‰ (1:60), for example, in the Velsen tunnel; use being made here of the kinetic energy of the train. A few other examples in Europe are:

Gotthard	: 27 ‰	DB main lines	: 12 ‰
Paris-Lyon (TGV)	: 35 ‰	DB secondary lines	: 40 ‰

With a few exceptions, such as the Flåmbahn in Norway with 55 ‰, the maximum value for adhesion railways lies between 30 ‰ and 40 ‰. Above this the use of a rack railway or a funicular must be considered.

Traffic Vmax [km/h]	Class 1 80 - 120			Class 2 120 - 200			Class 3 > 250				Class 4 250 -300	
							FS		DB		SNCF	
Value	Norm.	Max.	Exc.	Norm.	Max.	Exc.	Norm.	Max.	Norm.	Max.	Norm.	Max.
$\bar{h}_d$ [mm]	80	100	130	100	120	150	120	–	40	60	50	100
$\bar{a}_d$ [m/s²] plain track	0.52	0.65	0.85	0.67	0.79	0.98	0.79	--	0.26	0.40	0.33	0.67
$\bar{h}_d$ [mm]	60	80	120	60	80	100	--	--	--	–	50	100
$\bar{a}_d$ [m/s²] switches	0.40	0.52	0.79	0.40	0.52	0.67	--	–	--	–	0.33	0.67
$\bar{h}_e$ [mm]	50	70	90	70	90	110	100	–	50	70	–	110
$\bar{h}$ [mm] plain track	150	160	-	120	150	160	125	–	65	85	180	–
$\frac{\overline{dh_d}}{dt}$ [mm/s]	25	70	90	25	70	–	36	–	13	–	30	75
$\frac{\overline{da_d}}{dt}$ [mm/s³]	0.16	0.46	0.59	0.16	0.46	–	0.24	–	0.09	–	0.20	0.49
$\frac{\overline{dh}}{dt}$ [mm/s]	28	46	55	28	35	50	38	–	20	–	50	60
$\frac{\overline{da_v}}{dt}$ [mm/s³]	0.20	0.30	0.40	0.20	0.30	–	0.16	0.24	0.20	–	0.45	0.60

Table 3.1: Recommended admissible quasi-static accelerations and related values in curves

3.8 Alignment in mountainous areas

The alignment (or layout) in mountainous areas differs very much from that in flat or slanting land. Because of the great height differences in mountainous areas, the designer should always seek for a compromise between on the one hand the desires of:

- Slight gradients;
- Large curve radii;
- Minimal length-development.

And on the other hand the constructional and operational costs.

In mountainous areas main railway line gradients have been applied of 25 to 30 ‰ and curve radii of 300 m and less.
An important design variable is the ruling down-gradient: this the maximum gradient for a given locomotive, a given train weight, and a given speed. The choice of this gradient will compromise between investment costs and operational costs: investment costs in construction and purchase of locomotive, operational costs as a function of the number of trains, the running time, and the energy. One should realise that locomotives are not only meant for the difficult piece of alignment using a ruling down-gradient. On the rest of the line the power should also be utilised (in terms of speed); otherwise the rest power is unused and hence less profitable.

The traction vehicle on gradients should be able:

- To move the train with the desired speed;
- To start moving the train on the gradient after (unexpected) standstill.

During the start one is dependent on the available adhesion force, which in turn depends on the slipperiness of the track (rainfall, leaves). The adhesion force can be enlarged through measures at the traction vehicle (constant driving, anti-wheel slip system, and sand sprinkling devices).

In descending, the train should be capable to:

- Maintain a certain speed without thermal overburdening of the braking system;
- Come to a standstill within a certain distance (on very steep gradients the vehicle's safety brakes are required besides the standard brakes).

Originally railways have been built in mountainous areas under accurate consideration of the mass balance, equilibrium between cutting, and embankment with minimisation of the transport distance. Moreover, one strove for minimal profiles with tunnels, cuttings, and embankments. Nowadays minimally 6 m is necessary for a single track to obtain a proper ballast bed. At a number of locations the original track is being shortened and diverted. This are most wanted in places with an unfavourable alignment (detours, leakage's, avalanche danger) and where the costs of inspection, maintenance, and renewal are too high.

Concerning the layout design in mountainous areas, important differences in height are concentrated on as well as the easiest way to deal with the light gradients. The railway line is mostly built in (river) valleys to achieve easy access to the villages as well and to establish a connection with possible traffic over water. As a consequence, many curves are necessary in the alignment and there are risks with respect to floods and landslides.
An alternative strategy to cross a mountain range is to follow a mountain ridge at right angles to the ridge. Then simple gradients can be applied with minimal excavation.

Tunnels are expensive constructions in the alignment. Moreover, there are special risks involved in construction and operation: geologic surprises may be encountered. Nevertheless, tunnels are very attractive: a shorter route may be realised with less curves compared to an alignment where the field is followed accurately. And there is no hindrance of snow or falling rock.

Railway tunnels should have gradients of at least 3 ‰ with respect to the highest point in the middle of the tunnel. If the deepest point lies in the middle of the tunnel, provisions must be made to collect the water. If there is a big height difference between the two tunnel mouths, the risk exists of a chimney effect which may cause frost problems. In tunnels provisions should be installed to carry the mountain pressure, to secure loose stones, and to enable drainage.

There are several solutions to overcome big height differences:

- Detours;
- Terminal loops (inside or outside the tunnel);
- Pushing up humps;
- Zigzag track;
- Rack railway;
- Cable railway.

An example of the application of helical tunnels is the Flåmsbane in Norway. A level difference of 200 m was overcome in a straight line with a length of 1500 m using helical tunnels (with a minimal curve radius of 150 m!), in addition to which the track was lengthened by 4 km (steepest gradient 55 ‰).

In order to overcome local high level differences so-called push up humps can be applied. An extra locomotive is added to the train on the spot (before or after the train). The advantage is a saving in construction costs. A drawback is that extra locomotives and infrastructure are necessary (storage facilities for locomotives), which are then used ineffectively. Moreover, time-loss occurs when coupling and de-coupling.

Zigzag track, rack, and cable railways are not suitable for an efficient operation of a main line, because there is too much loss of time.

3.9 Computer-aided track design

During the nineties, manual track design has to a large extent been replaced by a sophisticated computer-aided track design. Various software available on the market like MXRAIL, INRAIL, AutoCIVIL, etc. help track designers with most of the complex problems they face while designing, taking away the tedious activities and allowing them to concentrate on the more creative parts. The scope and level of sophistication of these software products changes so often as their possibilities are constantly expanded, that it makes them almost impossible to categorize. However, it can be said that while initially most of the software packages were meant for general use, i.e. dealing with both road and rail design, recently the software has become more specialized and this has resulted in different modules being developed specifically for road, railway, site development, renewals, surveying, groundwork, etc.

The possibilities of these software packages include the design of three-dimensional alignments representing plain line track, simple turnouts, diamond crossings, crossovers, and complex multi-lead layouts. They can be used for improvements of track alignment, railway stations, and goods yards as well as light rail, high-speed rail, and freight rail design.

They provide utilities for coordinate geometry (COGO), site and track design, rail manufacturing, reporting, viewing, and annotation.

They enable the user to produce curvature and difference diagrams, calculate vertical and horizontal alignments, meet critical site constraints, create customized turnouts and crossings and store them in a library, and generate fabrication data. The user is even provided with options for multilayer digital terrain modeling, subgrade design, earthworks computations, and drawing production. These software packages usually provide all the necessary tools for each phase of track design, from conceptual through final design, and ensure workflow integration with operations and maintenance.

The geometric position of the track and its components can be defined through the use of survey crews, aerial surveys, and graphic elements. The user can vary the format of the input fields when loading point and element data. Once the data is loaded, it can be evaluated and sorted to properly identify the track design which is to be optimized (see Figure 3.11).

Often this software is also capable of providing a regression analysis which makes it easy to design under highly constrained conditions. The use of regression analysis is particularly beneficial in realignment projects. This analysis results in best-fit straight, spiral, and circular track elements. The best-fit alignment can be modified to derive a best "redesign alignment." Offsets or throws from the existing survey data are then available for track repositioning.

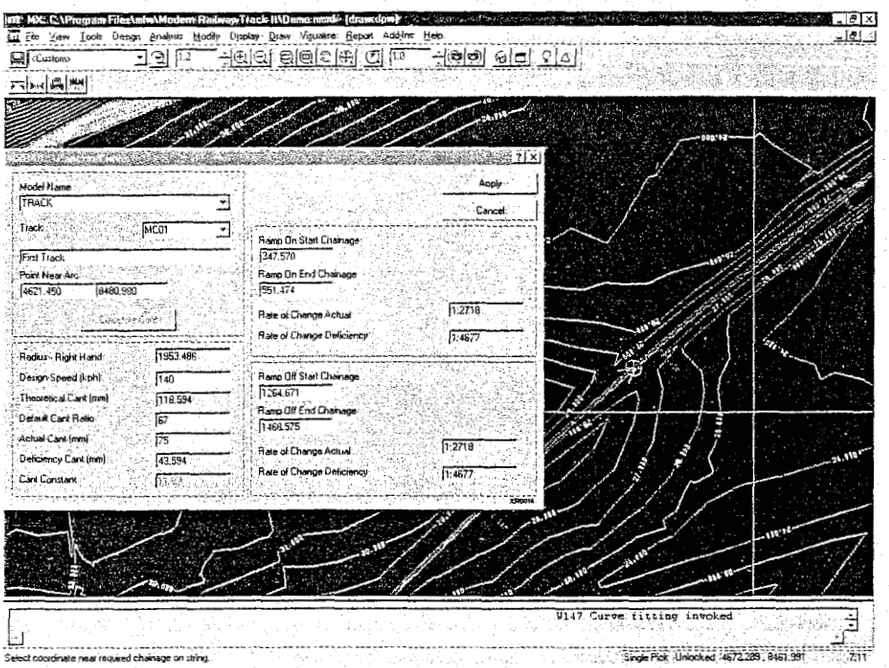


Figure 3.11: Designing superelevation using MXRail

When laying out horizontal geometry, usually several methods are available. The most basic method is to use COGO points to define an alignment. The other option is to use graphical placement of points of intersections (PIs) and curve sets, component placement, or track elements.

Using these systems, the user can also create slew (or difference) diagrams to help verify re-alignments. Slew diagrams represent the geometry of the elements that compose an alignment, such as straights, circular arcs, and spirals, and differences between the elements and survey points. The differences are juxtaposed as offsets that readily indicate the validity of the proposed geometry. The user can display slew diagrams graphically in the design file or display the results in a dialogue box.

In most of these software packages, the ability to freely place track elements is a key aspect of track design. Straights, circular arcs, and spirals which are the horizontal elements for building 3D track alignment can be defined, copied, and modified with precision. Often they provide an interactive environment for placing single elements or chains of elements. The tools for resolving tangency and coincidence between elements and for refining geometry are flexible, allowing the track designer to apply as many constraints as needed. Usually, track elements can be freely placed and then connected.

Perhaps the most important benefit resulting from using these tools is the capability to resolve complex 3D alignment geometry. A powerful track editor complements traditional alignment functions, such as component placement and traverse computations, by means of specialised techniques that automatically resolve the geometry of intermediate track elements. When using the track alignment editor, solutions can be constrained as much as needed for design conditions. Computed elements are usually displayed graphically, while providing a numeric description in a dialogue box. These techniques can also be used for vertical alignment design. In addition to the traditional PI design with parabolic curves, an ability to float straights and arcs is provided to the user. Once the vertical geometry is resolved, a single command allows the user to vertically lift the entire alignment, thereby increasing design productivity.

Efficient cross section generation and editing possibilities are also available. Cross sections can be extracted at any location along the design surface. Some software packages allow multiple cross sections to be cut along an alignment. All active surfaces and annotations can be displayed in cross sections, which can be reviewed individually or as part of a timed sequence.

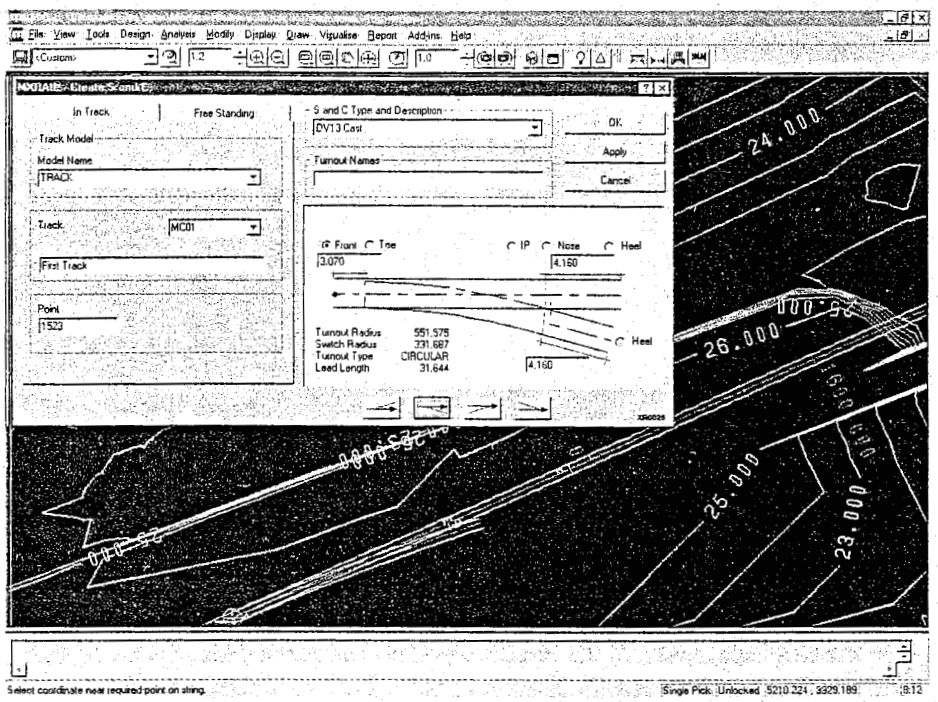


Figure 3.12: Designing switch using MXRail

The importance of turnouts to track design has also often been recognised. Therefore, these tools often enable the user to define geometry for standard turnouts and then store the turnouts in a library. Turnout types include single and double branch, single and double slip, and crossings.

The turnout feature includes a parametric definition of a basic turnout (overall length and distance from the theoretical PI to the end of the turnout) plus a geometric definition of the branch line from the point of switch to the end of the turnout. Once defined, turnouts may be placed on other track elements and they will adjust accordingly (see Figure 3.12).

Because turnouts are track elements, they can be designed to float or remain at fixed locations. For example, when locating a lead to serve an industry, the user can allow the turnout to float in order to determine the best location. Geometry for the lead is usually constrained at the industry site and there are other design constraints, such as railroad standards for industrial tracks. The user can establish geometry at the site, then place a turnout on the mainline. Using the track alignment editor, you can allow the turnout to float on the mainline, fix the radii of the curves, and let their lengths vary. Software then resolves the geometry and locates the turnout at the appropriate position. If the position is undesirable, you can reject the fit and alter the geometry or turnout.

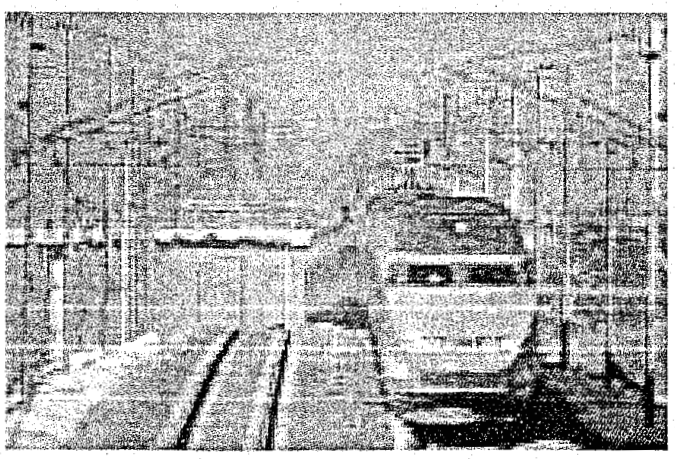


Figure 3.13: TGV train

3.10 PASCOM - software to estimate passenger comfort

As velocity of trains increases, requirements regarding the track geometry become very important when designing railway track. Currently used codes for track geometry are based on quasi-static models of a moving train and do not take the dynamic effects that play an important role in the train's behaviour into account, especially if the issue of high speed operations occurs. That is why a three-dimensional model of a track intended for the analysis of the dynamic behaviour of trains and passenger comfort has been developed at Delft University of Technology and implemented in the software called PASCOM. The model has been tested on several study cases and finally applied to the analysis of a high-speed track (Figure 3.13). The numerical results have shown that the dynamic behaviour of the vehicle depends not only on the cant deficiency/excess, but also on the geometry of super-elevation ramps and the type of transition curves. The model of the track and train as well as some illustrative examples are presented below.

3.10.1 Numerical model

PASCOM uses the 3-D mass-spring model of a train that is shown in Figure 3.14. The model consists of 3 masses representing the car body and two bogies, and spring - damper combinations related to the primary and secondary suspension systems. The dynamic behaviour of the model is analysed in the frequency domain. The interaction between the vehicle and track is described by the Frequency Response Functions (FRF) which, in general, characterise the relation between the input and output of a system. Here the input consists of the wheel displacements and non-compensated centrifugal forces acting on the masses, and the output consists of the displacements and rotations of the car body.

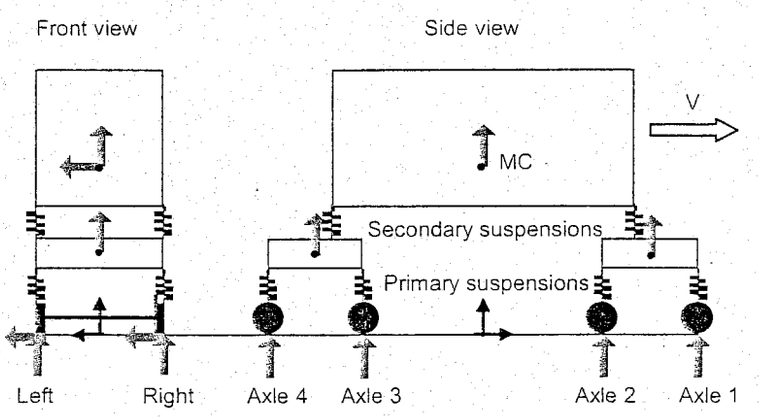


Figure 3.14: Vehicle model of PASCOM

To calculate the input to the vehicle system, the horizontal and vertical track geometry is represented by the curvature as a function of the distance. The position of the centre line of a track is described using the horizontal and vertical geometry representation. To describe the track plane, the angle of cant is also to be represented as a function of the distance. Based on the train speed and dimensions, the input to the model, i.e. the variation in height (as a result of cant) and non-compensated centrifugal forces (as a result of curvature), can be expressed as a function of time. Using the Fourier Transformation the input can be transferred to the frequency domain as follows:

$$\tilde{X}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt \qquad (3.32)$$

Figure 3.15 shows the response function of an ICE train moving with the speed of 300 km/h. The frequency of the car body in the vertical direction is approximately 0.6 Hz. The relation between train speed (v), wavelength (λ), and frequency makes it obvious that the vibrations resulting from track geometry irregularities which have a wavelength of 80-160 m, have a negative effect on the passengers' comfort on an ICE train moving at 200-300 km/h.

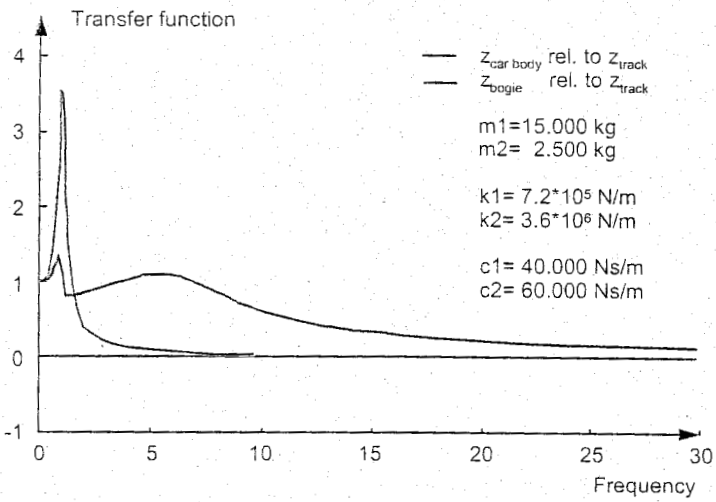


Figure 3.15: Response function of train (ICE) moving at 300 km/h

The output of the model is obtained by multiplying the spectrum of the input by the response function. The displacements and rotations of a car body as well as the acceleration required to estimate the comfort are then calculated in a time domain using the reverse Fourier Transformation. More information on the dynamic analysis in the frequency domain are discussed in Chapter 6. Two illustrative cases using the PASCOM model are described below.

3.10.2 Case 1: Investigation of dynamic effects

In order to demonstrate the effect of dynamic effects, the behaviour of two vehicles with and without suspensions moving on a curve track (Figure 3.16a) has been analysed. The vehicle without suspensions (or with infinitely stiff suspensions) represents the quasi-static model. The vehicle with suspensions is one of the ICE trains. Here, the theoretical cant has been used, i.e. the one when the passenger is not affected by the lateral accelerations. Clothoids have been used for both transition curves and super-elevation ramps resulting in a linear variation of cant. The results of simulation as function of the distance (s) along the curve are shown in Figure 3.17 and Figure 3.18. Figure 3.17 clearly shows that the behaviour of a rigid vehicle is completely determined by the geometry of a track. The lateral displacement and rotation of a car body can be derived from the angle of cant (Figure 3.16b). According to this model of a vehicle, a passenger can only feel the accelerations in the beginning and in the end of the transitional curves corresponding to the peaks in Figure 3.17c

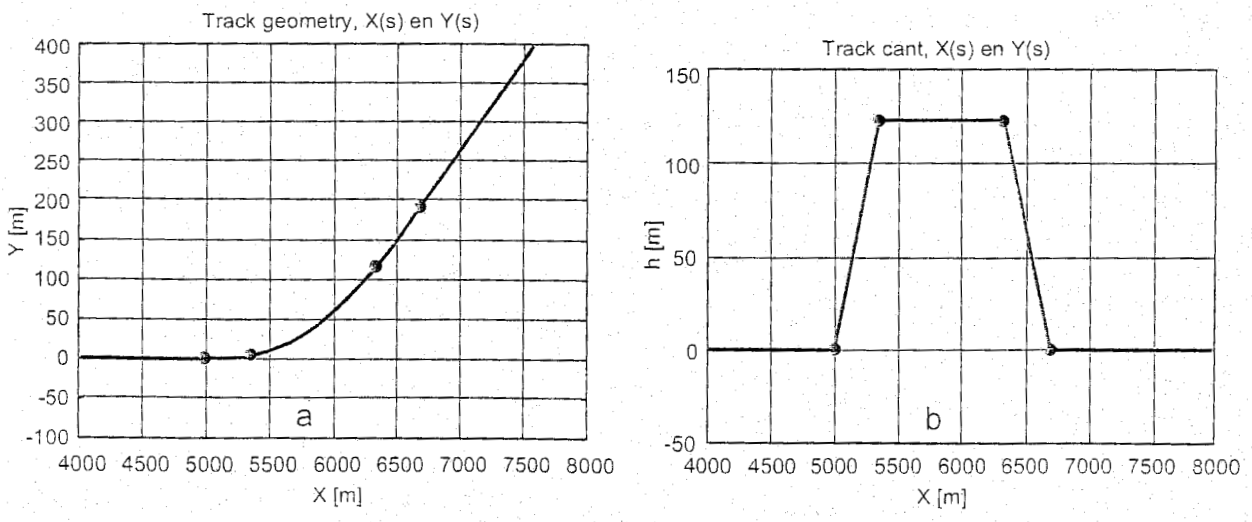


Figure 3.16: Track geometry: horizontal plane (a) and cant (b)

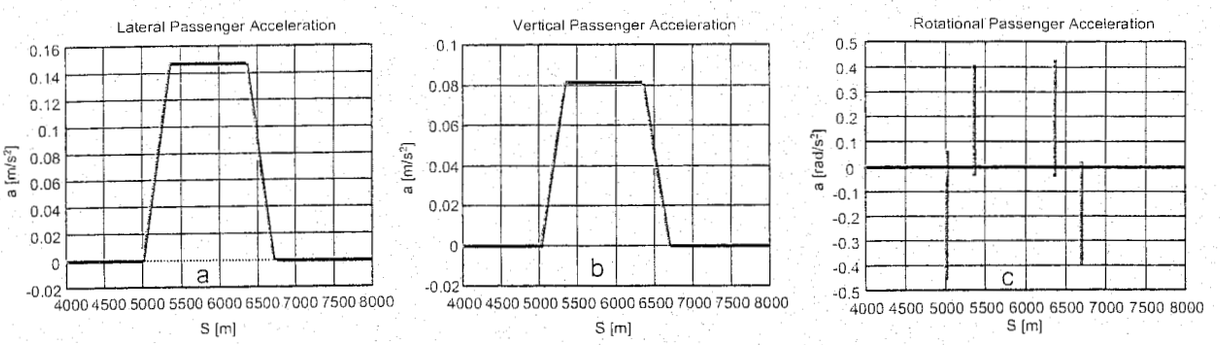


Figure 3.17: Results of simulation with 'rigid' train

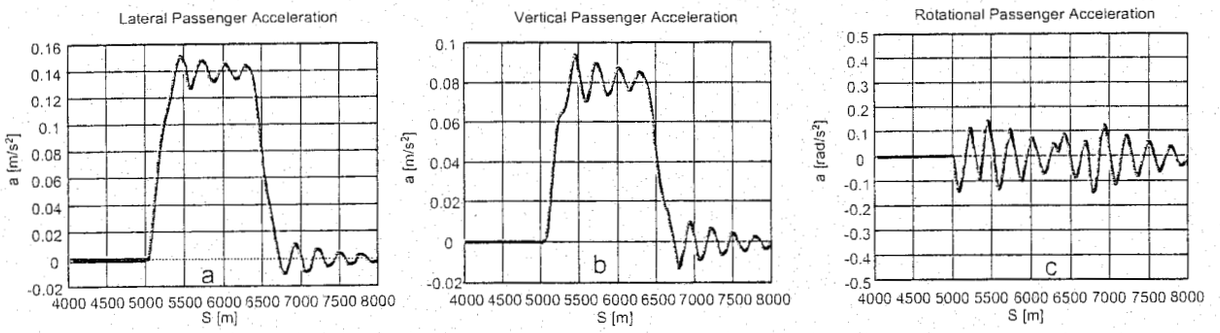


Figure 3.18: Results of simulation with ICE train

More realistic results are obtained using the model with suspensions (Figure 3.18). Since the theoretical cant has been used, the rotation of the coach and the track should be the same. However, due to inertia properties of the vehicle model and presence of suspensions, the coach starts to rotationally oscillate. The oscillations are damped out slowly and therefore a passenger experiences discomfort in the form of lateral, vertical, and rotational accelerations.

3.10.3 Case 2: Track HSL-Zuid (NL)

The PASCOM software has been used to investigate part of a real track which is part of the high-speed line HSL-Zuid (NL). To estimate the quality of the track geometry, an ICE train travelling at 300 km/h has been used. Figure 3.19 shows the resulting accelerations of a coach weighted using the ISO-2631 code [132] which shows that the level of comfort is within the prescribed limits $a<0.315$ m/s². This figure indicates that geometrical properties of a track, such as jumps in the track curvature and gradients of cant on super-elevation ramps, strongly affect the vehicle dynamics and ultimately the passenger's comfort. More information on the numerical model implemented in PASCOM as well as on estimating passenger comfort can be found in [280].

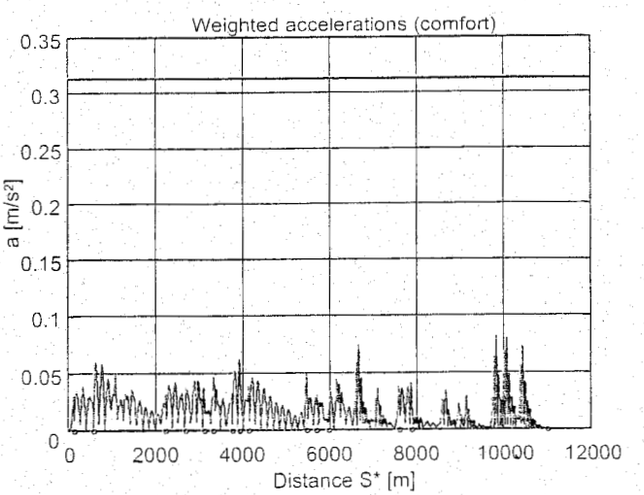


Figure 3.19: Accelerations of coach body travelling at 300 km/h weighted using ISO code. Track: part of high-speed line HSL-Zuid (NL), vehicle - ICE train.

4 TRACK LOADS

4.1 In general

The requirements for the bearing strength and quality of the track depend to a large extent on the load parameters:

- axle load: static vertical load per axle;
- tonnage borne: sum of the axle loads;
- running speed.

The static axle load level, to which the dynamic increment is added, in principle determines the required strength of the track. The accumulated tonnage is a measure that determines the deterioration of the track quality and as such provides an indication of when maintenance and renewal are necessary. The dynamic load component which depends on speed and horizontal and vertical track geometry also plays an essential part here.

4.2 Axle loads

The nominal axle loads applied to the track are as shown in Table 4.1.

	number of axles	empty	loaded
trams	4	50 kN	70 kN
light-rail	4	80 kN	100 kN
passenger coach	4	100 kN	120 kN
passenger motor coach	4	150 kN	170 kN
locomotive	4 or 6	215 kN	--
freight wagon	2	120 kN	225 kN
heavy haul (USA, Australia)	2	120 kN	250-350 kN

Table 4.1: Number of axles and weight per axle of several rolling stock types

With very high axle loads the number of rail defects increases considerably and the track requires far more maintenance.

4.3 Line classification

The UIC (International Union of Railways), which is the organisation for railway cooperation and which counts standardization among its tasks, makes a distinction between load categories according to UIC leaflet 700; these categories are shown in Table 4.2.

The NS network, in general, complies with category C_2; when building new lines C_4 or D_4 are applied.

Category	axle load [kN]	weight/m [kN/m]
A	160	48
B_1	180	50
B_2	180	64
C_2	200	64
C_3	200	72
C_4	200	80
D_4	225	80

Table 4.2: UIC load classification

4.4 Tonnages

Daily tonnage is used to express the intensity or capacity of rail traffic on a specific line. Figure 4.1 gives a summary of the daily tonnages on the NS network. The average daily tonnage is about 20.000 t. The most heavily loaded sections have a daily load of 60.000 t. Abroad, on what are known as heavy haul lines, daily tonnages of 300.000 t can occur.

Furthermore, all types of track deterioration features, such as increase in geometrical deviations, increase in rail fractures, and rail wear, can be very well expressed as a function of tonnage. This is often expressed as MGT = million gross tonnes (note: 1 MGT (US) = 8896 MN).

For the sake of dimensioning and maintenance of the permanent way, the track network is divided into classes determined by the equivalent tonnage defined in UIC leaflet 714 according to:

$$T_f = T_p \frac{V}{100} + T_g \frac{P_c}{18D} \qquad (4.1)$$

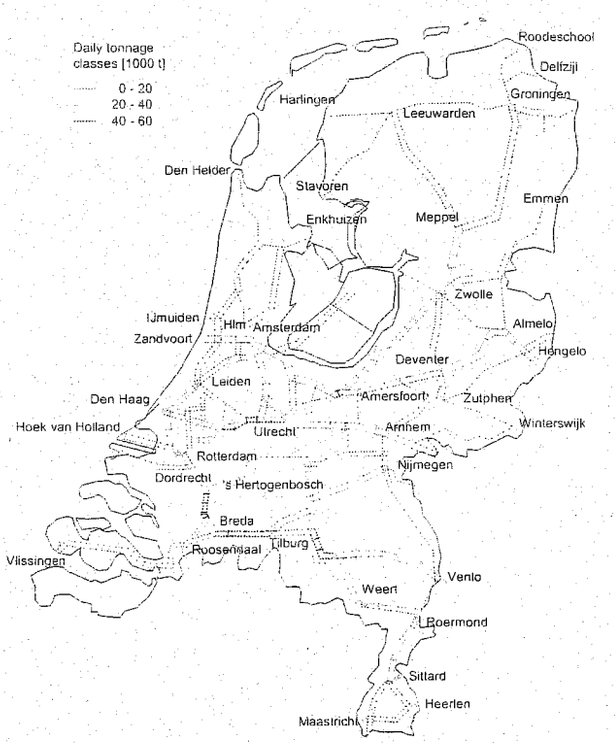


Figure 4.1: Daily tonnage map for NS network

in which:
T_p : Real load for daily passenger traffic;
T_g : Real load for daily freight traffic;
V : Maximum permissible speed [km/h];
D : Minimum wheel diameter [m];
P_c : Maximum axle load with wheels of diameter D [tonnes].

The groups used by the NS are globally speaking as follows:

class I 40.000 < T_f
class II 20.000 < T_f < 40.000
class III 10.000 < T_f < 20.000
class IV T_f < 10.000

4.5 Speeds

The maximum speed on a specific section is expressed in km/h. Line section speeds used by the NS are standardized at 40, 60, 80, 100, 130 and 140 km/h. Table 4.3 gives a survey.

	passenger trains	freight trains
branch lines	--	30-40 km/h
secondary lines	80-120 km/h	60-80 km/h
main lines	160-200 km/h	100-120 km/h
high speed lines*	250-300 km/h	
world record = 515.3 km/h (TGV-SNCF, May 1990)		

Table 4.3: Maximum speeds railways

Freight trains are allowed to run at a maximum of 100 km/h. Tests are currently under way to determine whether this speed can be increased to 120 km/h.

4.6 Causes and nature of track loads.

The forces acting on the track as a result of train loads are considerable and sudden and are characterized by rapid fluctuations. The loads can be considered from three main angles:

- vertical;
- horizontal, transverse to the track;
- horizontal, parallel to the track;

Generally, the loads are unevenly distributed over the two rails and are often difficult to quantify. Depending on the nature of the loads they can be divided as follows:

- quasi-static loads as a result of the gross tare, the centrifugal force and the centering force in curves and switches, and cross winds;
- dynamic loads caused by:
  - track irregularities (horizontal and vertical) and irregular track stiffness due to variable characteristics and settlement of ballast bed and formation;
  - discontinuities at welds, joints, switches etc.;
  - irregular rail running surface (corrugations);
  - vehicle defects such as wheel flats, natural vibrations, hunting.

In addition, the effects of temperature on CWR track can cause considerable longitudinal tensile and compressive forces, which in the latter case can result in instability (risk of buckling) of the track.

4.7 Vertical rail forces.

4.7.1 Total vertical wheel load

The total vertical wheel load on the rail is made up of the following components:

$$Q_{tot} = \underbrace{(Q_{stat} + Q_{centr} + Q_{wind})}_{\text{quasi-static forces}} + Q_{dyn} \qquad (4.2)$$

in which:

Q_{stat} : static wheel load = half the static axle load, measured on straight horizontal track;
Q_{centr} : increase in wheel load on the outer rail in curves in connection with non-compensated centrifugal force;
Q_{wind} : idem for cross winds;
Q_{dyn} : dynamic wheel load components resulting from:

- sprung mass 0 - 20 Hz;
- unsprung mass 20 -125 Hz;
- corrugations, welds, wheel flats 0 - 2000 Hz.

From the equilibrium consideration of the forces acting on the vehicle, as shown in Figure 4.2, the following can be deduced for each wheelset and a small cant angle:

$$Q_{centr} + Q_{wind} = G\frac{p_c h_d}{s^2} + H_w\frac{p_w}{s} \qquad (4.3)$$

with (3.4) it follows:

$$h_d = \frac{sv^2}{gR} - h \quad \text{(cant deficiency)} \qquad (4.4)$$

wherein:

G : weight of vehicle per wheelset;
H_w : cross wind force;
s : track width;
v : speed;
g : acceleration due to gravity;
R : curve radius;
h : cant;
p_c : vehicle centre of gravity distance;
p_w : distance of lateral wind force resultant.

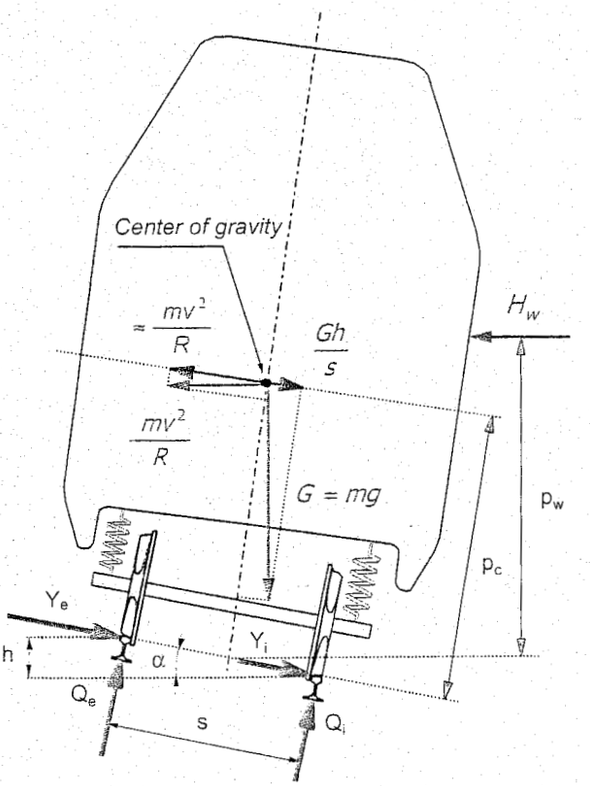


Figure 4.2: Quasi-static vehicle forces in a curve

The proportion of Q_{centr} is usually 10 to 25% of the static wheel load. With cross wind on the other side H_w is negative, which is important for vehicles standing still in curves.

The maximum wheel load usually occurs at the outer rail ($h_d > 0$) and is equal to:

$$Q_{e\,max} \approx ½\,G + G\frac{p_c h_d}{s^2} + H_w\frac{p_w}{s} \qquad (4.5)$$

In view of the large number of load repetitions, the dynamic wheel load can be considered as a fatigue load. This means that it is not the greatest wheel load which is the determining factor for the bearing strength, but the mean square of all the wheel loads. For a rough estimate of the dynamic wheel load, the static wheel load may be multiplied with a dynamic amplification factor.

About 2 m in front of and behind a wheel the track bends in an upward direction. This means that upwardly directed forces are also acting on the fastenings and sleepers.

4.7.2 Tilting risk.

Tilting of the vehicle in curves can occur in the following instances:

– over the outer rail at high speeds in case of a great cant deficiency and cross wind; due to off-loading of the inner rail;

– over the inside rail in the case of lightweight empty wagons standing still or wagons with a displaced freight load on track with high cant and with a cross wind in an unfavourable direction.

To determine safety against tilting, an equilibrium consideration has to be made taking load and a shift of the centre of gravity due to the suspension into account. A criterion for tilting can be, for example, that load-relieving of the critical wheel by up to 60% of the static value is allowed.
The notion of tilting as a risk should not be confused with tilting as a mechanism used intentionally in coaches to artificially increase the cant in curves.

4.8 Lateral forces on the rail.

4.8.1 Total lateral wheel load

The total horizontal lateral force exerted by the wheel on the outer rail is:

$$Y_{tot} = \underbrace{(Y_{flange} + Y_{centr} + Y_{wind})}_{\text{quasi-static forces}} + Y_{dyn} \qquad (4.6)$$

in which:

Y_{flange}: lateral force in curve caused by flanging against the outer rail;
Y_{centr} : lateral force due to non-compensated centrifugal force;
Y_{wind} : idem for cross wind;
Y_{dyn} : dynamic lateral force component; on straight track these are predominantly hunting phenomena.

If it is assumed that Y_{centr} and Y_{wind} act entirely on the outer rail, the equilibrium consideration per wheelset in Figure 4.2 gives:

$$Y_{e\,max} \approx G\frac{h_d}{s} + H_w \qquad (4.7)$$

The actual situation is considerably more complex because of the presence of several coupled wheelsets, the various positions the vehicle can assume in a curve, and the adhesion forces between the wheel and the rail. The total lateral force cannot therefore be predicted with great reliability. In practice, however, measuring methods have been developed for determining the total vertical and horizontal forces on the rail.
There are also measuring wheelsets which can measure these quantities directly. Currently, a new technique is being used to calculate Y and Q forces from measured geometry. More details are given on this in Chapter 16.

4.8.2 Derailment risk

As stated above, derailment can occur if the Y/Q ratio increases in value because of high lateral forces Y acting on the high rail or low wheel loads Q in the case of unloaded wheels.

In Figure 4.3 the situation is drawn where the forces are acting on the rail and where flange climbing is about to begin. From the equilibrium conditions the normal force N and the tangential force S in the contact area can be expressed as:

$$N = Y\sin\beta + Q\cos\beta \qquad (4.8)$$

$$S = Q\sin\beta - Y\cos\beta \qquad (4.9)$$

where β = flange angle.

Apparently, flange climbing can be prevented or stopped if the shearing force satisfies this relationship:

$$fN \leq S \tag{4.10}$$

where f = friction coefficient.

Introducing the formula:

$$f = \tan\phi \tag{4.11}$$

the insertion of (4.8), (4.9) and (4.11) in (4.10) results in the following condition:

$$\frac{Y}{Q} \leq \frac{\sin\beta - \tan\phi\cos\beta}{\tan\phi\sin\beta + \cos\beta} = \tan(\beta - \phi) \tag{4.12}$$

Tests [203] have shown that derailment can occur if the Y/Q ratio over a distance of more than 2 m is greater than 1.2. For this reason the following value is usually retained as the criterion for safety against derailment:

$$\frac{Y}{Q} < 1.2 \tag{4.13}$$

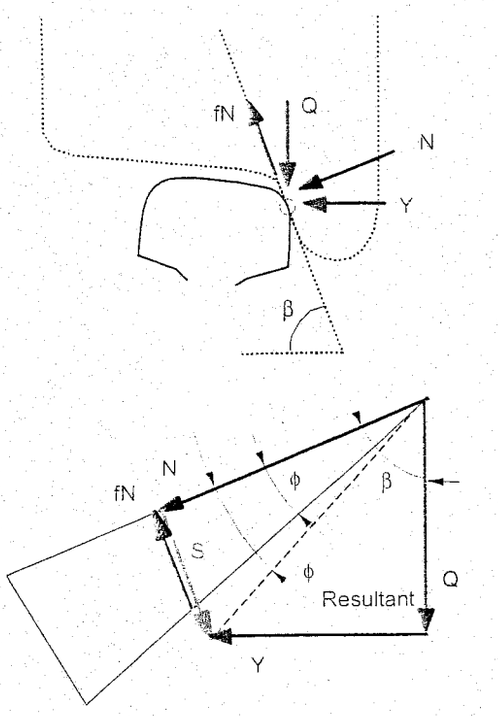


Figure 4.3: Quasi-static vehicle forces in a curve

Often, a maximum wear angle of ½π - β = 32° is accepted, resulting in β = 60°. For instance, with β = 60° and f = 0.15 it follows from (4.12) that Y/Q = 1.26, which value is in accordance with the international accepted value mentioned above.

Apparently, if the wheel flange or railhead is lubricated the situation regarding derailment is considerably improved because there is less friction.

4.8.3 Lateral force on the track

The total lateral force H on the track can be assessed as the sum of the Y-forces multiplied by a dynamic amplification factor:

$$H = DAF \cdot \left(G\frac{h_d}{s} + H_w \right) \tag{4.14}$$

This total lateral force exerted by the wheels on the track must be resisted by means of:

- resistance to lateral displacement of the sleepers in the ballast bed;
- horizontal stiffness of the track frame (5 to 10%).

In the horizontal direction the resistance of the track is limited. High lateral forces can cause the sleepers to move in the ballast bed, possibly causing permanent deformation.

A practical value for the lateral resistance required of a loaded track in order to guarantee stability was determined in the fifties by SNCF and is known as the Prud'homme formula which reads as follows:

$$H_{tr} > 10 + \frac{P}{3} \tag{4.15}$$

in which:

H_{tr} : minimum lateral force [kN] which the track should be able to resist without lateral deformation;
P : axle load [kN].

In general, the empirical coefficients appearing in the formula (here '10' and '1/3') are dependent on the type of track and its maintenance condition. In this case it concerns shovel-packed track, crushed stone, and wooden sleepers. Although measurements in tamped track with concrete sleepers suggested a higher value for the first coefficient, formula (4.15) is generally adopted as design standard. As far as the rolling stock is concerned, it is required that the horizontal wheel load H_{rs} exerted by a vehicle is restricted according to:

$$H_{rs} < 0.85\left(10 + \frac{P}{3}\right) \tag{4.16}$$

When assessing whether the actual lateral load is less than the value allowed, formula (4.16) is used as a comparison in order to be absolutely certain.

4.9 Longitudinal forces.

4.9.1 Causes

Horizontal longitudinal forces occur in the track as a result of:

- temperature forces, especially in CWR track. These forces can be considered as a static load;
- accelerating and braking;
- shrinkage stresses caused by rail welding;
- track creep.

4.9.2 Temperature forces

The change in length of a released rail as a result of a change in temperature is:

$$\Delta l = \alpha \ell \Delta T \tag{4.17}$$

in which:

α : linear expansion coefficient of rail steel;
ΔT: change in temperature (defined as $\Delta T = T_{actual} - T_{initial}$);
ℓ : original rail length.

This situation does not occur in the case of fixed rails, because in such a case the rail encounters a longitudinal resistance which opposes the axial displacements. This resistance is produced by the friction forces between rails and sleepers and between sleepers and ballast bed. In Chapter 7 this subject is discussed in more detail.

4.9.3 Track creep

This phenomenon concerns the gradual displacement in the running direction of either the rails relative to the sleepers or of the rails plus sleepers relative to the ballast bed. On single track with two-way traffic there is less creep. On gradients the track creeps downwards regardless of the traffic direction. Creep has the following disadvantages:

- increase in CWR forces;
- too large or too small expansion gaps in jointed tracks;
- non-uniform rail creep resulting in misalignment of the sleepers due to which horizontal bending moments are exerted on the rail;
- displacement of sleepers resulting in disturbance of the stability of the track in the ballastbed.

The cause of creep is to be found in the bending wave movement of the track close to the wheel. Increasing elongation of the rail foot caused by the approaching wheel load produces a slight forward displacement of that part of the track, because the shearing resistance at this point is less than under the loaded section. As the wheel load moves away, the section of track behind the wheel load shifts forward slightly for the same reasons owing to the decreasing elongation of the rail foot. Creep is eliminated by using fastenings with a sufficient clamping force and ballast with an adequate shearing resistance.

4.9.4 Braking load

When dimensioning structures, braking loads must be taken into account. Many railways assume 25% of the train weight for this axial load. In Chapter 7 this subject is discussed in more detail.

4.10 Influence of higher speeds and increased axle loads

4.10.1 Speed

Higher speeds and higher axle loads play an important role when competing with other forms of transport. It is hardly surprising that a great deal of research is being undertaken in this field the world over. This subject has already been given some consideration in the preceding chapters.

Apart from causing rolling stock problems and difficulties with pantographs, high speeds involve keeping the forces on the track and the car body accelerations at an acceptable level. These variables can be affected by track geometry. Vehicle reactions are a result of interaction with the track. The response variables consist not only of quasi-static components which occur in curves, but also of a dynamic component.

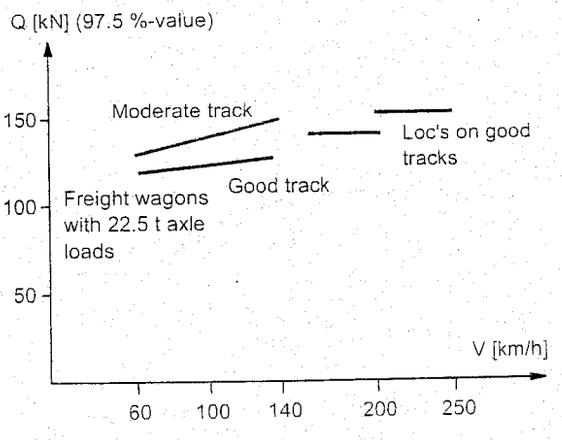


Figure 4.4: Q-forces at high speed

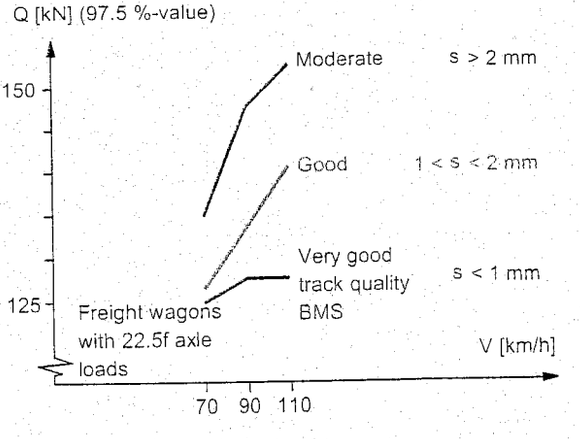


Figure 4.5: Q-forces at 22.5 t axle load

For the various high speed projects extensive series of measurements have been carried out to check whether the limit values are exceeded or not.

A summary of the Q-force measurements carried out by DB [145] is presented in Figure 4.4. The 97.5%-value of the wheel load due to locomotives reaches 150 kN at 250 km/h. Wagons with an axle load of 22.5 t on straight track at 120 km/h also produce similar dynamic loads.

4.10.2 Increase in axle loads

In 1983 a large-scale European measurement campaign was carried out under the auspices of ORE committees D 161 and B 12. A specially formed measuring train was used consisting of wagons with type Y 25 and DB 65 bogies. The NS recording car with the BMS system was also included in the train. Using measuring wheelsets Q and Y, force measurements were carried out for 20 and 22.5 t axle loads on different quality track. These measurements formed part of extensive research into the effect of raising the axle load from 20 to 22.5 t [220].

The dynamic wheel loads for a nominal axle load of 22.5 t for various track qualities expressed in mm standard deviation (s) of BMS are presented in Figure 4.5. The track quality is subdivided into 3 classes: 0 - 1 mm: very good, 1 - 2 mm: good, more than 2 mm: moderate. This classification applies in the following to both level and alignment. Figure 4.5 clearly shows the dominating influence of track quality.
The dynamic axle load increase $\Delta\Sigma Q$ regarding 20 t and 22.5 t at speeds of 70, 90, and 110 km/h is shown in Figure 4.6.

The dynamic wheel loads are also measured in a variety of curves. The 99.85%-values for lateral wheel-rail forces Y_a, under 20 t and 22.5 t, are gathered together in Figure 4.7.

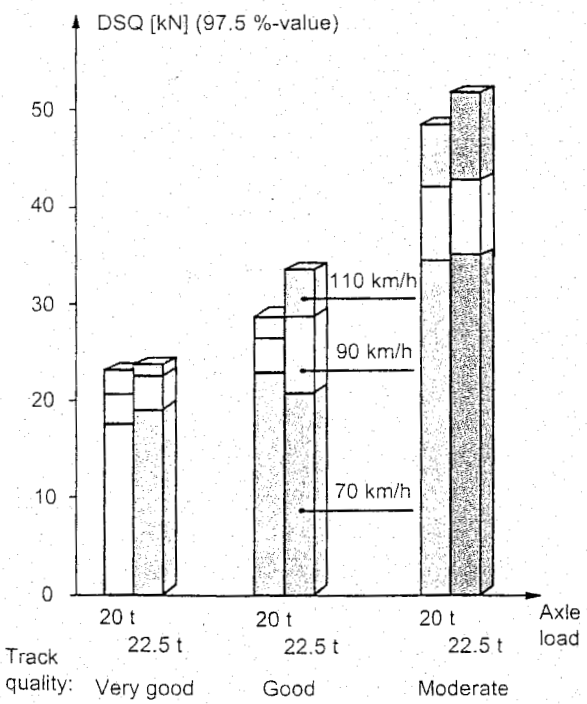


Figure 4.6: Dynamic component of wheel load versus axle load, speed and track quality

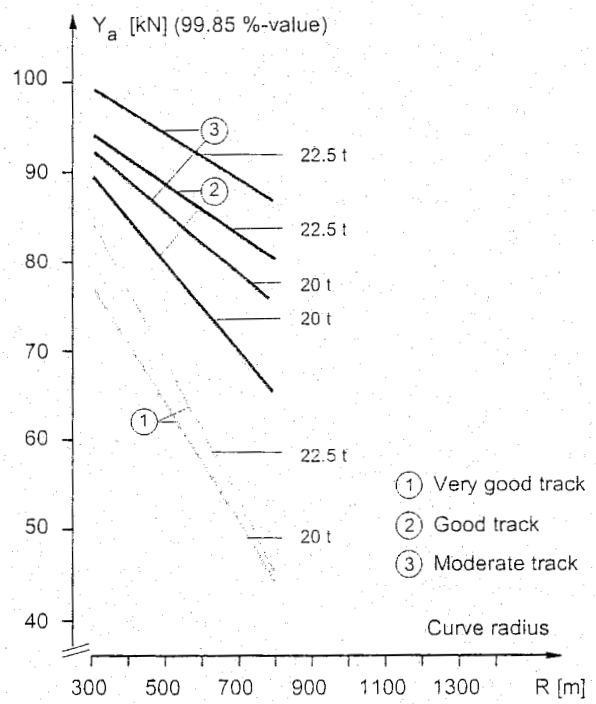


Figure 4.7: Dynamic Y-forces in curves

Stability of the track depends on the total lateral track load, especially that part which is present over a length of at least 2 m. This portion is referred to as ΣY_{2m} and is obtained from the sum of Y-left and Y-right, which implies that contributions with a wavelength of less than 2 m are filtered out.

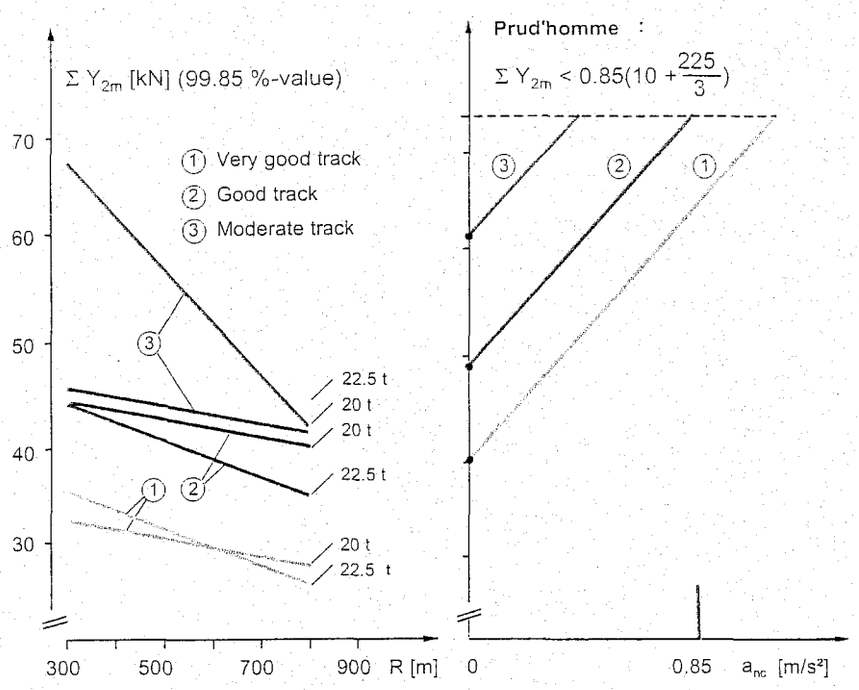


Figure 4.8: Dynamic lateral track load in curves

Figure 4.8 again shows the 99.85%-values for various parameters. An upper limit is also given for lateral load as a function of non-compensated lateral acceleration. A value of 0.85 m/s^2 would also seem to be acceptable for 22.5 t axle loads on very good and good tracks. In the case of moderate tracks the Prud'homme limit is reached sooner.

The DB measurements published in [145] give a maximum Q force of 170 kN. As can be seen in Figure 4.9, these values are reached in curves when non-compensated lateral acceleration is 0.85 m/s^2.

Conclusions of D 161 research on 22.5 t axle loads

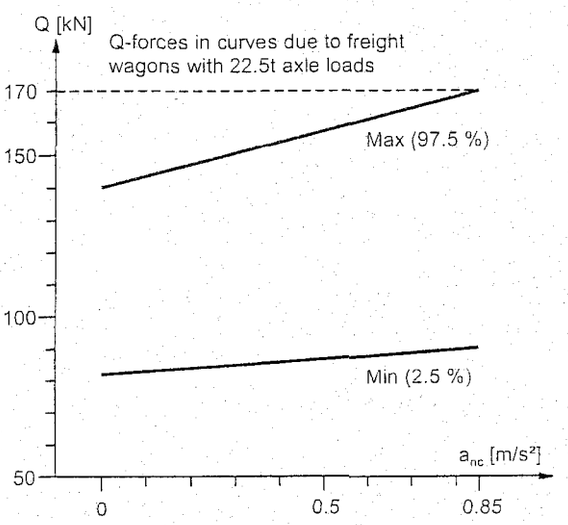


Figure 4.9: Dynamic Q-force in curves measured by ORE D 161/B12

The conclusions of the research undertaken by the ORE D 161 committee into the effects of raising the axle load from 20 t to 22.5 t can be summarized as follows [220].

For most of the parameters examined the increase remained less than the increase of 12.5% in nominal axle load. Accelerations in the track increase by 2 to 10%. The increase in dynamic Q-force component is of the order of 6%. The ballast stress is almost the same as for an axle load of 20 t when NP 46 is used instead of UIC 54. In general the lateral load turns out to be affected only slightly.

The above ORE research revealed that speed and the quality of track geometry were the most important parameters. As can be deduced from figure Figure 4.6, the dynamic component of the Q-force for an axle load of 20 t on poor quality track is considerably greater than that for an axle load of 22.5 t on good track supplemented by the static increase of 12.5 kN.

Expected increase in track maintenance costs

The parameter which is considered decisive in effecting the requirements for tamping, rail renewal, and component damage is the vertical load. The ORE Committee D141 developed a method for quantitative calculations as described in [211].

In this approach damage/deterioration is assumed to be a function of the power of the load according to:

$$E = kT^{\alpha}P^{\beta}V^{\gamma} \tag{4.18}$$

in which:
E = deterioration since renewal or last maintenance operation;
T = tonnage;
P = total axle load (static + dynamic);
V = speed;
k, α, β, γ = constants.

The shape of the function E in relation to tonnage and axle load is represented schematically in Figure 4.10. The equivalent tonnage for a 22.5 t axle load equals:

$$T_{eq\ 22.5} = \left[\frac{P_{20}}{P_{22.5}}\right]^{\beta/\alpha} T_{20} \tag{4.19}$$

$$= \left[\frac{E_{20}}{E_{22.5}}\right]^{1/\alpha} T_{20} \tag{4.20}$$

The costs are inversely proportional to the tonnage borne at the moment maintenance takes place. Consequently, the cost ratio $K_{22.5}/K_{20}$ amounts to:

$$\frac{K_{22.5}}{K_{20}} = \left[\frac{P_{22.5}}{P_{20}}\right]^{\beta/\alpha} \tag{4.21}$$

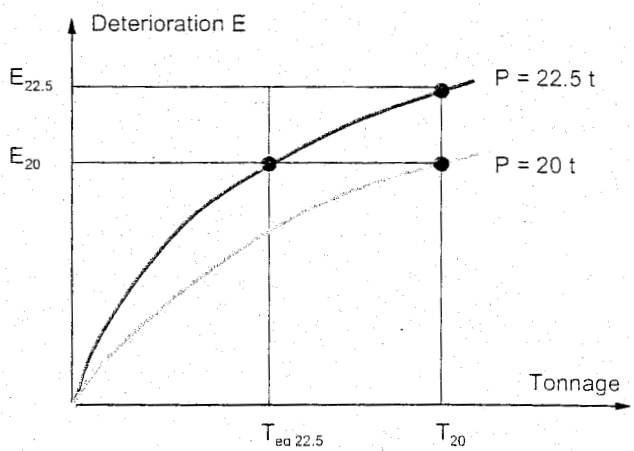


Figure 4.10: Deterioration law according to ORE D 141

In practice only a limited fraction of axles have loads of 20/22.5 t. As a result of this, the cost ratio for two-axle load distributions should be calculated by:

$$\frac{K_2}{K_1} = \frac{\sum_i n_{2i} P_i^{\beta/\alpha}}{\sum_i n_{1i} P_i^{\beta/\alpha}} \tag{4.22}$$

in which:
n_{2i}, n_{1i} = fraction per axle load class;
P_i = average load level per axle load class;

The α and β factors have been determined empirically by ORE D 141 and ORE D 17. In [220] the following factors have been used.

Phenomenon	α	β	β/α
rail fatigue	3	3	1
rail surface defects	1	3.5	3.5
fatigue of other components	3	3	1
track geometry deterioration	1	3	3

Table 4.4: Values of deterioration factors

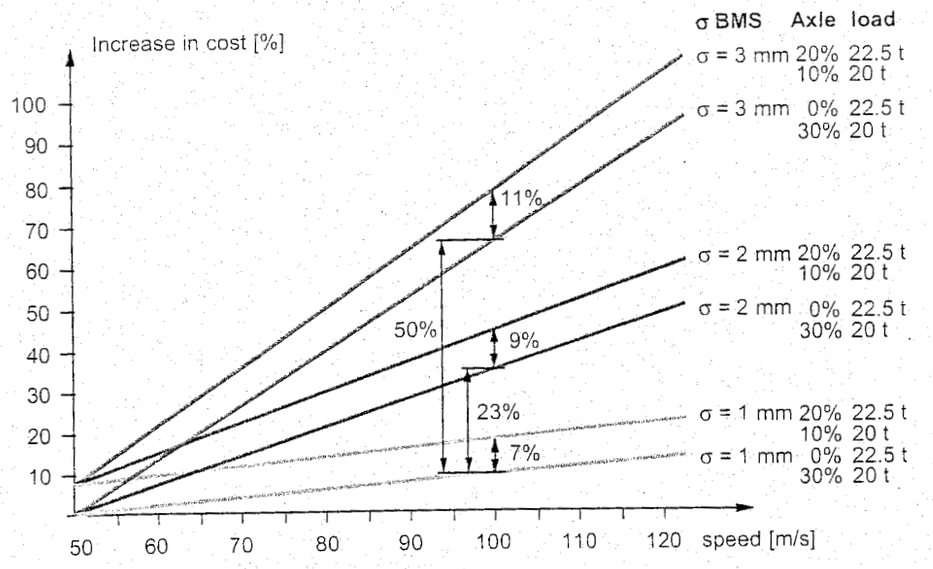


Figure 4.11: Increase in maintenance costs in relation to track quality, axle load, and speed

In [220] calculations of cost increase have been made for several axle load distributions. One of these examples is presented in Figure 4.11. The reference distribution consists of 30% axles of 20 t. The cost increase is computed if the 30% of 20 t axle loads is converted into 10% of 20 t and 20% of 22.5 t axle loads, taking track quality and running speed into account.

Track quality appears to be the most important factor. For a track quality between 1 and 2 mm standard deviation BMS, the order of magnitude of the cost increase ΔK according to Figure 4.11 can be summarized as follows:

σ_{BMS}	= 1 mm	→	σ_{BMS}	= 2 mm	→ ΔK = 23 %
V	= 80 km/h	→	V	= 100 km/h	→ ΔK = 15 %
P	= 20 t	→	P	= 22.5 t	→ ΔK = 8 %

International approval of 22.5 t axle loads at 100 km/h

For the purpose of requiring international approval of bogie wagons with 22.5 t axle loads and speeds of up to 100 km/h, further analyses were carried out by the UIC on the Europe Tour measurements and approval measurements were undertaken by DB an SNCF. Based on this, the conclusion was drawn that for the boundary conditions applied here, i.e. R > 450 m and maximum cant deficiency of 130 mm (0.9 m/s^2), a speed of 100 km/h will produce no problems provided σ_{BMS}-alignment remains in the 0 - 25 m band below 2 mm. Under these circumstances the lateral load lies at least 20% below the Prud'homme limit.

The running characteristics of the rolling stock are characterized by the W_z-factor according to Sperling. The reader is referred to Chapter 16 for the definition of this factor. The permissible value is 4.5, whereas in the measurements previously mentioned W_z values of the order of 3.0 were found.

4.11 Wheel flats

By far the largest dynamic loads applied to the track by vehicles are those which arise from irregularities on the wheel such as wheel flats. Normal dynamic forces are linked to the track geometry.
A location with bad track geometry will experience high forces from most vehicles. These high forces will contribute to track damage and geometry deterioration at that location. Wheel flat forces differ fundamentally from those discussed previously.

Wheel flats apply the same force randomly all along the track with no correlation to the track geometry. However, the occasional high load applied to the rail or component can cause failure and can significantly contribute to fatigue damage. Figure 4.12 shows a characteristic example of the force between rail and sleeper during passage of a wheel with a flat. In extreme situations these forces can increase to six times the static value.

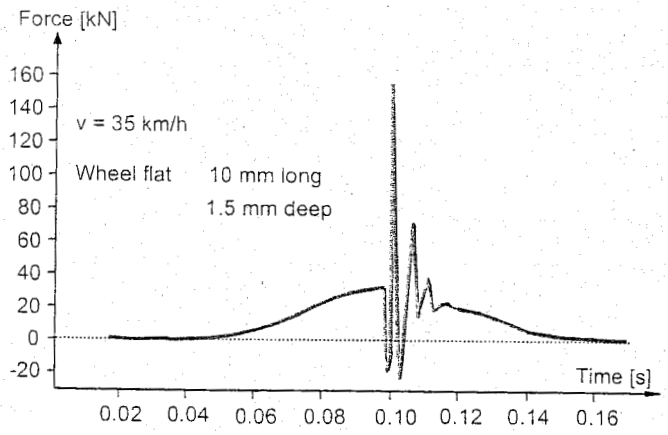


Figure 4.12: Force between rail and sleeper during passage of a wheel with a flat

During the D161 studies described in [219], a number of calculations were made with a model developed by BR to investigate the effect of different parameters on the magnitude of wheel flat forces [273]. Figure 4.13 shows some typical examples of calculated wheel flat forces.

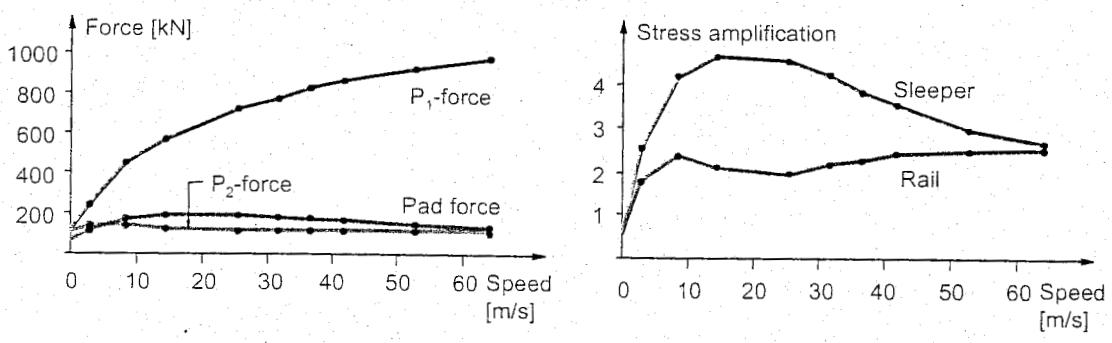


Figure 4.13: Typical examples

These results reveal that the forces at frequencies above 500 Hz, referred to as P_1-forces, increase continuously with speed, whilst forces at frequencies below 100 Hz, referred to as P_2-forces, are more or less independent of speed. The P_1-forces, associated with the stiff Hertzian contact spring, are important as far as wheel/rail contact stresses are concerned. If rail and sleeper strains are considered, the medium frequencies associated with the pad stiffness and the low frequencies associated with the ballast spring are important.

The calculation results are confirmed by the test measurements carried out by ORE D 161 [219]. Figure 4.14 summarizes these test results by plotting the wheel load in the 0 - 100 Hz band as a function of the flat depth for tracks with timber and concrete sleepers. The largest forces were found at a speed of 30 km/h. According to UIC leaflet 510-2, flats on wheels with a diameter of 1000 - 630 mm should be restricted to a length of 60 mm and a depth of 0.9 - 1.4 mm. In this range the dynamic wheel load increases by approximately 30 kN/mm for timber sleepers and 50 kN/mm for concrete sleepers.

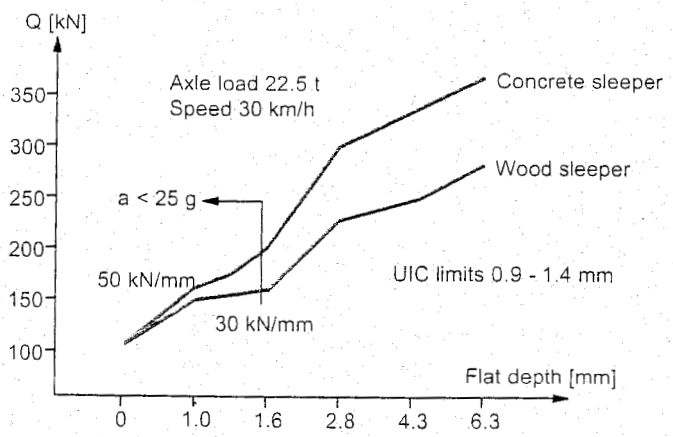


Figure 4.14: Summary of D161 investigations into the effect of wheel flats on the wheel load

However, the higher frequency forces rise continuously with speed and very large dynamic railhead forces can be recorded. The results of measurements on BR and the American Railroads have shown that there may be a chance of more than 1 in 10.000 of obtaining wheel forces as high as 300 kN.

4.12 Forces due to bad welds

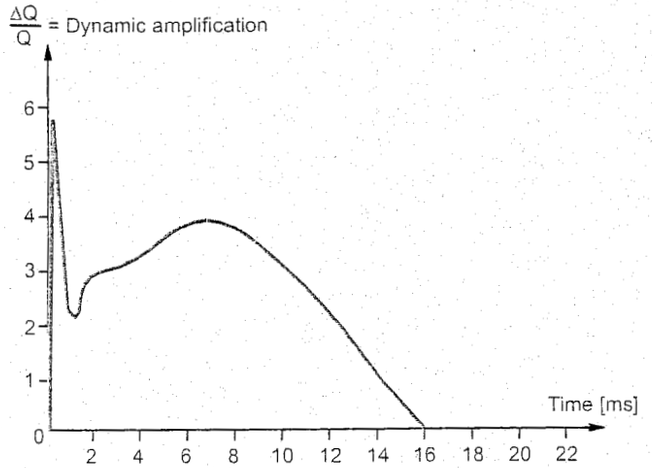


Figure 4.15: Dynamic wheel load during passage over a

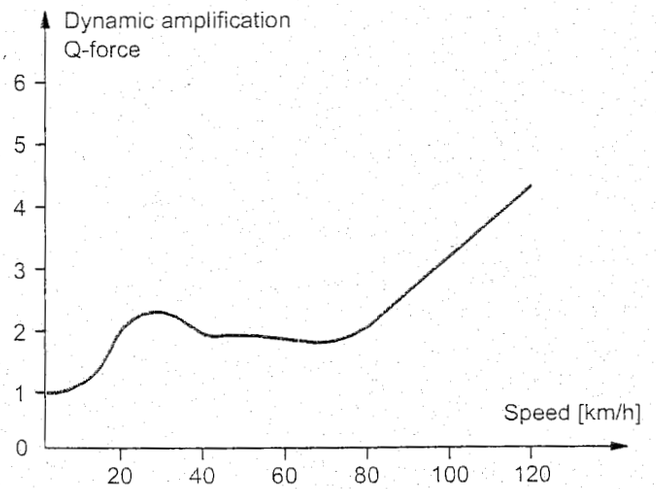


Figure 4.16: Dynamic Q-force magnification versus speed during passage over a weld

Even relatively small vertical deviations in weld geometry can cause huge dynamic forces on passage of a wheel. Steps, i.e. misalignments, are especially responsible for this, but so are kinks, normally expressed as the versine on a 1.2 m base. In order to confine these dynamic loads, geometrical deviations in welds should be limited to a few tenths of a millimetre.

Figure 4.15 shows an example of load distribution as a function of time during wheel passage over a poor weld. The behaviour is similar to that of wheel flats. After a sharp peak of some milliseconds, i.e. the P_1-force mentioned earlier which only has a local influence on the wheel/rail contact stress, a second, much broader, peak occurs which penetrates the whole construction.

The dynamic amplification of the Q-force during passage over a poor weld is presented in Figure Figure 4.16 as a function of speed. These values were obtained from BR measurements and calculations. Dynamic amplifications of the vertical rail load of up to 400% have been found on the NS network.

High frequency dynamic loads due to poor welds, corrugation, and wheel flats are very detrimental to the track. Concrete sleepers in particular are very susceptible to these loads. Delft University investigated load transfer from rail to sleeper for ORE Committee D 161, with special emphasis on the properties of rail pads [219].

Figure 4.17 shows sleeper strain versus time due to an impact load for normal and soft pads. The strain peak in the case of soft pads is about 50% lower than for normal pads.

Transfer functions have also been determined between vertical rail acceleration and vertical sleeper acceleration. The results, summarized in Figure 4.18, reveal that the NS concrete sleeper assembly is susceptible to loads in the 20 - 300 Hz frequency band. The pad hardness seems to have little influence, neither on this frequency band nor on the modulus of the transfer function. The little difference in sleeper strain according to Figure 4.17 should therefore be explained by a lower dynamic load between wheel and rail.

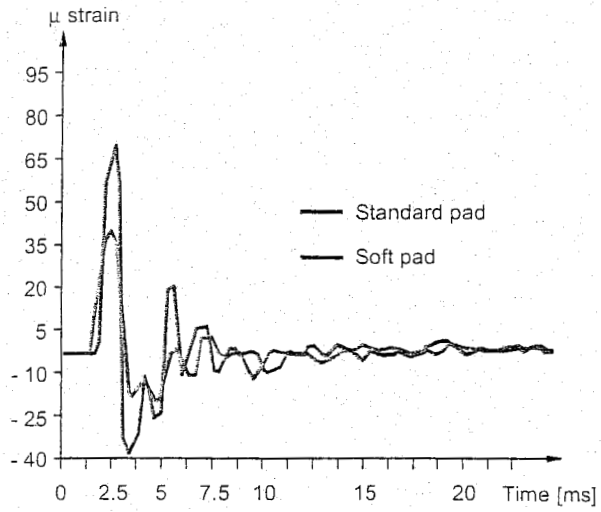


Figure 4.17: Sleeper strain versus time due to an impact load

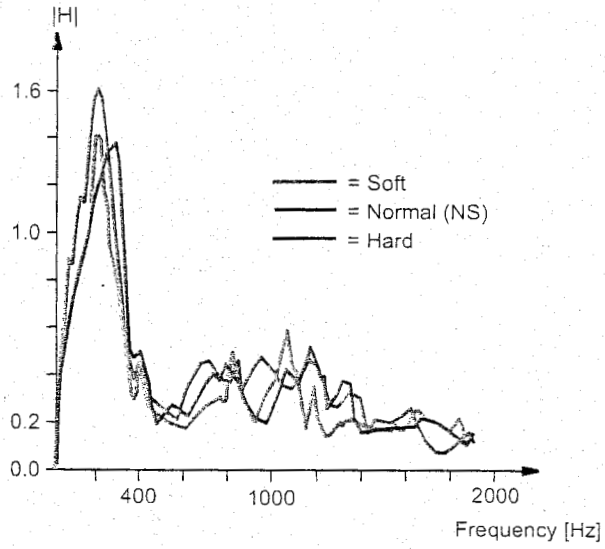


Figure 4.18: Estimated transfer function for rail pads

4.13 Axle box accelerations

The results of the D 161 calculations show the importance of very short wavelengths in generating the significant vertical loads applied to the track.
These short wavelengths of less than 0.5 m are not normally recorded by most track geometry cars, but the loads they produce can be important in causing both deterioration of the geometry and damage to components. These loads are well represented by axle box accelerations. From D 161 measurements, carried out during the Europe tour, the following conclusions on the forces from the unsprung mass of the wheelset were drawn:

- the standard deviation of the axle box accelerations ranged from 6 m/s² to 20 m/s²;
- as anticipated, the track geometry as measured by recording cars only has a secondary effect on the level of acceleration and thus on force.

5 STATIC TRACK DESIGN

5.1 Introduction

The subject of this chapter is track dimensioning, the main point of which is to ensure that the track structure is suitable for the loads it has to carry and the resultant stresses and deformations. Conventional track calculation is limited to quasi-static loading of the track structure, schematized as an elastically supported beam. To the static load is added a dynamic increment. Details on rail stresses as a result of contact pressure have been given earlier. Fatigue and high frequency loads at welds or caused by wheel flats are dealt with in chapter 6 on dynamic track design.

5.2 Supporting models

5.2.1 Winkler support model

Conventional track consists basically of two parallel continuous beams, the rails, which are fixed at regular intervals onto sleepers supported from below and from the side by a medium which cannot be deformed, the ballast bed. In turn, the ballast bed rests on a formation which also cannot be deformed [292]. In elementary calculations it is usually presupposed that the Winkler hypothesis applies to track support; this hypothesis was formulated in 1867 and reads: at each point of support the compressive stress is proportional to the local compression. This relation is illustrated in Figure 5.1 and can be written as:

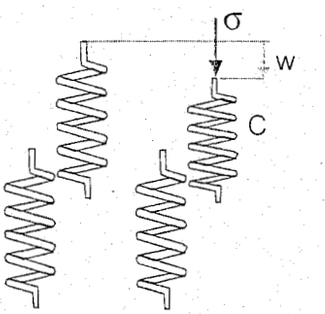


Figure 5.1: Winkler support model

$$\sigma = Cw \qquad (5.1)$$

in which:
σ = local compressive stress on the support [N/m^2];
w = local subsidence of the support [m];
C = foundation modulus [N/m^3].

5.2.2 Discrete rail support

Let us consider the situation of a discretely supported rail (Figure 5.2). Between the vertical force $F(x_i)$ on a support number at $x = x_i$ with effective rail support area A_{rs} and the deflection $w(x_i)$, the following relation exists according to Winkler:

$$F(x_i) = CA_{rs}w(x_i) = k_d w(x_i) \qquad (5.2)$$

Hence the spring constant of the support is:

$$k_d = CA_{rs} \qquad (5.3)$$

Determining the spring constant in a railway track with a homogeneous support is relatively simple using the equilibrium condition:

$$k_d = \frac{\Sigma F}{\Sigma w} = \frac{Q}{\Sigma w} \qquad (5.4)$$

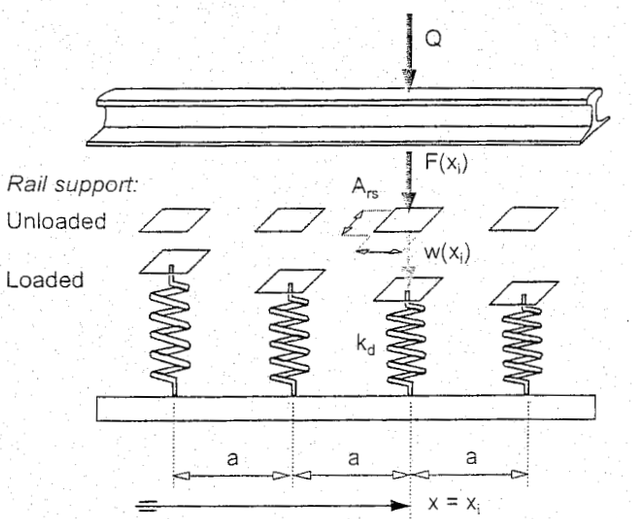


Figure 5.2: Discrete elastic support model

with:
Q = vertical wheel load (with given value) on the rail [N]
Σw = summation of all (measured) significant deflections in the vicinity of the load [m]

Note: This spring constant k_d is solely a property of the support and should not be confused with the total spring constant of the track, as 'seen' by the wheel:

$$k_{tot} = \frac{Q}{w_{max}} \tag{5.5}$$

From equations (5.5) and (5.6) follows:

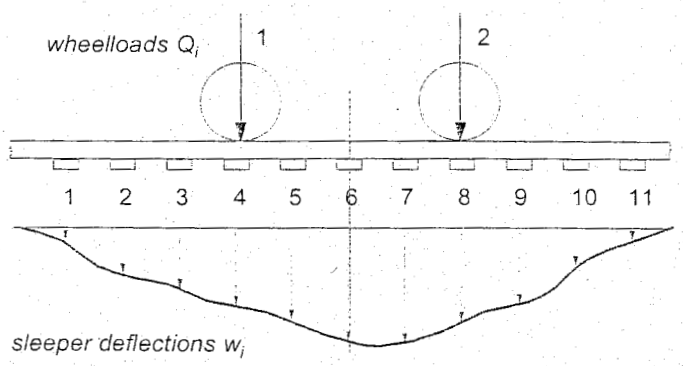
$$k_d = \frac{w_{max}}{\Sigma w} k_{tot} \tag{5.6}$$

which always shows that $k_d < k_{tot}$, as was to be expected.
The mean value of the contact pressure on a discrete rail support is according to Figure 5.2:

$$\sigma_{rs}(x_i) = \frac{F(x_i)}{A_{rs}} \tag{5.7}$$

5.2.3 Exercise: Spring constant determination

Problem 1:
To determine the mean spring constant k_d of a rail support of an existing track section, the rail is loaded by a bogie with two equal static wheel loads $Q_1 = Q_2 = 100$ kN. The sleeper spacing a = 55 cm. Sleeper deflections w_j are measured at 11 successive sleepers as indicated in the table. Deflections outside this area can be neglected.



sleeper no. j	sleeper deflection w_j [mm]
1	0.25
2	0.64
3	1.02
4	1.27
5	1.52
6	2.03
7	2.16
8	1.78
9	1.40
10	0.76
11	0.50

Question 1a:
Consider a general situation where the rail is loaded by a number of different wheel loads Q_i and the measured sleeper deflections in the selected area are w_j. Derive the following general formula for the mean value of the spring constant of a support:

$$k_d = \frac{\Sigma Q_i}{\Sigma w_j}$$

Question 1b:
Calculate with this formula the spring constant k_d using the given bogie load and the accompanying set of measurement deflections.

Question 1c:
Is this result for k_d representative for the material behaviour of the support? How can a better assessment of the value of the spring constant be obtained?

Answers:
1a $\Sigma Q_i = \Sigma F_j = k_d \Sigma w_j \Rightarrow k_d = \Sigma Q_i / \Sigma w_j$.
1b $\Sigma Q_j = 100 + 100 = 200$ kN; $\Sigma w_j = w_1$ t/m $w_{11} = 13.33$ mm $\Rightarrow k_d = 200.10^3/13.33 = 15004$ N/mm = 15 kN/mm.
1c Sleepers might not all be fully supported ('hanging sleepers'); due to the bridging effect of the rail, a lower value for k_d would be found compared to the value if individual supports would be tested.

Note 1: The assumed linear load-deflection relationship is an approximation. However, the supporting layer is essentially non-linear due to locally 'hanging' sleepers and stress-dependent elasticity moduli. Moreover, the relationship may vary locally between supporting points due to non-homogeneous material behaviour. A better assessment follows from the difference between measurements at different load levels.

Note 2: In section 5.3.7 a less elaborate way of determining the k_d value is discussed.

5.2.4 Continuous rail support

In this case a distributed load p(x) between rail and the supporting structure will exist (Figure 5.3) which is, according to Winkler, proportional to the deflection function w(x):

$$p(x) = kw(x) \qquad (5.8)$$

where:
k = foundation coefficient [N/m/m]

The foundation coefficient can be interpreted as a spring constant per unit length.
In this case the contact pressure on the continuous rail support is:

$$\sigma_{rs}(x) = \frac{p(x)}{b_c} \qquad (5.9)$$

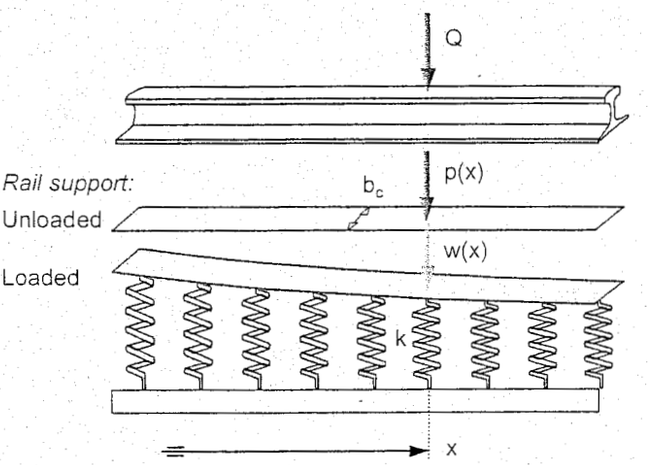


Figure 5.3: Continuous elastic support model

in which b_c is the width of the supporting strip under the rail.

5.2.5 Approximation of discrete rail support

A discretely supported rail structure can be analysed by numerical means. However, for a static assessment and simple configurations the continuous theory, to be discussed in Section 5.3, may also be used as an approximation method for discretely supported rail structures.
The equivalence with the actual discretely supported track then follows from:

$$k = \frac{k_d}{a} \qquad (5.10)$$

where:

a = spacing between centres of discrete supports (Figure 5.2).

5.3 Beam on elastic foundation model

5.3.1 Solution of the differential equation

Let us consider an infinitely long rail (CWR track) with bending stiffness EI which is continuously supported by an elastic foundation with foundation coefficient k and loaded according to Figure 5.4 by a wheel load Q at x = 0. This beam calculation was first proposed by Zimmermann [297].

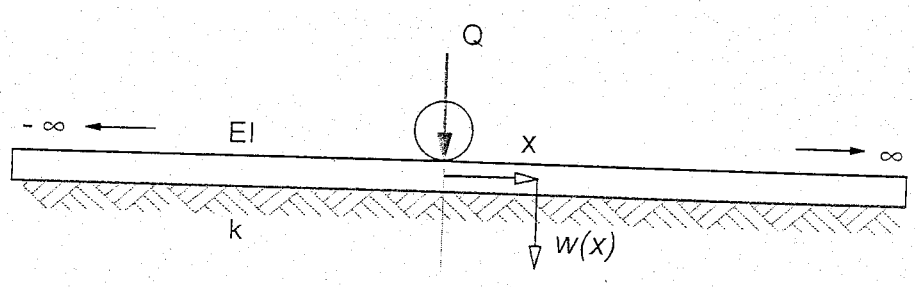


Figure 5.4: Infinite beam on elastic foundation model

where:
Q = wheel load [N];
EI = bending stiffness rail [Nm²];
k = foundation coefficient [N/m/m];
w(x) = rail deflection at x [m].

To derive the formula for the deflection w(x) of the beam, we first write down the equilibrium conditions of a beam element (Figure 5.5). Since there is only one variable (x) here, differentials are indicated by 'd' (rather than '∂').

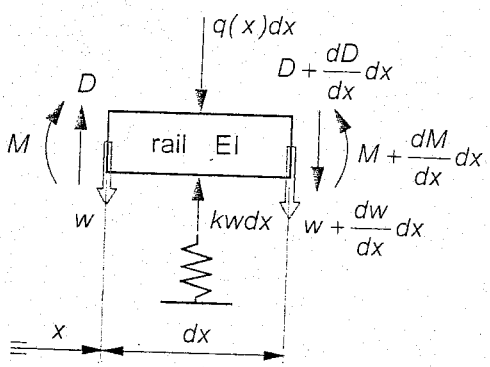


Figure 5.5: Beam element model (static)

Equilibrium requires:

$$q dx + \frac{dD}{dx} dx = k w dx \qquad (5.11)$$

$$D dx = \frac{dM}{dx} dx \qquad (5.12)$$

The constitutive equation is:

$$M = -EI \frac{d^2 w}{dx^2} \qquad (5.13)$$

From these equations the differential equation of the problem can be derived:

$$EI \frac{d^4 w}{dw^4} + kw = q(x) \qquad (5.14)$$

Since we deal only with point loads, the distributed load q(x) will not be considered here (q ≡ 0). The discrete wheel load Q will be introduced later as boundary condition.

After introduction of the short notation for spatial derivatives to x it follows:

$$EIw^{IV} + kw = 0 \qquad (5.15)$$

The boundary conditions for the part x > 0 are:

$$w(\infty) = 0; \qquad w'(0) = 0; \qquad w'''(0) = \frac{Q}{2EI} \tag{5.16}$$

After substitution of an exponential function for the deflection, the solution of the problem is:

$$w(x) = \frac{QL^3}{8EI}\eta(x) = \frac{Q}{2kL}\eta(x) \tag{5.17}$$

The distributed reaction force of the foundation follows from equation (5.17):

$$p(x) = kw(x) = \frac{Q}{2L}\eta(x) \tag{5.18}$$

The bending moment in the beam follows from equation (5.13):

$$M(x) = \frac{QL}{4}\mu(x) \tag{5.19}$$

The quantity L in these equations is the so-called characteristic length determined by:

$$L = \sqrt[4]{\frac{4EI}{k}} \tag{5.20}$$

Furthermore, two shape functions are present:

$$\eta(x) = e^{-x/L}\left[\cos\frac{x}{L} + \sin\frac{x}{L}\right] \quad x \geq 0 \tag{5.21}$$

$$\mu(x) = e^{-x/L}\left[\cos\frac{x}{L} - \sin\frac{x}{L}\right] \quad x \geq 0 \tag{5.22}$$

The functions $\eta(x)$ and $\mu(x)$ determine the form of the elastic line and the moment distribution which are shown in Figure 5.6. The left portion ($x < 0$) of these lines results from symmetry considerations. These lines may also be used as influence lines for determining deflection and bending moment for $x = 0$ resulting from nearby wheel loads. The expressions [5.21] and [5.22] represent heavily damped harmonic waves with wavelength $2\pi L$. They are therefore also a good tools for the approximation of finite beams with a central wheel load, provided the length of the beam is greater than $2\pi L$ [94].

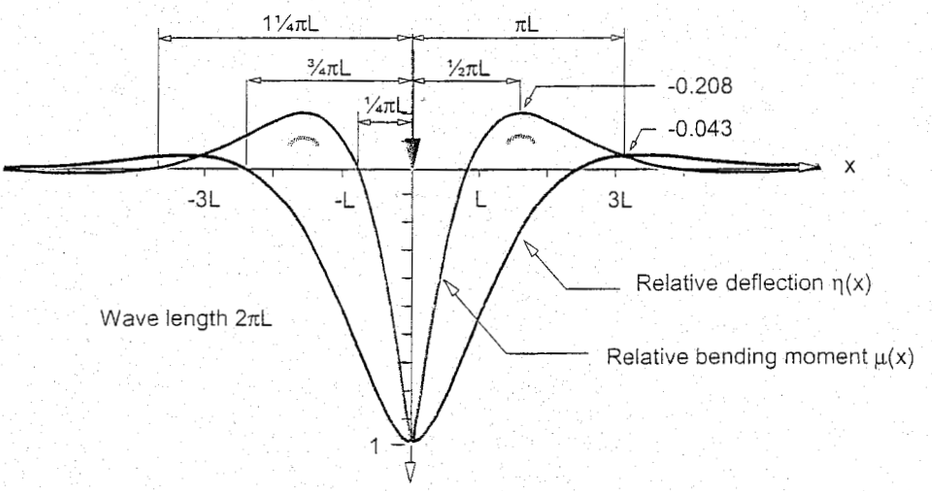


Figure 5.6: Relative deflection and bending moment in CWR track due to a concentrated vertical load

Note: A comparison of the beam of infinite length with an imaginary infinitely rigid beam (Figure 5.7) with a length of 2L shows that the maximum value (for x = 0) of both the deflection and the bending moment happens to be the same for both beams:

$$w_0 = \frac{Q}{2kL} \tag{5.23}$$

$$p_0 = kw_0 \tag{5.24}$$

$$M_0 = \frac{QL}{4} \tag{5.25}$$

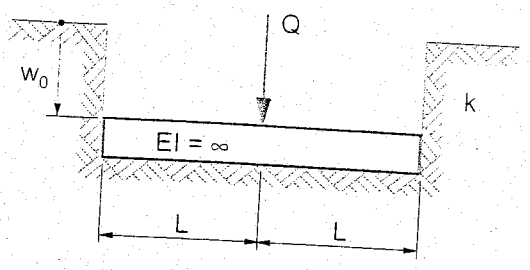


Figure 5.7: Infinitely rigid beam on elastic foundation

5.3.2 Several wheel loads

If there are several wheel loads, as in the case of bogies, then the resulting deflection and bending moment are found by means of superposition. At point x = 0 the following is found:

$$w_0 = \frac{1}{2kL}\sum_i Q_i \eta(\ell_i) \tag{5.26}$$

$$p_0 = kw_0 \tag{5.27}$$

$$M_0 = \frac{L}{4}\sum_i Q_i \mu(\ell_i) \tag{5.28}$$

In which ℓ_i is the distance of wheel load Q_i to point x = 0.

To illustrate the smoothing effect of a load system on the bending moment, the three cases (a), (b), and (c) in Figure 5.8 are examined. In case (a) the maximum moment for load Q is:

$$M = \frac{QL}{4} \tag{5.29}$$

In case (b) the total load is 2Q. The distance ℓ, for which M is a minimum, amounts to $\ell = 0.5\pi L$ so that according to equations 8.16 and 8.19 the maximum moment becomes:

$$M = \frac{QL}{4}[\mu(0) + \mu(\pi L/2)] = 0.79\frac{QL}{4} \tag{5.30}$$

In case c the total load is 3Q. The distance ℓ for which all moments are equal is $\ell = 1.12$ L, in which:

$$M = \frac{QL}{4}[\mu(0) + 2\mu(1.12L)] = 0.70\frac{QL}{4} \tag{5.31}$$

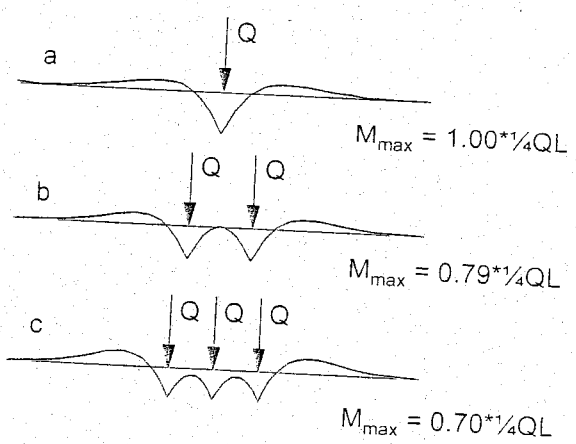


Figure 5.8: Smoothing effect of a load system on the maximum bending moment

These findings indicate that, in the case of several wheel loads Q, the maximum bending moment in the beam may be significantly less than in the case of one single wheel load Q. This positive effect only slightly applies to maximum deflection or compressive stress on the foundation, which are determined by the function $\eta(x)$.

5.3.3 Two-axle bogie

Consider the situation where an infinitely long track is loaded by a two-axle bogie according to Figure 5.9 with a half wheel-base ℓ = 1.25 m. The characteristic length L is fixed at 100 cm. UIC 54 rail with a bending stiffness of EI = 4.93 kNcm^2 is used. With the help of the η-line and the μ-line, both deflection and bending moment for the rail are determined at points A and B.

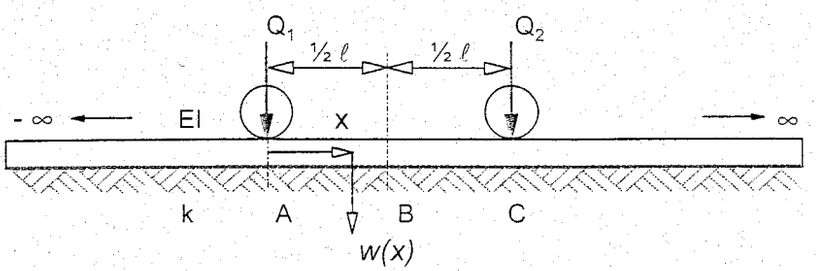


Figure 5.9: Bogie load

For a single wheel load Q = 125 kN the maximum deflection and the maximum bending moment are:

$$w_{max} = \frac{QL^3}{8EI} = 0.32\ cm \tag{5.32}$$

$$M_{max} = \frac{QL}{4} = 3125\ kNcm \tag{5.33}$$

For point A this reads:

$$w_A = \frac{QL^3}{8EI}(1+\eta_C) = 0.32\ cm \tag{5.34}$$

$$M_A = \frac{QL}{4}(1+\mu_C) = 2766\ kNcm \tag{5.35}$$

For point B the values become:

$$w_B = 2\frac{QL^3}{8EI}\eta_A = 0.23\ cm \tag{5.36}$$

$$M_B = 2\frac{QL}{4}\mu_A = -1125\ kNcm \tag{5.37}$$

This calculation example demonstrates that negative bending moments of the order of 40% of the maximum positive bending moments may occur in the rail.

5.3.4 Negative deflection

The relative deflection line in Figure 5.6 shows that the compressive stress should be negative for 2.36L < x < 5.50L. This is, however, physically impossible on ballasted track as ballast material cannot absorb tensile force. It is generally assumed that the dead weight of the track compensates the negative deflection. It is possible to verify theoretically that adequate compensation occurs, provided

Note: A comparison of the beam of infinite length with an imaginary infinitely rigid beam (Figure 5.7) with a length of 2L shows that the maximum value (for x = 0) of both the deflection and the bending moment happens to be the same for both beams:

$$w_0 = \frac{Q}{2kL} \qquad (5.23)$$

$$p_0 = kw_0 \qquad (5.24)$$

$$M_0 = \frac{QL}{4} \qquad (5.25)$$

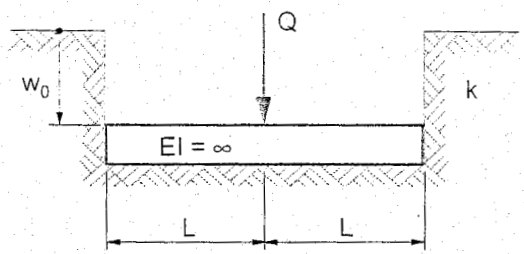


Figure 5.7: Infinitely rigid beam on elastic foundation

5.3.2 Several wheel loads

If there are several wheel loads, as in the case of bogies, then the resulting deflection and bending moment are found by means of superposition. At point x = 0 the following is found:

$$w_0 = \frac{1}{2kL}\sum_i Q_i \eta(\ell_i) \qquad (5.26)$$

$$p_0 = kw_0 \qquad (5.27)$$

$$M_0 = \frac{L}{4}\sum_i Q_i \mu(\ell_i) \qquad (5.28)$$

In which ℓ_i is the distance of wheel load Q_i to point x = 0.

To illustrate the smoothing effect of a load system on the bending moment, the three cases (a), (b), and (c) in Figure 5.8 are examined. In case (a) the maximum moment for load Q is:

$$M = \frac{QL}{4} \qquad (5.29)$$

In case (b) the total load is 2Q. The distance ℓ, for which M is a minimum, amounts to $\ell = 0.5\pi L$ so that according to equations 8.16 and 8.19 the maximum moment becomes:

$$M = \frac{QL}{4}[\mu(0) + \mu(\pi L/2)] = 0.79\frac{QL}{4} \qquad (5.30)$$

In case c the total load is 3Q. The distance ℓ for which all moments are equal is $\ell = 1.12$ L, in which:

$$M = \frac{QL}{4}[\mu(0) + 2\mu(1.12L)] = 0.70\frac{QL}{4} \qquad (5.31)$$

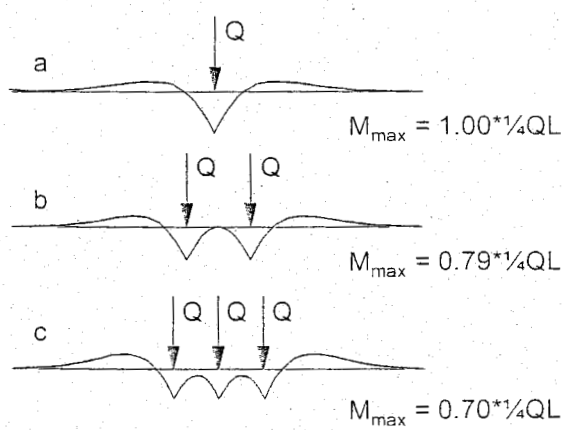


Figure 5.8: Smoothing effect of a load system on the maximum bending moment

These findings indicate that, in the case of several wheel loads Q, the maximum bending moment in the beam may be significantly less than in the case of one single wheel load Q. This positive effect only slightly applies to maximum deflection or compressive stress on the foundation, which are determined by the function η(x).

5.3.3 Two-axle bogie

Consider the situation where an infinitely long track is loaded by a two-axle bogie according to Figure 5.9 with a half wheel-base ℓ = 1.25 m. The characteristic length L is fixed at 100 cm. UIC 54 rail with a bending stiffness of EI = 4.93 kNcm^2 is used. With the help of the η-line and the μ-line, both deflection and bending moment for the rail are determined at points A and B.

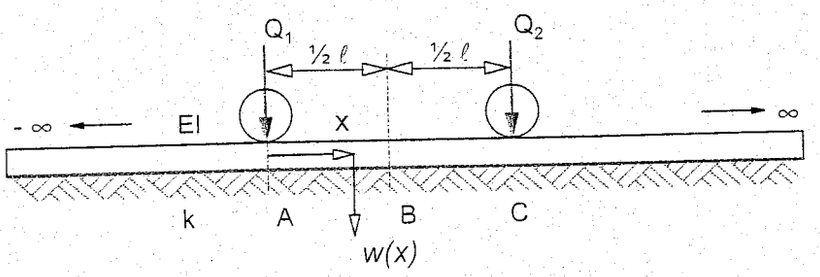


Figure 5.9: Bogie load

For a single wheel load Q = 125 kN the maximum deflection and the maximum bending moment are:

$$w_{max} = \frac{QL^3}{8EI} = 0.32\ cm \tag{5.32}$$

$$M_{max} = \frac{QL}{4} = 3125\ kNcm \tag{5.33}$$

For point A this reads:

$$w_A = \frac{QL^3}{8EI}(1+\eta_C) = 0.32\ cm \tag{5.34}$$

$$M_A = \frac{QL}{4}(1+\mu_C) = 2766\ kNcm \tag{5.35}$$

For point B the values become:

$$w_B = 2\frac{QL^3}{8EI}\eta_A = 0.23\ cm \tag{5.36}$$

$$M_B = 2\frac{QL}{4}\mu_A = -1125\ kNcm \tag{5.37}$$

This calculation example demonstrates that negative bending moments of the order of 40% of the maximum positive bending moments may occur in the rail.

5.3.4 Negative deflection

The relative deflection line in Figure 5.6 shows that the compressive stress should be negative for 2.36L < x < 5.50L. This is, however, physically impossible on ballasted track as ballast material cannot absorb tensile force. It is generally assumed that the dead weight of the track compensates the negative deflection. It is possible to verify theoretically that adequate compensation occurs, provided

the weight per length satisfies the following equation:

$$q \geq \frac{1}{2e^{\pi}} \frac{Q}{L} = 0.022 \frac{Q}{L} \tag{5.38}$$

Modern track with concrete sleepers satisfies this condition. Further theoretical study [149] has shown that the effect of negative deflection on the maximum positive deflection and maximum bending moment is very small.

5.3.5 Beam with hinge (jointed track)

The influence of a rail joint on deflection and bending moment can also be checked with the help of the elastically supported bending beam theory. The results for a wheel load Q on a hinge at x = 0 are presented in Figure 5.10.

In this case the boundary conditions for the part x > 0 become:

$$w(\infty) = 0; \qquad w''(0) = 0; \qquad w'''(0) = \frac{Q}{2EI} \tag{5.39}$$

After substitution of an exponential function for the deflection, the solution of the problem is:

$$w(x) = \frac{QL^3}{4EI} e^{-x/L} \cos\frac{x}{L} = \frac{Q}{kL} e^{-x/L} \cos\frac{x}{L} \tag{5.40}$$

Apparently, the maximum deflection is twice the value for the homogeneous rail (5.17).

The rail ends at the hinge; the maximum total dip angle 2α between the two rail ends is:

$$2\alpha = 2*|w'(0)| = \frac{2Q}{kL^2} \tag{5.41}$$

The absolute maximum moment appears at x = πL/4:

$$|M_{max}| = 1/4\sqrt{2}e^{-\pi/4}QL = 0.161QL \tag{5.42}$$

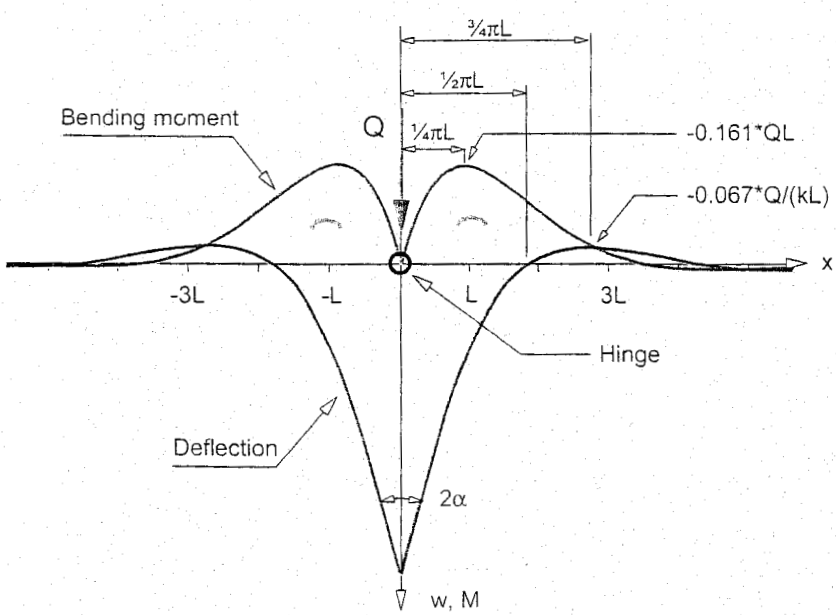


Figure 5.10: Deflection and bending moment in an elastically supported

5.3.6 Alternative expressions for characteristic length L

Based on equations (5.3) and (5.10), other useful expressions for equation (5.20) can be formulated for the characteristic length of the track:

$$L = \sqrt[4]{\frac{4EI}{k}} = \sqrt[4]{\frac{4EIa}{k_d}} = \sqrt[4]{\frac{4EIa}{C_{rs}A_{rs}}} = \sqrt[4]{\frac{4EIa}{C_{bs}A_{bs}}} = \sqrt[4]{\frac{4EIa}{C_{sb}A_{sb}}} \tag{5.43}$$

in which:
EI = bending stiffness of rail [Nm²];
k = foundation coefficient of continuous support [N/m/m];
k_d = spring constant of discrete support [N/m];
a = spacing between centers of discrete supports [m];
C_{rs} = foundation modulus at contact area between rail and sleeper [N/m³];
C_{bs} = foundation modulus at contact area between baseplate and sleeper [N/m³];
C_{sb} = foundation modulus at contact area between sleeper and ballast bed [N/m³];
A_{rs} = contact area between rail and sleeper for half sleeper [m²];
A_{bs} = contact area between baseplate and sleeper for half sleeper [m²];
A_{sb} = contact area between sleeper and ballast bed for half sleeper [m²].

Note that the spring constant of the discrete support k_d is, independent of any surface area at the support, in contrast with the foundation modulus C.

5.3.7 Fast determination of vertical elasticity constants

Based on the longitudinal beam theory using a single wheel load, the spring constant of a discrete support k_d can be obtained from the equations (5.10), (5.17), and (5.20):

$$k_d \approx \frac{a}{4}\sqrt[3]{\frac{Q^4}{EIw_{max}^4}} \tag{5.44}$$

This result may be used to determine the spring constant using a relatively simple experiment [72].

While the spring constant k_d and the foundation coefficient k are simply related by (5.10), the total spring constant of the track can be found from (5.5), (5.17), and (5.20):

$$k_{tot} \approx \frac{8EI}{L^3} = 2\sqrt{2}\sqrt[4]{k^3EI} \tag{5.45}$$

5.3.8 Order of magnitude of elasticity constants

Table 5.1lists the various quantities used to express the elasticity of the rail supporting structure together with their units. Moreover, global values are given which correspond to the qualifications 'poor' and 'good' to characterize the condition of the foundation.

Quality of track support	Unit	Poor	Good
Foundation modulus C	[N/mm³]	0.02	0.20
Spring constant k_d	[N/mm]	5.5	55
Foundation coefficient k	[N/mm²]	9	90
Characteristic length L	[m]	1.30	0.70

Table 5.1: Order of magnitude of elasticity constants

5.4 Double beam model

The Zimmermann model is sufficient to understand the general behaviour of the track structure. In some cases a more accurate description is necessary, for instance when the individual behaviour of rail pads is investigated or if the effect of a slab foundation under the track is to be assessed.

A more realistic model is the double beam model. For discrete supports with relatively high spacing a two-layer model can be used based on FEM techniques. For continuously supported track a two-layer model is useful, which actually is a generalization of the standard Zimmermann model. Both models, loaded for instance by two equal vertical forces, are drawn in Figure 5.11 with the characteristic deflection lines.

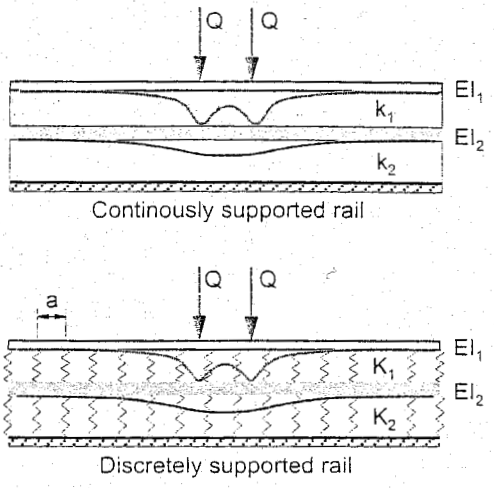


Figure 5.11: Double beam model (static)

The upper beam in each model is the rail, whereas the other beam represents the continuity in the supporting structure which is mostly due to a slab track or other reinforcing structure.

The unknown stiffness parameters are the effective bending stiffness EI_2 and the foundation coefficients k_1 and k_2 (or spring constants kd1 (K_1) and kd2 (K_2) in the discrete case and support spacing a). It may be assumed that the bending stiffness EI_1 of the rail is known.

The model for the continuously supported rail is relatively simple and will be presented here. Without a distributed load on the rail the differential equations for this model are:

$$EI_1 \frac{d^4 w_1}{dx^4} + k_1(w_1 - w_2) = 0$$
$$EI_2 \frac{d^4 w_2}{dx^4} + (k_1 + k_2)w_2 - k_1 w_1 = 0 \qquad (5.46)$$

The analytical solution for the deflections for both the infinitely long rail (w_1) and the likewise modelled slab (w_2) can be easily obtained (See [115]).

An example is given in Figure 5.12. The following values were used:

$Q_1 = Q_2 = 35$ kN;
$EI_1 = 549$ kNm² (Nikex rail);
$EI_2 = 80000$ kNm².

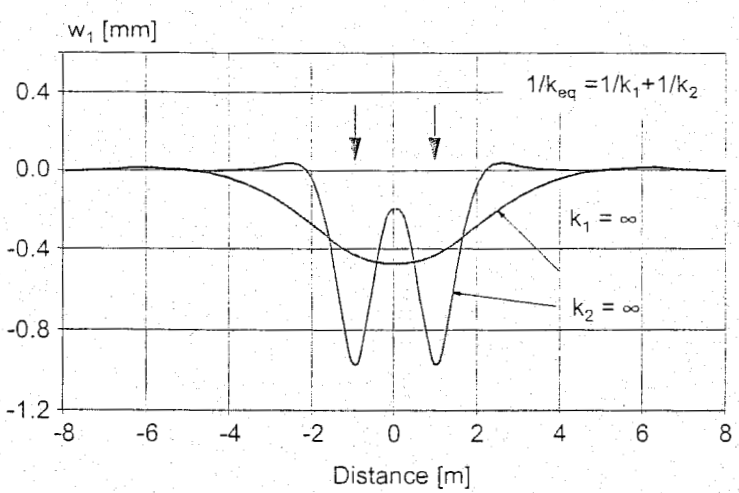


Figure 5.12: Rail deflection depending on ratio k_1 and k_2

The equivalent foundation coefficient:

$$\frac{1}{k_{eq}} = \frac{1}{k_1} + \frac{1}{k_2} \qquad (5.47)$$

was kept constant in this example ($k_{eq} = 34483$ kN/m²) while two extreme cases for k_1 and k_2 were considered. This theory was used in experiments to back-calculate the parameters from recorded field measurements using curve-fitting techniques.

5.5 Pasternak foundation model

As we have seen in Section 5.2 the simplest representation of a continuous elastic foundation has been provided by Winkler who assumed a base consisting of closely spaced, independent linear springs. Such a foundation is equivalent to a liquid base. The relation between the pressure and the deflection of the foundation surface is then:

$$p = kw \tag{5.48}$$

wherein $k \approx k_d/a$, according to equation (5.10).

In order to describe the deformation of a beam on elastic foundation more accurately, various kinds of interaction between the springs of the Winkler foundation should be considered. For instance, Pasternak assumes the existence of shear interactions between the spring elements. This may be accomplished by connecting the ends of the springs to the beam consisting of incompressible vertical elements which then only deforms by transverse shear.

In Figure 5.13 the classical rail element on a Winkler foundation is extended with a shear element representing the interaction between the springs. The shear element is connected to the rail element. The distributed load on the rail is assumed to be zero. For the derivation of differential equation in this more general case, we consider the equilibrium equations and the constitutive equations.

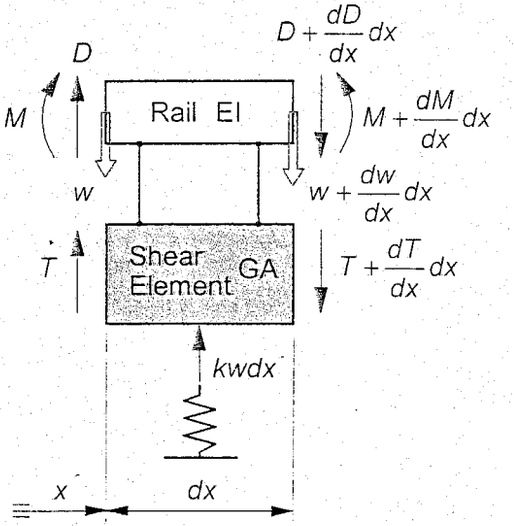


Figure 5.13: Pasternak foundation model for beams

The equilibrium requires:

$$\frac{dD}{dx}dx + \frac{dT}{dx}dx = kwdx \tag{5.49}$$

$$Ddx = \frac{dM}{dx}dx \tag{5.50}$$

Note that the T-forces in the shear element cannot produce bending moments.

The constitutive equation regarding the bending moment and the deformation reads:

$$M = -EI\frac{d^2w}{dx^2} \tag{5.51}$$

Taking the basic relation between shear stress and strain into consideration, a linear relation between the shear force and the shear angle is found:

$$T = GA\frac{dw}{dx} \tag{5.52}$$

where:

G = shear modulus of the foundation [N/m²];
A = cross section area for shear forces of the foundation [m²].

Combining equations (5.49) to (5.52) gives the differential equation in the vertical deflection w(x):

$$EI\frac{d^4w}{dw^4} - GA\frac{d^2w}{dx^2} + kw = 0 \qquad (5.53)$$

This result may be compared to the Zimmermann solution discussed under section 5.3, where the beam is loaded by a vertical concentrated force Q at x = 0.

With the boundary conditions for the part x > 0:

$$w(\infty) = 0; \qquad w'(0) = 0; \qquad w'''(0) = \frac{Q}{2EI} \qquad (5.54)$$

and the abbreviations:

$$a = \tfrac{1}{2}\sqrt{4\beta^2 - \gamma}; \qquad b = \tfrac{1}{2}\sqrt{4\beta^2 + \gamma}$$
$$\beta = \sqrt[4]{\frac{k}{4EI}} = (1/L); \qquad \gamma = \frac{GA}{EI} \qquad (5.55)$$

the solution of (5.53) reads:

$$w(x) = \frac{Q}{8EIab\beta^2}e^{-bx}(a\cos ax + b\sin ax) \qquad x \geq 0 \qquad (5.56)$$

The bending moment follows from (5.51) and (5.55):

$$M(x) = \frac{Q}{4ab}e^{-bx}(a\cos ax - b\sin ax) \qquad x \geq 0 \qquad (5.57)$$

Equations (5.56) and (5.57) are the generalized forms of the corresponding Zimmermann equations given by (5.17), (5.19), (5.21), and (5.22).

In Figure 5.14 the relative deflection $\eta(x)$ and relative moment curves $\mu(x)$ are drawn for three values of the parameter γ [m^{-2}] while the parameter β is kept constant on the value 1 m^{-1}. For γ = 0, the corresponding Zimmermann results (5.21) and (5.22) appear.

The effect of value $\gamma > 0$ is noticeable in a lower maximum value and the 'stretching' of the curves as compared to the Zimmermann curves.

Note that the addition of a longitudinal pulling force on the rail or the addition of a distributed rotational resistance would give the same form of differential equation and hence the same form of solution.

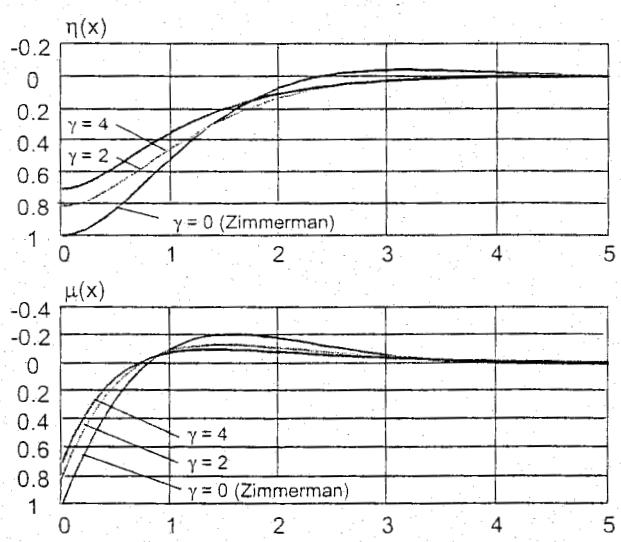


Figure 5.14: Relative rail deflection and moment for a Pasternak foundation model

The maximum value of $\eta(x)$ and $\mu(x)$ occurs at x = 0 and can be calculated from (5.56) and (5.57) as:

$$\eta_{max} = \mu_{max} = \sqrt{\frac{\beta^2}{\beta^2 + 1/4\gamma}} \qquad (5.58)$$

5.6 Rail stresses

Generally, the rails are loaded by vertical, longitudinal, and lateral forces. These loads lead to bending stresses caused by the vertical train load, normal stresses which are mainly determined by temperature effects on CWR track, residual stresses due to roller straightening during manufacture, contact stresses in the rail head, and stresses caused by high frequency impact loads.

5.6.1 Stresses in rail foot centre

It is known that under repeated loading, fatigue fractures may occur starting in the centre of the rail foot, partly as a result of the high internal tensile stresses at this point, as indicated in Figure 10.13. The combination of these static normal stresses and dynamic bending tensile stresses under the influence of vertical wheel load occurring in the rail foot centre, determines the rail strength. Lateral force on the rail or eccentricity of the vertical wheel load has no effect on this.

5.6.2 Dynamic amplification factor

It is common practice to carry out a strength or fatigue calculation for a static load system, or often a single wheel load, using the longitudinal beam theory, according to which the dynamic effects are taken into account by a speed coefficient or dynamic amplification factor.

The effect of running speed on load is in reality highly complex because of the dynamic interaction between vehicle and track. In view of the nature of the load, it is also more correct to carry out a fatigue calculation.

Several simple formulas have been proposed which aim to assess the dynamic rail stress. These formulas are a rough approximation of reality because the geometric quality of the track and the mechanical characteristics of the track and the vehicle are not sufficiently taken into account. One such empirical calculation scheme, which has been accepted by European railway companies, was developed by Eisenmann [94] and is based on the following observations and assumptions:

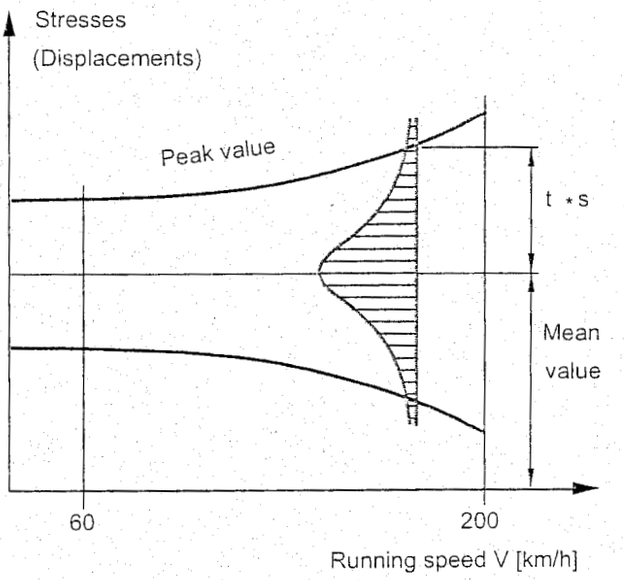


Figure 5.15: Normal distribution of measured bending stresses

- Measurements on which this empirical method is based have shown that the stresses in the rail foot, from a statistical point of view, have a normal distribution as illustrated in Figure 5.15;
- The mean value is independent of running speed V (studied up to 200 km/h) and can be determined with sufficient accuracy using Zimmermann's longitudinal beam calculation;
- The standard deviation is dependent on the running speed V and the state of the track.

The Eisenmann scheme to determine the DAF (Dynamic Amplification Factor) is dependent on the train speed, the track quality, and chosen factor t and reads as follows:

DAF- formula:

$$DAF = 1 + t\varphi \qquad \text{if } V < 60 \text{ km/h}$$

$$DAF = 1 + t\varphi\left(1 + \frac{V - 60}{140}\right) \qquad \text{if } 60 \le V \le 200 \text{ km/h} \qquad (5.59)$$

Probability	t	Application		Track condition	φ
68.3 %	1	Contact stress, subgrade		Very good	0.1
95.4 %	2	Lateral load, ballast bed		Good	0.2
99.7 %	3	Rail stresses, fastenings, supports		Bad	0.3

t = multiplication factor of standard deviation which depends on the confidence interval. Since the rail is so important for safety and reliability of rail traffic a value of 3 is recommended as the chance of exceeding the maximum calculated stresses is only 0.15 %;
φ = factor depending on track quality;
V = train speed [km/h].

Figure 5.16 gives the graphical presentation of this approach showing the range of the DAF running from 1.1 to 2.8.

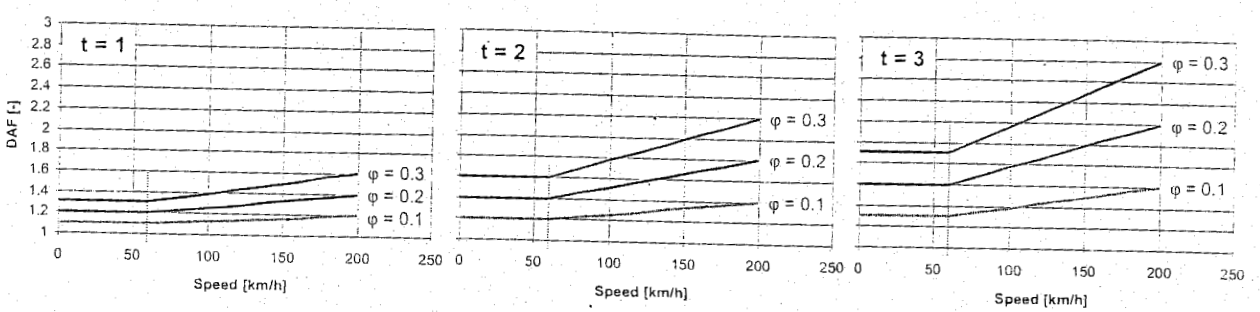


Figure 5.16: DAF values as function of the train speed for 3 track qualities and 3 probability levels

Example: presupposing V = 200 km/h for average (good) track quality (φ = 0.2) and t = 3, the calculated DAF = 2.2. Hence an increment of 120 % is added to the calculation according to Zimmermann.

5.6.3 Maximum bending stress in rail foot centre

According to the Eisenmann method, the greatest expected dynamic bending tensile stress in the rail foot centre can be determined from:

$$\sigma_{max} = DAF \cdot \sigma_{mean} \qquad (5.60)$$

The mean value of the rail bending stress follows from (5.19):

$$\sigma_{mean} = \frac{QL}{4W_{yf}} \qquad (5.61)$$

DAF = dynamic amplification factor, according to (5.59);
Q = effective wheel load [N];
L = characteristic length [m], according to (5.43);
W_{yf} = section modulus, relative to the rail foot [m³] (see Table 5.2).

Q is the effective wheel load, which is equal to the nominal wheel load, multiplied by a factor 1.2 taking account of the increase in wheel load in curves because of cant deficiency or excess.

The constant tensile stress is the sum of the internal residual stress in the rail foot centre and the tensile stress caused by low temperatures. These stresses are about 100 N/mm^2 in each case. The permissible values derived from this for the bending tensile stress in the rail foot centre are presented in Table 5.3, which is valid for long welded (CWR) vignole rails.

5.6.4 Stresses in the rail head

The concentrated load between wheel and rail causes a shear stress distribution in the rail head, as depicted in Figure 5.18, with at some depth a maximum which may give rise to a fatigue fracture in the rail head. The contact problem is most serious in the case of high wheel loads or relatively small wheel diameters. According to the Hertz theory, the contact area between two curved elastic bodies, such as wheel and rail head, is generally ellipsoidal and the contact stress distribution is semi-ellipsoidal (see Chapter 2). Eisenmann [94] has devised a simplified calculation method for the wheel-rail contact problem based on the following considerations:

Measurements have proven that for wheel diameters between 60 and 120 cm a simplified two-dimensional calculation suffices. All curve radii in the mathematical formulation of the contact problem are assumed to be infinitely large except the curve radius r of the wheel. The contact area then becomes rectangular and the contact stress distribution has the form of a semi-elliptical cylinder as indicated in Figure 5.19.

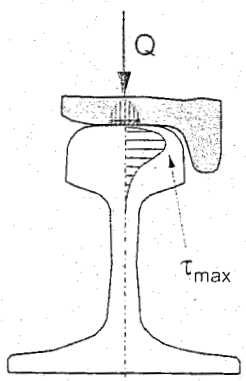


Figure 5.18: Shear stress distribution in rail head

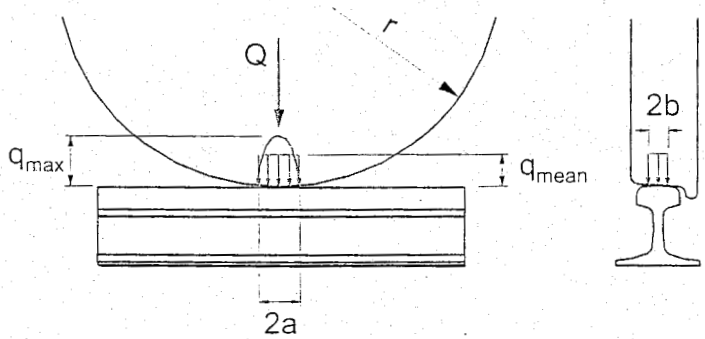


Figure 5.19: Assumed contact distribution between wheel and rail according to Eisenmann

If the wheel load Q is distributed evenly over the contact area with a width of 2b, the mean contact stress can be derived from [94]:

$$q_{mean} = \sqrt{\frac{\pi E}{64(1-\nu^2)}\frac{Q}{rb}} \qquad (5.63)$$

in which:
Q = effective wheel load;
r = wheel radius;
2b = width of wheel/rail contact area;
E = modulus of elasticity;
ν = Poisson's ratio.

With E = 210.000 N/mm^2, ν = 0.3, b = 6 mm the practical formula is obtained:

$$q_{mean} = 1374\sqrt{\frac{Q}{r}} \qquad (5.64)$$

where Q [kN], r [mm] and q_{mean} [N/mm^2].

Maximum shear stress in the rail head

The stress state in the rail head can be approximated sufficiently by using the Boussinesq half space theory. It has been shown by means of experimentation [94] that since the theory of elasticity does not hold well in all cases, the mean contact stress distribution according to [5.63] can be applied. Maximum shear stress in the rail head that determines the permissible wheel load or wheel radius can be estimated as:

$$\tau_{max} \cong 0.3 q_{mean} \tag{5.65}$$

Maximum shear stress occurs across the rail. Longitudinally the shear stress decreases with the occurrence of bending stresses. In the case of (5.64) maximum shear stress may be expressed as:

$$\tau_{max} = 412 \sqrt{\frac{Q}{r}} \tag{5.66}$$

with Q [kN], r [mm], and τ_{max} [N/mm^2].
If, for example Q = 100 kN, r = 400 mm, this results in τ_{max} = 206 N/mm^2.

It follows from equation (5.66) that the shear stress is not proportional to the load. Shear stress does depend on the wheel radius, but not on the bending stiffness of the rail or the state of the track. The calculated stress does not include an increment factor because the stress state has a very local character. Besides, the contact area changes location from wheel to wheel.

This simplified calculation model can be used to calculate that the maximum shear stress occurs at a depth of 4 - 6 mm in the rail head, which is comparable to the defects observed in practice known as shelling.

In tight curves with rail lubrication, it is, however, necessary to base the calculation on the general elliptical contact stress distribution. Here the maximum lies at a depth of 2 - 4 mm beneath the running surface. Maximum shear stress is then approximately 50% higher.

Permissible shear stress

On the basis of the Von Mises yield criterion, the permissible shear stress may be expressed as:

$$\bar{\tau} = \frac{\bar{\sigma}}{\sqrt{3}} \tag{5.67}$$

in which $\bar{\tau}$ is the permissible tensile stress. On account of the fatigue nature of the load, the permissible tensile stress should, according to tests, be fixed at 50% of the tensile strength σ_t of rail steel, thus resulting in:

$$\bar{\tau} \cong 0.3 \sigma_t \tag{5.68}$$

for permissible shear stress. The permissible effective wheel load is obtained from (5.66) and (5.68) and reads:

$$\bar{Q} = 4.9 \cdot 10^{-7} r \sigma_t^2 \qquad \text{(repetitive wheel load)} \tag{5.69}$$

with: r [mm], σ_t [N/mm^2] and Q [kN].

Occasional high wheel loads or small wheel radii

Occasional high wheel loads or small wheel radii do not cause fatigue fracture, but are responsible for plastic deformation of the rail head. If the yield point is set at 65% of the tensile strength, in the case of (5.67) the permissible shear stress becomes:

$$\bar{\tau}_{inc} = 0.38\ \sigma_t \tag{5.70}$$

and in the case of (5.66) the permissible effective incidental wheel load is:

$$\bar{Q} = 8.3 \cdot 10^{-7} r\ \sigma_t^2 \qquad \text{(incidental wheel load)} \tag{5.71}$$

The admissible shear stresses are presented in Table 5.4.

5.6.5 Rail stresses due to a combined Q/Y load

The interaction of an eccentric wheel load Q and a high horizontal lateral force Y may result in considerable bending stresses at the edges of the rail head and rail foot as shown in Figure 5.20, points A to D. This situation arises mainly in tight curves or switch curves under heavy locomotives. The number of such changes in load condition is usually limited and therefore no fatigue calculation is required. There is nevertheless the possibility of plastic deformation if the yield point is exceeded.

In order to determine the bending stress distribution in the cross-section, the load is first decomposed according to Figure 5.20 after which the cross-section bending stress is determined for each of the partial loads. After this the bending stresses are summated. To determine the partial stresses use is made of another Eisenmann [94] method which is based on results obtained by Timoshenko [265]:

Partial load I

The rail head is loaded centrally by the effective wheel load Q. According to the elementary beam theory, the bending stresses σ_{BH} and σ_{BF} occur at the indicated locations on the rail.

The linear stress distribution is, however, disturbed by the special form of the rail cross-section. Even if the rail lies on a rigid support, the load Q produces bending stresses in the rail head because the rail web indents. This so-called Timoshenko effect causes an additional tensile bending stress on the lower side of the rail head $\Delta\sigma_{1H}$.

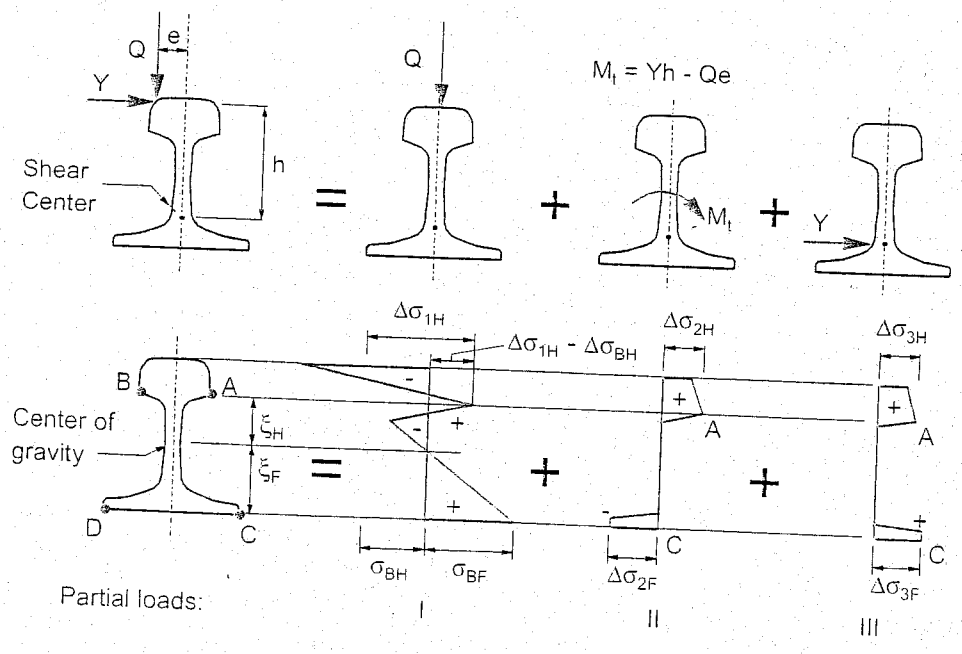


Figure 5.20: Load decomposition for calculation of rail stresses

Partial load II

The torsion moment M_t is thought to act on the rail at the mid-point of a sleeper bay. At the sleeper position the rail is assumed to be clamped. The pure torsion and constrained warping result in the bending stresses: $\Delta\sigma_{2H}$ and $\Delta\sigma_{2F}$ [265].

Partial load III

As a result of a horizontal load at the shear centre, tensile bending stresses occur at the indicated points. Assuming that the rails, with a length equal to the sleeper spacing a, are supported at the ends in the horizontal direction, the stresses are: $\Delta\sigma_{3H}$ and $\Delta\sigma_{3F}$.

Composition

In view of certain assumptions when calculating $\Delta\sigma_{1H}$ to $\Delta\sigma_{3F}$, these values are about 10% too high in comparison with the measured values. The total bending stress in the rail head and rail foot at the points indicated in Figure 5.20 is therefore:

$$\sigma_{H\,A,B} = -\sigma_{BH} + 0.9\,[\Delta\sigma_{1H} \pm \Delta\sigma_{2H} \pm \Delta\sigma_{3H}] \qquad (5.72)$$

$$\sigma_{F\,C,D} = +\sigma_{BF} + 0.9\,[\quad \mp \Delta\sigma_{2F} \pm \Delta\sigma_{3F}]$$

The upper sign belongs to the first mentioned index of the stresses.

Observations

No increment factor is applied to the stresses $\Delta\sigma_2$ to $\Delta\sigma_3$ because of their very local character. Apart from the stresses (5.72) due to traffic load, the following stresses and related maximum values must also be taken into consideration:

- *residual stresses:*

$$\sigma_{A,B} = -40\ \text{N/mm}^2$$

$$\sigma_{C,D} = -60\ \text{N/mm}^2$$

- *temperature stresses* in CWR track over entire cross-section:

$$\sigma_{A,B,C,D} = \pm 100\ \text{N/mm}^2$$

- *bending stresses* due to horizontal bending in tight curves:

$$\sigma_{A,B} = \pm \frac{Eb}{2R} = \pm 25\ \text{N/mm}^2 \qquad (5.73)$$

$$\sigma_{C,D} = \pm \frac{EB}{2R} = \pm 50\ \text{N/mm}^2 \qquad (5.74)$$

The relationships between stress and force are determined numerically for a number of rail profiles and are given in Table 5.2.

5.6.6 Rail tables

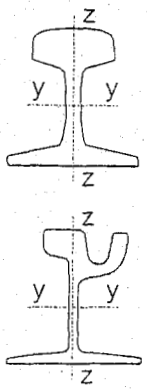
Rail profile	Rail section		S41	S49	NP46	UIC54	UIC60	Ri60
	Height	h_r [mm]	138	149	142	159	172	180
	Head width	b_h [mm]	67	67	72	70	72	113
	Foot width	b_f [mm]	125	125	120	140	150	180
	Area	A [cm²]	52.7	63.0	59.3	69.3	76.9	77.1
	Mass/meter	m [kg/m]	41.3	49.4	46.6	54.4	60.3	60.5
	Moment of inertia	$I=I_y$ [cm⁴]	1368	1819	1605	2346	3055	3334
	Moment of inertia	I_z [cm⁴]	276	320	310	418	513	884
	Section modulus	W_{yh} [cm³]	196	240	224	279	336	387
	Section modulus	W_{yf} [cm³]	200.5	248	228	313	377	355
	Section modulus	W_z [cm³]	44.2	51.2	52	60	68	135

Table 5.2: Rail dimensions and strength data

Tensile strength σ_t [N/mm²]	Yield stress σ_y [N/mm²]	Constant stresses		σ_f [N/mm²] Incidental loading	σ_f [N/mm²] Repeated loading
		Residual stress [N/mm²]	Temperature stress CWR track [N/mm²]		
700	450	220	100	450	~55*
900	580			580	220*
* from Smith diagram					

Table 5.3: Admissible dynamic stress range at the rail foot centre for as-rolled rails (refer to Figure 5.17)

Tensile strength σ_t [N/mm²]	τ_{max} [N/mm²] Incidental loading	τ_{max} [N/mm²] Repeated loading
700	260	200
900	340	260

Table 5.4: Admissible shear stress in the rail head

Modulus of elasticity:	E = 210·kN/mm²	Linear expansion coefficient:	$\alpha = 1.15 \cdot 10^{-5}$/°C
Poisson's constant:	ν = 0.3	Density:	ρ = 7850 kg/m³

Table 5.5: Rail steel properties

5.7 Sleeper stresses

Contrary to the rail, the sleeper may neither be considered as a beam of infinite length nor as a short infinitely stiff beam. To ensure stability in the ballast bed the sleeper is only supported at the ends under the rails. Therefore, for calculation purposes, it is normal to presuppose that the contact force is distributed evenly over the contact surface as shown in Figure 5.21.

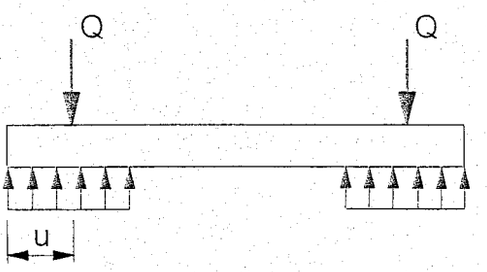


Figure 5.21: Assumed contact stress distribution on a sleeper

Using the dynamic amplification factor DAF (5.59), the maximum bearing force on a (single) discrete rail support due to the wheel load is:

$$F_{max} = DAF \cdot F_{mean} \tag{5.75}$$

wherein the mean value can be found using (5.2), (5.10), (5.17), and (5.20):

$$F_{mean} \approx \frac{Qa}{2L} = \frac{Q}{2}\sqrt[4]{\frac{k_d a^3}{4EI}} \tag{5.76}$$

DAF = dynamic amplification factor (for sleepers or other rail supports: t = 3);
Q = effective wheel load [kN];
a = sleeper spacing [m];
k_d = half support stiffness [kN/m];
EI = single rail bending stiffness [Nm²].

The sleeper spacing proves to have a relatively great influence on the support force. For the load situation depicted in Figure 5.21 the maximum moment is:

$$M_{max} = \frac{1}{4} F_{max} u \tag{5.77}$$

In unfavourable cases F_{max} may be of the same order of magnitude as the effective wheel load Q. Assuming Q = 125 kN, a = 60 cm, L = 70 cm (low value), V = 100 km/h, φ = 0.20, t = 3, then F_{max} = 95 kN.

Contact pressure between rail and sleeper

The mean contact pressure between rail and sleeper on the most heavily loaded sleeper can approximately be set at:

$$\sigma_{rs} = \frac{F_0 + F_{max}}{A_{rs}} \tag{5.78}$$

F_0 = total pretensioning force of fastening on rail support [N];
A_{rs} = effective rail support area of rail support [m²].
Note: In case of an intermediate baseplate between rail and sleeper: A_{rs} = baseplate area.

Permissible contact pressures between rail and sleeper are:

- Softwood sleepers:: $\sigma_{rs} \leq 1.0 - 1.5$ N/mm²;
- Hardwood sleepers::: $\sigma_{rs} \leq 1.5 - 2.5$ N/mm²;
- All concrete supports:: $\sigma_{rs} \leq 4$ N/mm².

5.8 Stresses on ballast bed and formation

5.8.1 Introduction

The ballast bed and formation are conceived as a two-layer system. The vertical stresses on the ballast bed and on the formation which are due to wheel loads will be considered as the determining stresses for the load-bearing capacity of the layer system. Overloading of the ballast bed causes rapid deterioration of the quality of the track geometry. Overloading of the formation raises the material in the ballast bed, especially in the case of materials susceptible to moisture. This phenomenon is known as pumping.

5.8.2 Vertical stress on ballast bed

The compressive stresses, which the sleepers exert on the ballast bed, are considered evenly distributed. The material from which the sleeper is made thus plays no role here. To determine the mean values of the stress the calculation is again based on Zimmermann's theory, whereas the dynamic amplitude is taken into account by Eisenmann's increment factor (5.59). The maximum stress between sleeper and ballast bed under a wheel load Q is expressed as:

$$\sigma_{sb\,max} = DAF \cdot \sigma_{sb\,mean} \tag{5.79}$$

in which, following equation (5.76):

$$\sigma_{sb\,mean} = \frac{F_{mean}}{A_{sb}} \approx \frac{Qa}{2LA_{sb}} = \frac{Q}{2A_{sb}}\sqrt[4]{\frac{k_d a^3}{4EI}} \tag{5.80}$$

DAF = dynamic amplification factor (for ballast: t = 2);
Q = effective wheel load [kN];
a = sleeper spacing [m];
A_{sb} = contact area between sleeper and ballast bed for half sleeper [m²];
k_d = half support stiffness [kN/m];
EI = single rail bending stiffness [Nm²].

Permissible contact pressure on the ballast bed: $\sigma_{sb} \leq 0.50$ N/mm²

Note: sleeper rotation can give rise to high local edge pressure. Sometimes this is taken into account by introducing an increment factor.

It can be gathered from equation (5.80) that sleeper spacing and the extent of the support area have a relatively important influence on the mean stress. A high value for the foundation coefficient leads to high values for the stress on the ballast bed. In certain cases, for example in the case of a ballast bed on a structure, the foundation modulus should be lowered. A heavier rail profile has a positive effect in this respect. Use of UIC 54 instead of NP 46, on the basis of (5.80), leads to a stress reduction in the ballast bed of about 10%.

Using (5.62) the rail foot stress can be calculated, and using (5.61) the stress between sleeper and ballast bed. For Q=170 kN, C_{sb} = 100 N/cm³, a = 60 cm, A_{sb} = 2850 cm², and making use of the rail data in Table 5.2, L results in 857 mm and the following data are obtained:

	Rail		Ballast	
	σ_r [N/mm²]	Ratio	σ_r [N/mm²]	Ratio
UIC 60	97	–	0.21	–
UIC 54	110	13 %	0.22	7 %
NP 46	137	41 %	0.25	18 %

Table 5.6: Stresses resulting from Q = 170 kN

A heavier rail profile has a great influence on rail stress reduction. The effect on ballast stress is approximately half of the effect on rail stress.

The relation between vertical stress and deterioration in the quality of track geometry is still ambiguous. On the basis of the AASHO Road Test for road structures, it is nevertheless assumed that:

Decrease in track geometry quality = (increase in stress on ballast bed)^m

in which m = 3 to 4. A 10% higher stress on the ballast bed thus leads to a 1.2 to 1.5 times faster reduction in track geometry quality and a proportional increase in maintenance.

5.8.3 Vertical stress on formation

In order to calculate the maximum vertical stress on the formation the contributions of the various sleepers have to be superimposed. Figure 5.22 shows the stress pattern on the ballast bed along the length of the track. For each sleeper the stress is assumed to be evenly distributed over the sleeper surface. The magnitude of this stress beneath the various sleepers caused by wheel load Q is:

$$\sigma_i = \sigma_{max}\eta(x_i) \qquad (5.81)$$

in which:

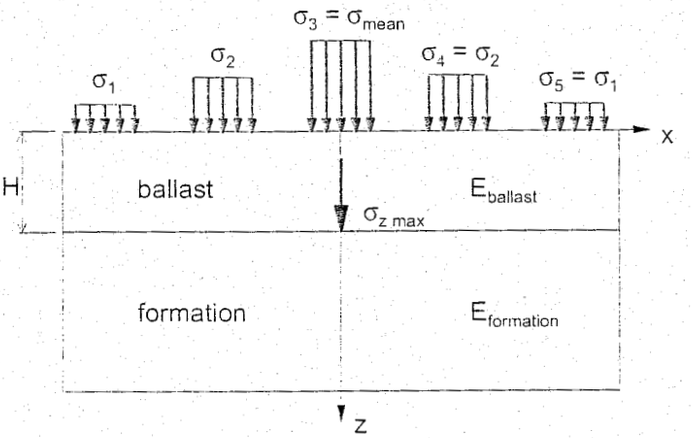


Figure 5.22: Ballast bed and formation represented as two-layer system

$$\sigma_{max} = DAF \cdot \frac{Qa}{2LA_{sb}} \qquad (5.82)$$

$$\eta(x_i) = e^{-x_i/L}\left[\cos\frac{x_i}{L} + \sin\frac{x_i}{L}\right] \qquad x_i \geq 0 \qquad (5.83)$$

To determine the vertical stress on the formation the value of factor t = 1 can be taken, as adjacent sleepers cannot all be subjected to an unfavourable load at the same time.

The evenly distributed stresses per sleeper are then replaced by equivalent strip loads covering the sleeper width. Assuming this, the problem can be described by the known two-dimensional stress distribution for a plane strain situation.

5.8.4 Odemark's equivalence method

Ballast bed and formation create a two-layer system as illustrated in Figure 5.22. Based on the given strip loads on the ballast bed, the thickness of the ballast bed, and the elasticity constants for both layers the maximum vertical stress on the formation can now in principle be determined. With the help of Odemark's equivalence method [58] much can be simplified as the two-layer system can be converted into a single-layer system. The maximum vertical stress on the formation in the actual two-layer system then correlates with the maximum vertical stress in the equivalent half space at a distance from the surface:

$$H_e = 0.9H\sqrt[3]{\frac{E_{ballast}}{E_{formation}}} \qquad (5.84)$$

in which:
H_e : equivalent ballast depth;
H : actual ballast bed depth under the sleeper;
$E_{ballast}$: modulus of elasticity of ballast bed;
$E_{formation}$: modulus of elasticity of formation.

If the ratio between $E_{ballast}$ and $E_{formation}$ is 3 and the ballast depth is 30 cm the equivalent depth is 39 cm.

The vertical stress in a half space loaded by an evenly distributed strip load according to Figure 5.23 can be determined using the two-dimensional theory of elasticity [265]. This compressive stress reads:

$$\sigma_{zi} = \sigma_i f(x_i) \tag{5.85}$$

in which:

$$f(x_i) = \frac{1}{\pi}\left[\alpha_1 - \alpha_2 + \frac{1}{2}(\sin 2\alpha_1 - \sin 2\alpha_2)\right] \tag{5.86}$$

$$\alpha_1 = \operatorname{atan}\frac{x_i + b/2}{H_e} \tag{5.87}$$

$$\alpha_2 = \operatorname{atan}\frac{x_i - b/2}{H_e}$$

In this manner, the contributions to the maximum vertical stress on the formation can be determined for each strip load according to:

$$\sigma_{zmax} = \sum_i \sigma_{zi} \tag{5.88}$$

Only a few of the strip loads in the vicinity of the maximum load need be taken into consideration because of the decrease in strip load according to (5.83) and the load spreading under a strip load according to (5.86).

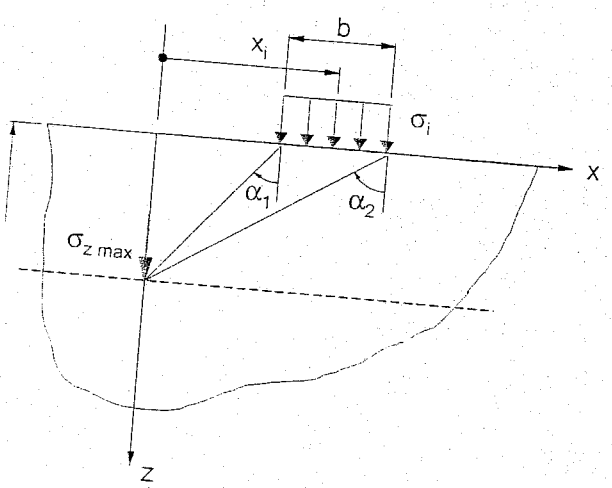


Figure 5.23: Stress due to striploads on half space

Permissible contact pressure on the formation:

The permissible compressive stress on the formation can be established using the following empirical formula according to Heukelom and Klomp [58]:

$$\bar{\sigma}_z = \frac{0.006 E_{v2}}{1 + 0.7 \log n} \qquad (5.89)$$

in which:
E_{v2} : modulus of elasticity taken from the second load step in a plate loading test;
n : number of load cycles.

The table below contains the permissible stresses according to (5.89) for 2 million cycles, with different values for E_{v2}. The table also indicates the order of magnitude of the foundation modulus C when using a 30 cm deep ballast bed with E = 150 N/mm².

Classification	E_{v2} [N/mm²]	C [N/mm³]	$\bar{\sigma}_z$ [N/mm²] n=2.10^6
Poor	10	0.03	0.011
	20	0.04	0.022
Moderate	50	0.07	0.055
Good	80	0.09	0.089
	100	0.11	0.111

Table 5.7: Permissible stresses on formation

When laying main line track, DB demands an E_{v2}-modulus of at least 120 N/mm² just beneath the ballast bed. If the measured values from the plate loading tests do not comply with this, an intermediate layer, referred to as sub-ballast layer with a depth according to Figure 5.24, must be inserted.

ORE Committee D117 compiled a "Design Handbook" [209] on the basis of a major series of measurements and calculations. Figure 5.25 gives the desired depth for various types of foundation consisting of a 25 cm ballast layer and one or more intermediate layers. In addition to the values recommended by D117, the curves used by DB, SNCF, and SBB are likewise plotted, as are those for heavily loaded track based on the CBR index derived from Alias [3].

Figure 5.25 also indicates the various scales used in practice for classification of formation, such as CBR (California Bearing Ratio), E_{v2}, and the classes according to UIC leaflet 719R. The UIC classification is presented in Table 5.8.

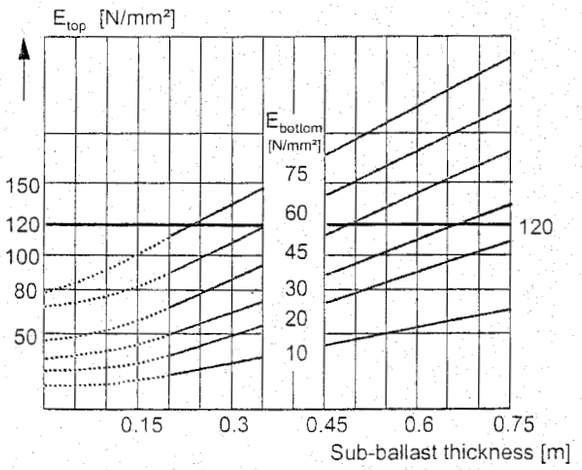


Figure 5.24: Thickness of sub-ballast layer according to DB standards

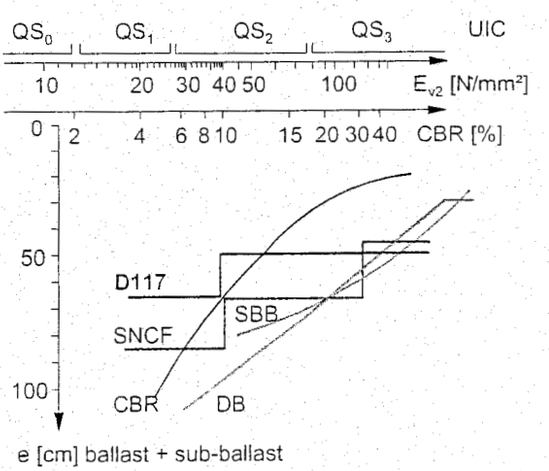


Figure 5.25: Data from ORE D117 'Design Handbook', RP28

5.9.3 Sleeper

Problem 4:
The design parameters of a new railway track structure are:

- Sleeper spacing : a = 60 cm;
- Characteristic length of the track : L = 90 cm;
- Repetitive wheel loads : Q = 125 kN;
- Train speed : V = 100 km/h;
- Track condition : 'good'; (t = 3);
- Baseplates on soft wooden sleepers are applied.

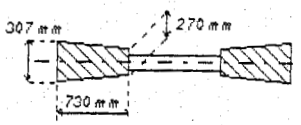
Question 4a:
Use the DAF method to calculate the maximum value of the support load between rail and sleeper F_{max}.

Question 4b:
If a uniform distribution of the compressive stress on the sleeper may be supposed to exist, calculate this stress σ_{bs}, using the following data:
- Clamping force per rail : F_0 = 20 kN;
- Area baseplate/sleeper : A_{bs} = 550 cm².

Question 4c:
What measures could be taken regarding the sleeper if the maximal compressive stress would be higher than the allowable value σ_{bs} = 1.5 N/mm²?

5.9.4 Ballast bed

Problem 5:
The following loading and structure data are given:

– One repetitive wheel load : Q = 100 kN;	– Foundation modulus : C_{sb} = 0.15N/mm³;
– Extra wheel load in curves : 20 %;	– Admissible stress on ballastbed : 0.3 N/mm².
– Maximum train speed : V = 100 km/h;	– Probability factor : t = 2;
– Rail profile : UIC 54;	effective area of twin-block sleeper 
– Sleeper spacing : a = 60 cm;	
– Half sleeper area on ballastbed: A_{sb} (see figure);	
– Track condition : 'good';	

Question 5a:
Calculate the maximum compressive stress on the ballast in straight track and check with the allowable value.

Question 5b:
Calculate the maximum compressive stress on the ballast in a curve and check with the allowable value.

Question 5c:
Consider a change in relevant track parameters to satisfy the compressive stress limit. Quantify in each case the relative effect on the maximal compressive stress (use a formula showing all relevant track parameters).

5.9.5 Temperature effects

Problem 6:
A very long rail is subjected to a temperature increase with respect to the neutral temperature ΔT. Due to the shear forces, acting through the fastenings on the rail, the free expansion of the rail is counteracted. In the analytical approach the shear forces may be replaced by an equivalent continuous shear resistance distribution.

Consider the case that the shear resistance distribution τ [N/m] is a linear function of the displacement u, which itself is a function of the x coordinate:

$$\tau = ku(x)$$

Question 6a:
What is the differential equation for this case?

Question 6b:
Derive the formula for the axial rail displacement as a function of x, accounting for the boundary conditions. The rail may be supposed to be a semi-infinite long beam (½ CWR track).

Question 6c:
Give the expression for the axial normal force in the rail.

Question 6d:
Which phenomenon can occur in the track at high temperatures?

Answers to exercises 2 to 6:
2a. extra wheel load is due to non-compensated lateral acceleration causing a force couple leading to wheel load increase on the high rail.
2b. $L = (4EIa/k_d)^{1/4} = (4\times21000\times2346\times60/120)^{1/4} = 99.6$ cm.
2c. $\sigma_{max} = \sigma_{gem}\times DAF$:
$\sigma_{gem} = QL/(4Wyf) = 1.2\times100\times99.6/(4\times313) = 9.55$ kN/cm² = 95.5 N/mm².
DAF = 1+ tφ (1+(V-60)/140) = > = 1+3×0.2(1+(140-60)/140) = 1.94.
=> $\sigma_{max} = 95.5\times1.94 = 185$ N/mm².
2d. max. fatigue stress (185) < admiss. stress $\sigma_f = 220$ N/mm² (from table)
2e. w_{max} prop. to L^3,prop. to $1/k_d^{3/4}$ => $(k_{d1}/k_{d2})^{3/4} = (1/2)^{3/4} = 0.59$ reduction;
σ_{max} prop. to L,prop. to $1/k_d^{1/4}$ => $(k_{d1}/k_{d2})^{1/4} = (1/2)^{1/4} = 0.84$ reduction;
p_{max} prop. to 1/L,prop. to $k_d^{1/4}$ => $(k_{d2}/k_{d1})^{1/4} = (2)^{1/4} = 1.19$ amplification.

3a. $\tau_{max} = 412\times(Q/r)^{1/2} = 412\times(110/430)^{1/2} = 208$ N/mm².
3b. τ_{max} becomes √2 higher, the other two have no effect;
3c. $Q_{max} = 4.90\times10^{-7}\times430\times700^2 = 103$ kN (admissible value exceeded).
3d. fatigue crack starting in the rail head; may cause shelling of rail head.
3e. - better rail steel $\sigma_t = 900$ N/mm² => $Q_{max} = (900/700)^2\times103 = 170$ kN;
- greater wheel radius => $r_{min} = 110/103\times430 = 459$ mm;
- load spreading over more wheels.
3f. $Q = Q_{inc\,max} = 8.29/4.90\times103 = 174$ kN.

4a. $F_{max} = F_{mean}\times$ DAF:
$F_{mean} = Qa/(2L) = 125\times60/(2\times90) = 41.7$ kN.
DAF = 1+3×0.2×(1+(100-60)/140) = 1.77.
$F_{max} = 41.7\times1.77 = 73.8$ kN.
4b. $\sigma_{bs} = (F_0 + F_{max})/A_{bs} = (73.8+20)/550 = 0.171$ kN/cm² = 1.71 N/mm² > 1.5 N/mm².
4c. Concrete instead of wooden sleepers; smaller sleeper spacing; increase area baseplate.

5a. $\sigma_{sb\,max} = \sigma_{sb\,mean}\times$ DAF
$\sigma_{sb\,mean} = F_{max}/A_{sb} = Qa/(2LA_{sb})$
$C_{sb} = 0.15$ kN/cm³ $A_{sb} = 73/2\times(30.7+27) = 2106$ cm²
$L = (4EIa/(C_{sb}A_{sb}))^{1/4} = (4\times21000\times2346\times60/(0.15\times2106))^{1/4} = 78.2$ cm.
$\sigma_{sb\,mean} = 100\times60/(2\times78.2\times2106) = 0.0182$ kN/cm² = 0.182 N/mm².
DAF = 1+2×0.2×(1+(100-60)/140) = 1.51.
$\sigma_{sb\,max} = 0.182\times1.51 = 0.275$ N/mm² (is lower than admissible value).
5b. in the curve: $Q_{eff} = 1.2\times100 = 120$ kN.
$\sigma_{max} = 1.2\times0.275 = 0.33$ N/mm² (now higher than admissible value).
5c. $\sigma_{sb\,max} = Q\times(C_{sb}a^3/(64\times EI\times A_{sb}^3))^{1/4}\times$ DAF
- V lower => DAF and $Q_{sb\,max}$ lower => effective, but not attractive;
- C_{sb} lower, small effect, prop. to $C^{1/4}$;
- a lower, effective, prop. to $a^{3/4}$, perhaps problems with maintenance;
- A_{sb} higher, effective, prop.to $A^{3/4}$, longer sleepers;
- I higher, small effect, prop. to $I^{1/4}$.

6a. $d^2u/dx^2 - ku/EA = 0$
6b choose origin of axle system and rail end.
Take as solution: $u = c\times\exp(-\mu x)$, flattening for x => ∞, c and μ are constants;
Insertion in the differential equation gives: $\mu^2 = k/EA$;
Substitution in the normal force gives: $N = EA[-\mu c\times\exp(-\mu x) - \alpha\Delta T]$;
With boundary cond. N(0) = 0 follows $c = -\alpha\Delta T/\mu$
The desired expression becomes: $u(x) = -\alpha\Delta T/\mu\times\exp(-\mu x)$
6c. Substitution of the constants gives: $N = EA\alpha\Delta T(\exp(-\mu x) - 1)$
6d. Horizontal deflection, possible buckling

5.10 Computer models

5.10.1 GEOTRACK program

If more detailed information is wanted on the behaviour and magnitude of local stresses and strains in the track components, the ballast layer, the sub-ballast layer and the subgrade, more complex computer models are necessary based on finite elements. GEOTRACK is a static program based on a three-dimensional multi-layer system with a track panel on top. Because it is an elastic program creep effects or settlements cannot be accounted for.
Figure 5.26 shows an example of a track model in GEOTRACK (because of symmetry only one half of the track is drawn).

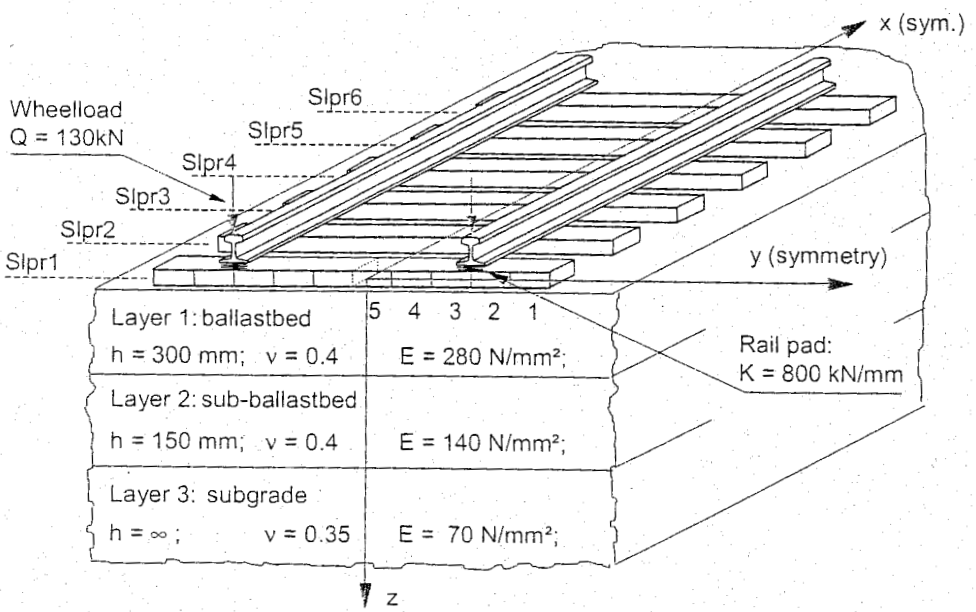


Figure 5.26: Example modelling rail track structure using GEOTRACK

In order to compare normal ballast track with some alternative track forms, GEOTRACK was used to test three cases as indicated in Figure 5.27. The input parameters are given in Table 5.9, while the main results are summarized in Table 5.10.

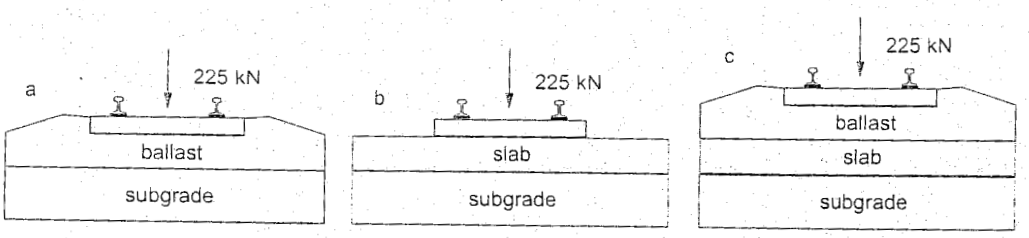


Figure 5.27: Comparison between various track forms

The following quantities were determined:

F_{max} : maximum rail support load;
w_{max} : maximum rail deflection;
σ_{max} : maximum vertical stress on subgrade.

The following conclusions may be drawn:

- The effect of a ballast bed thicker than 30 cm or slabs thicker than 20 cm is limited;
- Generally the presence of a slab will reduce the rail deflection, but higher rail support forces may be the result, especially in the case of normal soil;
- The effect of slabs thicker than 20 cm only reduces soil stresses significantly.

Input parameters	Type	Weight	E [N/mm²]	ν	k	Mass [t/m³]
axle load = 225 kN						
rail	UIC 54	54 kg/m	$2.1\ 10^5$	0.3		
sleeper	wood	100 kg	$1.0\ 10^4$	0.25		
rail pad stiffness = 1420 kN/m						
ballast			150	0.3	1	2.0
concrete slab			28500	0.2	1	2.5
normal subgrade			75	0.47	0.75	1.9
weak subgrade			8.7	0.49	0.75	1.1

Table 5.9: Input parameters for GEOTRACK program

	normal subgrade			weak subgrade		
Track structure	K_{max} [kN]	w_{max} [mm]	σ_{max} [N/cm²]	K_{max} [kN]	w_{max} [mm]	σ_{max} [N/cm²]
a1. ballast track 300 mm	49.2	1.54	8.97	37.3	7.58	3.72
a2. ballast track 400 mm	50.0	1.52	7.10	40.9	7.24	3.24
b1. slab track 200 mm	81.2	0.92	3.59	48.2	4.37	1.17
b2. slab track 300 mm	83.0	0.75	2.07	44.3	3.09	0.62
b3. slab track 400 mm	83.2	0.63	1.38	45.9	2.39	0.41
c1. ballast 300 on slab 200 mm	56.4	1.22	2.69	46.7	4.40	1.03
c2. ballast 300 on slab 300 mm	57.0	1.08	1.66	48.1	3.33	0.62
c3. ballast 300 on slab 400 mm	57.2	0.98	1.17	50.1	2.69	0.35

Table 5.10: Summary of results of GEOTRACK program

5.10.2 The ANSYS program

For the calculation of static forces especially within the ballastless track structures general-purpose commercial finite element software packages like ANSYS can be used. The calculation capabilities of ANSYS are quite sufficient for this kind of purposes and friendly Graphical User Interface (GUI) considerably facilitates the work.

ANSYS is a finite element analysis software which enables engineers to perform the following tasks:

- Build computer models or transfer CAD models of structures, products, components or systems;
- Apply operating loads or other design performance conditions;
- Study physical responses, such as stress levels, temperature distributions, or the impact of electromagnetic fields;
- Optimize a design early in the development process to reduce production costs;
- Do prototype testing in environments where it otherwise would be undesirable or impossible (for example, biomedical applications).

Features of the multi-purpose ANSYS program include finite element analysis capabilities for virtually all engineering disciplines - structural, mechanical, electrical, electromagnetic, electronic, thermal, fluid an biomedical.

Structure of the ANSYS program

The ANSYS program is organized into two basic levels (Figure 5.28):

- Begin Level;
- Processor (or Routine) Level.

The Begin Level acts as a gateway into and out of the ANSYS program. It is also used for certain global program controls such as changing the job-name, clearing (zeroing out) the database, and copying binary files. When the user first enters the program he is at the Begin Level.

At the Processor Level, several processors are available. Each processor represents a set of functions that perform a specific analysis task. For example, the general preprocessor is where the user can build a model, the solution processor is where he can apply loads and obtain the solution, and the general postprocessor is where the evaluation of the solution results is being performed. A general additional postprocessor also enables the user to evaluate solution results at specific points in the model as a function of time.

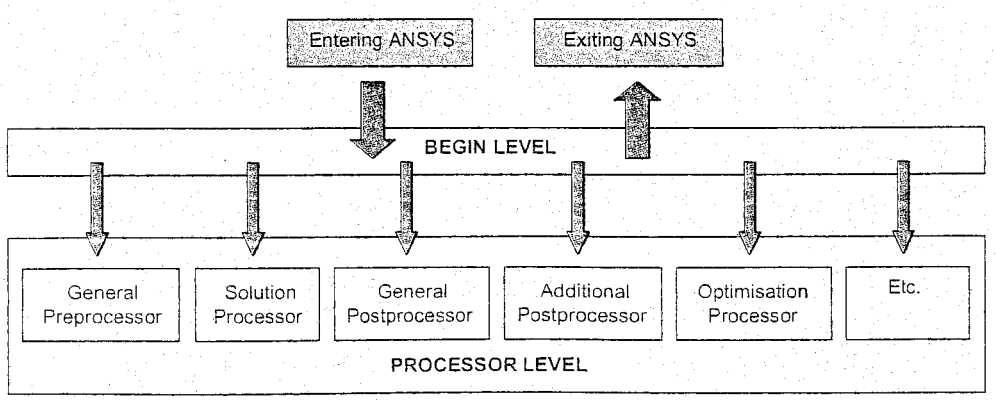


Figure 5.28: Structure of the ANSYS program

There are basically several ways of using or working with ANSYS.

- The first and the easiest one is running ANSYS in Interactive Mode, i.e. via the Graphical User Interface (GUI) in Figure 5.29. GUI gives the user easy, interactive access to program functions, commands, documentation and reference material. An intuitive menu system helps users navigate through the ANSYS program. User can input data using a mouse, a keyboard, or a combination of both.

- Another option is using ANSYS in a Batch Mode. In a Batch Mode, the user submits a file of commands to the ANSYS program (a script file). This script file may have been generated by previous ANSYS session (using the log file), by another program, or by creating a command file with the use of editor. Batch mode feature allows the user to run ANSYS in the background while doing some other work on a computer. Also, running ANSYS in a Batch Mode is extremely useful in cases when ANSYS is coupled with other programs for the purposes of performing more thorough analyses, and especially when this process is to be automatically performed and repeated many times as in case of multiple-iteration optimization processes.

- The third way is utilizing the Input Window. This is a window where the user can provide input to the program by directly typing in commands. It also shows prompts for functions that involve graphical picking. The Input Window also includes the History Buffer, which contains all previously entered commands and prompt responses, which can be then scrolled back and forth and re-executed.

- Finally, utilizing ANSYS Log Files represents another way of working with ANSYS. ANSYS program records every command it executes, whether typed in via Input Window or directly executed either by a function in the GUI or by pre-defined script file. These commands are recorded in two places: the session log file and the internal database command log. The session log file is a text file which is saved in the user's working directory. The database command log is saved in the ANSYS database (in memory). The user can copy the database log to a file at any time by choosing proper menu option or by issuing the proper command. Both files are command logs that can be used again as input (script files) to the ANSYS program.

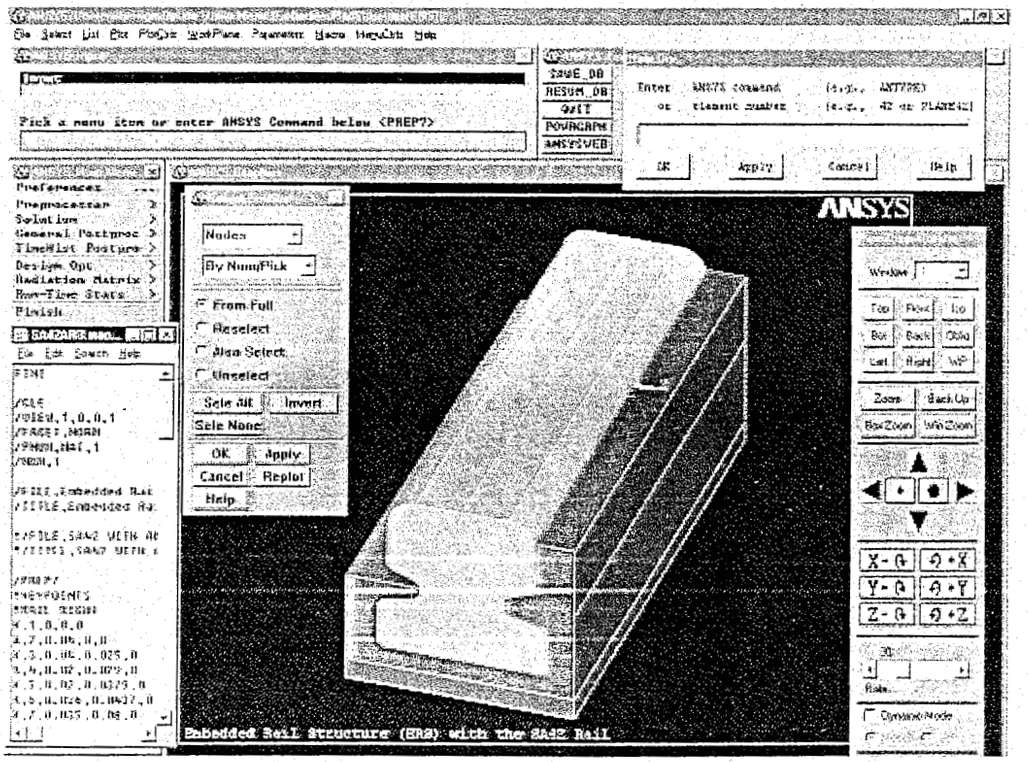


Figure 5.29: ANSYS Graphical User Interface (GUI)

5.11 Two Case ERS designs

5.11.1 Testing of the UIC54 ERS

Several case studies were performed at the Technical University Delft using the ANSYS software on ballastless track structures with a special concentration on the Embedded Rail Structure (ERS). Initial calculations of ERS using the ANSYS program were made with the "standard" ERS, i.e. the ERS with UIC 54 rail, Edilon prefab elastic strip nr. 102, and Edilon Corkelast compound VA60, as shown in Figure 5.30. The aim of this investigation was to devise and calibrate a numeric FEM-based model of ERS, which could be reliably used in future to reduce the number of long and expensive laboratory tests.

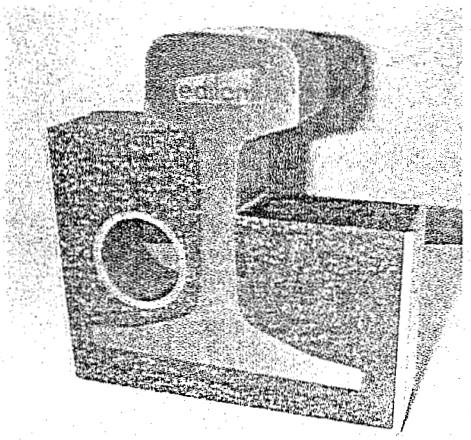


Figure 5.30: Standard UIC 54 Embedded Rail Structure

Test	Angle	Load	Type of Load
1a	0°	V=*P*	Static
1b	0°	V=*P*	Static & dynamic
2	22°	V=*P*; H=0,4•*P*	Static & quasi-static
3	31°	V=*P*; H=0,6•*P*	Static & quasi-static
4	0°	V=*P*	Static

Table 5.11: Tests determining the elasticity

The calibration of the model was performed by comparing the obtained numeric calculation results with the ones previously obtained in laboratory. Applied loads complied with the NS regulations regarding testing of the track elasticity of Embedded Rail Structures, i.e. with the loads applied in the vertical, sloped (22 and 31 degrees), and longitudinal direction, as shown in Table 5.11 and in Figure 5.31. Some of the essential results of this study were:

- The ANSYS FEM calculation could reliably describe the results of laboratory tests on stiffness and strength of ERS, hence it could be used instead of multiple testing of intermediate designs. Thus, the laboratory work can be restricted to only performing tests on the final design.

- The obtained results not only staid within a 5% margin, but were even less compared to the laboratory tests. This means that the ANSYS FEM calculations could be successfully used to reduce the laboratory tests. However, calibration of the FEM with a corresponding laboratory test still remains a necessary prerequisite in order to use it.

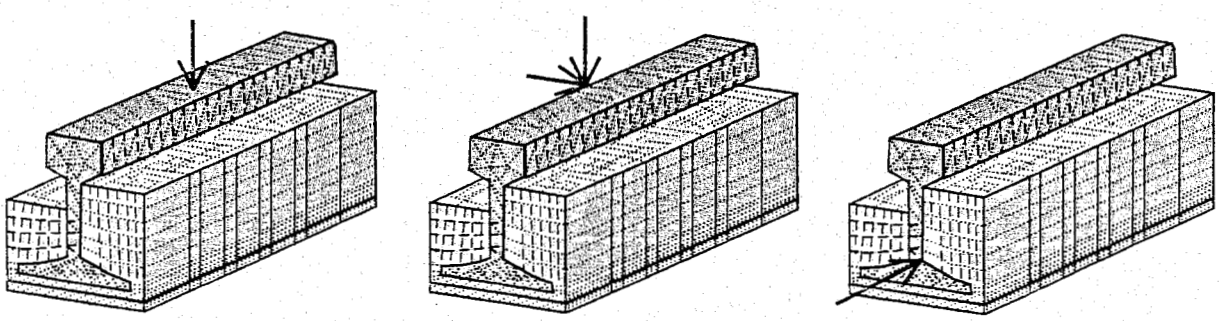


Figure 5.31: Testing with different loading cases

– In the Standard ERS with UIC54 rail, the elastic pad could not be replaced by a softer compound material due to the resulting high lateral displacements.

The successful completion of the UIC54 ERS prompted the continuation of the research, this time directed towards the innovative ERS structures, e.g. the ones with SA37 and SA42 rails and subsequently those with even more non-standard rail types and various compound characteristics. The reason for taking this direction was the firm belief that the standard UIC54 ERS offered a lot of opportunities for further optimization. The first two objectives were to investigate the possibility of replacing the elastic strip with a softer type of compound and reducing the total amount of the compound, both of which would reduce the cost of a structure and simplify the construction.

5.11.2 Testing of the SA42 ERS

The first step in testing SA42 ERS was to perform the verification of the FEM numeric model. In order to verify the numerical model, the available testing results of the 250 mm long sample of ERS with a low-noise rail SA42, compound VA90, and elastic strip fc6-sp have been used, Figure 5.32.

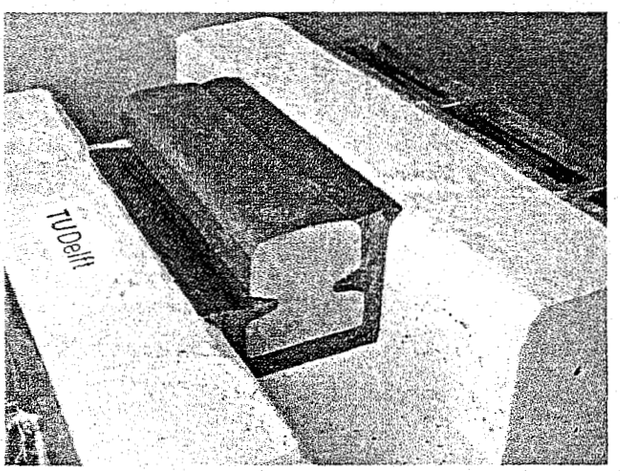


Figure 5.32: Laboratory testing of ERS at Delft University of Technology

A 3-D model of this structure created in ANSYS is shown in Figure 5.33a. E-moduli of the compound and strip have been determined by fitting the responses of the numerical model into the results of the laboratory tests. One of these responses is the vertical displacement of the rail which should be 2.1 mm if a vertical load of 30.4 kN is applied.

SA42, VA90 and fc6-sp sample length 250 mm	Vertical displacement at loading point [mm]			Maximum Von Mises stresses in compound [N/mm²]		
Loading types	2-D results	3-D results	Ratio	2-D results	3-D results	Ratio
Vertical load 30.4 kN	2.29	2.11	1.08	3.98	4.01	0.99
Angular load (22 degrees) Vertical component 30.4 kN Lateral component 12.2 kN	2.70 lateral: 0.40	2.48 lateral: 0.41	1.09 lateral: 0,99	4.97	4.89	1.02

Table 5.12: 2-D and 3-D calculation results of SA42 ERS in ANSYS

Although the 3-D model can accurately describe the experiment, due to high computational costs it was practically impossible to use it in the optimization process where multiple evaluations of the structural responses are required. That is why a 2-D model has been created (Figure 5.33b) assuming that in the middle of the ERS sample the model satisfies the plane strain condition. The results of the 2-D model have been compared with those obtained using the 3-D model. They are collected in Table 5.12 which shows that a difference of about 10% exists between the calculated vertical displacements of the 2-D and 3-D models.

This comparison has proven to be very stable regardless of the geometry of ERS, the type of elements, and the mesh density used in the numerical model. A comparable relationship was also obtained for the ERS model and the UIC54 rail. The obtained ratio was then used as correction-factor during the optimization in order to estimate the displacements of the 3-D model based on the results of the 2-D model.

6 DYNAMIC TRACK DESIGN

6.1 Introduction

When dealing with track mechanics most of the problems are related in one way or another to dynamics. The dynamic interaction between vehicle and track can be described reasonably well in the vertical direction using mathematical models. Figure 6.1 gives an example of such a model made up of a discrete mass-spring system for the vehicle, a discretely supported beam to describe the track, and a Hertzian spring acting in the wheel/rail contact area.

Dynamic behaviour occurs in a fairly wide band ranging from very low frequencies of the order of 0.5-1 Hz for lateral and vertical car body accelerations to 2000 Hz as a consequence of geometrical irregularities in rails and wheel treads. The suspension system between wheelset and bogie is the first spring/damper combination to reduce vibrations originating from the wheel/rail interaction and is therefore called primary suspension. The reduction of the vibrations of lower frequency is dealt with in the second stage between bogie and car body and is called secondary suspension. This terminology can be applied to the track part of the model in the same way. The railpad and railclip represent the primary suspension of the track and the ballast layer or comparable medium represent the secondary suspension of the track.

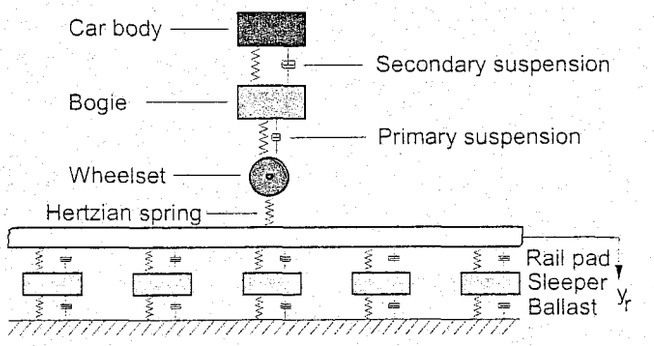


Figure 6.1: Dynamic model of vehicle-track interaction

Actual dynamic calculation is, however, extremely complex and is by no means generally accessible. Most analyses are limited to quasi-static considerations. Real dynamic problems are for the most part approached in a very pragmatic way by carrying out measurements.

In this chapter attention is given to the basic ingredients of the dynamic behaviour of railway track. Section 6.2 deals with some fundamental aspects. The 1-mass spring system, presented in Section 6.2.2, can be regarded as the most elementary system with the aid of which a number of practical problems can be considered. Extensions can be made in two directions: the construction can be enhanced to a multi degree of freedom system, and the load can be made more complex in terms of impact loads, and loads with a random character.

In Section 6.3 the track is modelled with relatively simple beam models consisting of the beam on an elastic foundation, a double beam, and a discretely supported track structure. The transfer function between track load and track displacement is discussed. Also the effect of a moving load running on the track is considered, as the track is considered to be infinitely stiff.

Track and rolling stock should in fact not be considered separately, but as one consistent system. For this reason the interaction between vehicle and track is introduced here without going into all the details required for a full treatment of this complex matter. After the introduction of the Hertzian spring, the physics of which were discussed earlier in Chapter 2, the transfer function between wheel and rail is derived in Section 6.4. This relationship plays an important role when interpreting track recording car data.

In Section 6.5 a concept is developed from which the relevant vehicle reactions can be calculated in real time using transfer functions based on track geometry measured independently of speed. A transfer function represents the contribution made by a geometry component to a vehicle reaction in the frequency domain. Geometry components include cant, level, alignment, and track gauge, and vehicle reactions include Q forces, Y forces, and horizontal and vertical vehicle body accelerations.

The theory of spectral analysis and the MISO method based on it are dealt with in Section 6.6. Estimates obtained from both models and MISO are compared. In Section 6.7 the Vehicle Response Analysis (VRA) system used for on-line response calculations is discussed. The relationship between Sperling's Ride Index and ISO weighted accelerations is discussed in Section 6.8.

Finally, in Section 6.9 and Section 6.10 a number of more advanced dynamic models are presented. A short description is given of a number of models, which were developed recently at the TU Delft, covering the dynamic effects of high-speed interaction, wave propagation in soft soils, and the effects on the dynamic vehicle response and of stiffness transitions near bridges and other fixed installations.

6.2 Dynamic principles

6.2.1 In general

When considering dynamic aspects of track one should realise that dynamics is in fact the interaction between load and structure. Loads vary in time and the way this happens determines the character of the load. Generally speaking, distinctions can be made between periodic loads, impact loads, and stochastic loads.

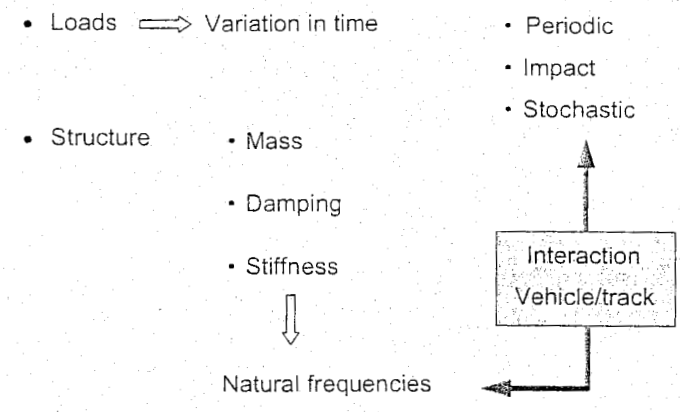


Figure 6.2: Dynamic aspects

Structures are characterized by their frequency response function which is governed by mass, damping, and stiffness. These parameters determine the natural frequencies of the structure, in other words, those frequencies in which the structure likes to vibrate (Figure 6.2). If the loads contain frequency components corresponding to the natural frequencies of the structure, large dynamic amplifications may occur. The general term used for this is resonance.

Dynamic behaviour is governed by the law of momentum, which says that force equals the change of momentum, or:

$$F = \frac{d(mv)}{dt} \tag{6.1}$$

When the mass is constant this equation transforms into Newton's law which implies that force equals mass times acceleration:

$$F = ma \tag{6.2}$$

6.2.2 One-mass spring system

The simplest dynamic model is the so-called discrete one-mass spring system with mass m, spring constant k, and damping c presented in Figure 6.3. This single degree of freedom system is governed by the following differential equation for the vertical displacement w(t):

$$m\ddot{w} + c\dot{w} + kw = F(t)^{1} \tag{6.3}$$

1. Note that due to conventions the symbols m, c, and k may also be used later for continuous systems as well.

The solution of (6.3) may be found in the time domain using the convolution integral involving the unit impulse response function $h(\tau)$, which is the response to a unit impulse input on the system, resulting in:

$$w(t) = \int_0^t F(\tau)h(t-\tau)d\tau \qquad (6.4)$$

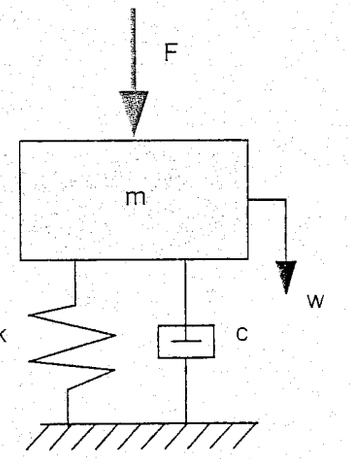


Figure 6.3: One-mass spring system

Frequency response function

From (6.4) it is obvious that the steady-state response calculation in the time domain is quite complicated. A better way to solve (6.3) is, even in this simple case, to use the Fourier transform[1] to determine the so-called frequency response function H(f) or transfer function[2] of the system. This function describes the relationship between response and excitation in the frequency domain. When the force as function of time is known, a Fourier transform can be made and the response then simply follows from a multiplication of the transformed force by the transfer function. In the frequency domain the Fourier transform of the displacement w(x) is:

$$W(f) = H(f)F(f) \qquad (6.5)$$

For the response to a unit impulse load F(t) = δ(t) on the one-mass system the Fourier transform results in the solution:

$$H(f) = \frac{W(f)}{F(f)} = \frac{1/k}{1-\frac{f^2}{f_n^2}+i2\zeta\frac{f}{f_n}} \qquad (6.6)$$

in which:

$$f_n = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \qquad (6.7)$$

$$\zeta = \frac{c}{2\sqrt{km}} \qquad (6.8)$$

are the *undamped natural frequency* and the *damping ratio*, respectively.

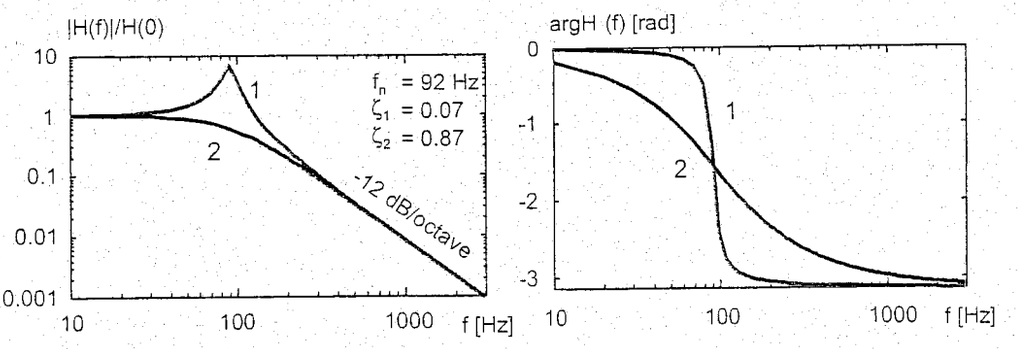


Figure 6.4: Flexibility function of a second order: two damping cases

1. In section 6.6 a thorough treatment is given of the principle of the Fourier transform technique.
2. Note that the *frequency response function* H(f), defined as the *Fourier transform* of h(t), is formally a special case of the *transfer function* H(p) = H(a+ib), defined as the *Laplace transform* of h(t), with a = 0 and b = ω = 2πf.

The static value of (6.6) is found by setting f = 0:

$$H(0) = 1/k \tag{6.9}$$

Writing (6.6) in complex polar notation gives the frequency response function in terms of a gain factor |H(f)| and a phase factor ϕ(f) as follows:

$$H(f) = |H(f)|e^{-i\phi(f)} \tag{6.10}$$

where:

$$H(f)| = \frac{1/k}{\sqrt{\left(1-\frac{f^2}{f_n^2}\right)^2 + 4\zeta^2\frac{f^2}{f_n^2}}} \tag{6.11}$$

$$\phi(f) = -atan\left[\frac{2\zeta\frac{f}{f_n}}{1-\frac{f^2}{f_n^2}}\right] \tag{6.12}$$

Other names for (6.11) are the modulus, absolute value, etc., while (6.12) is also called argument or angle. Because in this case the input is a force and the response is a displacement, the frequency response function is also called the *flexibility function*.

The example given in Figure 6.4 was obtained using (6.6) to (6.9), with m = 120 kg, k = 4.010^7N/m, and two damping values i.e. $c_1 = 1.010^4$ and $c_2 = 1.210^5$ Ns/m.

If $f >> f_n$ then |H(f)|/H(0) approaches the asymptote $1/(f/f_n)^2$ which characterizes a second order system. The slope of the asymptote is often expressed in dB/octave. In this case we have $20\log(1/(f/f_n)^2)$ = –40dB/decade or –12dB/octave. For high frequencies the phase ϕ approaches π rad.

Simple vehicle-track interaction

In case of vehicle-track interaction excitations are induced by irregularities in the wheel rail interface. In the most simplified form this can be considered as forced displacements, the principle of which is illustrated in Figure 6.5.

The equation of free motion reads in this case:

$$m\ddot{w} + c\dot{w} + kw = c\dot{y} + ky \tag{6.13}$$

or alternatively:

$$m\ddot{z} + c\dot{z} + kz = -m\ddot{y} \tag{6.14}$$

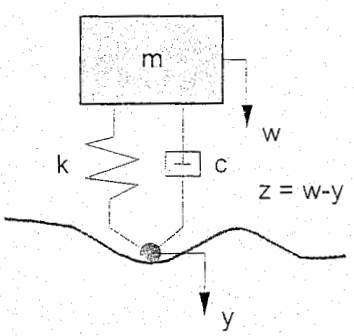


Figure 6.5: Forced displacement

This equation has the same form as (6.3) and thus the same solution techniques are applicable.

It is interesting, for instance, to determine the transformation function with y(t) as input and w(t) as output function, respectively.

behaviour, while experiments show that the elasto-plastic with softening behaviour is typical for well consolidated ballast. Both numerical models account for the effect of vertical forces by introducing the Mohr-Coulomb criterion. The element shows linear-elastic behaviour until the applied lateral load s has reached the peak value $s_{max} = F_p - s_v tan\varphi$ (s_v - vertical load, $tan\varphi$ - friction coefficient between ballast and sleeper) and corresponding peak displacement W_p. After that, the element begins to yield, that is the deformations increase without increasing the force s, approaching its limit value F_l as shown in Figure 4. The softening branch is approximated by the function

$$s = s_{max}[s_{lim} + (1 - s_{lim}) 2^{w/W_l}] \tag{7.77}$$

in which $s_{lim} = F_l / F_p$ (the limit value W_l is defined so that $s(W_l) = 0.5(F_p - F_l)$). The b-linear model can be considered as a particular case of the elasto-plastic with softening when $F_p = F_l$ and $W_p = W_l$.

The model has been intensely tested within the framework of the ERRI project D202 regarding analysis of CWR track stability and formulation of safety criteria [65]. One application of the model (Figure 7.43) in the investigation of tram track stability is presented in Section 7.5.8. Another application of the problem of determining temporary speed limits of trains on tracks after renewal or full maintenance is described in Chapter 13.5.

7.5.5 Vertical stability of track

One of the advantages of ballastless track structures such as an Embedded Rail Structure (ERS) is the absence (or very small chance) of lateral buckling. On the other hand, ERS can become unstable in the vertical direction. Two models for analysis of vertical buckling of embedded rail structure are shown in Figure 7.45. The model takes different types of vertical misalignment into account (Figure 7.45). A train in such models is represented by concentrated forces corresponding to axle loads applied to the rails. The vertical behaviour of elastic mixture (in which the rails are embedded) is modelled by the same spring element as has been used for modelling the lateral behaviour of ballast in the previous section (Figure 7.44).

It should be noted that the library of wizards is constantly extended and models of other track structures can easily be added to a standard version of CWERRI. The buckling mechanism of CWR tracks and safety criteria are discussed in the next section.

7.5.6 Buckling mechanism

Experiments and field observations have shown that, at a critical temperature increase $T_{B,MAX}$ (measured with respect to the neutral temperature), the track can, under certain conditions, buckle out into a new position C over a certain length. An additional temperature increase of the rails will gradually increase the deflection in the deformed region. The temperature increase versus the maximum lateral deflection is shown in Figure 7.46. In this case, an explosive buckling is assumed (a monotonous progressive response also exists).

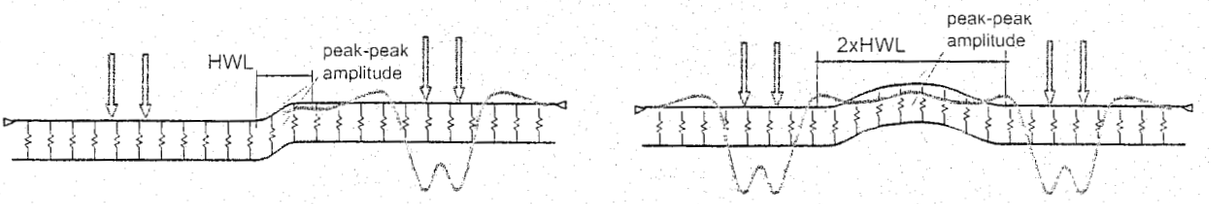


Figure 7.45: Models for vertical buckling analysis of embedded rail structure

7.5.4 Track lateral behaviour

Numerical models implemented in CWERRI can analyse the track behaviour in lateral direction as well. A model with which the lateral behaviour of a classical track can be analysed is shown in Figure 7.43.

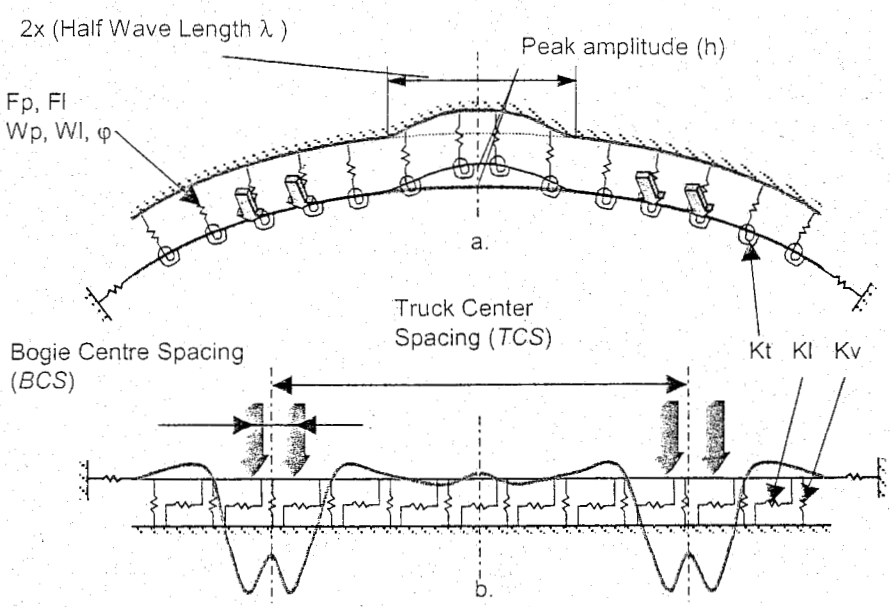


Figure 7.43: CWERRI model for stability analysis. Top (a) and side view (b).

The lateral behaviour is modelled by non-linear spring (parameters - Fp, Fl, Wp, Wl and φ), longitudinal, vertical and torsional behaviour - by linear springs (Kl, Kv, and Kt)
This model can be used for the analysis of straight as well as curved tracks, which means that the effects of centrifugal forces in rails caused by a moving train are taken into account. A track can have a horizontal misalignment in the middle approximated in the model by a sine function which is characterised by a half wavelength λ and peak-to-peak amplitude (h), as shown in Figure 7.43a.

The rails are modelled using 3-D elastic beam elements. In this model the linear longitudinal and rotational behaviour of a track is assumed and, therefore, finite elements have been used for ballast, fasteners, and linear spring (respectively stuffiness Kl and Kt in Figure 7.43). The beam elements representing the rails are supported by linear springs with the stiffness Kv, according to the Winkler foundation model. The static vertical deformations of the track due to four axle loads representing a train are denoted by a dashed line in Figure 7.43b.

Lateral behaviour of a classical track is to a large extent defined by lateral resistance of ballast. An experiment has shown that lateral behaviour of ballast is non-linear (e.g. a single sleeper test [298]). That is why an elasto-plastic spring element is used to model ballast behaviour in the lateral direction [282]. The element can describe two types of non-linear behaviour, namely elasto-plastic with softening or bi-linear, as shown in Figure 7.44. The bi-linear model is usually used to describe ballast lateral

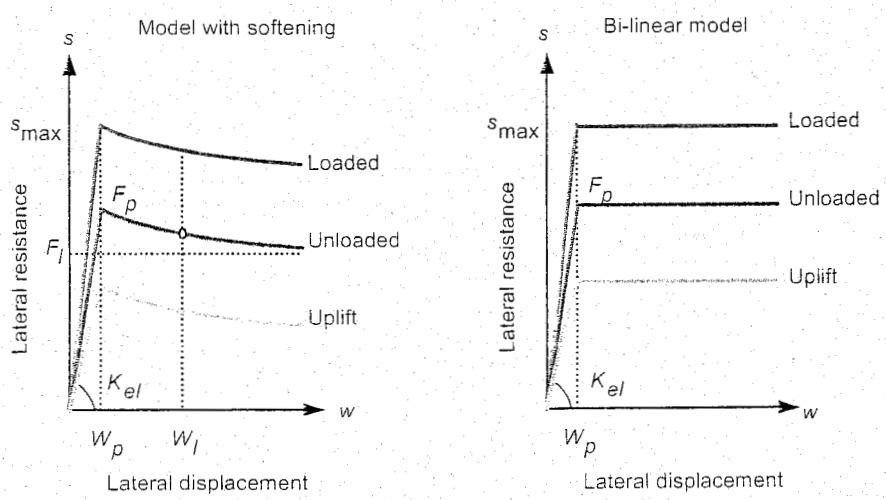


Figure 7.44: Model of lateral behaviour of ballast

By using the Fourier transformation of the response displacement w(t) for a unit impulse rail displacement y(t)=δ(t), in equation (6.13), we arrive at:

$$H_{y\to w}(\omega)=\frac{k+ci\omega}{k+ci\omega-m\omega^2} \qquad (6.15)$$

Using the expressions (6.7), (6.8) and using $\omega = 2\pi f$, we can then write:

$$H_{y\to w}(f)=\frac{1+i2\zeta\frac{f}{f_n}}{1-\frac{f^2}{f_n^2}+i2\zeta\frac{f}{f_n}} \qquad (6.16)$$

The magnitude and phase of this function are given in Figure 6.6 using the indicated values for f_n and ζ. This time linear scales are used. The relation between the forced frequency f and the wavelength λ of the rail roughness is $f = v/\lambda$ in which v is the speed of the vehicle. The λ-axis is also drawn in the figure assuming a speed v = 50 m/s (180 km/h).

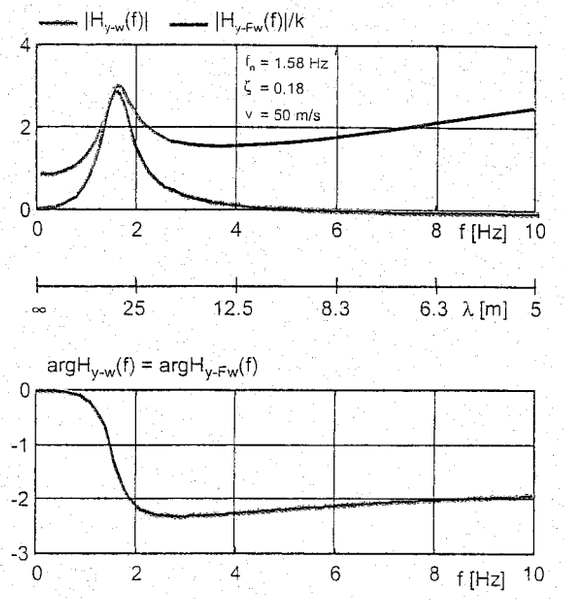


Figure 6.6: Transfer functions of vehicle model

Another interesting quantity, in this respect, is the transfer function for the wheel force F_w in the contact point. This force follows directly observing Figure 6.5:

$$F_w(t) = kz + c\dot{z} \qquad (6.17)$$

The transfer function in this case is therefore:

$$H_{z\to F}(\omega) = k + ci\omega \qquad (6.18)$$

To establish the transfer function with y as input, we also need the transfer function from equation (6.14):

$$H_{y\to z}(\omega)=\frac{m\omega^2}{k+ci\omega-m\omega^2} \qquad (6.19)$$

The transfer function between the input displacement y and the output F_w is now simply the product of (6.18) and (6.19):

$$H_{y\to F}(\omega)=\frac{m\omega^2(k+ci\omega)}{k+ci\omega-m\omega^2} \qquad (6.20)$$

Using the expressions (6.7), (6.8), and $\omega = 2\pi f$, we can write:

$$H_{y\to F}(f) = k\frac{f^2}{f_n^2}\frac{1+i2\zeta\frac{f}{f_n}}{1-\frac{f^2}{f_n^2}+i2\zeta\frac{f}{f_n}} \qquad (6.21)$$

This transfer function, divided by k, is also depicted in Figure 6.6.

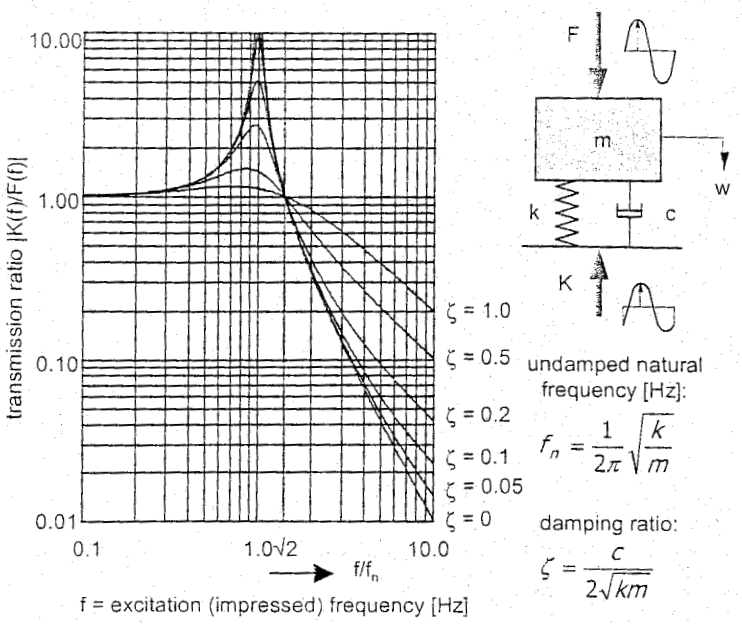


Figure 6.7: Principle of vibration reduction

If the transformation of the input function y(t) is obtained analytically or by means of measurements, the corresponding transformation for the displacement or the wheel force can easily be calculated in the frequency domain followed by back transformation to the time domain. To the result of the dynamic wheel load the static wheel load (= mg) should be added to get the total wheel force.

In Figure 6.7 an example is worked out in which the vibration transfer from a periodic load to the base of the 1-mass spring system is calculated. This is the situation encountered in vibration reduction problems regarding track in tunnels in densely populated areas. The lower the natural frequency of the 1-mass spring system the better the excitation frequency components are reduced or, in other words, are filtered out.

6.2.3 Wheel/rail forces due to dipped rail joints

The wheel rail forces, which act as dynamic loads to the track, follow from the interaction between vehicle and track. If high frequency phenomena in relation to loads caused by imperfections in wheel and rail geometry are the objective of the analysis, the so-called P_1 and P_2 forces have to be considered as displayed in Figure 6.8.

The forces are described by the following equations, as given by Jenkins [137] (see also Cope [37]):

$$P_1 = P_0 + 2\alpha v \sqrt{\frac{k_H m_{T1}}{1 + m_{T1}/m_u}} \qquad (6.22)$$

$$P_2 = P_0 + 2\alpha v \sqrt{\frac{m_u}{m_u + m_{T2}}} \left[1 - \frac{c_T \pi}{\sqrt{k_{T2}(m_u + m_{T2})}}\right] \sqrt{k_{T2} m_u} \qquad (6.23)$$

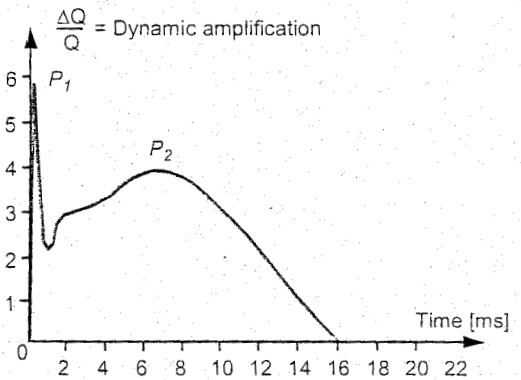


Figure 6.8: Wheel/rail impact forces

in which:

P_0	= static wheel/rail contact force (identical to Q_{st})	[kN]
2α	= total dip angle at joint	[rad]
v	= train speed	[m/s]
k_H	= linearised Hertzian contact stiffness (see also equation (2.12))	[N/m]
m_{T1}	= effective track mass for P_1 calculation	[kg]
m_u	= unsprung mass	[kg]
m_{T2}	= equivalent track mass for P_2 calculation	[kg]
k_{T2}	= equivalent track stiffness for P_2 calculation (identical to k_{tot} (5.5))	[N/m]
c_T	= equivalent track damping for P_2 calculation	[Ns/m]

An approximate formula is given by Alias [3]:

$$P \approx P_0 + \lambda 2\alpha v\sqrt{km_u} \tag{6.24}$$

in which:

λ = dimensionless factor depending on track damping (≤ 1);
k = track stiffness [N/m].

6.2.4 Track excitation origins

The excitations originated from track induced forces are summarized in Figure 6.9. For short waves in the order of centimetres rail corrugation, wheel irregularities, and weld imperfections are the most important factors. For such high frequencies (up to 2000 Hz) the Hertzian contact spring between wheel and rail should be taken into consideration.
Rail rolling defects have wavelengths in the order of 3 m. The natural frequency of bogies is in the order of 20 - 25 Hz, whereas the sprung mass has a lowest natural frequency between 5 and 0.7 Hz and therefore primarily influences ballast and formation.

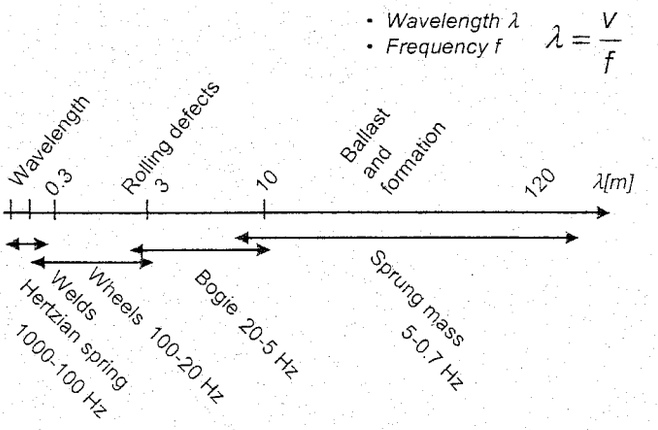


Figure 6.9: Track excitation in relation to frequency

6.3 Track modelling

6.3.1 General considerations

The type of interaction phenomenon to be described determines to a large extent the structure of the model and the degree of refinement. In the case under consideration here, the track is of prime importance. The quality of the track, however, is assessed on the basis of vehicle reactions.

To describe high frequency vertical vibrations associated with corrugations and poor quality welds use is made of a track model and a wheelset connected in the wheel/rail contact area by means of a Hertzian spring.

Vehicle reactions associated with low frequency track geometry are analysed on the basis of a number of very simple models in which track stiffness is not taken into consideration. Running characteristics and aspects relating to stability are totally disregarded in such analyses.

6.3.2 Transfer function between track load and track displacement

The track structure can be seen as a system consisting of rails which are elastically supported by means of rail pads by sleepers spaced at a fixed distance. The sleepers are supported by a damped elastic foundation consisting of ballast plus formation [102], [33].

Figure 6.10 shows three models which can be used to describe the track structure. The parameter values used in these models are shown in Table 6.1.

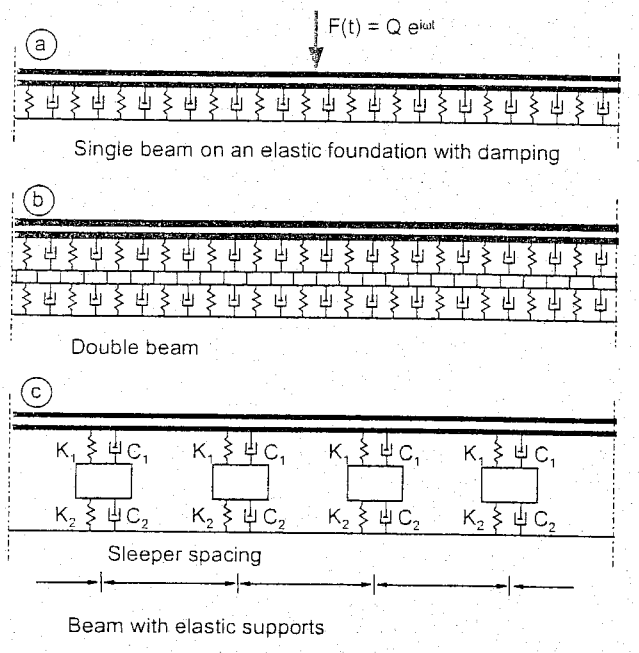


Figure 6.10: Different models for describing dynamic properties

EI	$= 4.5.10^6$	Nm^2	rail bending stiffness
M_w	= 350	kg	unsprung wheel mass
k_H	$= 1.4.10^9$	N/m	Hertzian contact spring stiffness
m	= 119	kg/m	track mass (single beam)
m_1	= 54.43	kg/m	rail mass (double beam)
m_2	= 157	kg/m	sleeper mass (double beam)
k	$= 4.10^7$	N/m^2	track stiffness (single beam)
k_1	$= 2.5.10^8$	N/m^2	pad stiffness (double beam)
k_2	$= 4.10^7$	N/m^2	foundation stiffness (double beam)
c	$= 1.2.10^5$	Ns/m^2	track damping (single beam)
c_1	$= 9.10^4$	Ns/m^2	pad damping (double beam)
c_2	$= 1.2.10^5$	Ns/m^2	foundation damping (double beam)

Table 6.1: Survey of different track parameter values used in the dynamic models

6.3.3 Beam on an elastic foundation

The simplest dynamic track model consists of a beam on an elastic foundation with a continuously distributed stiffness and damping, as shown in the drawing in Figure 6.10(a). The beam is considered an Euler beam in which the rotational inertia and shear force deformation are disregarded. The load exists of a concentrated load at x = 0 with a magnitude of:

$$F(t) = Qe^{i\omega t} \tag{6.25}$$

where

$$\omega = 2\pi f \tag{6.26}$$

The equilibrium conditions of a beam element are, considering Figure 6.11:

$$\frac{\partial D}{\partial x}dx = \left(m\frac{\partial^2 y}{\partial t^2} + c\frac{\partial y}{\partial t} + k\,y\right)dx \tag{6.27}$$

$$D dx = \frac{\partial M}{\partial x}dx \tag{6.28}$$

The constitutive equation reads:

$$M = -EI\frac{\partial^2 y}{\partial x^2} \tag{6.29}$$

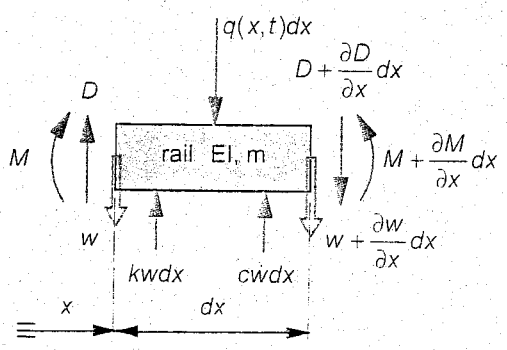


Figure 6.11: Beam element model (dynamic)

From equations (6.27) to (6.29) the differential equation of the problem can be derived:

$$EI\frac{\partial^4 y(x,t)}{\partial x^4} + m\frac{\partial^2 y(x,t)}{\partial t^2} + c\frac{\partial y(x,t)}{\partial t} + k\,y(x,t) = 0 \tag{6.30}$$

The solution of this differential equation can be found by splitting variables:

$$y(x,t) = w(x)e^{i\omega t} \tag{6.31}$$

Thus (6.30) can be rewritten in short notation for spatial derivatives to x:

$$EIw^{IV} + [k - m\omega^2 + i\omega c]\,w = 0 \tag{6.32}$$

This differential equation is very similar to the equation (5.15) for the statically loaded and elastically supported beam. However, the foundation coefficient k in (5.15) must be replaced by the complex coefficient k* according to:

$$k^* = k - m\omega^2 + ic\omega \tag{6.33}$$

or, in polar form:

$$k^* = \sqrt{([k - m\omega^2]^2 + c^2\omega^2)}\; e^{i\,\mathrm{atan}\frac{c\omega}{k-m\omega^2}} \tag{6.34}$$

The displacement w(x) is now also a complex quantity. The boundary conditions are the same as in (5.16) for the static case. For further considerations it is important to know the flexibility function $H_r(f)$ between displacement and load at x = 0 according to:

$$H_r(f) = \frac{w(0)}{Q} = \frac{1}{8EI\beta^3} \tag{6.35}$$

wherein:

$$\beta^4 = \frac{k^*}{4EI} \tag{6.36}$$

Using the substitutions for f_n (6.7) and ζ (6.8) and the formulae (6.34) and (6.36) the flexibility function according to (6.35) can now be written as:

$$H_r(f) = \frac{1}{2kL}\left\{\left(1 - \frac{f^2}{f_n^2}\right)^2 + 4\zeta^2\frac{f^2}{f_n^2}\right\}^{-\frac{3}{8}} e^{i\phi} \tag{6.37}$$

in which:

$$\phi = -\frac{3}{4}\,\mathrm{atan}\left[\frac{2\zeta\frac{f}{f_n}}{1 - \frac{f^2}{f_n^2}}\right] \tag{6.38}$$

In equation (6.37) L is the characteristic length (5.20). If f is zero (static case) (6.37) becomes identical to inverse total spring constant $1/k_{tot}$ according to (5.45):

$$H_r(0) = \frac{1}{2kL} = \frac{L^3}{8EI} \qquad (6.39)$$

The modulus of the transfer function is given by the term preceding the exponential function in formula (6.37), whereas (6.38) describes the argument.

Figure 6.12 shows the result obtained using (6.35) and (6.39) with the same values as used fore the simple beam.

The correlation between this solution and that of the 1-mass spring system Figure 6.4 is striking. Both transfer functions have the same form (for ζ = 0.87). However, compared to the 1-mass spring system, the order of the distributed system is 1.5 instead of 2. Therefore the slope of the asymptote in this case is $20\log(1/(f/f_n)^{3/2})$ = -30dB/decade or –9dB/octave.

The phase approaches $-3/4\pi$ instead of π for high frequencies.

The complex characteristic length λ is obtained from (6.35), defined by:

$$\lambda = \frac{1}{\beta} \qquad (6.40)$$

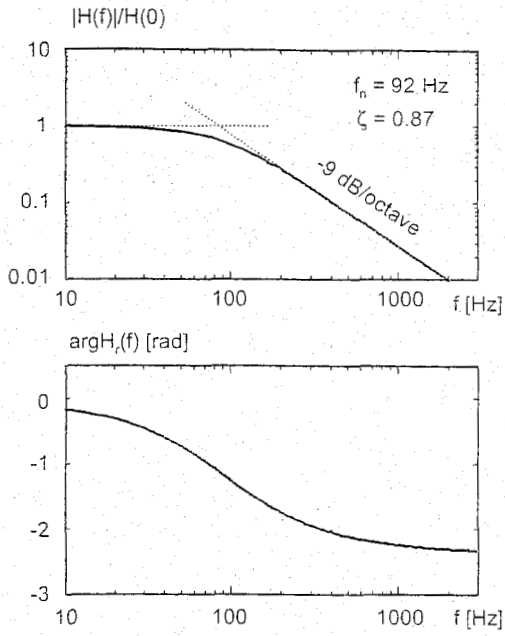


Figure 6.12: Transfer functions for track modelled as a single beam

For small values of the frequency λ approaches the characteristic length L. To find the actual wavelength λ must now be multiplied by 2π. The modulus of this natural wavelength λ_n is therefore:

$$\lambda_n = \left|\frac{2\pi}{\beta}\right| \qquad (6.41)$$

and is plotted in Figure 6.13. At 1500 Hz λ_n reaches a value of twice the sleeper distance (= 1.2 m). For higher frequencies the schematisation based on a continuous support is thus no longer permissible.

Double beam

When the track structure is schematised as a double beam, as shown in Figure 6.10(b), in which the top beam represents the rail and the lower beam the sleeper, the displacement as a function of the load will also be determined.

The upper beam, which represents the rail, has a distributed mass m_1 and a bending stiffness EI. The sleepers are represented as a continuous beam with a mass m_2 and a bending stiffness of zero. The connection between the two beams consists of the rail pads with a stiffness k_1 and a damping c_1. The ballast contributes to the model with a stiffness k_2 and a damping c_2.

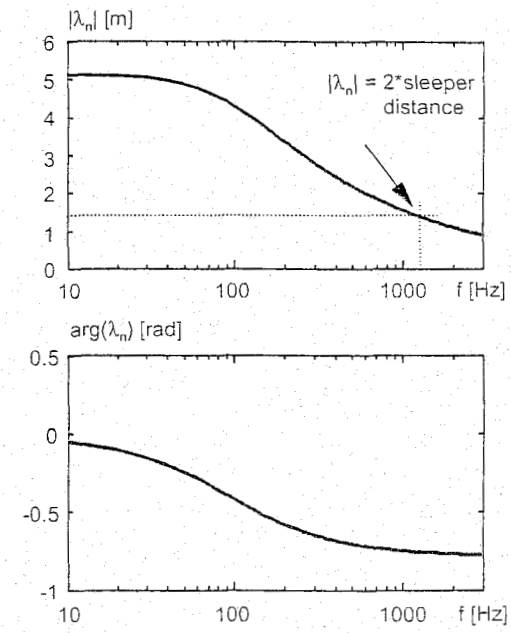


Figure 6.13: Natural wavelength versus frequency

The differential equation of the problem can be obtained in analogy with (6.30):

$$EI\frac{\partial^4 w(x,t)}{\partial x^4} + m\frac{\partial^2 w(x,t)}{\partial t^2} + c\frac{\partial w(x,t)}{\partial t} + k\,w(x,t) = 0 \qquad (6.44)$$

The moving wheel load $Q\delta(x\text{-}vt)$ will be accounted for later in (6.53) as a boundary condition.

After introduction of the new dimensionless variable s, indicating a moving set of variables:

$$s = \lambda(x - vt), \quad \textit{wherein:} \quad \lambda = \left(\frac{k}{4EI}\right)^{1/4} = \frac{1}{L} \qquad (6.45)$$

after substitution in (6.44):

$$\frac{\partial w}{\partial x} = \lambda\frac{dw}{ds};\ \frac{\partial w}{\partial t} = -\lambda v\frac{dw}{ds},\ etc \qquad (6.46)$$

for the homogeneous normal differential equation in the variable s it follows that:

$$\frac{d^4 w}{ds^4}(s) + 4\alpha^2\frac{d^2 w}{ds^2}(s) - 8\alpha\beta\frac{dw}{ds}(s) + 4\,w(s) = 0 \qquad (6.47)$$

in which:

$$\alpha = \frac{v}{2\lambda}\left(\frac{m}{EI}\right)^{1/2} \qquad \text{(dimensionless)} \qquad (6.48)$$

$$\beta = \frac{c}{2m}\left(\frac{m}{k}\right)^{1/2} \qquad \text{(dimensionless)} \qquad (6.49)$$

Note that the quantities in (6.48) and (6.49) have the following physical meaning:

λ = the inverse of the static characteristic length of the track (1/L);
α = the ratio between actual speed and critical speed (resonance);
β = the ratio between actual damping and critical damping (9.7).

Inserting $w = e^{\gamma s}$ in (6.47) results in the following algebraic equation:

$$\gamma^4 + 4\alpha^2\gamma^2 - 8\alpha\beta\gamma + 4 = 0 \qquad (6.50)$$

Note that the four roots will generally be complex. All roots are significant here because there is no symmetry as is the case in the static problem. In order to meet the condition $w(\pm\infty) = 0$, the solution must be expressed as:

$$s \geq 0 \quad w = A_1 e^{\gamma_1 s} + A_2 e^{\gamma_2 s} \quad \text{(roots } \gamma_1 \text{ and } \gamma_2 \text{ should possess a } \textit{negative} \text{ real part)} \qquad (6.51)$$

$$s < 0 \quad w = A_3 e^{\gamma_3 s} + A_4 e^{\gamma_4 s} \quad \text{(roots } \gamma_3 \text{ and } \gamma_4 \text{ should possess a } \textit{positive} \text{ real part)} \qquad (6.52)$$

The coefficients A_i are generally complex. The solution w(s), however, will be real to have physical significance.

The dynamic theory of double beams is analogous to the static theory discussed in section 5.4. The expressions for the rail deflection can be applied if the static degrees of stiffness k_1 and k_2 are replaced by their dynamic equivalents:

$$k_1^* = k_1 - m_1\omega^2 + ic_1\omega \tag{6.42}$$

$$k_2^* = k_2 - m_2\omega^2 + ic_2\omega \tag{6.43}$$

in much the same way as was done with the single beam theory.

By numerical evaluation of the expressions (6.35) and (6.39) with the appropriate values indicated in Table 6.1 (double beam) the modulus and argument of H_r are found. Both functions are illustrated in Figure 6.14. By way of comparison the functions for the single beam are also shown.

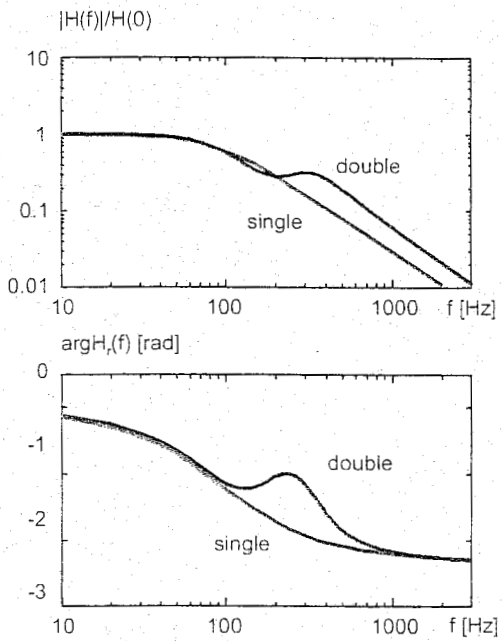


Figure 6.14: Transfer function for track modelled as a double beam

Beam on an elastic foundation (Moving load)

When considering dynamic track loads it is usually assumed that the load is stationary. The running speed does, however, have a certain influence on the dynamic interaction between vehicle and track because the point of application of the load moves with the running speed. Timoshenko [263] examined this problem for an undamped elastically supported rail. Later on Fryba [98] included the influence of damping as well.

In the following discussion the general case is considered based on a single beam on elastic foundation with damping according to Figure 6.10a. The beam is loaded by a moving vertical constant load. The appropriate model for this system is given in Figure 6.15.

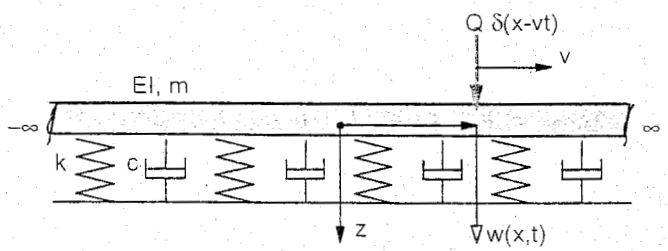


Figure 6.15: Beam model with moving load

Quantities:

EI	[Nm²]
m	[kg/m]
c	[Ns/m²]
k	[N/m²]
q	[N/m]
Q	[N]
v	[m/s]
x	[m]
t	[s]
w	[m]

The solution method chosen here is again directed to the use of complex calculi in order to obtain numerical results, rather than first explicitly deriving the analytical expression for the dynamic deflection.

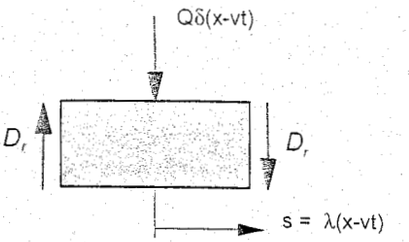


Figure 6.16: Transfer condition at s = 0

The transition conditions at $s = 0$ result in the following equations:

$$\begin{aligned} w_l &= w_r \quad \rightarrow A_3 + A_4 = A_1 + A_2 \\ \frac{dw_l}{ds} &= \frac{dw_r}{ds} \quad \rightarrow A_3\gamma_3 + A_4\gamma_4 = A_1\gamma_1 + A_2\gamma_2 \\ M_l &= M_r \quad \rightarrow A_3\gamma_3^2 + A_4\gamma_4^2 = A_1\gamma_1^2 + A_2\gamma_2^2 \\ Q + D_r &= D_l \quad \rightarrow \frac{Q}{EI\lambda^3} + A_3\gamma_3^3 + A_4\gamma_4^3 = A_1\gamma_1^3 + A_2\gamma_2^3 \end{aligned} \tag{6.53}$$

Note that the last condition in (6.53) can be obtained using equations (6.28), (6.29), and Figure 6.16.

Writing this system of equations in matrix notation gives:

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ \gamma_1 & \gamma_2 & -\gamma_3 & -\gamma_4 \\ \gamma_1^2 & \gamma_2^2 & -\gamma_3^2 & -\gamma_4^2 \\ \gamma_1^3 & \gamma_2^3 & -\gamma_3^3 & -\gamma_4^3 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = w_0 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 8 \end{bmatrix} \tag{6.54}$$

in which the factor $w_0 = Q/(8EI\lambda^3)$ represents the maximum static deflection. Apparently, if this factor is set to one, the solution corresponds with the dimensionless relative displacement $\eta(s)$ in the static case ($v = 0$). This will be useful to compare the dynamic solution ($v \neq 0$) with the static one for various values of α and β, and to determine the amplification factor.

After solving the matrix equation (6.54) the constants A_i are known and the relation between the deflection and the distance can be drawn. Figure 6.17 shows the characteristic wave shapes of the relative displacement of the beam for several values of α and β in the case of undamped ($\beta = 0$), lightly damped ($\beta = 0.1$), over critically damped systems ($\beta = 1.1$), static situations ($\alpha = 0$), subcritical velocities ($\alpha = 0.5$), critical velocities ($\alpha = 1$), and super critical velocities ($\alpha = 2$).

The fourth row in Figure 6.17 shows that the maximum amplitude of the displacements is moving behind the location of the load for super critical velocity. For the critical speed ($\alpha = 1$) and undamped case ($\beta = 0$) the wave amplitudes became infinite.

For a lightly damped system (middle column) a similar behaviour takes place. The wave shape calculated for $\alpha = 1$ shows large amplifications. For an over critically damped case ($\beta = 1.1$) the wave forms are asymmetric with respect to the load and show no amplifications anymore with respect to the static case.

In Figure 6.18 the ratio of the maximum deflection is given as function of the ratio of the load velocity and the critical velocity α for several values if the damping ratio is β. The equivalence to the frequency response function of the simple spring-mass system is striking. For small damping ratios the wave amplitude shows severe amplifications.

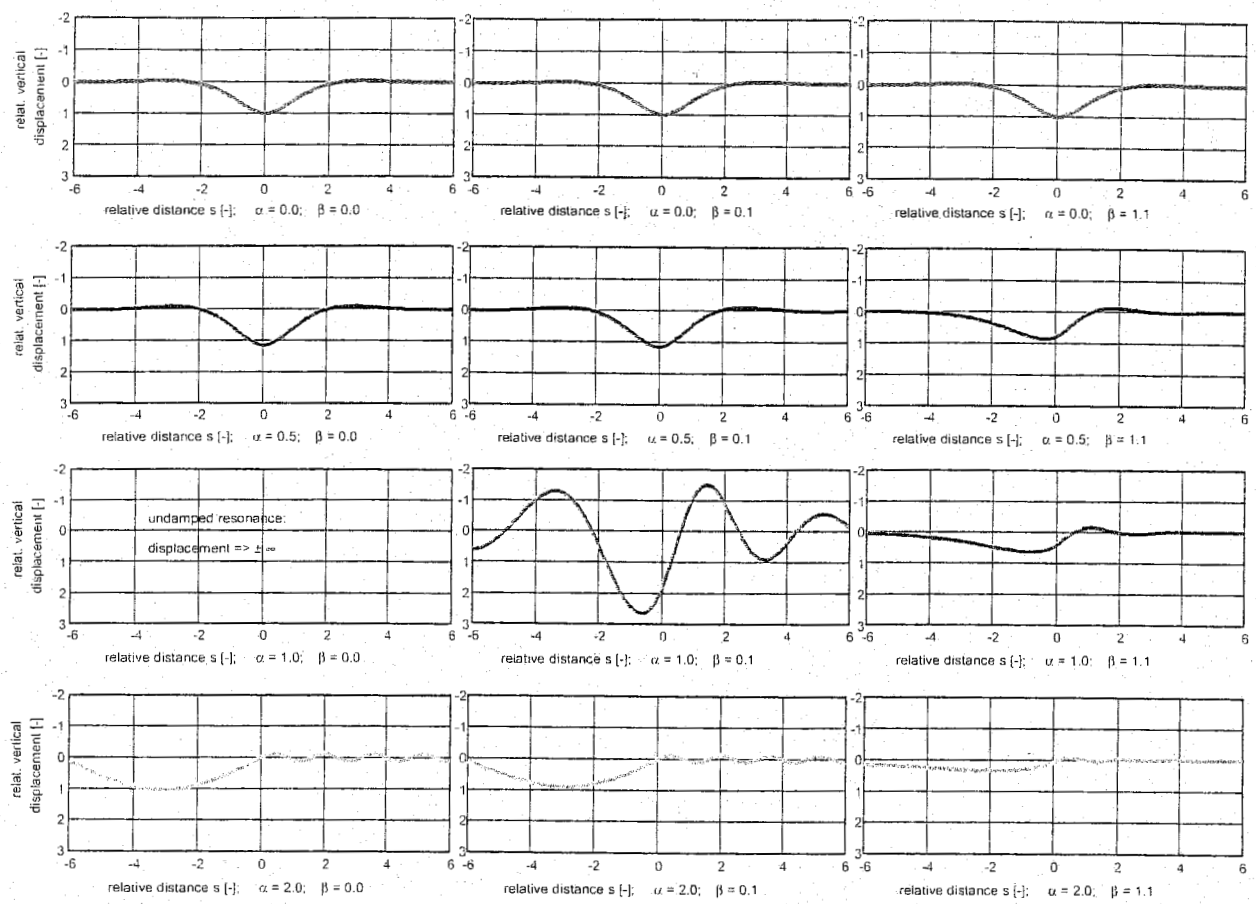


Figure 6.17: Wave shapes versus relative distance

Critical train speed

As can be seen from Figure 6.18, the critical speed or velocity is situated on or near the velocity ratio $\alpha = 1$. According to equation (6.48), it can be derived for the critical speed v_{cr} that:

$$v_{cr}^2 = \frac{2}{m}\sqrt{kEI} \qquad (6.55)$$

in which:
m = rail mass per length;
k = track stiffness;
EI = bending stiffness.

At conventional speeds the influence is negligible as these speeds are much lower than the critical speed v_{cr}. For instance, using the track parameters listed in Table 6.1 , the critical speed amounts to 475 m/s. A speed of 200 km/h thus corresponds with $\alpha = 0.12$. According to Figure 6.18 this would give a very low dynamic amplification and the effect of the load travelling speed can therefore be neglected.

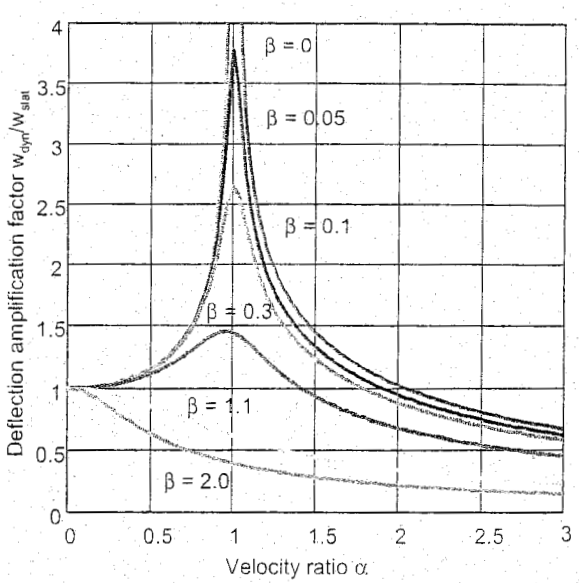


Figure 6.18: Dynamic amplification versus speed due to moving load

For tracks of good quality the critical speed lies far beyond the operating speed, but with poor soil conditions or other mass/spring configurations the critical speed can be so low that special measures are required. In case the train speed approaches the wave propagation speed, the soil may experience a liquefaction type of phenomenon as seen in Figure 6.19. An actual measurement in track on soft soil is shown in Figure 6.20.

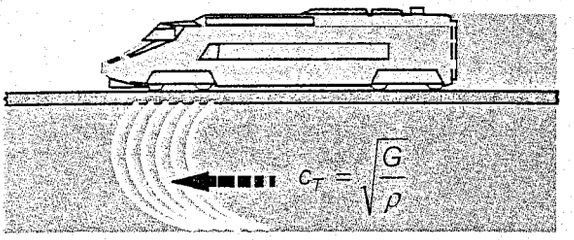


Figure 6.19: Wave propagation at high speed

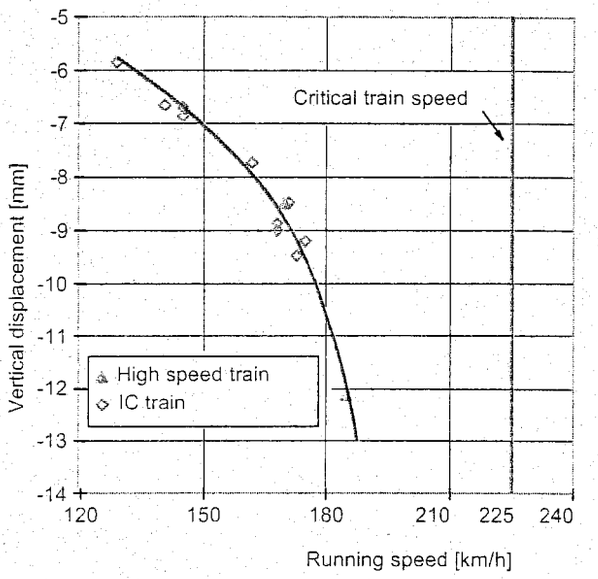


Figure 6.20: Actual measurement on soft soil

For the undamped case (left column of Figure 6.17) a simple formula exists [98] for the dynamic amplification:

$$\frac{w_{dyn}}{w_{stat}} = \frac{1}{\sqrt{\left|1 - \left[\frac{v}{v_{cr}}\right]^2\right|}} \qquad (6.56)$$

6.3.4 Discrete support

The model in Figure 6.10(c), in which the rail is supported in a discrete manner, gives the best approximation. Such an approach also lends itself to the application of standard element programs programs which will be discussed later in Section 6.9. These element method programs give great flexibility as regards load forms and support conditions.

6.4 Vertical wheel response

6.4.1 Hertzian contact spring

During vehicle/track interaction the forces are transmitted by means of the wheel/rail contact area. On account of the geometry of the contact area between the round wheel and the rail, the relationship between force and compression, represented by the Hertzian contact spring, is not linear as has already been discussed in Section 2.7. The relationship between force F and indentation y of the contact surface can be written as:

$$F = c_H y^{3/2} \qquad (6.57)$$

in which c_H [$Nm^{-3/2}$] is a constant depending on the radii and the material properties.

Since a description of the wheel/rail relationship using transfer functions requires that all components are linear, the Hertzian spring must also be linearised. This linearised value of the stiffness can be found by considering the relationship between the force and displacement increments around the static wheel load. The linearised Hertzian spring stiffness k_H is then:

$$k_H = \frac{dF}{dy} = \frac{3}{2} c_H^{2/3} F^{1/3} \tag{6.58}$$

Jenkins e.a. [137] determined the k_H value for old and new wheels as a function of the wheel diameter. For a wheel diameter of 1 m and a static wheel load of 75 kN, a k_H value of 1.4 10^9 N/m is found for new wheels and 1.6 10^9 N/m for old wheels (see also Section 4.11).

6.4.2 Transfer functions between wheel and rail

Figure 6.21 shows the model of a wheel which is connected to the rail by means of a Hertzian spring. From the equilibrium the following is obtained:

$$F_H + M_w \ddot{y}_w = 0 \tag{6.59}$$

with:

$$y_w = \hat{y}_w e^{i2\pi ft} = \hat{y}_w e^{i\omega t} \tag{6.60}$$

The transfer function of the wheel is obtained from (6.59) according to:

$$H_w(f) = \frac{\hat{y}_w}{F_H} = \frac{1}{M_w \omega^2} \tag{6.61}$$

In the following, the relationships between wheel displacement at axle box level and vertical rail geometry, as well as axle box acceleration and vertical rail geometry are examined. These relationships are important when analysing phenomena associated with corrugations and poor quality welds. These transfer functions also formed the basis of the calculations which were carried out when designing the BMS-2 system discussed in Chapter 16.

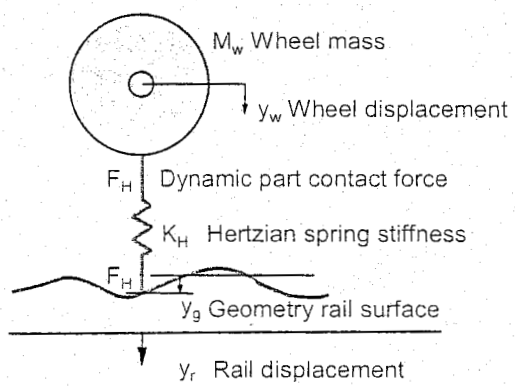


Figure 6.21: Hertzian spring force acting between wheel and rail

The relation between the interaction force F_H and the change in length of the Hertzian spring is determined by:

$$F_H = k_H [y_w - y_r - y_g] \tag{6.62}$$

in which:
y_w = vertical wheel displacement at the level of the axle box;
y_r = vertical rail displacement under the effect of F_H;
y_g = vertical rail geometry;
k_H = linearised stiffness of Hertzian spring;
F_H = dynamic component of wheel/rail force.

If (6.62) is transformed to the frequency domain and the Fourier transformations are indicated in capital letters, the expression can be written as:

$$Y_g(f) = Y_w(f) - Y_r(f) - F_H(f)/k_H \tag{6.63}$$

Using the previously derived transfer functions for the double beam in Section 6.3.3, wheel (6.61), and the rail (6.35), the wheel and rail displacements can be expressed in the wheel/rail force:

$$Y_w(f) = H_w(f)F_H(f) \tag{6.64}$$

$$Y_r(f) = H_r(f)F_H(f) \tag{6.65}$$

After substitution of both in (6.63), this expression becomes:

$$Y_g(f) = [H_w(f) - H_r(f) - 1/k_H]F_H(f) \tag{6.66}$$

$$= H_1(f)F_H(f) \tag{6.67}$$

The relation between wheel displacement $Y_w(f)$ and rail geometry $Y_g(f)$ is now obtained by substitution of (6.64) in (6.67), which results in:

$$Y_g(f) = \frac{H_1}{H_w}Y_w(f) \tag{6.68}$$

Furthermore, by differentiating the wheel displacement twice according to:

$$\ddot{Y}_w(f) = -\omega^2 Y_w(f) \tag{6.69}$$

and substituting this result together with (6.61) in (6.68), the relation between axle box acceleration and rail geometry becomes:

$$Y_g(f) = -M_w H_1(f)\ddot{Y}_w(f) = H_{\ddot{y}}(f)\ddot{Y}_w(f) \tag{6.70}$$

in which:

$$H_{\ddot{y}}(f) = -M_w\left[H_w(f) - H_r(f) - \frac{1}{k_H}\right] \tag{6.71}$$

This transfer function forms the basis of the measuring principle of BMS-2 [268] (see Chapter 16) and is illustrated in Figure 6.22 in which the moduli of the various contributions are plotted, as is the modulus of the resulting transfer function.

The contribution of the rail is calculated using the double beam model based on the appropriate data in Table 6.1 (double beam).

Figure 6.22 shows that the wheel produces by far the greatest contribution in the frequency band up to about 50 Hz. The rail is mainly responsible for the behaviour in the 50 to 1000 Hz band and the Hertzian spring determines the behaviour above 1000 Hz.

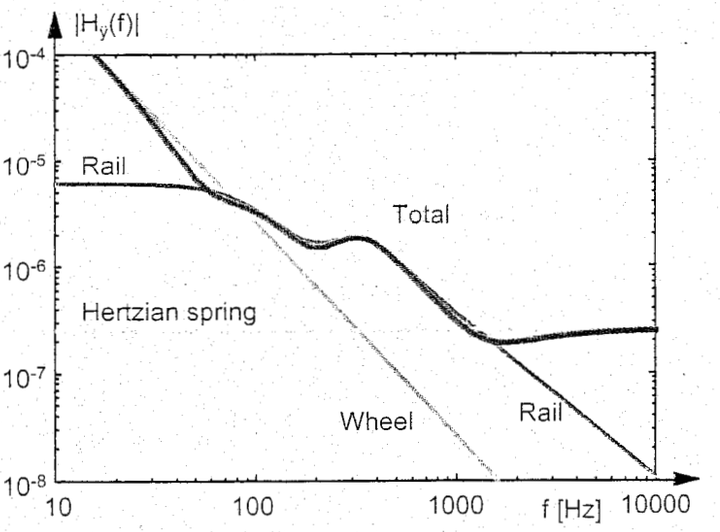


Figure 6.22: Transfer function between rail geometry and axle box acceleration

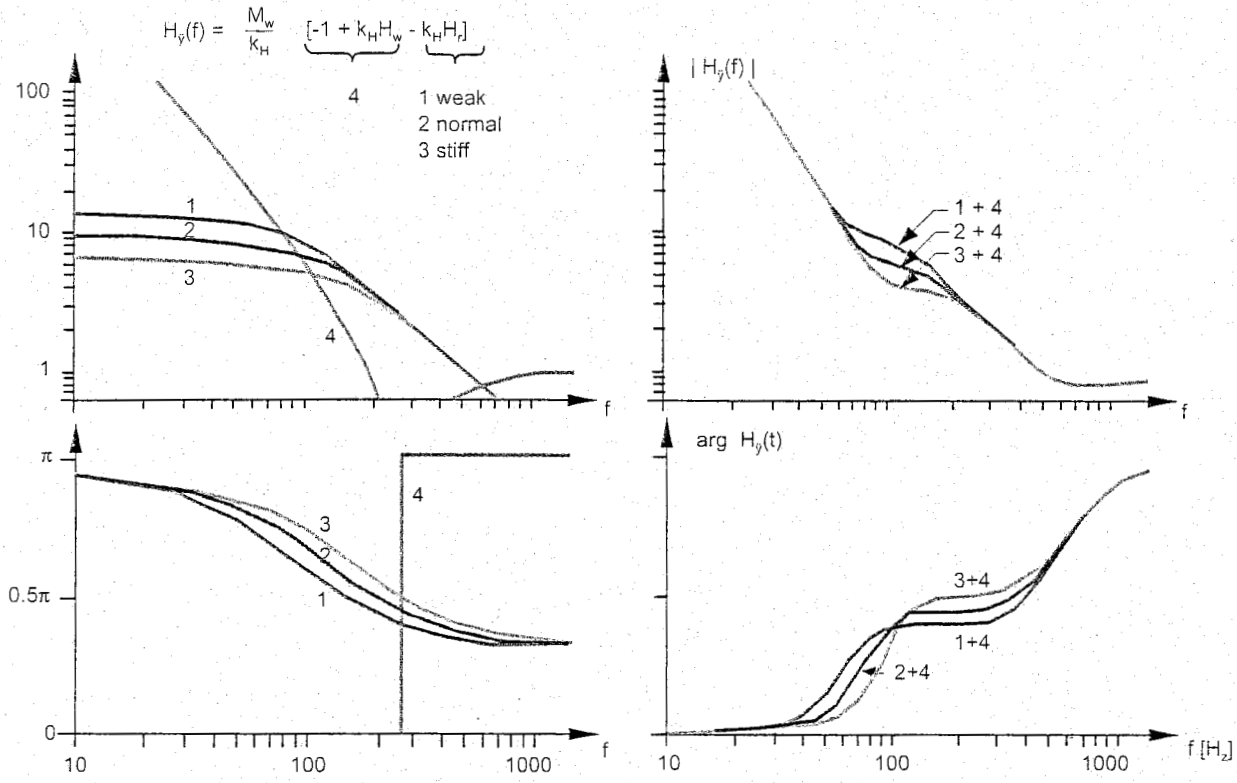


Figure 6.23: Influence of track stiffness on transfer functions $H_{\ddot{y}}(f)$

Since corrugations appear predominantly between 10 and 1500 Hz, it is clear that the track construction in particular has a very great influence. The question is whether, when measuring corrugations by means of axle box accelerations, the variation in track condition can be disregarded. This is examined by varying the track stiffness k_1. Figure 6.23 shows the various contributions made to the transfer function according to formula (6.42) for standard track with a stiffness k_1, for track with a low stiffness of 0.5 k_1, and for track with a high stiffness of 2 k_1. Differences due to the characteristics of the track only show up in the frequency band between 60 and 200 Hz. As a result of system damping due to half-space radiation, for which we refer the reader to [208] and [231], the effect remains limited. No special measures have therefore been taken in the BMS-2 system.

6.5 Linear vehicle model

6.5.1 Schematisation

Transfer functions between track geometry and vehicle reactions can be determined using mathematical models. Figure 6.24 shows a very simple model which describes the main dynamic response of the vehicle. This model can be used to calculate the various relations between track geometry (consisting of cant, level, and alignment) and vehicle response in the form of Q and Y-forces between wheel and rail, but also to calculate horizontal and vertical car body accelerations [269].

In the following the various transfer functions will be derived first. After a discussion of ISO filtering of the vehicle body acceleration signals to take account of human perception, a number of examples are given of transfer functions calculated based on the NS measuring coach.

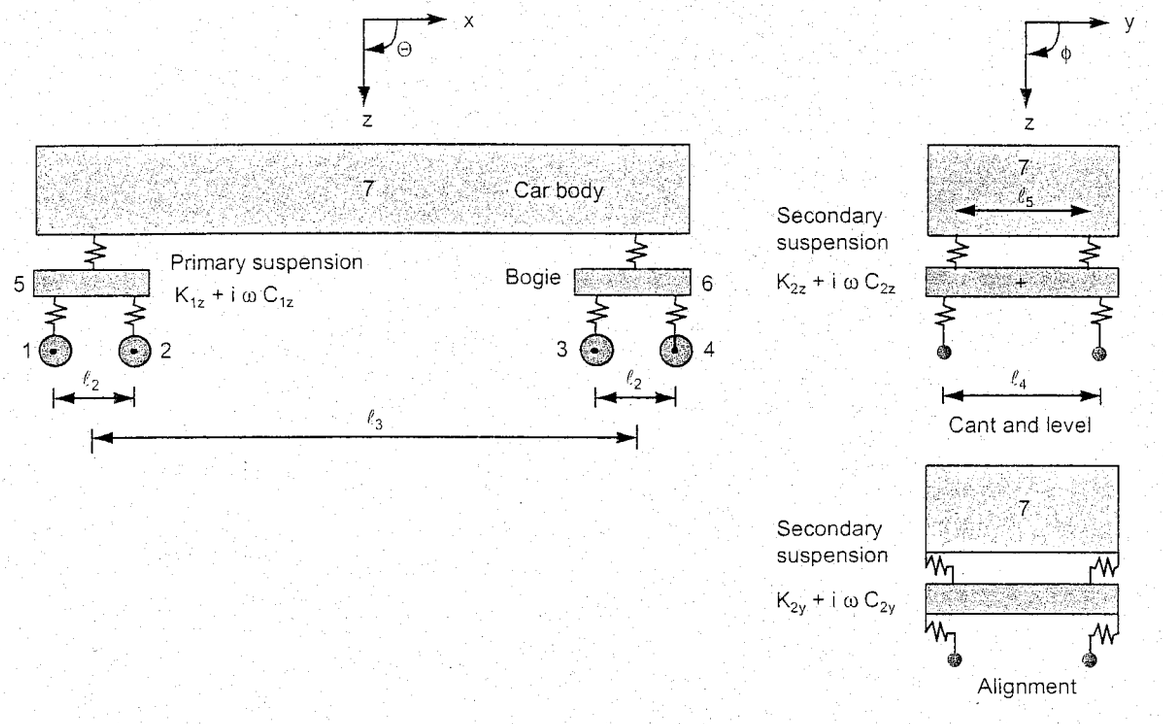


Figure 6.24: Mathematical model for describing vehicle/track interaction

6.5.2 Response to irregularities in level

When examining the response to an irregularity in level the z displacements and θ rotations are of importance. Furthermore, the input will appear four times in succession, i.e. on each of the axles.

The strategy followed here in determining the transfer functions can be summed up as follows:

- determine the response for four symmetrical and asymmetrical wheel displacements;
- combine these so that the response for a single displacement of a wheel is obtained;
- combine these 4 inputs from the separate wheels with their respective time lags.

Since the wheels can be assumed to be infinitely stiff, the inputs act directly on the primary suspension at points 1, 2, 3, and 4 in Figure 6.24. To start with, the mass of the wheelsets is disregarded. Its effect is taken into account later.

Level: case 1

The first mode of vibration is purely a translation characterized by:

$$z_1 = z_2 = z_3 = z_4 = v_1 = \hat{v}_1 e^{i\omega t} \tag{6.72}$$

If the amplitude is normalized to 1, the following amplitude vector is obtained for points 1 to 4:

$$\underline{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{6.73}$$

The equations of motion for this load case can be described in the frequency domain as:

$$-M_5\omega^2 Z_5 = 4F_1 - 2F_2 \tag{6.74}$$

$$-M_7\omega^2 Z_7 = 4F_2 \tag{6.75}$$

$$F_1 = (K_{1z} + i\omega C_{1z})(V_1 - Z_5) \tag{6.76}$$

$$F_2 = (K_{2z} + i\omega C_{2z})(Z_5 - Z_7) \tag{6.77}$$

After substitution of:

$$a = \tfrac{1}{2}M_5\omega^2 \qquad \textit{bogie} \tag{6.78}$$

$$b = \tfrac{1}{2}M_7\omega^2 \qquad \textit{car body} \tag{6.79}$$

$$c = K_{1z} + i\omega C_{1z} \qquad \textit{primary suspension} \tag{6.80}$$

$$d = K_{2z} + i\omega C_{2z} \qquad \textit{secondary suspension} \tag{6.81}$$

the transfer function H^*_{ij} can be calculated between the displacement amplitude v_1 and the response, i.e. the Q-force F_1 and the displacement Z_7. At a later stage the displacement Z_7 is converted to the acceleration $\ddot{Z}_7 = -\omega^2 Z_7$. The transfer function H_{ij} is defined as follows:

$$H^*_{1j} = \textit{Q force } F_1 \textit{ for } \hat{v}_j = 1 \tag{6.82}$$

$$H^*_{2j} = \textit{car body displacement } Z_7 \textit{ for } \hat{v}_j = 1 \tag{6.83}$$

$$H^*_{3j} = \textit{car body rotation } \Theta_7 \textit{ for } \hat{v}_j = 1 \tag{6.84}$$

$j = 1, 4$, index of the 4 forced modes of displacement.

For the first displacement mode the transfer functions are as follows:

$$H^*_{11} = F_1/V_1 = \frac{1}{n_0}\left[-acd + \frac{1}{2}abc - \frac{1}{2}bcd\right] \tag{6.85}$$

$$H^*_{21} = Z_7/V_1 = \frac{1}{n_0}2cd \tag{6.86}$$

$$H^*_{31} = \Theta_7/V_1 = 0 \tag{6.87}$$

where:

$$n_0 = 2cd - bc + \frac{1}{2}ab - ad - \frac{1}{2}bd \tag{6.88}$$

Level: case 2

With this displacement mode both bogies move asymmetrically according to:

$$z_1 = z_2 = -z_3 = -z_4 = v_2 = \hat{v}_2 e^{i\omega t} \tag{6.89}$$

The amplitude vector is:

$$\underline{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \tag{6.90}$$

As a result of the asymmetrical movement the car body only rotates and the z displacement is zero. The equations of motion are:

$$-M_5\omega^2 Z_5 = 4F_1 - 2F_2 \tag{6.91}$$

$$-J_{7y}\omega^2\Theta_7 = 2\ell_3 F_2 \tag{6.92}$$

$$F_1 = (K_{1z} + i\omega C_{1z})(V_2 - Z_5) \tag{6.93}$$

$$F_2 = (K_{2z} + i\omega C_{2z})\left(Z_5 - \frac{1}{2}\ell_3 Z\Theta_7\right) \tag{6.94}$$

With:

$$b_1 = \frac{2}{\ell_3^2} J_{7y}\omega^2 \tag{6.95}$$

the following transfer functions are found:

$$H_{12}^* = F_1/V_2 = \frac{1}{n_1}\left[-acd + \frac{1}{2}ab_1c - \frac{1}{2}b_1cd\right] \tag{6.96}$$

$$H_{22}^* = Z_5/V_2 = \frac{1}{n_1}c(2d - b_1) \tag{6.97}$$

$$H_{32}^* = \Theta_7/V_2 = \frac{4cd}{\ell_3 n_1} \tag{6.98}$$

with:

$$n_1 = 2cd + \frac{1}{2}ab_1 - b_1c - \frac{1}{2}b_1d - ad \tag{6.99}$$

Level: cases 3 and 4

In case 3 the displacement mode and amplitude vector take the following form:

$$z_1 = -z_2 = z_3 = -z_4 = v_3 = \hat{v}_3 e^{i\omega t} \tag{6.100}$$

The amplitude vector is:

$$\underline{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \tag{6.101}$$

For case 4 they are:

$$z_1 = -z_2 = -z_3 = z_4 = v_4 = \hat{v}_4 e^{i\omega t} \tag{6.102}$$

The amplitude vector is:

$$\underline{v}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \tag{6.103}$$

In neither case is the car body loaded. The primary suspensions vibrate independently of each other, symmetrically in case 3 and asymmetrically in case 4. Whenever a wheelset is being considered the following equations of motion apply:

$$-J_{5y}\omega^2\Theta_5 = 2\ell_2 F_1 \tag{6.104}$$

$$F_1 = (K_{1z} + i\omega C_{1z})\left(V_3 - \frac{1}{2}\ell_2\Theta_5\right) \tag{6.105}$$

With:

$$a_1 = \frac{2}{\ell_2^2} J_{5y}\omega^2 \tag{6.106}$$

the following transfer functions are found:

$$H_{13}^* = F_1/V_3 = \frac{-a_1 c}{a_1 - 2c} \tag{6.107}$$

$$H_{14}^* = F_1/V_4 = \frac{-a_1 c}{a_1 - 2c} \tag{6.108}$$

6.5.3 Combination of level results

The response due to the movement of one axle can be obtained by combining the above-mentioned results. For instance, by adding all the results the displacements for axles 1, 3, and 4 equal zero and axle 2 has a displacement of 4. It can clearly be seen that if the k-th axle undergoes a forced movement with an amplitude of 1, the response, i.e. the transfer function H_{ik}^+, can be derived as follows from the above results:

$$H_{ik}^+ = \frac{1}{4}\sum_{j=1}^{4} H_{ij}^* v_{jk} \tag{6.109}$$

In this equation v_{jk} is the k-th element of the amplitude vector v_j. The resulting transfer function between track geometry, in this case level z_t, and the respective response component can now be obtained by adding together the contributions from the 4 axles taking their lag into account. This results in the following expressions:

$$Q = H_1 Z_t \tag{6.110}$$

$$\ddot{Z}_c = (Z_7 - \ell_x\Theta_7)\omega^2 = H_2 Z_t \tag{6.111}$$

Here the transfer functions H_1 and H_2 have the following form:

$$H_1 = -\frac{1}{2}m_1\omega^2 + \sum_{k=1}^{4} H_{1k}^{+} e^{i\omega\tau_k} \tag{6.112}$$

$$H_2 = -\omega^2 \sum_{k=1}^{4} \{H_{2k}^{+} - \ell_x H_{3k}^{+}\} e^{i\omega\tau_k} \tag{6.113}$$

where:

$$\tau_1 = 0 \qquad \textit{reference axle} \tag{6.114}$$

$$\tau_2 = \ell_2 / v \tag{6.115}$$

$$\tau_3 = \ell_3 / v \tag{6.116}$$

$$\tau_4 = \frac{\ell_2 + \ell_3}{v} \tag{6.117}$$

ℓ_x = the distance between the centre of gravity of the vehicle body in the x-direction and the point to which the response relates;
v = running speed.

The first term in (6.112) represents the proportion of the wheelset mass in the Q-force.

6.5.4 Response to irregularities in alignment

The vehicle reactions to an alignment input show a strong resemblance to the reactions to a level input. The degrees of freedom change from z to y and from Q to ϕ. If the ϕ-rotation is disregarded the model used for level fits exactly. This assumption will be used here. The transfer function H_1, according to (6.112), gives the relation between the Y-force between wheel and rail and the alignment, whereas H_2, according to (6.113), represents the relation between horizontal vehicle body acceleration and alignment.

6.5.5 Response to irregularities in cant

In the case of excitation due to cant, only ϕ-rotations occur. As with the response to alignment, where the ϕ-rotations are disregarded, in the case of response to cant lateral spring reactions are not taken into consideration. Calculation of the relationships between cant and vehicle response is similar to that for level input. Here again the same 4 forced displacement modes are selected and applied to the degree of freedom ϕ, from which the effect of each axle can then be derived separately.

Cant: case 1

The mode of displacement is as follows:

$$\phi_1 = \phi_2 = \phi_3 = \phi_4 = v_1 = \hat{v}_1 e^{i\omega t} \tag{6.118}$$

The interaction forces for each axle are equal because of the symmetry. The reaction forces at the left

and right rail are in opposite directions. The equations of motions for this system are:

$$-J_{5x}\omega^2\Phi_5 = 2\ell_4 F_1 - \ell_5 F_2 \tag{6.119}$$

$$-J_{7x}\omega^2\Phi_7 = 2\ell_5 F_2 \tag{6.120}$$

$$F_1 = (K_{1z} + i\omega C_{1z})\left(V_1 - \frac{1}{2}\Phi_5\ell_4\right) \tag{6.121}$$

$$F_2 = (K_{2z} + i\omega C_{2z})(\Phi_5 - \Phi_7)\frac{1}{2}\ell_5 \tag{6.122}$$

With:

$$a = J_{5x}\omega^2 \qquad \textit{bogie} \tag{6.123}$$

$$b = J_{7x}\omega^2 \qquad \textit{car body} \tag{6.124}$$

$$c = (K_{1z} + i\omega C_{1z})\ell_5^2 \qquad \textit{primary suspension} \tag{6.125}$$

$$d = (K_{2z} + i\omega C_{2z})\ell_4^2 \qquad \textit{secondary suspension} \tag{6.126}$$

$$n = 2cd - bc + \frac{1}{2}ab - ad - y\frac{1}{2}bd \tag{6.127}$$

the transfer functions for this load case are:

$$H_{11}^* = F_1/V_1 = \frac{1}{\ell_5}\frac{1}{n}\left[-acd + \frac{1}{2}abc - \frac{1}{2}bcd\right] \tag{6.128}$$

$$H_{21}^* = \Phi_7/V_1 = \frac{1}{n}2cd \tag{6.129}$$

Cant: case 2

In this case the displacement mode is as follows:

$$\phi_1 = \phi_2 = -\phi_3 = -\phi_4 = v_2 = \hat{v}_2 e^{i\omega t} \tag{6.130}$$

In such a case the vehicle body will not rotate. The solution is obtained from load case 1 by putting $\phi_7 = 0$. The transfer functions for load case 2 are:

$$H_{12}^* = F_1/V_2 = \frac{1}{\ell_5}\frac{c(d-a)}{2c+d-a} \tag{6.131}$$

$$H_{22}^* = \Phi_7/V_2 = 0 \tag{6.132}$$

Cant: cases 3 and 4

$$\phi_1 = -\phi_2 = \phi_3 = -\phi_4 = v_3 = \hat{v}_3 e^{i\omega t} \tag{6.133}$$

$$\phi_1 = -\phi_2 = -\phi_3 = \phi_4 = v_4 = \hat{v}_4 e^{i\omega t} \tag{6.134}$$

In these cases both car body and bogie are not moving. The transfer functions are now:

$$H_{13}^{*} = \frac{F_1}{V_3} = -\frac{c}{\ell_5} \tag{6.135}$$

$$H_{14}^{*} = \frac{F_1}{V_4} = -\frac{c}{\ell_5} \tag{6.136}$$

6.5.6 Combination of cant results

Excitation due to cant produces two response components, i.e. F_1 and ϕ_7. Two outputs can be derived from ϕ_7 i.e.:

$$\ddot{Y}_b = \omega^2 \ell_z \Phi_7 \tag{6.137}$$

$$\ddot{Z}_b = -\omega^2 \ell_y \Phi_7 \tag{6.138}$$

Here ℓ_z and ℓ_y are the distances between the centre of gravity of the car body and the point to which the response relates in z and y directions.

Combination of the 4 load cases in order to obtain the input for one axle is achieved using the formula (6.109) in which:

$$H_{1j}^{*} = Q \text{ force} \qquad \text{for } \hat{v}_1 = 1 \tag{6.139}$$

$$H_{2j}^{*} = \Phi_7 \qquad \text{for } \hat{v}_1 = 1 \tag{6.140}$$

The response due to cant:

$$\Delta Z = Z_l - Z_r = -1.435\Phi_t \tag{6.141}$$

can now be determined as follows using (6.109):

$$Q = \left\{ J_{1x} \frac{\omega^2}{1.435} + \sum_{k=1}^{4} H_{1k}^{+} e^{i\omega\tau_k} \right\} \phi_t \tag{6.142}$$

The first term inside the brackets is the contribution of the unsprung mass of the wheelset. Bearing in mind (6.141), the transfer function H_1 defined as:

$$Q = H_1 \Delta Z \tag{6.143}$$

follows from the relation:

$$H_1 = -J_{1x}\left[\frac{\omega}{1.435}\right]^2 - \frac{1}{1.435}\sum_{k=1}^{4} H_{1k}^{+} e^{i\omega\tau_k} \tag{6.144}$$

For the car body accelerations according to (6.137) and (6.138) the relationships regarding cant input are:

$$\ddot{Y}_b = H_2 \Delta Z \tag{6.145}$$

where:

$$H_2 = -\frac{1}{1.435}\omega^2 \ell_z \sum_{k=1}^{4} H_{2k}^{+} e^{i\omega\tau_k} \qquad (6.146)$$

$$\ddot{Z}_b = H_3 \Delta Z \qquad (6.147)$$

in which:

$$H_3 = -\frac{1}{1.435}\omega^2 \ell_y \sum_{k=1}^{4} H_{2k}^{+} e^{i\omega\tau_k} \qquad (6.148)$$

6.5.7 ISO weighting of car body accelerations

In order to apply the transfer functions to the calculation of vehicle reactions from measured track geometry, it is necessary to filter the transfer functions. On the one hand, the purpose is to confine unlimited growth in the higher frequencies. On the other hand, the transfer functions for calculating car body accelerations are weighted according to the ISO characteristics which are also incorporated in the ride index meters.

The Q and Y-forces are filtered with a 6th order low-pass filter with the modulus of a Butterworth filter and a zero phase according to:

$$H(f) = \sqrt{\frac{1}{1+\left[\frac{f}{f_c}\right]^{12}}} \qquad (6.149)$$

in which the 3 dB point is at f_c. The value $1/f_c$ is set at 3 m. Figure 6.25 shows the modulus of this filter function.

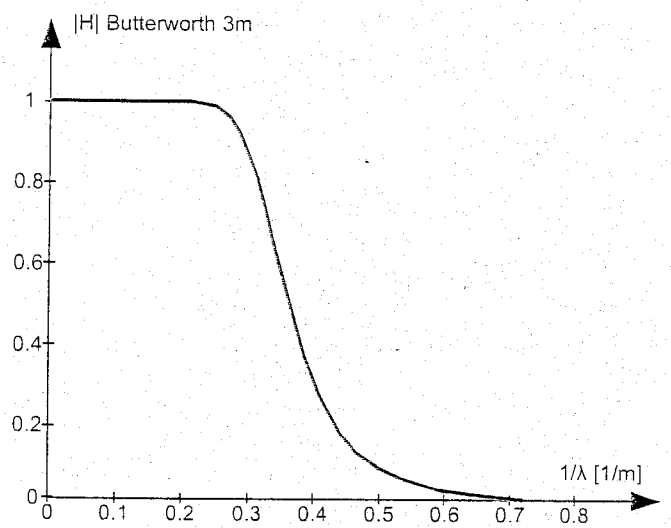


Figure 6.25: Butterworth low-pass filter applied to Q and Y forces

The ISO weighting is implemented using two filters for vertical and horizontal accelerations respectively. These transfer functions are as follows:

vertical:

$$H_v(s) = \frac{0.248\,s\,(0.065\,s+1)\,(0.01457\,s+1)}{(3.5\cdot 10^{-3}\,s^2+0.184\,s+1)\,(5.96\cdot 10^{-4}\,s^2+0.04096\,s+1)} \qquad (6.150)$$

horizontal:

$$H_h(s) = \frac{0.6\,s\,(0.0135\,s+1)\,(0.051\,s+1)}{(1.67\cdot 10^{-3}\,s^2+0.151\,s+1)\,(7.19\cdot 10^{-3}\,s^2+0.141\,s+1)} \qquad (6.151)$$

Here s is the complex frequency which in this application can be set at:

$$s = i\omega = i\,2\pi f = i\frac{2\pi}{\lambda} \qquad (6.152)$$

where λ = v/f = wavelength.

The moduli of the ISO filters are shown in Figure 6.26 for vertical weighting and in Figure 6.27 for horizontal weighting. The functions are calculated for speeds of 90, 120, 140, 160, and 200 km/h.

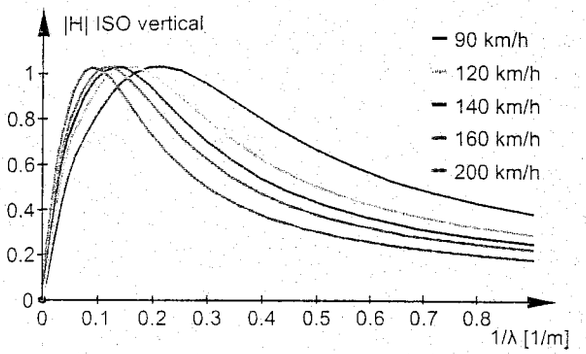


Figure 6.26: ISO characteristic applied to horizontal car body acceleration

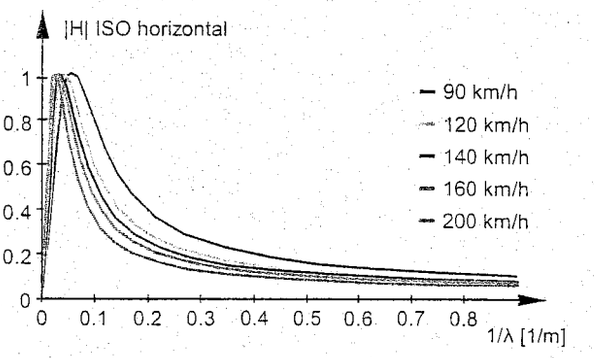


Figure 6.27: ISO characteristic applied to vertical car body acceleration

6.5.8 Calculated transfer functions for the NS measuring coach

The transfer functions derived above were calculated for an NS vehicle, i.e. the measuring coach in which the BMS track recording system is installed.

Table 6.2 summarizes the relevant parameters. The results are given in the form of the modulus of the transfer function and the argument. The unit impulse response function is also calculated.

Figure 6.28 shows the relationship between Q-force and level, Figure 6.29 between vertical car body acceleration and level, Figure 6.30 between Y-force and alignment, Figure 6.31 between lateral car body acceleration and alignment, Figure 6.32 between increase in Q-force and cant, and Figure 6.33 between lateral car body acceleration and cant

Primary suspension		**Secondary suspension**	
K_{1y}	$= 4.75\ 10^6$ N/m	K_{2y}	$= 0.18\ 10^6$ N/m
K_{1z}	$= 0.70\ 10^6$ N/m	K_{2z}	$= 0.41\ 10^6$ N/m
C_{1y}	$= 3.99\ 10^4$ Ns/m	C_{2y}	$= 1.50\ 10^4$ Ns/m
C_{1z}	$= 5.88\ 10^3$ Ns/m	C_{2z}	$= 2.20\ 10^3$ Ns/m
Unsprung mass			
m	= 1500 kg		
J_{1x}	= 730 kgm²		
Bogie frame		**Car body**	
M_5	$= 3.15\ 10^3$ kg	M_7	$= 3.37\ 10^4$ kg
J_{5x}	$= 2.02\ 10^3$ kgm²	J_{7x}	$= 5.24\ 10^4$ kgm²
J_{5y}	$= 2.02\ 10^3$ kgm²	J_{7y}	$= 7.67\ 10^5$ kgm²
J_{5z}	$= 3.56\ 10^3$ kgm²	J_{7z}	$= 7.36\ 10^5$ kgm²
Dimensions			
ℓ_2	= 2.56 m	ℓ_x	= -6.95 m
ℓ_3	= 15.66 m	ℓ_y	= 0.00 m
ℓ_4	= 2.78 m	ℓ_z	= 1.45 m
ℓ_5	= 2.00 m		

Table 6.2: Relevant parameters for the NS vehicle containing BMS

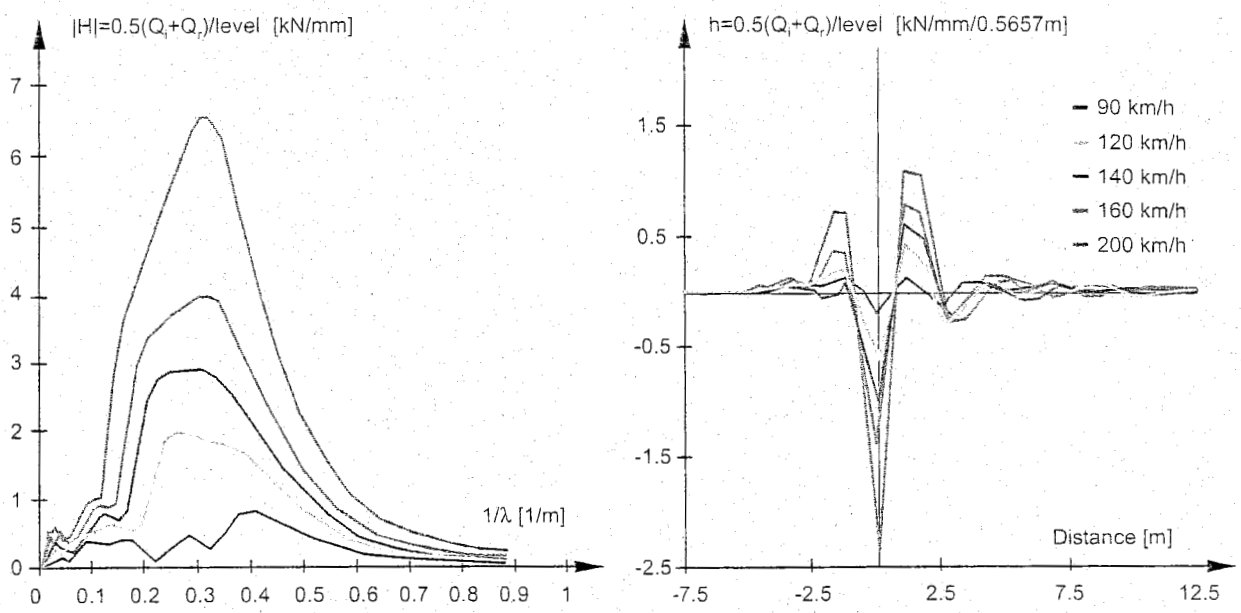


Figure 6.28: Relationships between Q-force and level

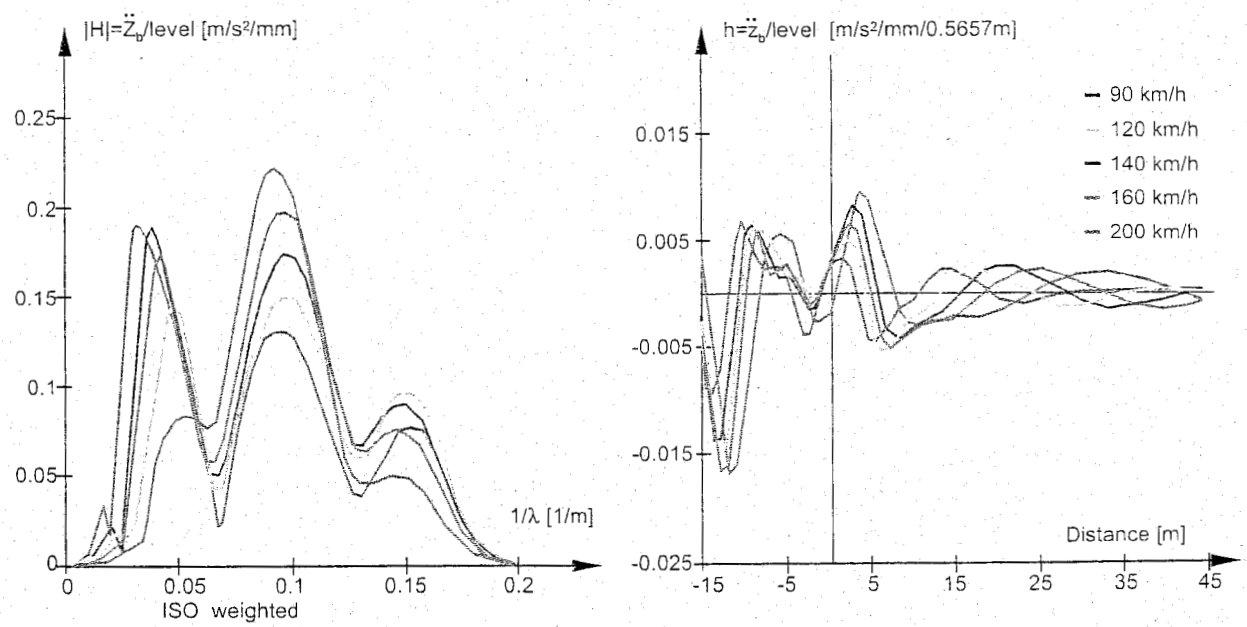


Figure 6.29: Relationships between vertical car body acceleration and level

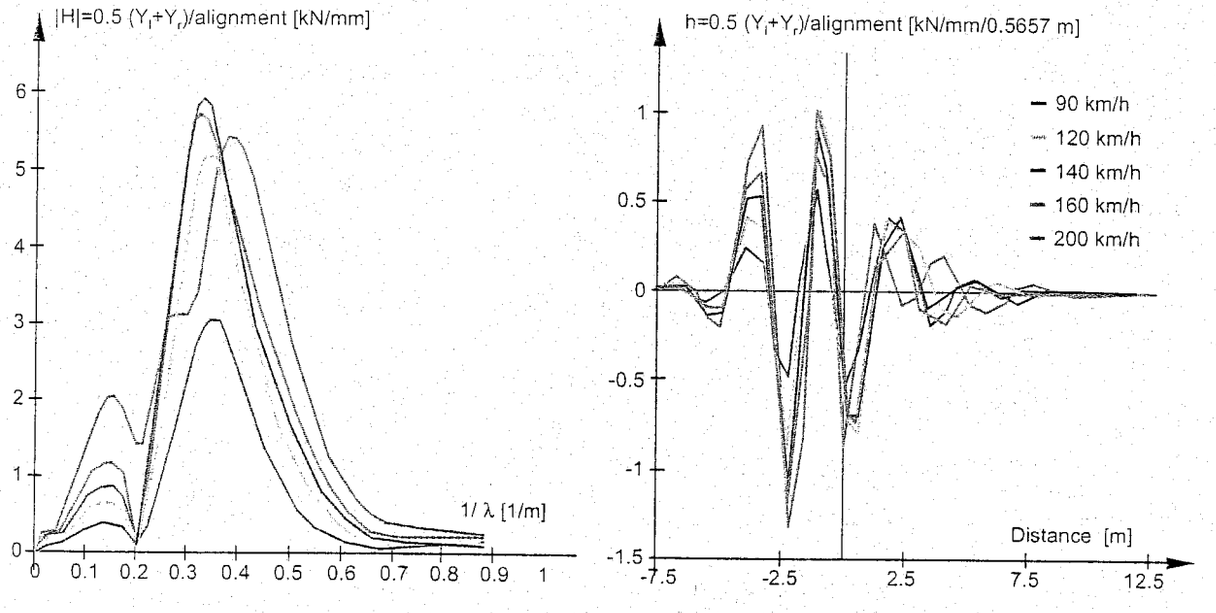


Figure 6.30: Relationships between Y-force and alignment

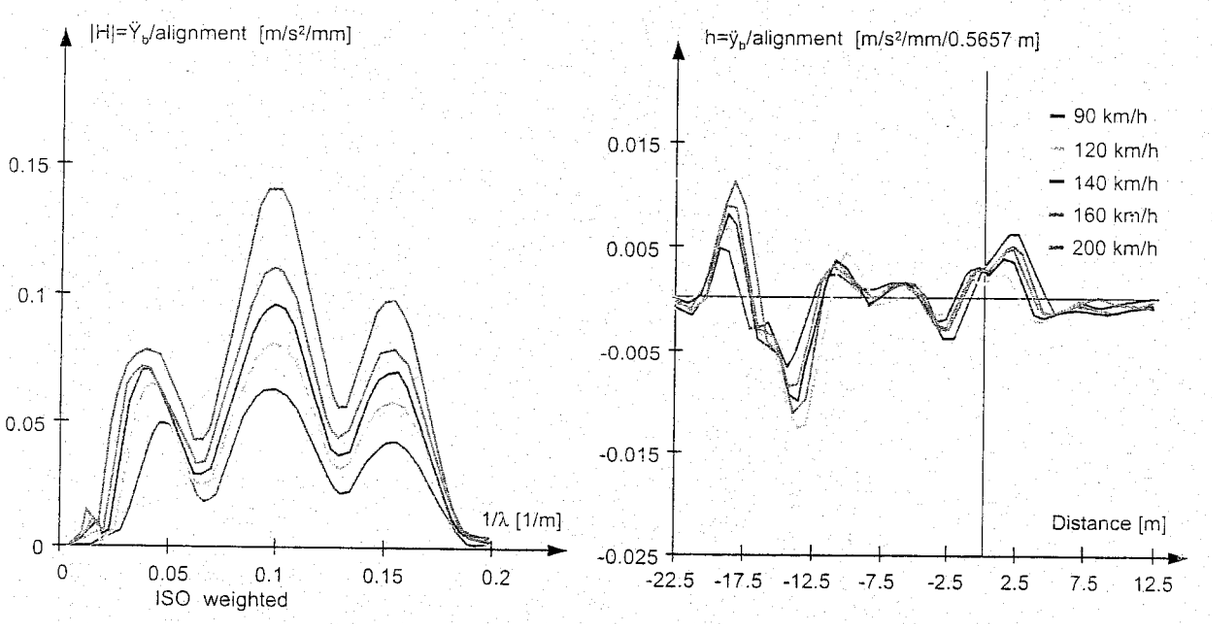


Figure 6.31: Relationships between lateral car body acceleration and alignment

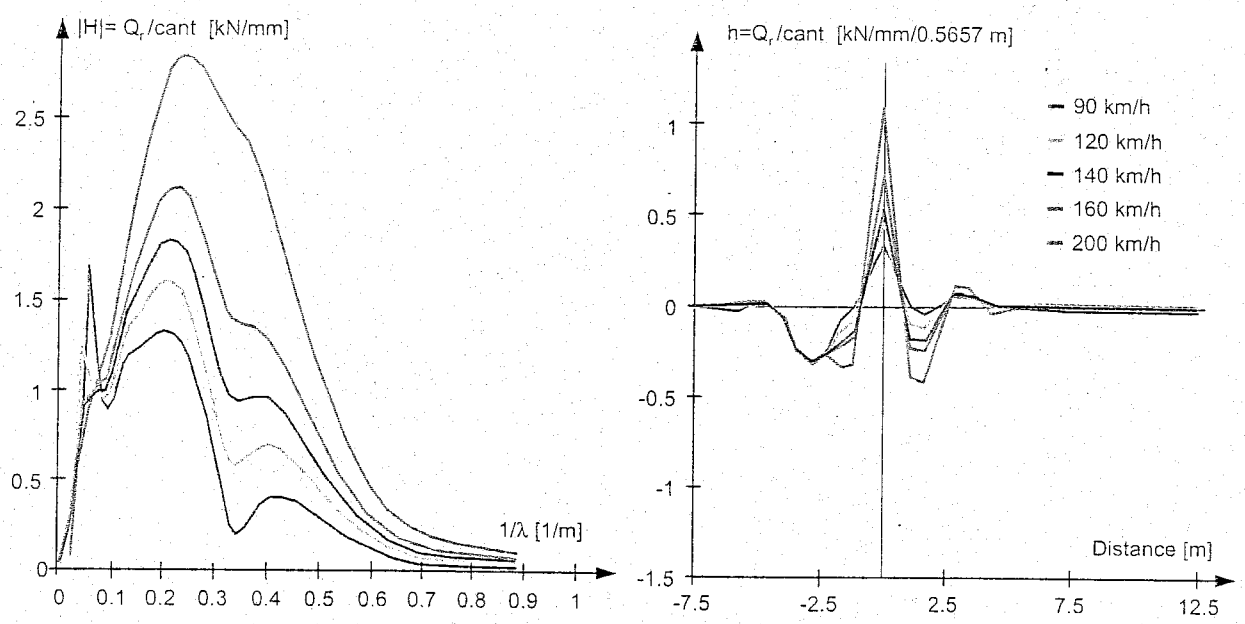


Figure 6.32: Relationships between increase in Q-force and cant

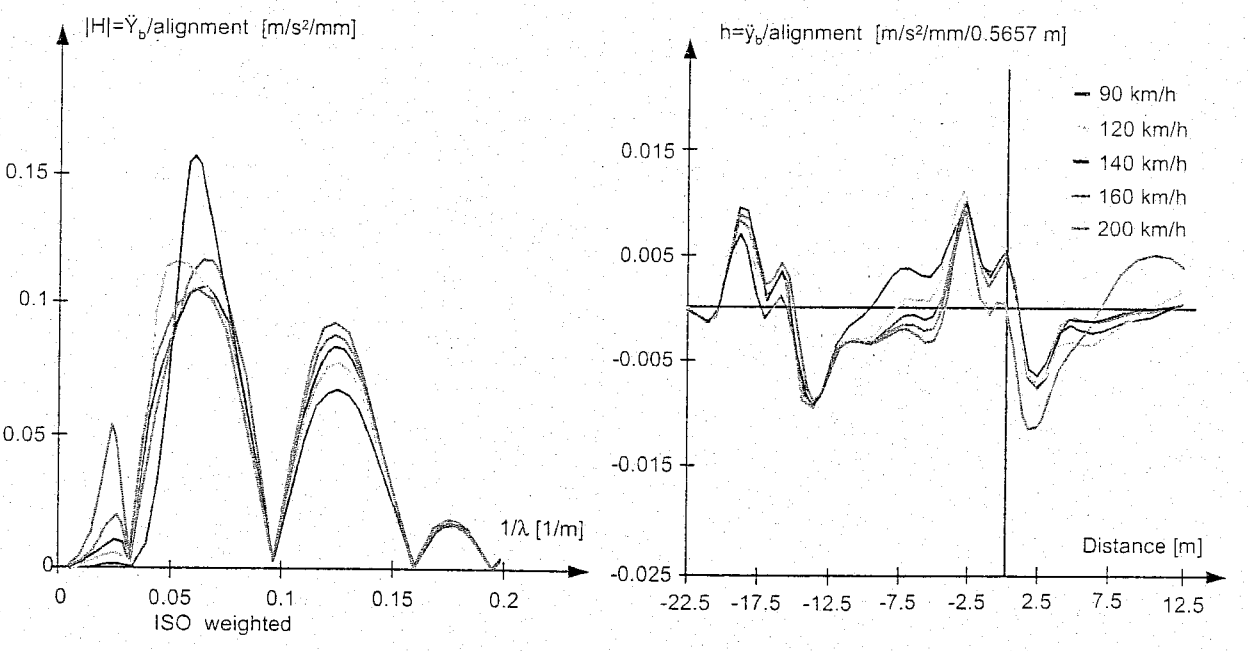


Figure 6.33: Relationships between lateral car body acceleration and cant

6.6 Estimate of transfer functions using measured data

6.6.1 General concept

The relationships between vehicle and track can be described in terms of transfer functions (see Section 6.5), for instance indicating how the various track geometry components contribute to a given response component of the vehicle. In the previous part the mathematical model approach based on a schematisation according to masses, springs, and dampers was discussed.

The method described here uses measured geometry signals as inputs and a corresponding response signals of the vehicle as output to establish transfer functions with the aid of the MISO method (Multiple Input Single Output) based on the theory of random signal analysis. The ORE Committee C 152, set up in 1979, has also dealt with this method and has meanwhile published reports [204], [205], and [74].

A number of examples are discussed concerning the estimate of transfer functions using recorded data, with special emphasis on the reliability aspect. In addition, a new concept for the dynamic measurement of Q and Y forces using measuring wheelsets is discussed.

6.6.2 Basic principles for 1 input and 1 output

Since the primary objective of the theory presented here consists of giving a survey of the main trends of the theory without entering into all sorts of minor details, no derivations will be discussed. As far as details and more basic considerations are concerned, reference is made, in the first instance, to the standard work by Bendat and Piersol [15] and to references [68], [69], [70], [74], and [78] which, in addition to practical implementations, include information on rail applications.

The theory of random signal analysis distinguishes between the time domain for dynamic processes (or the spatial domain for geometrical processes) and the frequency domain. The frequency is composed of the reciprocal time or the reciprocal distance for dynamic or geometrical processes, respectively. Although in the following text the magnitude t is used as time variable, this may be replaced by distance (x or s) too. Likewise, the frequency f may represent both the reciprocal time and the reciprocal distance. As a matter of fact, the variables time and distance are interlinked by the running speed.

If the signal x(t) denotes a magnitude in the time domain, the representation in the frequency domain is obtained by means of the so-called Fourier transformation. Provided that $\int_{-\infty}^{\infty}|x(t)|\,dt < \infty$ and consequently also $\int_{-\infty}^{\infty}|X(f)|\,df < \infty$, both transformations from and to the time domain read as follows:

$$X(f) = \int_{-\infty}^{\infty} x(t)\,e^{-i2\pi ft}\,dt \qquad (6.153)$$

$$x(t) = \int_{-\infty}^{\infty} X(f)\,e^{-i2\pi ft}\,df \qquad (6.154)$$

If these transformations are made digitally, this is done with the aid of the Fast Fourier Transformation (FFT) which is at present readily available in hardware.

If the 1-input-1-output model shown in Figure 6.34 is composed of a linear physically realisable system, the transfer function H(f) can be explicitly determined on the basis of the system parameters. For a measured input x(t), with the corresponding X(f), an output value Y(f) can be calculated for any f as follows:

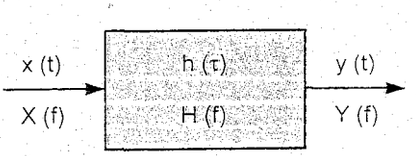


Figure 6.34: One-degree-of-freedom system

$$Y(f) = H(f)X(f) \qquad (6.155)$$

If, however, both input and output are measured, it might be wrongly inferred from (6.155) that the transfer function would follow from the quotient of output and input. However:

$$H(f) \neq \frac{Y(f)}{X(f)} \tag{6.156}$$

Both the real and the imaginary parts of the complex Fourier transformations have, in general, a rather irregular shape. Therefore, it is necessary to use quadratic spectral density functions, which must first be subjected to an averaging procedure, so as to obtain an acceptable statistical degree of reliability. Only after this may an estimate of H(f) be made. This will be discussed later on.

The relationship between input and output is described in the time domain as the convolution product of h and the input x according to (6.157). In the frequency domain this complicated procedure is reduced to a simple multiplication according to (6.155).

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau \qquad \textit{convolution} \tag{6.157}$$

$$Y(f) = H(f)X(f) \qquad \textit{multiplication} \tag{6.158}$$

In these expressions $h(\tau)$ represents the unit impulse response and H(f) the transfer function, interrelated as follows:

$$H(f) = \int_{-\infty}^{\infty} h(\tau)e^{-i2\pi f\tau}d\tau \tag{6.159}$$

$$h(\tau) = \int_{-\infty}^{\infty} H(f)e^{-i2\pi f\tau}df \tag{6.160}$$

From the Fourier transformations, spectral density functions may be deduced by multiplying the two moduli with each other and by subsequently dividing them by the record length T. This leads to the complex cross-spectrum $S_{xy}(f)$. If y is replaced by x, a real valued auto-spectrum $S_{xx}(f)$ is obtained. In (6.161), $\overline{X}$ stands for the complex conjugate of X.

$$S_{xy}(f) = \lim_{T\to\infty} \frac{1}{T}\overline{X}(f)Y(f) \tag{6.161}$$

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)y(t+\tau)d\tau \tag{6.162}$$

Here, too, an equivalent operation in the time domain exists and, thus, leads to the cross correlation function $R_{xy}(t)$ shown in (6.162). From the point of view of the calculation technique, this expression is very similar to the convolution process discussed before. By replacing y by x, the autocorrelation function R_{xx} is calculated.

In an absolutely analogue way, as in (6.159) and (6.160), $S_{xy}(f)$ and $R_{xy}(t)$ are interrelated by means of a Fourier transformation. These expressions are known in literature as the Wiener-Kintchine relationships which read as follows:

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau)e^{-i2\pi f\tau}d\tau \tag{6.163}$$

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} S_{xy}(f)e^{-i2\pi f\tau}df \tag{6.164}$$

An important feature of the auto-spectra is that they are symmetrical with respect to the line f=0, which is illustrated in Figure 6.35. Moreover, the area equals the variance according to:

$$\sigma_x^2 = \int_{-\infty}^{\infty} S_{xx}(f)df = 2\int_0^{\infty} S_{xx}(f)df \qquad (6.165)$$

The cross-correlation function is particularly useful in determining the shift between two signals. This displacement corresponds, in fact, to the place where maximum correlation occurs. Figure 6.36 illustrates this approach.

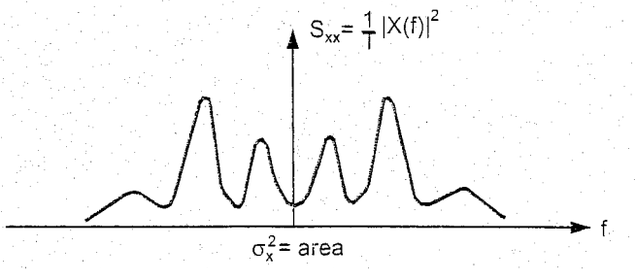


Figure 6.35: Symmetrical auto-spectrum

In quantifying the correlation between two signals, a consideration in the frequency domain is once more applicable, designated as the coherence $\gamma_{xy}^2(f)$, as well as a consideration in the time domain resulting in the correlation function $\rho_{xy}^2(\tau)$.
These expressions read as follows:

$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \qquad 0 \le \gamma_{xy}^2(f) \le 1 \qquad (6.166)$$

$$\rho_{xy}^2(\tau) = \frac{R_{xy}^2(\tau)}{R_{xx}(0)R_{yy}(0)} \qquad 0 \le \rho_{xy}^2(\tau) \le 1 \qquad (6.167)$$

In (6.167) $R_{xx}(0) = \sigma_x^2$ and $R_{yy}(0) = \sigma_y^2$. In addition, it should already be observed here that the coherence according to (15.136) only furnishes useful information if the spectra S_{xx}, S_{yy}, and S_{xy} have been averaged according to the rules to be discussed in section 6.6.4. This also applies to the formulae for estimating the transfer function H(f) which will be discussed now.

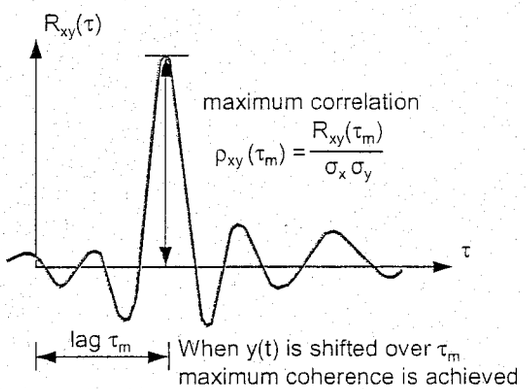


Figure 6.36: Lag to achieve maximum correlation

From the relationship between input and output, the following relation on a spectral level may be deduced:

$$S_{xy}(f) = H(f)S_{xx}(f) \qquad (6.168)$$

From this it follows for the transfer function H(f) that:

$$H(f) = \frac{S_{xy}(f)}{S_{xx}(f)} \qquad (6.169)$$

The coherence function $\gamma_{xy}^2(f)$ is obtained from:

$$\gamma_{xy}^2(f) = \frac{H(f)S_{xy}(f)}{S_{yy}(f)} \qquad (6.170)$$

A simulation study of the influence of non-correlated contributions to the output on the error in estimating the transfer function, published in [205], has shown that about 10% of non-correlated data in the output leads to an error of about 10% in the transfer function estimate, the coherence being reduced to about 0.8. Seen from this angle, whilst also allowing for other possible causes, only estimates of H(f) for which $\gamma^2_{xy}(f) > 0.85$ should be accepted.

6.6.3 Multiple input single output (MISO)

The model depicted in Figure 6.37 shows how the q inputs $x_i(t)$ produce, by means of q linear systems, q outputs $y_i(t)$ which together constitute the overall output according to:

$$y(t) = \sum_{i=1}^{q} y_i(t) \qquad (6.171)$$

The outputs $y_i(t)$ follow from the inputs by means of the convolution products:

$$y_i(t) = \int_{-\infty}^{\infty} h_i(\tau)x(t-\tau)d\tau \qquad (6.172)$$

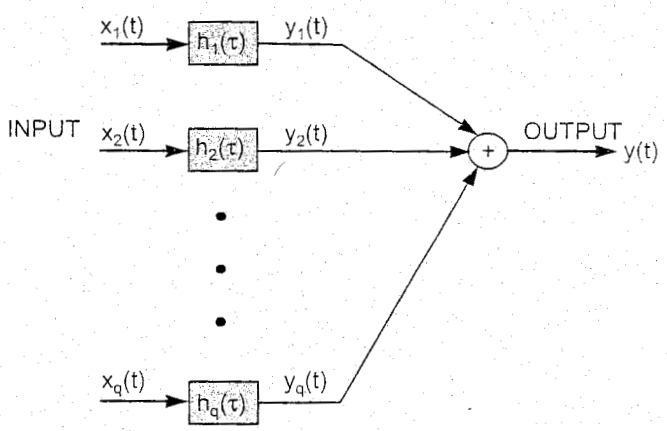


Figure 6.37: MISO model

Assuming the process to be stationary, auto- and cross-correlation functions may be deduced furnishing, by means of Fourier transformation, the following set of equations in the frequency domain:

$$S_{iy}(f) = \sum_{j=1}^{q} S_{ij}(f)H_j(f) \qquad (6.173)$$

in which according to (6.161):

$$S_{iy}(f) = \lim_{T\to\infty} \frac{1}{T} \overline{X}_i(f)Y(f) \qquad (6.174)$$

$$S_{ij}(f) = \lim_{T\to\infty} \frac{1}{T} \overline{X}_i(f)X_j(f) \qquad (6.175)$$

In matrix notation (6.173) reads as follows:

$$\{S_{xy}\} = [S_{xx}]\{H\} \qquad (6.176)$$

The generation and solution of these q complex equations is discussed in section 6.6.5. However, the formal solution can be written provisionally as follows:

$$\{H\} = [S_{xx}]^{-1}\{S_{xy}\} \qquad (6.177)$$

The reliability of the transfer functions thus estimated follows from the multiple coherence function $\gamma^2_{y\bullet x}(f)$ which depicts the ratio between the output spectrum calculated on the basis of (6.177) and the measured output spectrum according to:

$$\gamma^2_{y\bullet x}(f) = \frac{S_{yy\ calculated}}{S_{yy\ measured}} \qquad 0 \le \gamma^2_{y\bullet x}(f) \le 1 \qquad (6.178)$$

After substitution this leads to:

$$\gamma^2_{y\bullet x}(f) = \frac{\sum_{i=1}^{q} \bar{H}_i(f) S_{iy}(f)}{S_{yy}(f)} = \frac{\{\bar{H}\}^T \{S_{xy}\}}{S_{yy}(f)} \tag{6.179}$$

Summation in the numerator indicates the contribution of the various inputs to the output spectrum. In this case, too, it is necessary that for a reliable estimate $\gamma^2_{y\bullet x}(f)$ must lie in the interval

$$0.85 \leq \gamma^2_{y\bullet x}(f) \leq 1 \tag{6.180}$$

6.6.4 Statistical reliability

In this section some attention is paid to random errors ε_r and systematic errors (bias errors) ε_b. Systematic errors can be compensated for by correction or calibration according to:

$$X = \tilde{X}(1 - \varepsilon_b) \qquad -1 \leq \varepsilon_b \leq 1 \tag{6.181}$$

where X denotes the true value and $\tilde{X}$ the estimator. If the systematic error is negligibly small (ε_b < 0.02) and the random error is small (ε_r < 0.10), the 95% confidence interval for X is given approximately by:

$$\tilde{X}(1 - 2\varepsilon_r) \leq X \leq \tilde{X}(1 + 2\varepsilon_r) \tag{6.182}$$

In this section only some main trends are given. For a more detailed discussion see references [15], [16], [17], and [106].

Random errors

As stated before, spectral density functions must be averaged so as to keep the random error at an acceptably low value. This averaging procedure can be carried out in two different ways: either by averaging out the records (ensemble averaging), this number amounts to NSEC, or by combining a number of frequency components known as frequency smoothing, this number amounts to NA. The overall number of averaging operations n is thus:

$$n = NA * NSEC \tag{6.183}$$

The random error ε_r in spectral density functions is inversely proportional to $\sqrt{n}$ and follows from:

$$S_{xx} \rightarrow \varepsilon_r = \frac{1}{\sqrt{n}} \tag{6.184}$$

$$S_{xy} \rightarrow \varepsilon_r = \frac{1}{|\gamma^2_{y\bullet x}|\sqrt{n}} \tag{6.185}$$

The random error ε_r in the modulus of the transfer function H_i, determined by means of MISO, can be approximated according to:

$$|H_i| \rightarrow \varepsilon_{r|H_i|} = \sqrt{\frac{q}{n-q} F_{n_1, n_2, \alpha} \frac{1 - \gamma^2_{y\bullet x}}{1 - \gamma^2_{i\bullet x}} \frac{S_{yy}}{S_{xx}}} \tag{6.186}$$

where:
q = number of inputs;
n = number of averaging operations;
F = 100 α percentage point of the F-distributions, with $n_1 = 2q$ and $n_2 = 2(n-q)$;
$\gamma^2_{i \bullet x}$ = multiple coherence function between input x_i and the other inputs.

From the first term under the square root sign it becomes apparent that the number of averaging operations n must at least be equal to q + 1. In practice, the value of n will certainly have to be one order of magnitude higher. The random error in the argument of H_i follows from:

$$\arg H_i \rightarrow \varepsilon_r = \arcsin \frac{\varepsilon_{r|H_i|}}{|H_i|} \tag{6.187}$$

Bias errors

On the subject of estimating bias errors relatively little is known. For auto-spectra, a Taylor-series expansion for approximating the bias error may be applied. This leads to the following formula:

$$\varepsilon_b \cong \frac{B_e^2}{24} \frac{S''_{xx}}{S_{xx}} \cong -\frac{1}{3} \frac{B_e^2}{B_r^2} \tag{6.188}$$

where:
B_e = resolution band width = Δf = NA/L;
L = record length;
NA = number of frequency smoothing operations;
B_r = "half power point" band width ≈ 0.03 m^{-1} for track geometry spectra.

Summary of random and bias errors

Obviously, the requirements to keep the random error and the bias error small are sometimes incompatible since:

$$\varepsilon_b \therefore \frac{(NA)^2}{L^2} \tag{6.189}$$

$$\varepsilon_r \therefore \frac{1}{\sqrt{NA * NSEC}} \tag{6.190}$$

Averaging over frequencies soon leads to large bias errors whereas long records and a high value of NSEC are attractive, but soon lead to measuring problems. The following maximum values of random and bias errors could be adhered to:

$$\begin{aligned} \varepsilon_r &< 0.10 \\ \varepsilon_b &< 0.025 \end{aligned} \tag{6.191}$$

With $B_r \approx 0.03$ m^{-1}, this leads to the following choice of parameters:

$$\begin{aligned} NA &= 4 \\ NSEC &= 25 \\ L &= 500\ m \end{aligned} \tag{6.192}$$

The distance to be measured in this case is therefore 12.5 km. In [205] a minimum length of 10 km was applied.

6.6.5 Numerical aspects

The majority of numerical questions forms the subject of a detailed description in references [68], [69], and [74]. Despite this, some aspects will be explained in greater detail here.

Frequency smoothing

Smoothing of frequencies should be carried out symmetrically about the line f=0 so as to preserve the even and odd character of the real and imaginary parts, respectively. This averaging is undertaken separately for the real part and the imaginary one. If NA is odd, NA values are summed up and divided by NA. If, on the other hand, NA is even, NA + 1 points are considered of which the central NA - 1 points are entirely included and only 50% of the 2 boundary points. This sum is once more divided by NA. This is schematically shown in Figure 6.38.

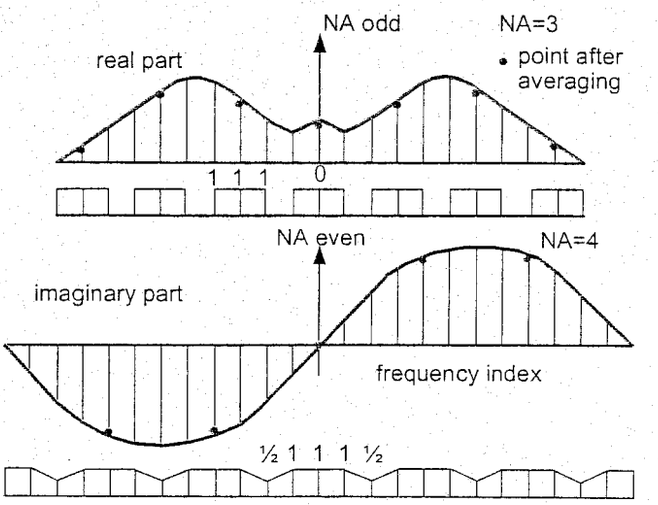


Figure 6.38: Principle of symmetrical frequency smoothing

The generation of MISO equations

A general routine has been written for generating spectra, enabling the cross spectrum of two sets of Fourier transformations to be determined according to:

$$G_{mn} = \frac{2}{L \cdot NSEC} \sum_{k=1}^{NSEC} \bar{X}_m^k X_n^k \qquad (6.193)$$

where:

X_m^k, X_n^k	are the Fourier transformations from series k of the ensemble;
$\frac{1}{NSEC} \sum_{k=1}^{NSEC}$	stands for "ensemble averaging";
L	is the record length.

The factor 2 denotes that what are known as one-sided spectra are used, owing to which only values for $f \geq 0$ are considered.

After the spectra thus established have been averaged out over the frequency components in accordance with the previously described procedure, the set of equations defined above in (6.176), which have now been subjected to the required averaging operations, are obtained:

$$\{\tilde{G}_{xy}\} = [\tilde{G}_{xx}]\{\tilde{H}\} \qquad (6.194)$$

This is a system of q complex equations, which after the substitutions:

$$\{\tilde{G}_{xy}\} = \underline{a} + i\,\underline{b} \qquad (6.195)$$

$$\{\tilde{H}\} = \underline{c} + i\,\underline{d} \qquad (6.196)$$

$$[\tilde{G}xx] = R + iQ \tag{6.197}$$

can be converted into the following system of 2q real equations:

$$\begin{bmatrix} \underline{a} \\ \underline{b} \end{bmatrix} = \begin{bmatrix} R & -Q \\ Q & R \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} \tag{6.198}$$

The sub-matrices have the following properties:

$$R^T = R \tag{6.199}$$

$$Q^T = -Q \tag{6.200}$$

$$Q_{ii} = 0 \tag{6.201}$$

The set (6.198) may now be written as:

$$\underline{y} = A\underline{x} \tag{6.202}$$

in which, by virtue of (6.199), (6.200), and (6.201), matrix A is symmetrical. This set of equations can be solved using the decomposition method of Crout-Cholesky briefly described in the following section.

Solving equations

In set (6.202) matrix A describes the inputs while the output only occurs in the $\underline{y}$ vector. In practice, a full series of output signals is usually involved, implying that the (6.202) system would then have to be solved just as many times. In the solution technique of Crout-Cholesky, the arithmetical operations are therefore split up into two parts, i.e. into one part which is independent of the output, thus only bearing on matrix A, and into one part affecting the whole system. The single-time operation on matrix A is designated to factorize and splits up this matrix into two triangular matrices and one diagonal matrix, according to:

$$A = U^T D\, U \tag{6.203}$$

where:
U = upper triangular matrix;
U_{ij} = 0 for i < j; U_{ii} = 1;
D = diagonal matrix; D_{ij} = 0 for i = j.

This is executed in the subroutine FACBAN. The solution proper is now carried out in two steps by means of the subroutine SYMBAN:

$$U^T\underline{z} = \underline{y} \tag{6.204}$$

$$D\, U\, \underline{x} = \underline{z} \tag{6.205}$$

In equation (6.204) vector $\underline{z}$ is solved from top to bottom, after which vector $\underline{x}$, the solution vector of set (6.202), is found by going through (6.205) from bottom to top.

6.6.6 Applications

The MISO applications described here are primarily confined to the field of interaction between vehicle and track. The method is aimed at determining the relationships between track geometry compo-

nents, serving as inputs, and a vehicle response magnitude, representing the output. The model describing the approach by means of MISO is depicted in Figure 6.39. The geometry components cant, level, alignment, and gauge constitute the input whereas so far only car body accelerations have been considered as output.

Within the scope of the ORE C 152 work program a great number of measurements were carried out, during which the track geometry and the vehicle response were recorded on magnetic tape. During evaluation of these measurements the coherence turned out to be much too low in various cases. In most cases this could be ascribed to problems in the signal-to-noise ratio. For example, on some measured line sections the quality was so high that car body accelerations were barely measurable. If, in addition, the signals were not fully amplified before being recorded onto tape, little else but noise remains for analysis.

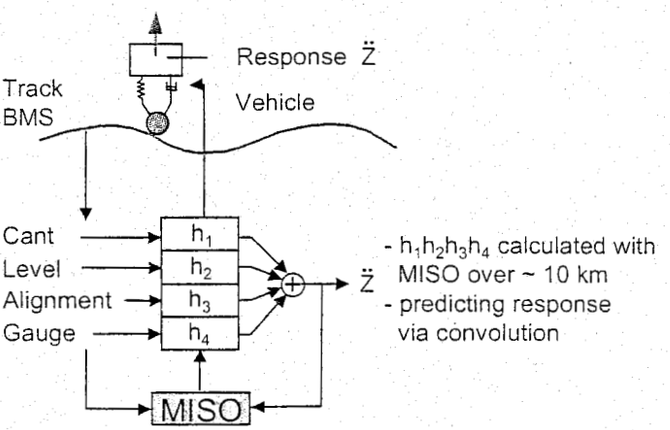


Figure 6.39: MISO model for estimating transfer functions between track geometry and vehicle reactions

Another important factor is the frequency range within which the measured signals fall. If this range is different for input and output, the correlation between the two cannot be expected to be good either. Such problems occur, for example, if low frequency car body accelerations which have a wavelength of 30 - 40 m in the higher speed ranges are compared with track geometry measured with a conventional track recording system capable only of measuring wavelengths up to 20 - 25 m.

Conversely, however, problems will also arise if an attempt is made to relate vehicle reactions of high frequency, such as axle box accelerations and dynamic Q and Y forces, to the track geometry measured in the waveband between 0.5 and 25 m. In this case the geometry will have to be high-pass filtered to remove the long waves with relatively high energy.

Some examples

The NS track recording car

The NS recording car, in which the BMS system is installed, is fitted with Y-32 bogies which have a very linear spring characteristic.

Within the scope of the C 152 studies, a MISO analysis was applied to this recording car with the vertical body acceleration furnished by BMS which is considered to be the vehicle reaction.

The results are given in Figure 6.40 to Figure 6.45.

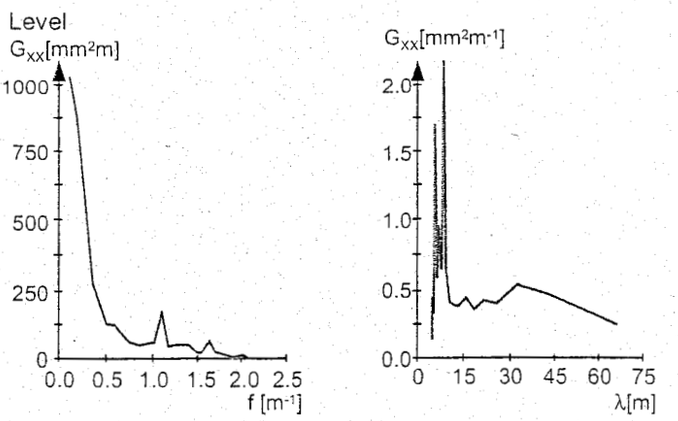


Figure 6.40: Geometry spectra recorded by BMS

Figure 6.40 shows the power spectrum of the track geometry component "level" as a function of the spatial frequency f [m^{-1}] and the wavelength λ [m]. The response spectra at 80 and 120 km/h are shown in Figure 6.41. These spectra were calculated using 20 records of 500 m length and a frequency smoothing factor NA = 4, so that the bias and random errors remain within the limits mentioned in (6.191).

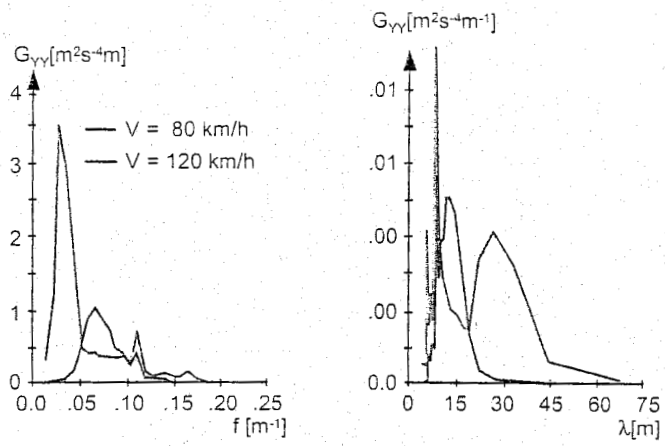


Figure 6.41: Vertical car body acceleration spectra measured on the NS recording car

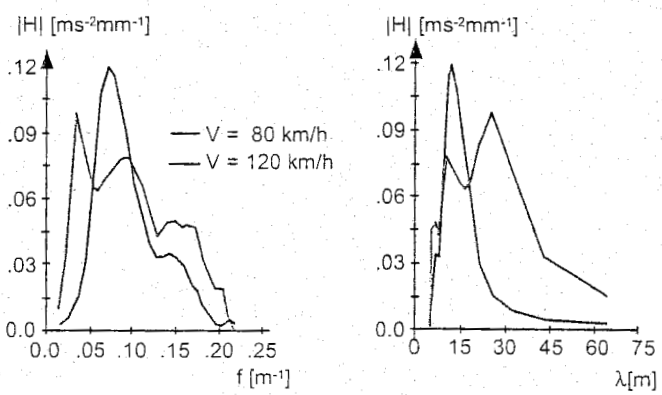


Figure 6.42: Transfer functions between level and vertical car body acceleration estimated for the NS recording car

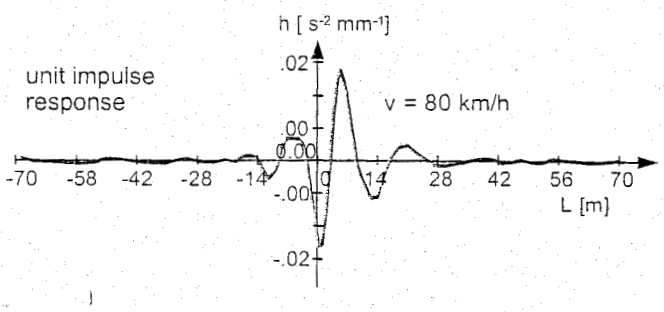


Figure 6.43: Unit impulse response function based on the data in Figure 6.42

As an example Figure 6.42 shows the transfer functions H_2 between level and vertical car body response; the other transfer functions are negligibly small.

These are, in fact, the primary results of the MISO analysis according to (6.177). In agreement with formula (6.160), these functions have been Fourier transformed so as to obtain the unit impulse response functions h_i of which h_2 is illustrated in Figure 6.43. In this case, too, all other h-functions can be neglected, implying that only the level contributes to the vertical acceleration.

The degree to which the transfer function values are reliable is shown by the multiple coherence function $\gamma^2_{y\bullet x}(f)$ depicted in Figure 6.44. As stated before, the $\gamma^2_{y\bullet x}(f)$ value should be higher than 0.85 if practical applications are to be made possible. On further analysis of the results, it appears that the $\gamma^2_{y\bullet x}(f)$ value only meets this requirement in the frequency bands where the measured signals contain enough energy. This is rather obvious and it also explains, perhaps in a different way, why long measuring sections should be chosen, preferably with maximum possible variation in the geometry spectra.

To complete the sequence of computations, the response is once more calculated as a function of the distance covered by means of the convolution principle according to (6.157), using the unit impulse response functions previously obtained, and is compared with the response signal originally measured. Figure 6.45 contains a graphical representation of the calculated and measured signals; the similarity between the measured and the calculated response is remarkable.

To quantify the deviation between the two signals, the RDS value (relative difference between standard deviations) is determined for each 200 m sub-section. The RSD value is defined as:

$$RDS = \frac{\sigma_{measured} - \sigma_{calculated}}{\sigma_{measured}} \qquad (6.206)$$

This value roughly conforms to the value $1-|\gamma^2_{y\bullet x}(f)|$, with respect to which the mean value of $|\gamma^2_{y\bullet x}(f)|$ must be imagined over the area in which the energy in the signal is concentrated. In fact, RSD denotes the error in standard deviation if the latter has been calculated using the transfer functions obtained from MISO.

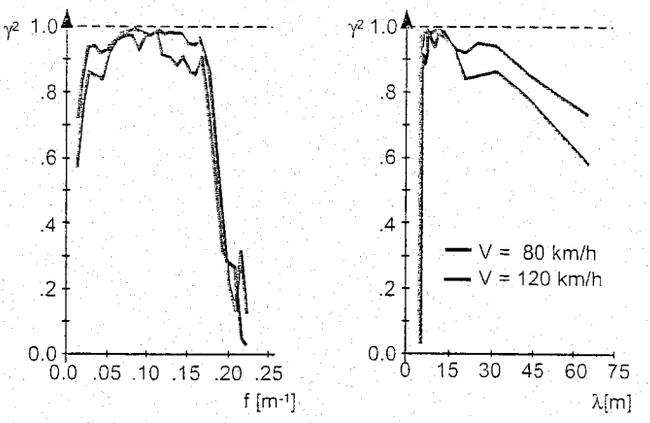


Figure 6.44: Multiple coherence for the estimates in Figure 6.42

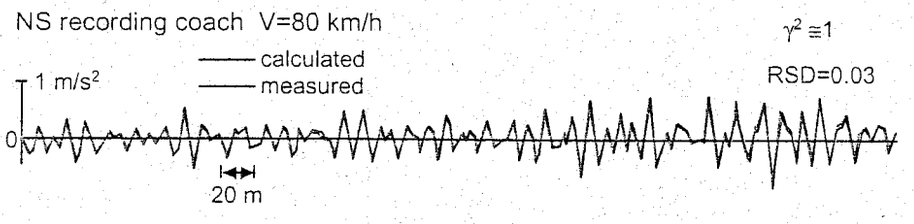


Figure 6.45: Comparison between measured and calculated response based on the estimated transfer function in Figure 6.42

Results of other measurements

As part of the C 152 study mentioned above, various other measurements were analysed using MISO. Some characteristic results will now be highlighted.

The data contained in Figure 6.46 and Figure 6.47 refer to SNCF measurements of vertical car body accelerations of a CORAIL coach. The relationship between these measurements and the track geometry recordings made by the Mauzin car have been established. Figure 6.48 shows the coherence to be very poor with a maximum value of 0.6.

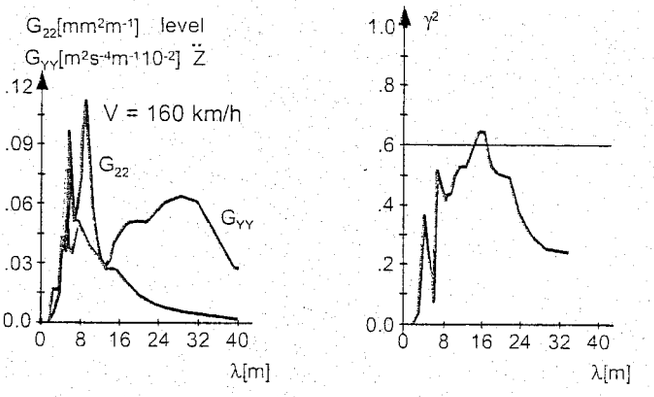


Figure 6.46: Example of bad coherence for transfer estimate regarding SNCF Corail coach

The factors mentioned earlier concerning compatibility in frequency range and proper signal amplification are, in all probability, responsible for this. The calculated and measured responses according to Figure 6.49 show, in a different way, that in such a case the results are useless for practical application.

The last example concerns measurements on two-axle wagons of KS design taken by the Hungarian State Railways (MAV). Despite the rather non-linear spring characteristic of this type of rolling stock, the coherence turns out to be fairly high. Figure 6.50 shows the coherence vertically with a peak value of 0.95. The response calculated according to Figure 6.49 consequently conforms very well to the measured vehicle response. Considered laterally, the coherence shown in Figure 6.50 turns out to be somewhat lower with a peak value of about 0.8. As can be seen from Figure 6.51, the agreement between the measured and calculated response figures is nevertheless quite reasonable.

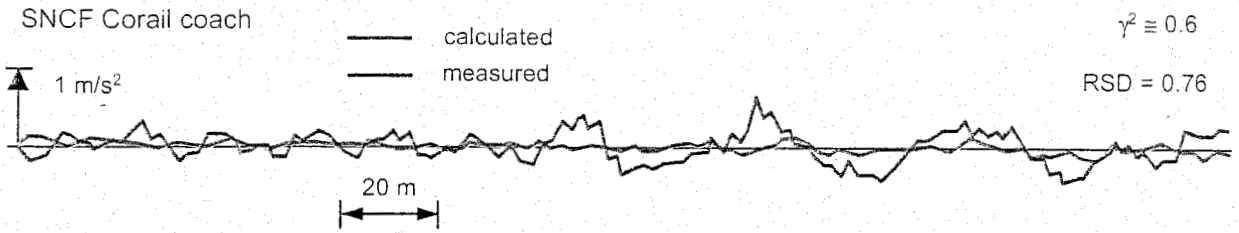


Figure 6.47: Comparison between measured and calculated response based on the data in Figure 6.46

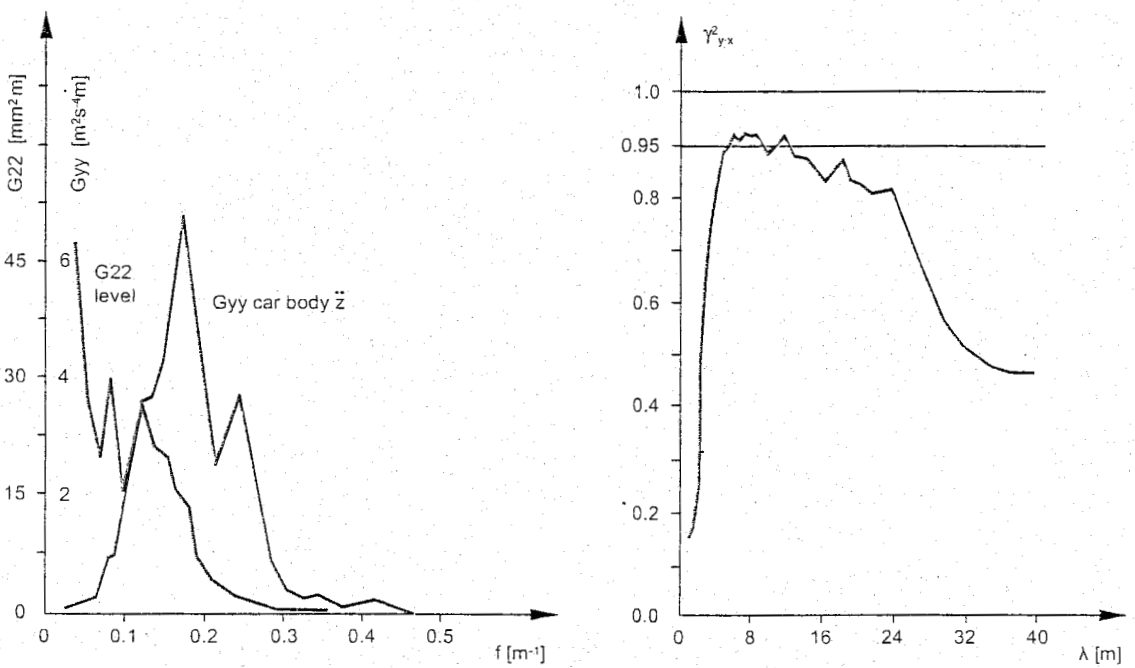


Figure 6.48: Example of vertical response for MAV freight car

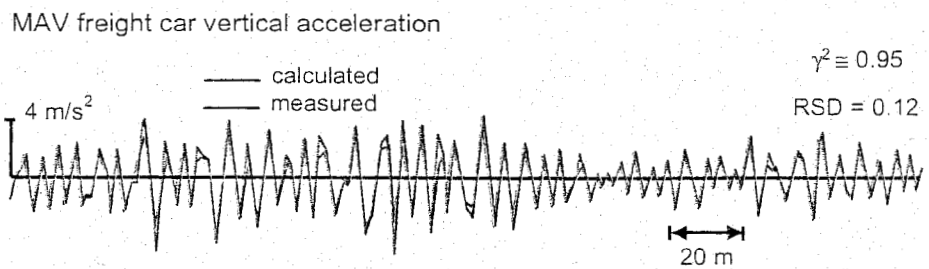


Figure 6.49: Comparison between measured and calculated response based on the data in Figure 6.48

In 1986 a short BMS measuring car campaign was carried out in Sweden to provide SJ with geometry data for vehicle model input. On this occasion SJ also recorded Q and Y forces on a locomotive with the aid of measuring wheelsets. With these data and the geometry recorded by BMS some MISO analyses were carried out. These results are summarized in Figure 6.52. In the vertical direction the multiple coherence is sufficiently high for wavelengths shorter than 8 meter. The transfer function increases progressively with frequency. This reflects the strong influence of the unsprung mass which contributes to the transfer function in proportion to the square of the frequency. In the lateral direction the coherence is somewhat lower. The shape of this estimate has a peak at about 16 m wavelength, remains more or less constant between 13 and 14.5 m and then increases progressively for shorter waves. Although the theoretical transfer functions presented in Figure 6.28 and Figure 6.30 refer to much lower unsprung masses, the shapes resemble the estimated transfer functions quite well.

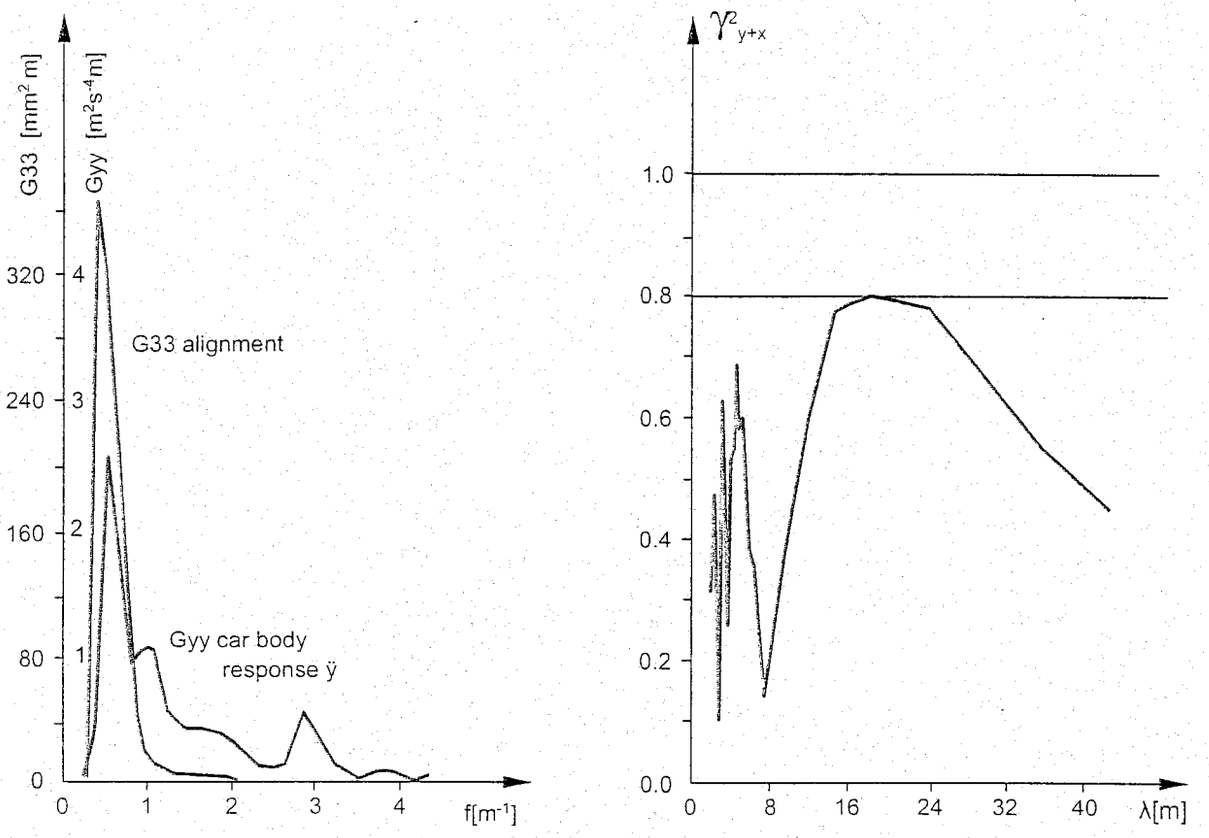


Figure 6.50: Example of lateral response for MAV freight car

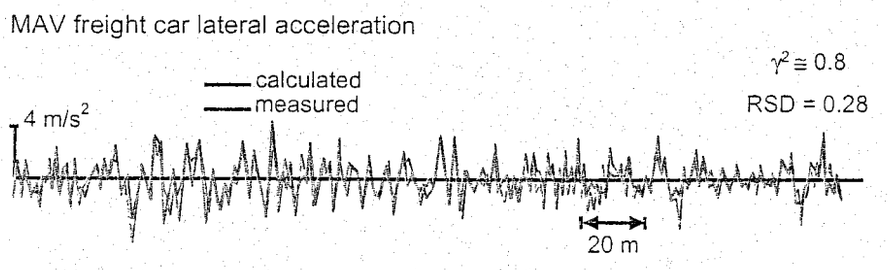


Figure 6.51: Comparison between measured and calculated response based on the data in Figure 6.50

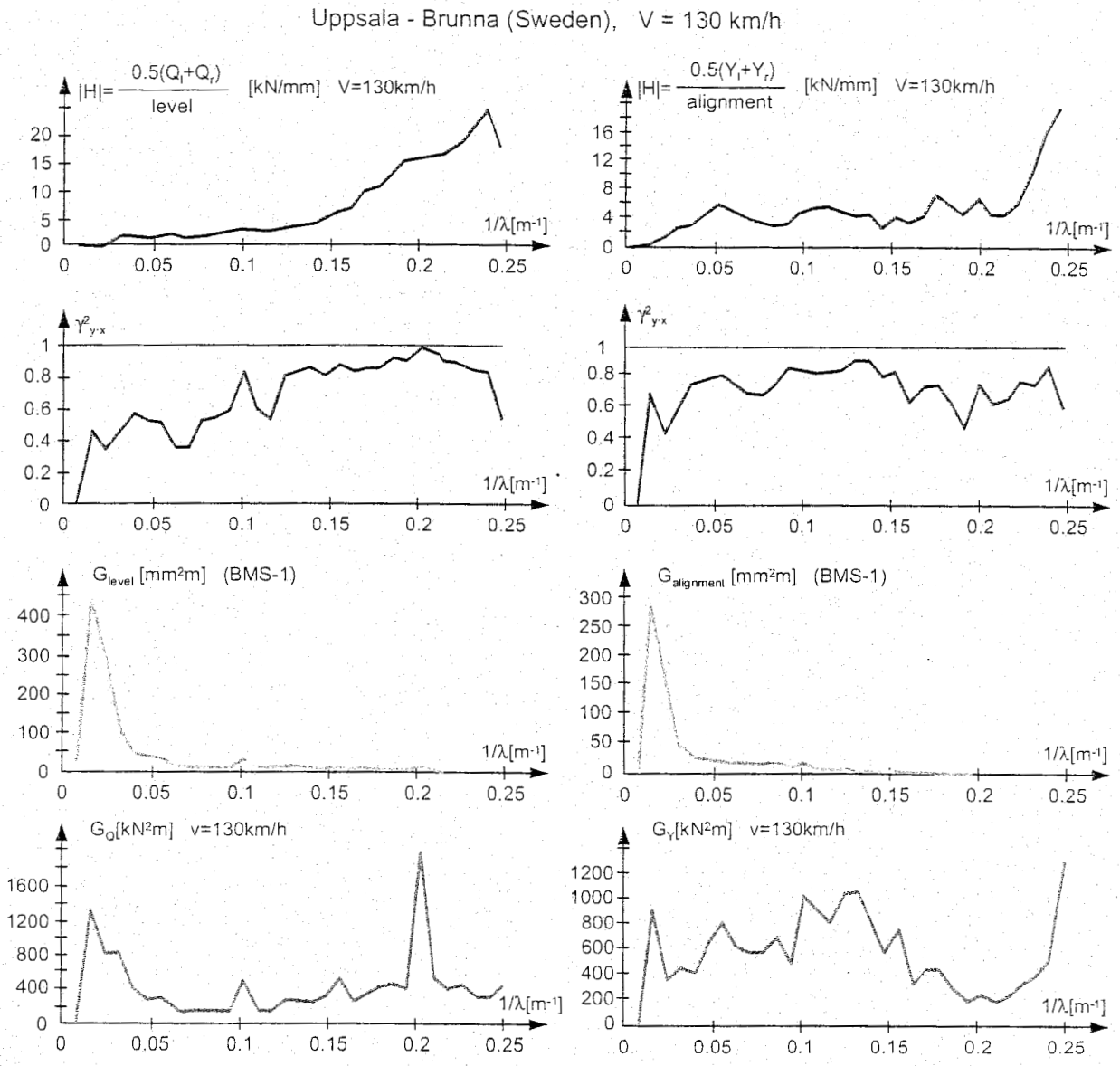


Figure 6.52: Estimated transfer function for the Q and Y forces of an SJ locomotive

Measuring wheelsets in order to determine Q and Y forces

When measuring Q and Y forces with the aid of measuring wheelsets, use is made of strain gauges from which the respective forces are deduced. The required relationships between strains and forces are determined experimentally on a roller test rig. Here, the variation in strain versus distance covered is calculated for one force component which is then kept constant. This is in fact a static calibration which restricts Q and Y force measurements to quasi-static phenomena [228]. Such a system gives incorrect values for real dynamic phenomena at frequencies between 2 and 150 Hz. As a matter of fact, strain gauges do furnish a signal in this frequency band, but the transfer functions between strains and forces are unknown. These transfer functions could be estimated using MISO during calibration on the roller test rig.

Static calibration requires that the number of strain measuring points are equal to the number of forces to be measured, since the experimental test arrangement only enables the relation between force and strain to be determined; the relationship sought in the reverse direction then follows from this by matrix inversion. The dynamic relationships in the frequency domain offer a similar picture, though all the magnitudes are now frequency-dependent.

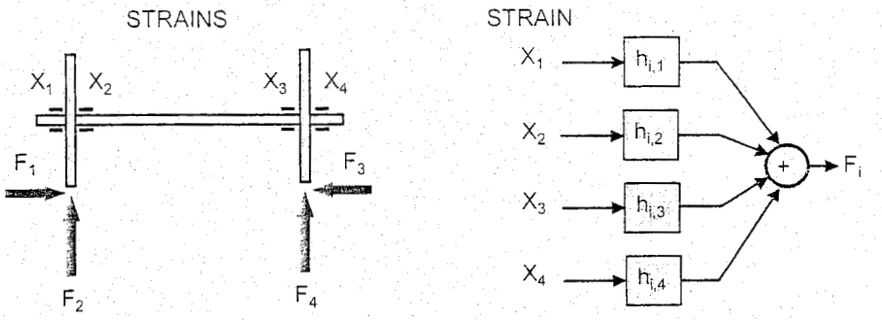


Figure 6.53: Principle of force measuring by means of a wheelset

Figure 6.53 shows the arrangement of the strain gauges and the forces acting on the measuring wheelset. The conventional quasi-static approach remains invariably valid and must be considered separately from the dynamic calibration.

In order to estimate the transfer functions between the strain gauges and each force component, the forces exerted by the test rig on the wheels must be varied randomly in the frequency range of application. If sampling is carried out as a function of distance or revolution, these realisations should be repeated at different running speeds. Using these measurement series, the different transfer functions can be estimated for each speed by means of MISO by considering the strains as input and each force as output, as indicated in Figure 6.53.

During actual measurements the Q and Y forces can be derived from the strains in real time by applying the estimated transfer functions. These should be continuously adapted to the actual running speed by interpolation.

6.6.7 Comparison between transfer functions estimated by MISO and calculated with models

In section 6.5 a simple vehicle model was discussed which was used to derive transfer functions, amongst other things, between level and vertical car body acceleration. These results, presented in Figure 6.29, have been compared with the estimated transfer functions using MISO, depicted in Figure 6.42. Both results for the same speed and running direction are displayed in Figure 6.54. Clearly there is quite good agreement between both functions.

The practical advantage of the model approach is that additional estimates for different speeds can easily be made, whereas in the case of MISO a test run should be carried out for each speed. MISO has of course the big advantage that all important natural frequencies are automatically taken into account.

In practice, a hybrid solution may be feasible in which a MISO estimate is made for one or two speeds. After adapting the model parameters in such a way that the MISO estimates are approximated as well as possible, the transfer values at other speeds can be obtained by means of calculation.

6.7 Vehicle response analysis in real time

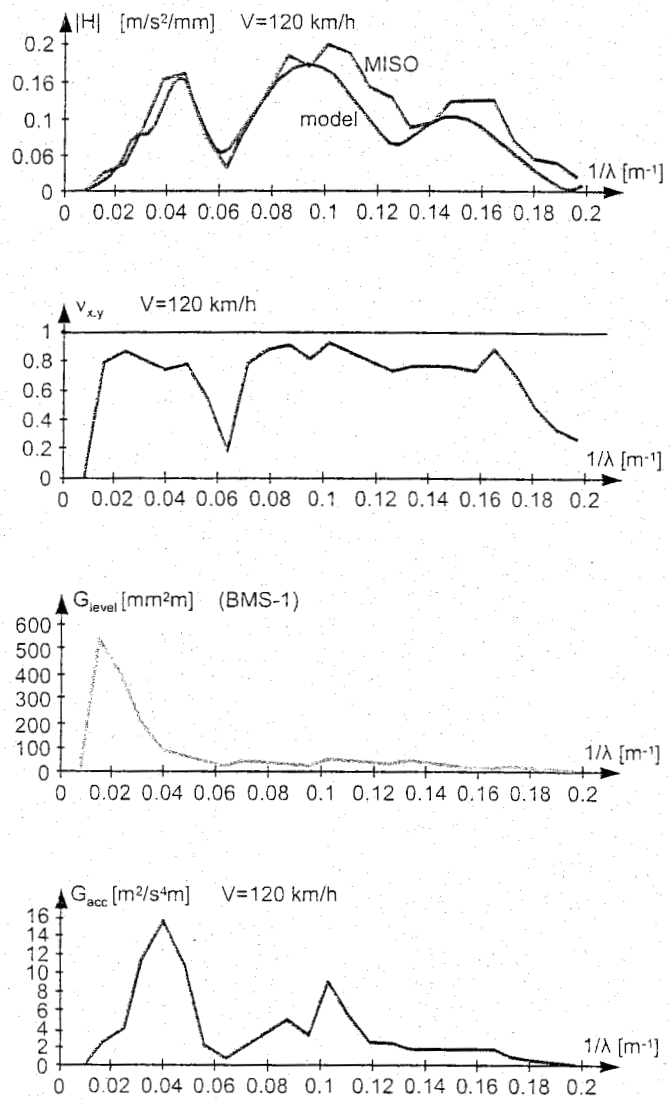


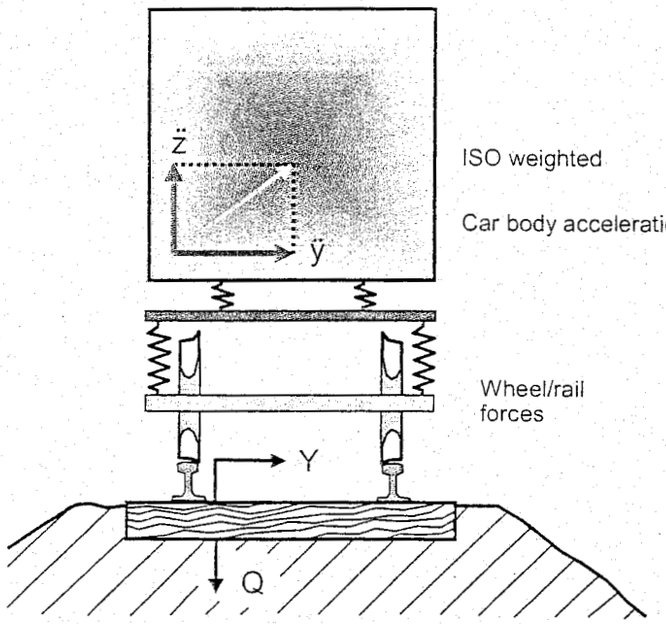
Figure 6.54: Transfer function between level and car body acceleration for NS recording car estimated by means of MISO and calculated with models

In 1988 the computers of the NS track recording system BMS were replaced by one powerful computer system. Progress in both information technology and railway vehicle dynamics, as described earlier in this chapter, allowed the development of a Vehicle Response Analysis (VRA) system using a state-of-the-art Array Processor in combination with a central minicomputer system.

The central minicomputer is a Digital PDP 11/53 minicomputer connected to an Analogue Devices AP 500 Array Processor. The software for the PDP 11/53 was developed under Digital's RT11 real time operating system. The combination of a modern PDP 11/53 under RT11 with the AP 500 Array Processor gives a state-of-the-art real time signal-processing system.

The sampling interval of the 5 analogue BMS-1 signals is 0.25 m. At the maximum speed of the measuring vehicle of 45 m/s (= 162 km/h) the sampling rate is 180 Hz. For BMS-2 the 8 channels are sampled at a constant rate of 250 Hz. While the on-line sampling is performed continuously for both systems, all 200 m block calculations have to be done by the system within 4.4 seconds.

From the BMS-1 signals cant, level, alignment, and gauge in the 0 - 25 m waveband and alignment in the 0 - 70 m waveband VRA calculates the vehicle reactions indicated in Figure 6.55. For 3 types of rolling stock and 5 speeds the horizontal and vertical car body acceleration, weighted according to ISO 2631, and the vertical and horizontal track load are computed.



Calculation of vehicle reactions
in real time for:
- 3 vehicles: locomotive
  passenger coach
  freight car
- 5 speeds: 40-160 m/h

Calculations via array processor:

- Fourier transform of 4 geometry signals
- Computation of vehicle reactions via 240 transfer functions
- Determination of largest variance per speed range: these decisive 12 values are printed
- Back transformation of selected 12 responses
- Vectorial summation of car body accelerations
- Recording resulting in 9 response signals
- Determination of exceedences

Figure 6.55: Vehicle reactions calculated by VRA

The calculation flow in VRA is shown in Figure 6.56. The track geometry data of a 600 m block, represented by a 1024 point complex array, is transformed to the frequency domain using the FFT (= Fast Fourier Transformation) capabilities of the Array Processor. All communication between the computer and the Array Processor is performed using a high-speed DMA (= Direct Memory Access) channel.

The 600 m record is composed of the current 200 m section preceded by a section of 350 m and followed by a section of 50 m. The section of 350 m should cover the attenuation of the vehicle, i.e. the length of the unit impulse response. As far as BMS-1 is concerned, the longest wavelength in the analysis is 70 m and consequently the preceding section should be a number of times this wavelength. The section of 50 m takes the part of the vehicle ahead of the measuring point into account.

At the systems power-up, the transfer functions of three representative railway vehicles at five different speeds are loaded into the Array Processor. The frequency domain responses are calculated by multiplying the transformed signals with the corresponding transfer functions.

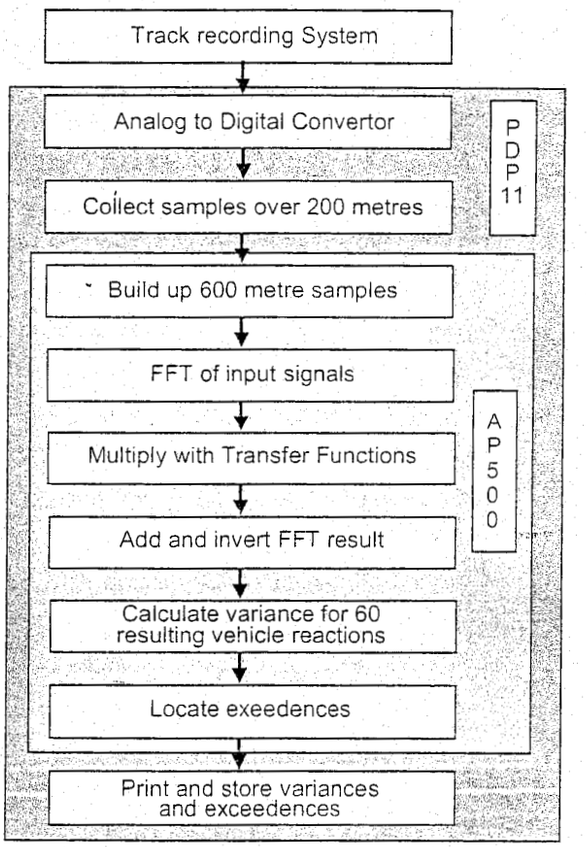


Figure 6.56: Calculation flow VRA

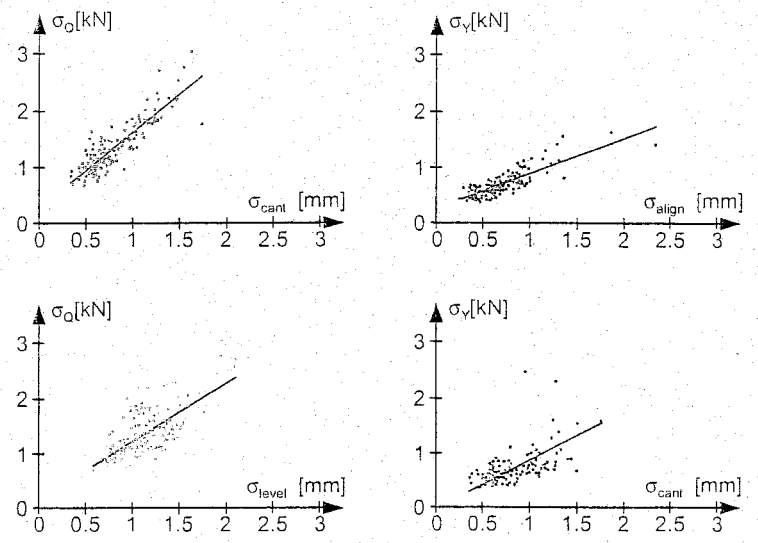


Figure 6.57: VRA-calculated forces versus track geometry per 200 m

These results are then transformed back into the time domain, i.e. per vehicle and response component 5 signals for 5 different speeds. Next the variances for the current 200 m section, and also the force and acceleration levels which are exceeded are calculated. Per vehicle and response component the maximum is taken for the 5 speeds. Both variances are added for the car body accelerations.
By taking the square root, standard deviations are obtained and these are printed, if desired, after having been normalized.

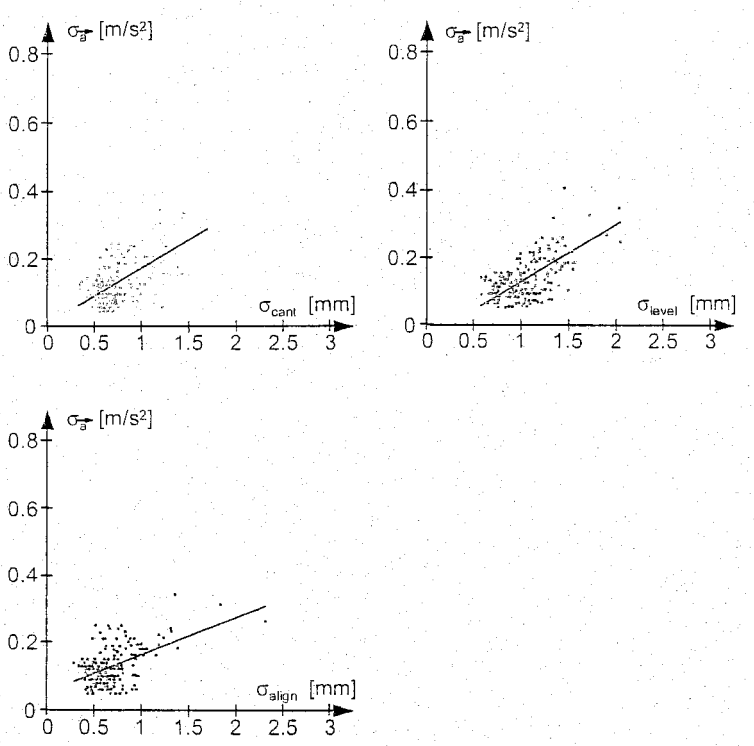


Figure 6.58: VRA-calculated accelerations versus track geometry per 200 m

At the time VRA was completed not all the required transfer functions were available. Although the final aim is to provide the Array Processor with transfer functions estimated by means of MISO, it was decided that the transfer functions calculated for the model of section 6.5. would first be implemented. Regarding this implementation Figure 6.57 shows a comparison between track geometry and vehicle reactions produced by VRA.

Per 200 m the standard deviations of Q-force and vertical car body acceleration are plotted against the standard deviation of level. Figure 6.58 presents the same information in the lateral direction.

Obviously, large variations in vehicle reactions are found for the same geometry values. It will be evident that the dynamic responses of railway vehicles are by far a better criterion for the actual maintenance status of a railway track than the pure track geometry. From this point of view it is expected that VRA will contribute to significantly better control of the track maintenance process.

6.8 Relation between Sperling's Ride Index Wz and ISO-weighted accelerations

The Wz Ride Index introduced by Sperling is the classical way of evaluating vehicle ride quality and passenger comfort [100]. Using modern analysing tools, the Wz factors are evaluated by multiplying the power spectral density function of the acceleration $G_a(f)$ with the squared modulus of the weighting function H(f) and integrating this result over the frequency range of interest to obtain the variance in that band according to:

$$\sigma^2 = \int_0^F G_a(f)|H(f)|^2 df \tag{6.207}$$

Finally, the Wz factor follows from:

$$Wz = [\sigma^2]^{1/6.67} \tag{6.208}$$

In this expression σ has the units cm/s^2. For passenger comfort in the vertical direction:

$$H(f) = H_{CV} = 0.588\left[\frac{1.911\,f^2 + (0.25\,f^2)^2}{(1 - 0.277\,f^2) + (1.563\,f - 0.0368\,f^3)^2}\right] \tag{6.209}$$

In lateral direction:

$$H(f) = H_{CL} = 1.25\,H_{CV} \tag{6.210}$$

For vehicle ride quality, the weighting in vertical and lateral directions is the same and reads:

$$H(f) = H_{RV} = H_{RL} = 1.14\left[\frac{[(1 - 0.056\,f^2)^2 + (0.645\,f^2)^2](3.55\,f^2)}{[(1 - 0.252\,f^2)^2 + (1.547\,f - 0.00444\,f^3)^2](1 + 3.55\,f^2)}\right]^{1/2} \tag{6.211}$$

The transfer functions are represented as a function of frequency in Figure 6.59 and Figure 6.60. The evaluation scales for the Wz factors were constructed based on vibration tests on people and were supplemented by other test results.

The assessment can be summarized as follows:

Wz passenger comfort		Wz vehicle ride comfort	
1	Just noticeable	1	Very good
2	Clearly noticeable	2	Good
2.5	More pronounced but not Unpleasant	3	Satisfactory
3	Strong, irregular, but still tolerable	4	Acceptable for running
3.25	Very irregular	4.5	Not acceptable for running
3.5	Extremely irregular, unpleasant, Annoying; prolonged exposure Intolerable	5	Dangerous
4	Extremely unpleasant; prolonged Exposure harmful		

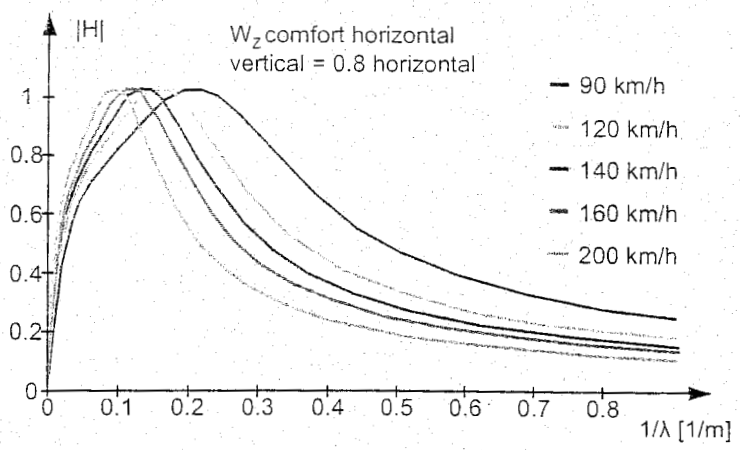


Figure 6.59: Transfer functions used to determine Sperling's ride index Wz for comfort

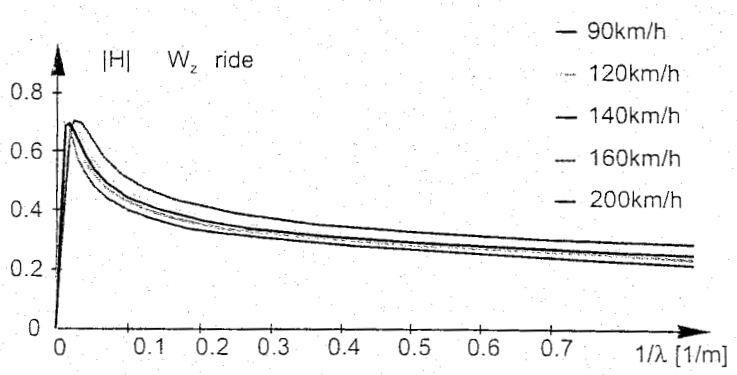


Figure 6.60: Transfer functions used to determine Sperling's ride index Wz for ride

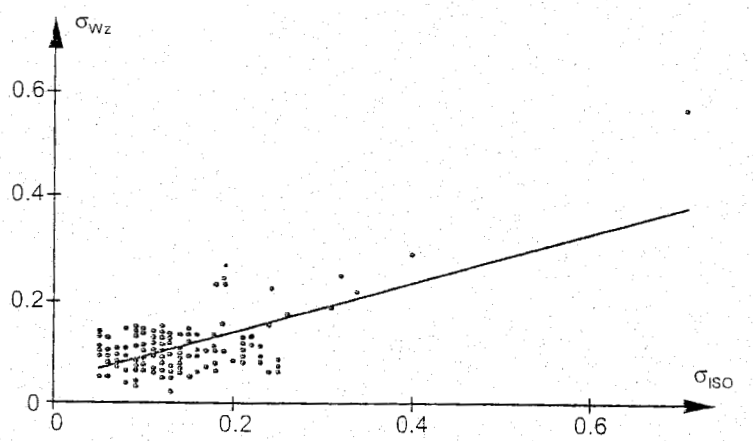


Figure 6.61: Relationship between Wz and ISO obtained from VRA simulations

To get a better understanding of how the ISO-weighted car body accelerations should be interpreted, the Sperling weighting functions were also stored temporarily in VRA. In this way it was possible to produce standard deviations per 200 m section for the ISO-weighted accelerations and the accelerations weighted according to the Sperling transfer functions for comfort. The latter have subsequently been treated according to (6.208).

Figure 6.61 shows the results obtained per 200 m section with the corresponding regression lines. It should be stated here that in the vertical direction the ISO and Wz weighting functions are quite similar. In the lateral direction Wz weighting deviates strongly from the ISO characteristic. In fact, the Wz weighting is not appropriate for evaluating passenger comfort in the lateral direction. For this purpose the ISO transfer function is now commonly used.

6.9 Applications of advanced dynamic models

6.9.1 Introduction

In the preceding chapters we discussed how to model and analyse dynamic track problems. Many methods were related to single degree of freedom (SDOF) systems, sometimes more complicated systems were discussed. These methods are valuable in that, with the limited support of some computer power, they are easy to handle in a design process. The disadvantage is that the structure and loadings are sometimes oversimplified and much experience is required to interpret, if still possible, the results in a meaningful way.

The application of much computer power makes it possible to apply 3D finite element method (f.e.m.) models. However, the use of 3D f.e.m. models is very time consuming, both of computer power and manpower. The time necessary to prepare, to execute, and to interpret one single problem is prohibitive in a design process. Unknown material properties and lack of homogeneity limit the meaningfulness of a 3D analysis.

On the other hand, the support of computer power makes it possible to apply more sophisticated models than were used before. Here an example is given of such a model, called RAIL, in addition to the models discussed in section 6.3. The position of RAIL in the field of track models is indicated in Figure 6.62.

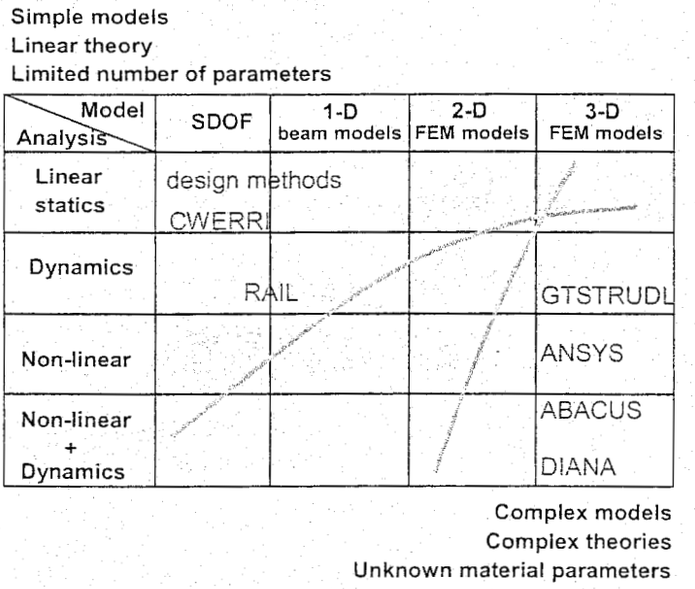


Figure 6.62: Overview of available software and railway track models for static and dynamic calculations

6.9.2 The RAIL-model

To assess the vertical dynamic behaviour of railway track due to moving railway vehicles, an integrated model called RAIL was developed at Delft University of Technology. The model was built up of two structures, namely the moving train and the railway track.

The coaches were schematised by rigid bodies, pivoted at the bogies by springs and dampers. Bogies and wheelsets were also modelled as rigid masses connected with springs and dampers. The vehicle model and the track model are schematically depicted in Figure 6.63. Only vertical forces and displacements etc. are considered.

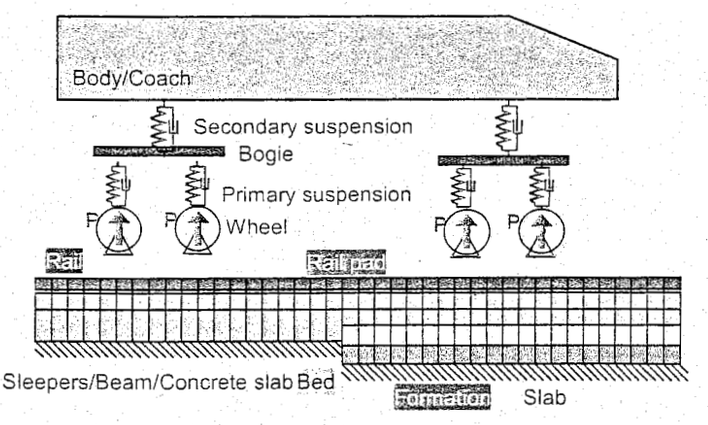


Figure 6.63: Model of moving vehicle on track structure

Rail and track irregularities are one of the most important sources of dynamic loads generated by a moving train. In the RAIL model the loads were introduced by means of Hertzian springs travelling along a sine-shaped or irregularly shaped rail and track surface as explained in Figure 6.64.

The equations of motion were formulated on the basis of the discussed structural models and were evaluated according to a direct integration procedure, based upon the concept of time-space elements. During a time step Δt the train moves over a distance Δx, which is dependent on the train velocity v, and 'crosses' the time-space elements as indicated in Figure 6.65.

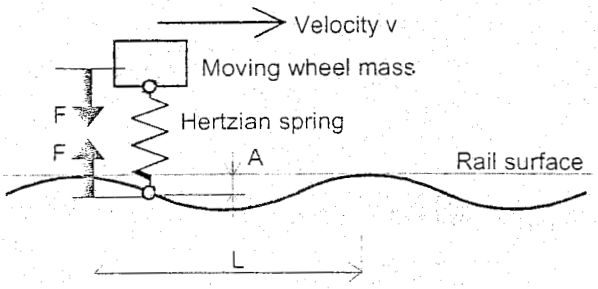


Figure 6.64: Modelling of wheel/rail loads

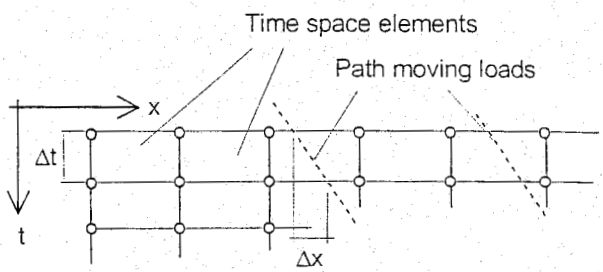


Figure 6.65: Calculation principle of moving trains

6.9.3 A comparison of several different track types

In order to study the effect of several different track types maximum elastic rail deflection, wheel rail contact force, and vertical body accelerations were investigated. In the underlying example it is shown how the properties of both vehicle and track are integrated into one complete model that takes the full interaction between the two structures into account.

The investigated structures consist of a Thalys train (special type TGV) which travels at different speeds on either a classic ballast track or an embedded rail structure (ERS). The ERS track is supported either by a rigid slab or a flexible slab which is itself discretely supported by a rigid pile foundation. The loads between track and vehicle are introduced by a sine-shaped surface deformation of the rail. Different wavelengths are investigated for their impact on the results. Three subjects are investigated: the elastic displacement of the rail head under the wheels, the contact forces between wheel and rail, and, finally, the vertical accelerations of the body, referring to the passenger coach or the locomotive.

The entire model is built up of two structures, namely, the moving train and the railway track. Each structure is modelled separately. During the dynamic analysis, that is the numerical integration procedure, the interaction of the two structures will be taken into account.

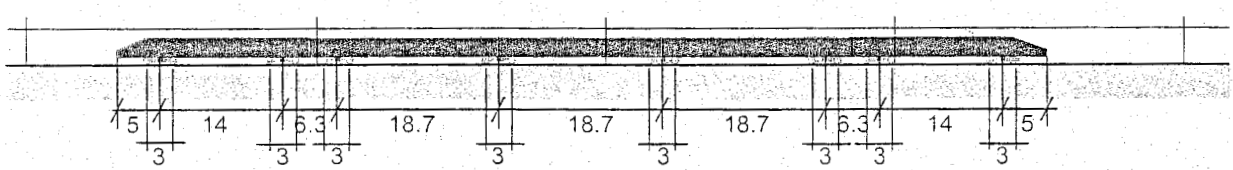


Figure 6.66: Dimensions of a (short) Thalys train as applied in the model

The train modelled here (short Thalys, see Figure 6.66) consists of five vehicles. It is typical for Thalys that bogies between two coaches are shared. The total length between the first wheel and the last wheel is 99.7 m. Rigid bodies, supported at the bogies by springs and dampers, schematise the coaches. Springs, dampers, and rigid bodies also model bogies and wheels. Only vertical forces and displacements are considered.

Regarding the Thalys we applied the following data:

- Mass (per item): coach 27140 kg, locomotive 54280 kg, bogie 2791 kg, wheel 1013.5 kg;
- Primary suspension spring (per wheel): K=1150 kN/m, C=2.5 kNs/m;

- Secondary suspension spring (per bogie): K=600 kN/m, C=4 kNs/m;
- Wheel radius: 0.42 m;
- To the coach next to the locomotive an additional lumped mass of 13570 kg. This equals all static wheel loads to 85 kN.

For our analyses we investigated the following track structures (Figure 6.67):

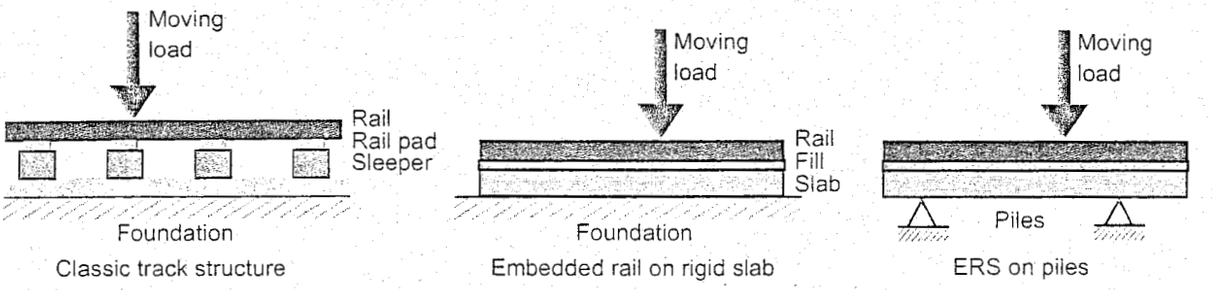


Figure 6.67: Analysed track structures

a A classic ballast track, using sleepers, rail pads, and an elastic ballast bed.

b An embedded rail structure (ERS) in which the rail, embedded in a fill material, has been put into a through in the rigid supporting concrete slab.

c A comparable embedded rail structure (ERS) with a more flexible concrete slab and supported by concrete piles (Figure 6.68).

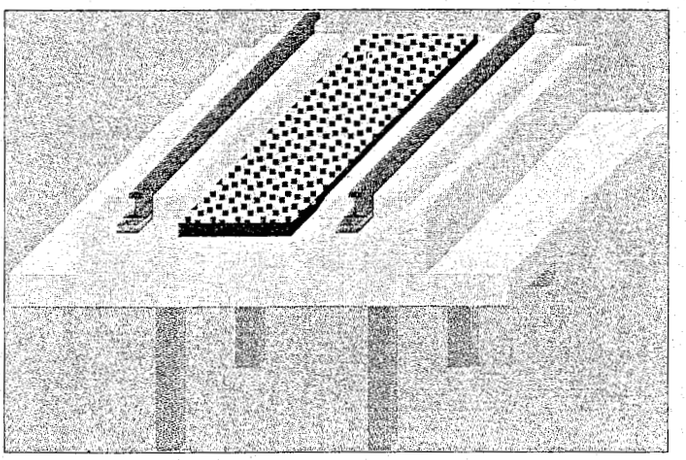


Figure 6.68: ERS with piles as foundation

For the analysis of the track we only took one rail into consideration assuming both rails and both sides of the train would show equal and synchronous behaviour.

The following data have been used for the classic ballast track:

- Rail: UIC-60;
- Rail pads (per rail): K=100 10^3 kN/m, C=15 kNs/m;
- Sleeper (per rail): mass 150 kg, spacing 0.60 m, width 0.15 m;
- Ballast bed (distributed per rail):
  k=180 10^3 kN/m^2, c=82 kNs/m^2 or per sleeper: K=27 10^3 kN/m, C=12.3 kNs/m.

For the ERS we applied:

- Rail: UIC-54;
- Fill material: k=52.5 10^3 kN/ m^2, c=4.98 kNs/ m^2;

- Slab rigidly supported (structure b);
- Slab supported on piles (structure c): A=0.765 m^2, I_{zz}=0.02746 m^4;
- Concrete properties: E=3.1 10^7 kN/m^2, ν=0.30, ρ=2500 kg/m^3;
- Pile distances: 3.0 and 6.0 m.

We considered a track length of 150 m for the analysis.

The surface geometry of the rail is one of the most important sources of load applications to a moving train. The extremes occur with short surface waves and high speeds. We considered three surface shapes $z_i(x)$ given by the algebraic formula:

$$z_i(x) = A_i\left(1 - \cos\frac{2\pi x}{L_i}\right) \tag{6.212}$$

where we substituted the following data:

- shape 1: L_i =3.0 m, A_i =1.5 mm;
- shape 2: L_i =12.5 m, A_i =3.0 mm;
- shape 3: L_i =50.0 m, A_i =6.0 mm.

The moving train has been analysed with regards to the speeds v=30 m/s, 60 m/s, and 90 m/s. The analysis has been performed with 1000 rail track elements of 0.15 m each and time steps Δt = 0.0015s. Dependent on the rail surface wavelengths, we applied 1600 time steps (2.40 s) for shape 1 and 4000 time steps (6.00 s) for shape 3.

The railway dynamics problem has been formulated by the equations of motion and the initial conditions. This means that, starting at t=0, a steady state solution will be obtained after damping of the initial disturbances. This requires some time.

This can be seen in a comparable calculation of a moving Thalys on an ERS at 3 different speeds. In Figure 6.69 the vertical body accelerations of the first Thalys vehicle are depicted for shape 1 rail surface geometry. The time-invariant period appears after almost 1 second of simulation. Moreover, the eigen period of the vehicle body is clearly present and lasts almost 1.5s.

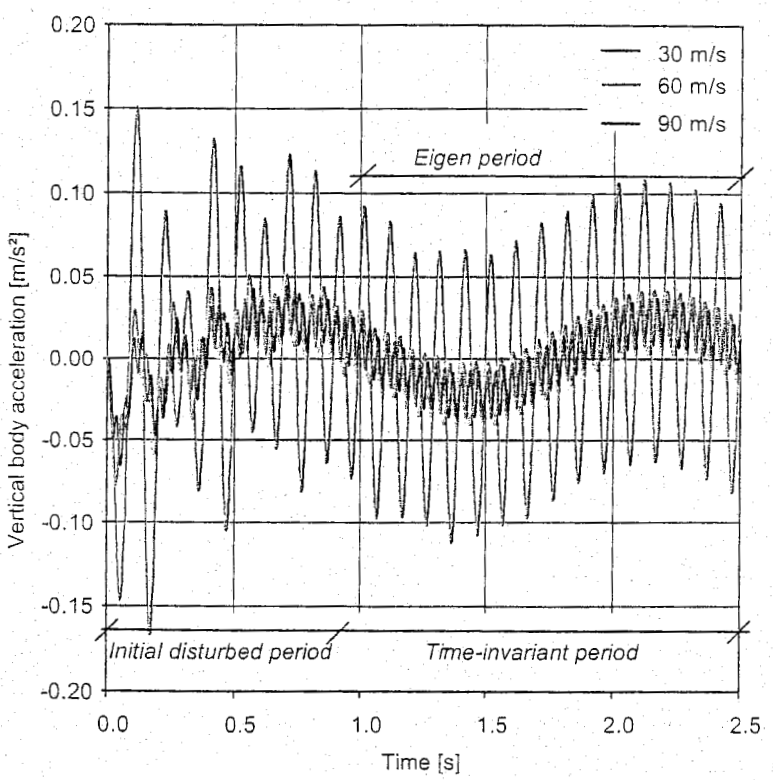


Figure 6.69: Vertical body accelerations of the first vehicle of Thalys at 3 different speeds

To study the velocity parameter and the dynamic response following our numerical process, the results for displacements and contact forces between t=1.8 s and t=2.4 s regarding the Thalys travelling at three different speeds on an ERS are shown in Figure 6.70 and Figure 6.71. As the speed increases, the variation in displacements and in contact forces also increases.

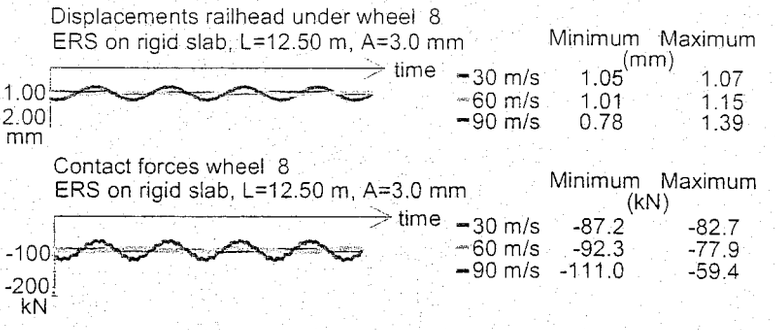


Figure 6.70: Displacements and contact forces between t=1.8s and t=2.4 s

When comparing different track structures, ERS on a continuously supported slab or a slab with small pile distance (3.00 m) performs much better than the classic ballast track or an ERS track on a slab with large pile distance. Continuously supported ERS shows a factor 2 between static and maximum dynamic contact forces, while the classic ballast track and an ERS on a widely spaced pile foundation (6.00 m) show a factor 3. For the elastic displacements of the rail head under the wheels, the size of the variations is even more distinct: here the classic ballast track shows the poorest performance as well.

The performed analyses confirm, of course, many observations that are made by intuition or by simple hand calculations. The analyses make it possible to compare the performance of a classic ballast track with respect to the ERS track, and the effect of pile spacing on an ERS track support. Certainly with respect to the contact forces and the rail head displacements, the ERS track performs better than the classic ballast track.

Besides full train simulations, a simpler approach is also available within RAIL. Moving loads or moving lumped mass models can be a suitable simplification of a train while running over predefined surfaces and structures of rail and track (see Figure 6.72). This gives e.g. vertical body accelerations in Figure 6.73, which are comparable to those already shown earlier in Figure 6.69. Please note that the track structure is different from that previously mentioned.

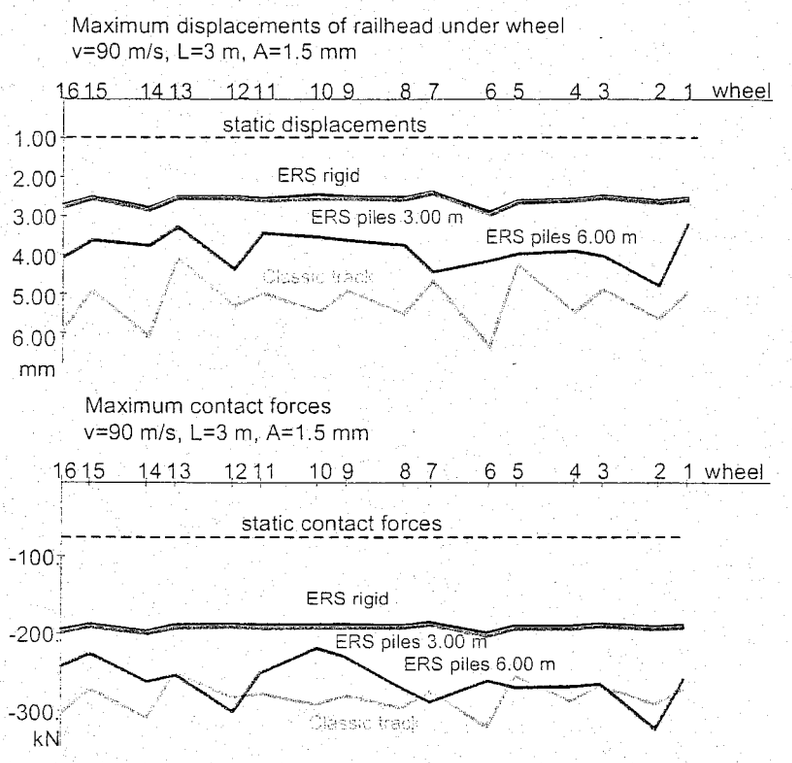


Figure 6.71: Maximum displacements and contact forces

It is obvious that not all of the possible applications of the integrated model of track and vehicle have been explored completely. It is also obvious that the integrated model offers many new possibilities for an integrated investigation of track and vehicle properties. For example, further applications of the RAIL program can be found in Chapter 13 and Chapter 14.

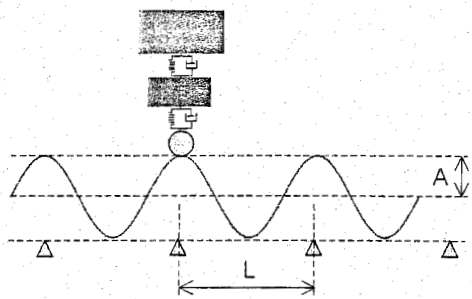


Figure 6.72: Simplified lumped mass vehicle model in RAIL

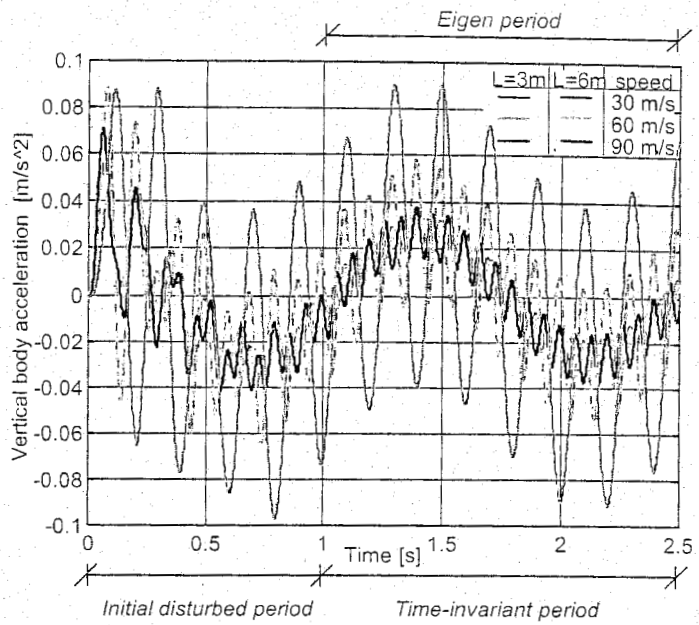


Figure 6.73: Vertical body (upper lumped mass) accelerations at different speeds on L = 3.0 m and L = 6.0 m and A = 3.0 mm rail surface geometry

6.9.4 Transitions in railway track on embankments and bridges

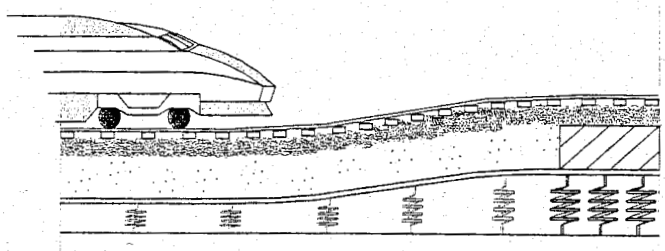


Figure 6.74: Principle of train response at a transition

Experience has shown that the transition between bridge and plain track (see Figure 6.74) often causes problems. Immediately after the track is laid, in particular, the plain track is subject to relatively high subsidence, at a rate different to that of the bridge. These transitions have been shown to cause problems, both in theory ([254], [127]) and in practice.

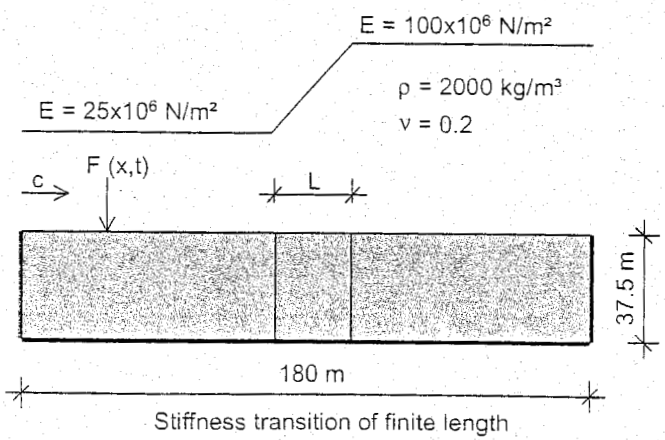


Figure 6.75: Stiffness transition of finite length

For the theoretical solution, finite element routines are available since a certain time, enabling to model and calculate the behaviour of trains at a transition.

The change in stiffness causes increased dynamic forces, the extent of which is determined by speed, stiffness ratio, damping and the length of the transition. In [254] a study is described into the effects of changes in vertical stiffness on the dynamic response.

Figure 6.75 defines the stiffness transition between a soft soil (clay) and a stiff soil (sand), with a linear stiffness transition of length L. Figure 6.76 shows the dynamic amplification factor for two situations. If the ratio between train speed and Rayleigh speed is low, an asymptotic D.A.F. value of 1.21 is reached after a relatively short length of about 8 m. If the train speed is close to the critical speed, a much larger asymptotic value of 2.51 is found, which is reached at a transition length beyond 25 m. So it obvious to stay far from the critical speed. The change in track geometry resulting from subsidence thus causes increased dynamic forces.

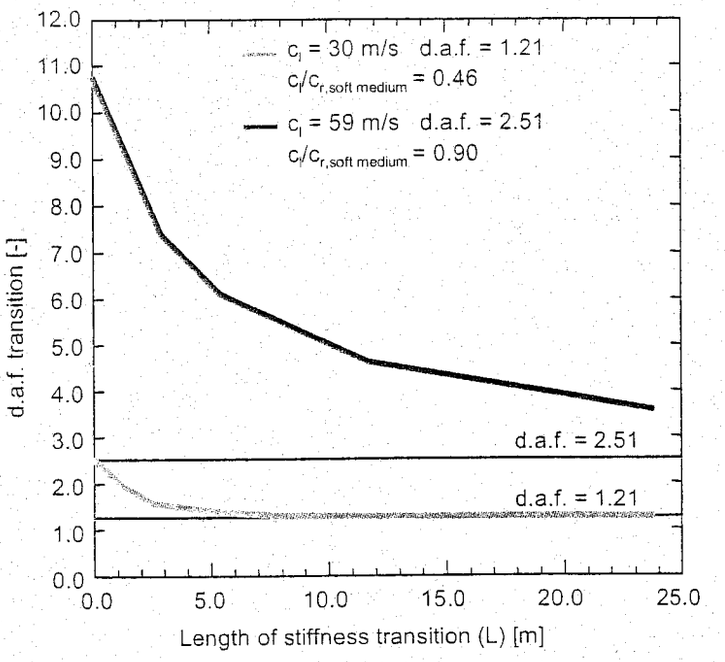


Figure 6.76: Dynamic amplification within the stiffness transition

Like large changes in subgrade stiffness, changes in the vertical alignment of the track lead to increasingly pronounced vertical accelerations in the vehicle, which may mean that the criteria for passenger comfort or maximum dynamic track force are no longer met. A marked increase in forces leads to accelerated deterioration of the track geometry and hence to additional maintenance.

The dynamic effects at the transition between bridge and plain track is a typical part of research into the problems of high speed rail systems. TU Delft has carried out a study [243], which have revealed that the effects of a change in height are generally more significant than those of a change in stiffness. Figure 6.77 shows an example of the car body accelerations calculated for a discrete event (change in height) of 30 mm, over a length of 30 m. Parametric studies have been carried out on the basis of the model described in [127], enabling accelerations and vertical wheel loads to be derived for various lengths of event and changes in height, at 300 km/h. Figure 6.78 shows the differences in height that are acceptable for HSL rolling stock for a given length of event, on the basis of a maximum permissible dynamic wheel force Q_{dyn} of 170 kN and a vertical car body acceleration a_v of 1.0 m/s^2.

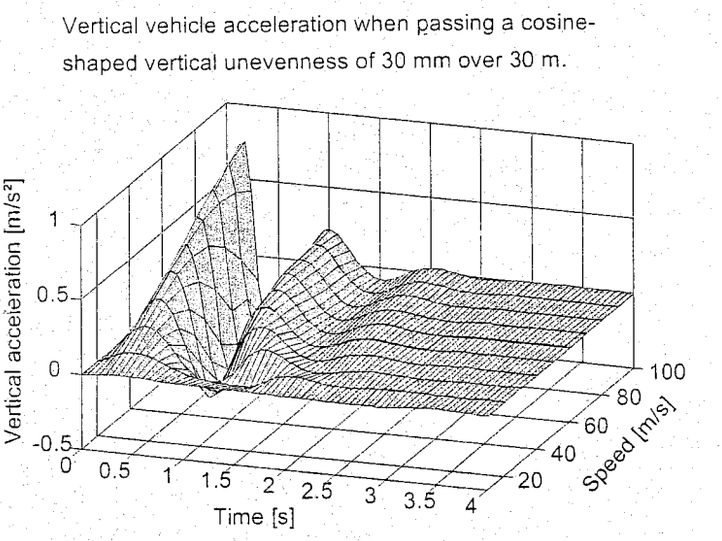


Figure 6.77: Calculated car body acceleration at transition

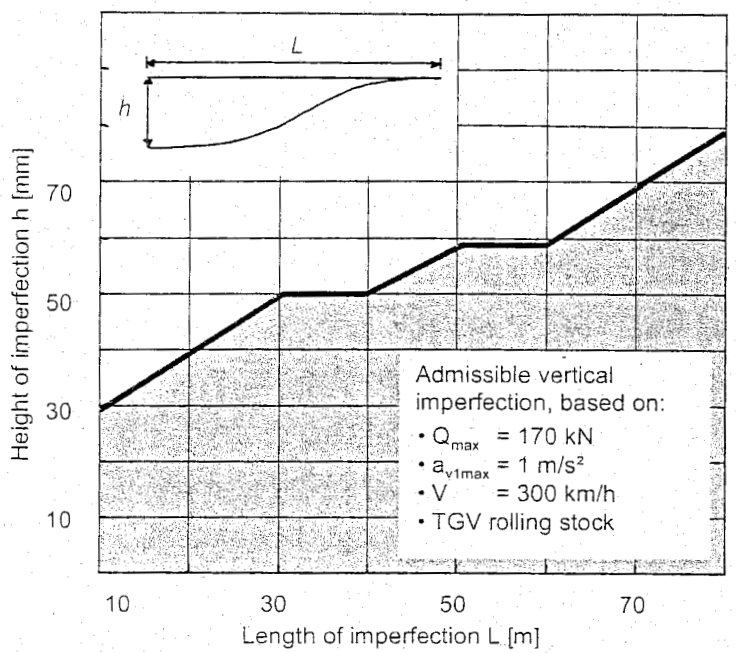


Figure 6.78: Design graph for a transition structure

6.10 Track response to a moving axle load

In railway transport, the continuous demand for capacity increase as well as for more extensive mobility networks automatically requires critical assessment of the effectiveness of currently operating and newly built railway lines. In order to meet the high standards regarding effectiveness, in many countries train cruising speeds of 200 km/h and more have been employed. As a result of such high train velocities, the track response in these high-speed lines has a typically dynamic character.

In order to reliably predict track safety and track deterioration, it is necessary to develop and utilise appropriate track models that cover the relevant vibration effects and wave propagation phenomena. In this section, the influence of the train velocity on the dynamic response of a railway track is considered. Accordingly, some models are discussed that consider the elasto-dynamic wave propagation under a moving load. These models provide further insight into the dynamic behaviour of a railway track, revealing typical effects that can only be captured by means of advanced analyses.

6.10.1 Track response at the critical train velocity

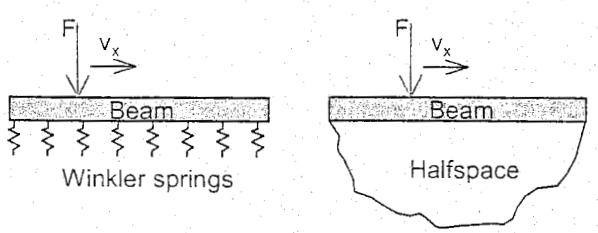


Figure 6.79: Beam-spring configuration and beam-halfspace configuration subjected to load F that moves with velocity v_x.

The analysis of the instantaneous response of a homogeneous railway track to a moving train axle commonly occurs by employing moving load models. In these models the train axle is modelled by a force of a constant (or a harmonically varying) amplitude F that moves with a velocity v_x over a configuration representing the railway track. The compound system of rails, sleepers, and ballast, known as the superstructure, is thereby often modelled by a beam, while the supporting subgrade, known as the substructure, is modelled by either discrete Winkler springs ([263], [52], [98]) or by a continuous halfspace ([95], [154], [50], [254], [163], [148]), see Figure 6.79.

In the current section it is illustrated what kind of results can be obtained from a moving load model and how these results relate to railway practice. First, the model presented in [255] will be discussed which considers a Timoshenko beam supported by a two-dimensional halfspace and subjected to a moving load with a constant force magnitude. Here, the Timoshenko beam simulates the elastic bending and shearing behaviour of the superstructure, while the halfspace simulates the elastic response of the subgrade. It is reasonable to assume that the structure behaves elastically, since the response of a railway track during the passage of an individual train axle is almost completely reversible (see for instance [247]).

In order to analyse the wave propagation in the beam-halfspace system, in [254] the corresponding boundary value problem has been elaborated on. This elaboration was based on combining the equations of motion of the system with the boundary conditions. By searching for harmonic solutions, the characteristic waves in the system have been computed where two body waves (compression wave, shear wave) can be distinguished and a wave that propagates along the surface of the track system (surface wave).

The surface wave commonly conveys the largest part of the energy generated by the train. When the velocity of the surface wave is of a similar magnitude as the velocity of the train, the generated energy remains close to the train. This leads to accumulation of energy under the train as time progresses. This phenomenon can be identified as resonance. Because this resonance may result in large track amplifications, the train's velocity is often designated as 'critical' when resonance occurs.

Although several critical velocities may be identified for a specific track configuration (see for instance, [50], [254]), the lowest critical velocity is the most important since this is the first one to be met by an accelerating train vehicle. As confirmed by 'in-situ' measurements, rail deflections can increase to more than three times the static deflection when a train reaches the lowest critical velocity ([126], [164]). It is easily understood that for reasons of safety and limitation of train/track deterioration such track amplifications can not be tolerated.

The main features of a train that reaches the critical regime will be illustrated by assuming the halfspace in Figure 6.79 to be softer than the Timoshenko beam. Accordingly, the railway superstructure is considered to be supported by a relatively soft formation of clay or peat. In the model [254], the load starts to move from zero velocity and accelerates up to a velocity larger than the lowest critical velocity.

The analysis of the response has been performed by means of a finite element model with dimensions b x h = 180 m x 37.5 m, see Figure 6.80. The moving character of the load is simulated by means of a set of discrete pulses which act successively on the element nodes at the surface along which the load is supposed to propagate. Furthermore, the half-infinite character of the halfspace is simulated by means of viscous damping elements that are connected to the artificial model boundaries (energy-absorbing boundary). By providing the damping elements with dynamic impedances that are similar to that of the adjacent continuous medium, the energy of the waves which encounters the artificial model boundaries is (almost completely) absorbed. For more details on the finite element model, see [254].

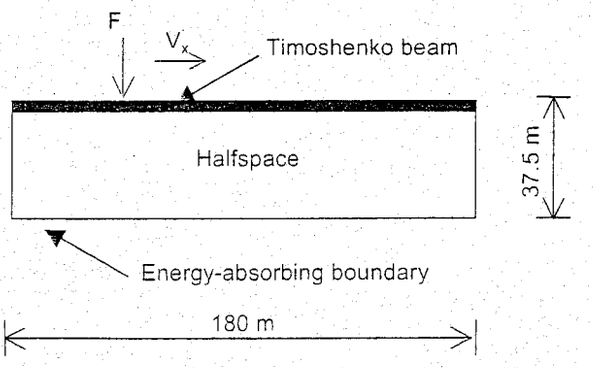


Figure 6.80: Moving load F subjected to a Timoshenko beam-halfspace configuration. Waves arriving at artificial model boundaries are absorbed.

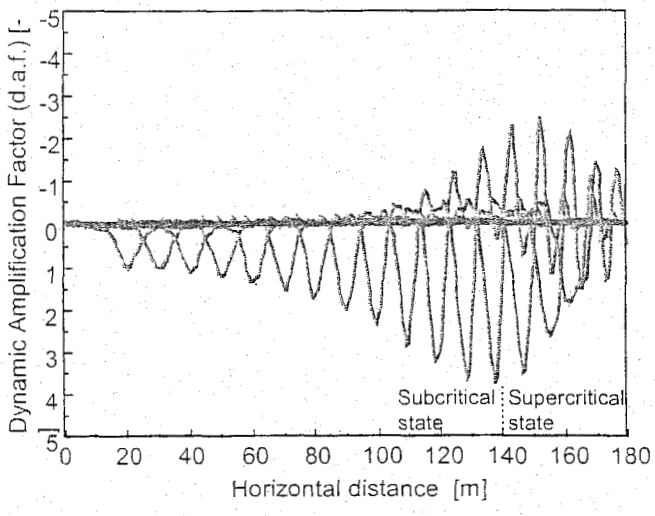


Figure 6.81: Stroboscopic development of the dynamic amplification factor (d.a.f.) under an accelerating load. Timoshenko beam is supported by a relatively soft half-space.

In Figure 6.81, the stroboscopic development of the dynamic amplification of the normal stress in vertical direction, σ_{zz}, is depicted. The dynamic amplification relates to a point at 5.68 m below the surface, and has been computed by dividing the dynamic stress by its static counterpart. The total horizontal distance of 180 m plotted on the horizontal axis reflects the total width of the finite element model. The load starts to move at a horizontal distance of 20 m from the left-side model boundary at which the dynamic amplification equals 1.0, thus corresponding to a static response.

Due to the fact that the load velocity increases with increasing horizontal distance, the dynamic amplification grows in a monotonic manner. At a horizontal distance of 140 m from the left-side model boundary, a maximum of about 4 times the static response is reached. At this stage, the load velocity has approached the shear wave velocity c^s of the halfspace and the system behaves critically. After this critical stage has passed the response becomes super critical.

Obviously, in the supercritical range the system response decreases under increasing load velocity. This is because the load propagates faster than the energy transmitted by the radiated surface waves. Correspondingly, the energy can no longer accumulate under the load.

When the load velocity is strongly super critical, the amplitude of the response appears to be of the same order of magnitude as that of the static response. The characteristic behaviour sketched in Figure 6.81 is, actually, similar to that of an aeroplane passing the sound barrier. This results from the fact that the phenomenon of an aeroplane catching up with a sound wave is completely analogue to that of a train catching up with surface waves.

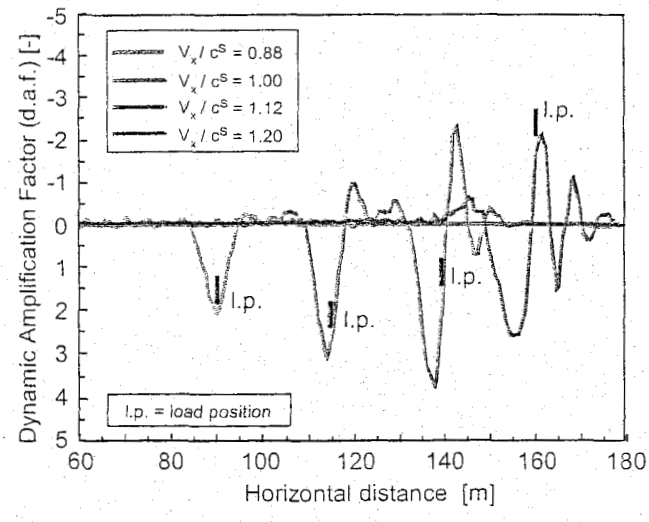


Figure 6.82: The dynamic amplification factor (d.a.f.) under an accelerating load in the velocity range 0.88 $c^S < v_x < 1.20$ c^S. Timoshenko beam is supported by a relatively soft half-space.

In order to study the response at the critical stage in more detail, in Figure 6.82 the dynamic amplification has been plotted for a load velocity (v_x) in between 0.88 and 1.20 times the shear wave velocity (c^s) of the halfspace. Obviously, at v_x=0.88c^s the response is still approximately symmetric. For larger load velocities, however, the response becomes increasingly asymmetric. This is the result of the appearance of oscillating Mach waves upon reaching the critical regime. These Mach waves have also been monitored during 'in-situ' track deflection measurements [164], thereby revealing a maximum upward deflection of 9 mm in front of the front train axle and a maximum amplitude of 12 mm behind the front train axle.

For obvious reasons, the generation of Mach waves may have a detrimental effect on both the track and the train. It may, in addition, cause the train to derail. This can be motivated from the supercritical response at $v_x = 1.20c^s$ which shows that the response directly below the load position (l.p.) acts in a direction opposite to the loading direction. In other words, the train axle is lifted up by the surface waves.

The analysis above has illustrated that in the case of a railway track built on a relatively soft subgrade, the critical regime is reached when the train velocity is near the shear wave velocity of the subgrade. The magnitude of the shear wave velocity is determined by material properties, as computed from

$$c^s = \sqrt{\mu/\rho} \tag{6.213}$$

where μ is the shear modulus and ρ is the density of the material.

For a subgrade of soft clay or peat, the shear wave velocity commonly lies in between 150 and 250 km/h. 'In-situ' track deflection measurements in Great-Britain [126] and Sweden [164]) confirmed that the track response may become critical in this velocity range. For a railway track resting on a subgrade of relatively soft clay, the measurements demonstrated that the track response is amplified to more than three times the static response when the train reaches a velocity of about 200 km/h.

Track amplifications are decreased when the soft subgrade is replaced by a stiffer structure such as a sand embankment of considerable thickness. Accordingly, the shear wave velocity of the supporting subgrade can be increased to a value that lies outside the velocity domain of the high-speed train vehicle. An alternative solution is to disconnect the railway superstructure from the soft subgrade by means of a pile foundation. The soft subgrade then will not affect the track response anymore, since the piles convey the generated waves into the stiff soil layer that supports the piles. Another option is to leave the soft subgrade as it is and limit the train velocity to a level at which the track amplifications are acceptable. However, this option is certainly not preferred as it may considerably increase the transportation time on the specific railway line.

As already pointed out in Figure 6.79, moving load models may be divided into two categories: beam-spring models and beam-halfspace models. Because the analysis of a beam-halfspace model is considerably more complicated than the analysis of a beam-spring model, there may be a strong preference to use beam-spring models when examining wave propagation phenomena in a railway track. Nonetheless, it should be realised that a discrete spring support is not able to transmit waves. As a consequence, incorporation of the commonly accepted spring properties for modelling the track subgrade results in an inadequate description of the track dynamics. In fact, the dynamic amplification predicted by a beam-spring model then becomes significant only for velocities far beyond the range of operational train speeds. This may lead to the misconception that the dynamic track amplification caused by the train is generally negligible.

6.10.2 Dynamic response of a ballast layer

For a railway track supported by a soft subgrade the wavelength of the surface waves is usually relatively long, typically ranging from 5 to 20 meters for a peat or clay formation. These long wavelengths are caused by the soft nature of the subgrade. When the superstructure is supported by a stiff substratum, such as a rock formation, a concrete bridge, or a concrete tunnel, the waves which propagate at the surface of the track are not necessarily restricted to the domain of long wavelengths.

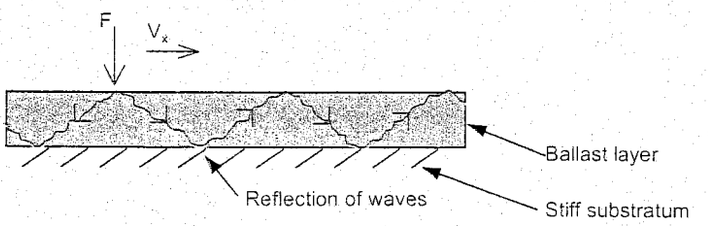


Figure 6.83: Ballast layer supported by a stiff substratum. Multiple reflection of waves.

This is because a stiff substratum reflects body waves of all wavelengths back into the superstructure, see Figure 6.83. Consequently, a ballast layer on a concrete bridge or tunnel acts as a waveguide, conveying waves of both long and short wavelengths.

The shorter waves in the ballast layer can have a wavelength of the order of magnitude of the ballast particle size which may perturb the individual particles. In order to model relative motions by the ballast particles, it is necessary to incorporate the particle size into a mechanical track model. This can be done either by employing continuum models that are derived from the micro-mechanical particle behaviour ([256], [257]), or by using discrete particle models ([258], [259]). In these models the particles are assumed to have an ideal spherical shape. The interaction with neighbouring particles is prescribed at particle contact points by means of a contact law. Moving load analyses carried out with these models have demonstrated that a ballast material consisting of large particles increases the intensity of the wave radiation, especially when the damping capacity of the ballast is low. Hence, to suppress such effects, the damping capacity of a ballast material should be sufficiently high.

The damping capacity of ballast generally depends on two effects: the inter-particle friction and the distribution of particle sizes. As far as the first effect is concerned, when the friction between the particles is high (i.e. a coarse-grained ballast) a relatively large amount of energy is dissipated at the particle contacts, thus causing the ballast to have a high damping capacity. The second effect stems from the fact that a wide particle size distribution excludes a dominant appearance by large particles that may act as a resonator. Accordingly, a ballast material with a random distribution of various particle sizes has a better damping capacity than a ballast material which consists of identical particles of a relatively large size.

In general, dynamic amplifications are not only caused by the train velocity, but also by impact loads such as those generated by the sleeper distance effect or by other track irregularities. In the case of a ballast layer supported by a stiff substratum, the energy transmitted by the track vibrations remains for a large part inside the ballast layer as a the result of (multiple) wave reflections at the stiff substratum, see Figure 6.83.
Consequently, reflected waves may interfere with other (reflected or non-reflected) waves. This causes the amplitude of the response either to increase or to decrease depending on the motion characteristics of the interfering waves. When superimposing the dynamic amplification by load impacts on that by the train speed, it may appear that the track response becomes 'critical' at a considerably lower velocity than the critical velocity which relates to the effect of train speed only ([98]; [257], [259]).

The extent to which this occurs depends on the characteristics of the impact loading (e.g. frequencies and duration of the impact loading) as well as on the geometry and material properties of the track structure (e.g. thickness and stiffness of the ballast layer). An adequate way of reducing the dynamic amplification by load impacts is to apply ballast mats between the ballast layer and the stiff substratum. In fact, the damping characteristics of the ballast mats will reduce the wave reflection depicted in Figure 6.83.

6.10.3 Stiffness transitions

Stiffness transitions emerge when a relatively soft substratum of clay or peat changes into a relatively stiff substratum of sand or concrete (or vice versa). Such transitions appear nearby concrete railway bridges or railway tunnels, but also when the characteristics of the natural soil formation (i.e. the substructure) change abruptly. Stiffness transitions form the basis for the emergence of differential settlements, which, once initiated, may grow considerably as time progresses.

The growth rate of differential settlements is governed by the dynamic properties of the train, i.e. the mass and velocity of the train in combination with its spring and damping characteristics. The mechanism of growing differential settlements can be explained as follows. A stiffness difference increases the dynamic loading on the track. As a consequence of the increased dynamic loading, differential settlements may emerge. Differential settlements activate additional train vibrations which cause the dynamic loading to become larger the next time the train passes the stiffness transition. Accordingly,

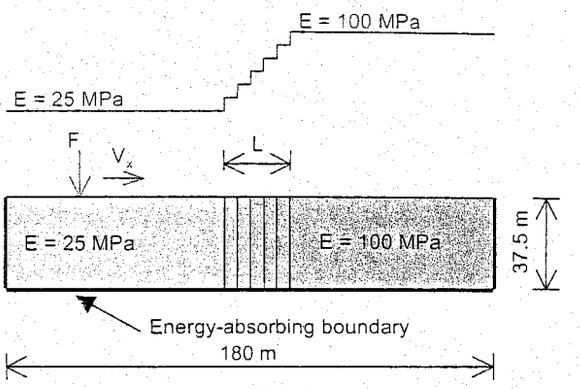


Figure 6.84: Stiffness transition L in an elastic halfspace comprising a relatively soft medium (E = 25 MPa) and a relatively stiff medium (E = 100 MPa).

the differential settlements grow further and the above mechanism repeats itself.

Since differential settlements may cause excessive track deterioration, they should be controlled as much as possible. Therefore, it is of practical importance to understand how the dynamic response of a stiffness transition, which forms the basis for the emergence of settlement differences, is influenced by a passing train axle. This is demonstrated by means of a finite element model, see Figure 6.84, that has been discussed previously in [253].

The model consists of an elastic halfspace comprising a relatively soft material of E=25 MPa (e.g. clay) and a relatively stiff material of E=100 MPa (e.g. sand), in which E represents the Young's modulus of the materials.

The stiffness transition L between the soft and the stiff material occurs in a stepwise manner, with a step-size equal to the transition length divided by 0.75 m (= element size). The modelling of the moving load and the absorption of the wave energy at the artificial model boundaries occurs in the same manner as explained for the model in Figure 6.80.

For an ideal homogeneous halfspace, the critical velocity is equal to the Rayleigh wave velocity c^r which has a magnitude close to that of the shear wave velocity [1]).

The influence of the load velocity v_x on the dynamic stress amplification at the transition L has been examined by considering a relatively low load velocity and a relatively high load velocity. Here, the low load velocity v_x=30.0 m/s is about half the critical velocity of the soft medium, while $v_x/c^{r,\ \text{soft medium}} = 0.46$. The high load velocity v_x=59.0 m/s is very near the critical velocity of the soft medium, $v_x/c^{r,\ \text{soft medium}} = 0.90$.

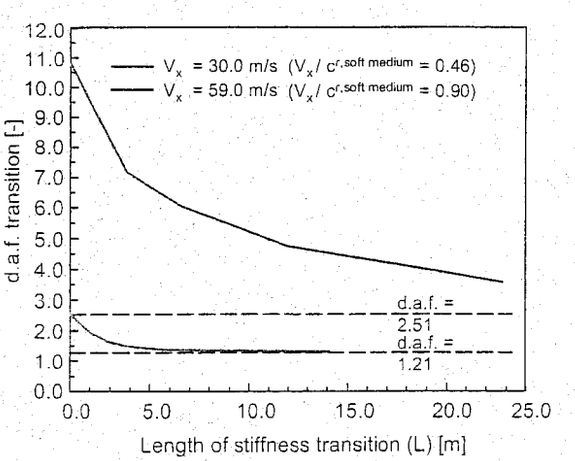


Figure 6.85: Dynamic amplification factor (d.a.f.) versus the length of the stiffness transition. Moving load with a relatively low velocity ($v_x/c^{r,\ \text{soft medium}} = 0.46$) and a relatively high velocity ($v_x/c^{r,\ \text{soft medium}} = 0.90$).

The effect of the transition length L on the dynamic stress amplification has been depicted in Figure 6.85. Clearly, for both load velocities the dynamic amplification factor (d.a.f.) at the stiffness transition decreases with an increasing transition length. As for the previous analysis, the dynamic amplification factor has been computed by dividing the dynamic stress σ_{zz} at 5.68 m below the surface by its static counterpart.

For the high load velocity v_x=59.0 m/s, the dynamic amplification appears to be significantly larger than for the low load velocity v_x=30.0 m/s The depicted horizontal lines d.a.f. = 1.21 and d.a.f. = 2.51 represent the dynamic amplification for v_x=30.0 m/s and v_x=59.0 m/s in case the halfspace would be homogeneous, with E=25 MPa. Hence, these lines act as asymptotes to which the dynamic amplification factor approaches when the transition length increases. Obviously, for the low load velocity v_x=30.0 m/s the asymptote is reached much faster than for the high load velocity v_x=59.0 m/s.

From the current analysis two important conclusions can be drawn. Firstly, in order to adequately reduce the dynamic amplification at a stiffness transition, the length of the transition should be sufficiently large. At a certain stage, a further increase of the transition length will no further reduce the dynamic amplification. Secondly, railway lines constructed for fast(er) trains require a large(r) transition length to reduce the dynamic track amplification. Apart from these two effects there is another effect that influences the dynamic amplification: the magnitude of the stiffness difference. Trivially, an increasing stiffness difference increases the dynamic amplification at the stiffness transition so that a large stiffness difference should be bridged by large transition length L. The quantitative effect of the magnitude of the stiffness difference on the dynamic amplification is exemplified in [253] by means of finite element analyses and energy considerations.

6.10.4 Brief discussion

The models presented in this section have provided insight into some important features of the dynamic response of a railroad track to a moving axle load. It is possible to extend the models in several ways in order to provide a more realistic representation of an 'in-situ' railway system. Here, one may think of replacing the moving load by a mass-dashpot-spring system that accounts for the train vehicle dynamics, or of adding the influence of ground water and frictional sliding to the ballast and subgrade behaviour.

7 TRACK STABILITY AND LONGITUDINAL FORCES

7.1 Introduction

In conventional non-welded tracks the rails are connected by means of joints to allow for length changes caused by temperature fluctuations. Using joints prevents the development of axial forces and the consequent risk of track buckling at high temperatures. However, the penalty for this is the care for maintenance-intensive joints which generate high dynamic loads during train passage. These loads are responsible for many problems like rapid deterioration of vertical track geometry, plastic deformation of the rail head, dangerous rail cracks as well as damage to sleepers and fastenings. These problems increase progressively as speed increases. As a rule, joints have a very considerable negative effect on the service life of all track components.

Tracks with continuous welded rails (CWR) do not possess the above drawbacks. Owing to the absence of joints the quality of the track geometry is better by an order and this results in a substantial decrease in the total life cycle cost. CWR does not, however, only have advantages. As was pointed out in Chapter 5, the stresses resulting from the plane strain situation may be of the order of 100 N/mm² and should be added to the residual rail stresses and bending stresses caused by train loads which are of the same order of magnitude. Temperature stresses especially are responsible for failure of welds with small imperfections at low temperatures. On the other hand, lateral stability should be sufficiently great to resist compression forces developing at temperatures above the neutral temperature of 25°C, as buckling may otherwise occur as, for example, illustrated in Figure 7.1. The principle of this phenomenon is sketched in Figure 7.2 showing the compressive forces and the resistance forces on the track and the resulting typical buckling shape.

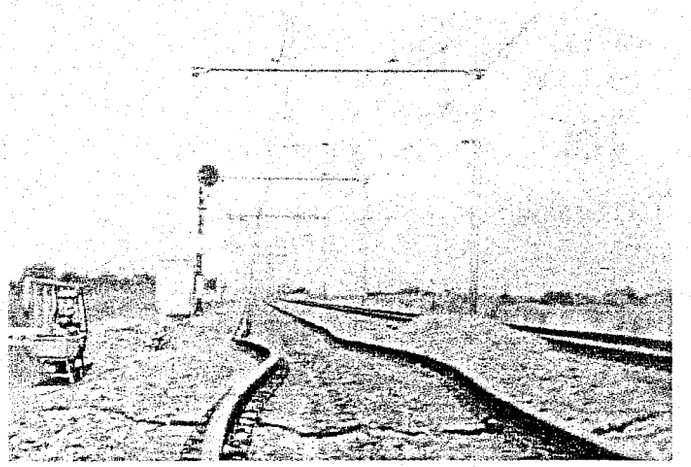


Figure 7.1: Example of track buckling

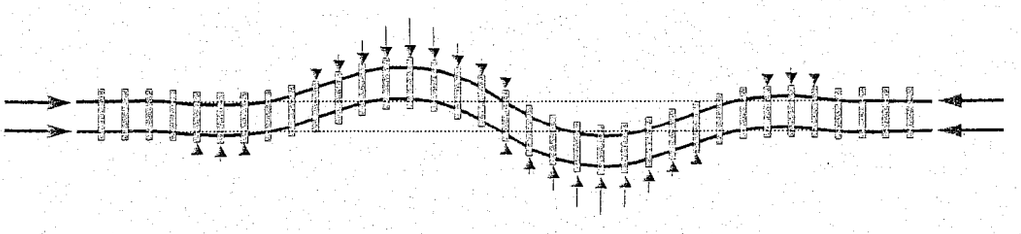


Figure 7.2: Typical buckling shape

On bridges and viaducts the deformation regime deviates from the plain track situation. The rails follow the construction which can undergo large displacements with respect to the adjacent track. Without adequate measures this would result in high rail stresses. To avoid these stresses expansion joints are applied.

This chapter is devoted to track stability and track longitudinal problems which, in the case of compression forces, are strongly interrelated. For both fields analytical and finite element modelling approaches are presented with examples. The last section discusses recently developed advanced models which describe safety considerations about track buckling or deal with more general or complicated track systems.

7.1.1 Straight track and elastic lateral resistance

In order to analyse this basic problem the track is modelled as a straight bending beam with a length L[1]. The beam is loaded by axial compressive forces and is elastically supported by a laterally distributed resistance which is proportional to the lateral deflection:

$$\tau = \beta v \qquad (7.1)$$

where ß is a constant and v(x) = lateral deflection.

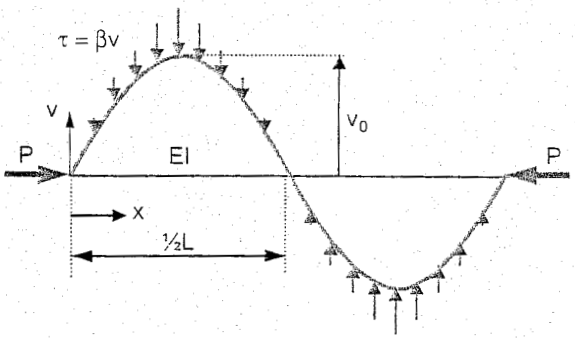


Figure 7.3: Buckling of straight beam with elastic lateral resistance

In view of the treatment of the rail buckling problem, a continuous sinusoidal shape will be assumed. Only one wavelength need to be considered in Figure 7.3. The energy method will be used to solve the problem because it not only gives the equilibrium equation, but also the nature (stable or unstable) of the equilibrium. The total energy consists of three parts:

1. Bending energy of the beam:

$$U_{beam} = \tfrac{1}{2}\int_0^L EIv''^2 dx \qquad (7.2)$$

where v"2 is d2v/dx2.

2. Deformation energy lateral pressure, with (1-1):

$$U_{press} = \tfrac{1}{2}\int_0^L \beta v^2 dx \qquad (7.3)$$

3. Work done by the compressive force:

$$W_P = \tfrac{1}{2}\int_0^L Pv'^2 dx \qquad (7.4)$$

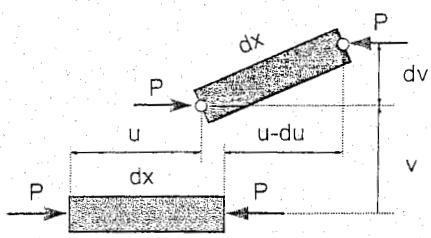


Figure 7.4: Rail element in buckling

While the first two energies are obvious, the last one may need some explanation. Referring to Figure 7.4 the elementary energy can be written as:

$$dW_P = Pdu = P(dx - \sqrt{dx^2 - dv^2}) = Pdx(1 - \sqrt{1 - v'^2}) \qquad (7.5)$$

where P is assumed to be constant and axial deformation is neglected. Linearisation gives:

$$dW_P = \tfrac{1}{2}Pv'^2 dx \qquad (7.6)$$

which explains equation (7.4).

The total energy is then:

$$U_{tot} = U_{beam} + U_{press} - W_P \qquad (7.7)$$

1. Not to be confused with the same symbol L used for the characteristic length of the track

Giving as solution:

$$v(x) = v_0 \sin\frac{2\pi x}{L} \tag{7.8}$$

the total energy per wavelength becomes, after inserting (7.8) in (7.2), (7.3), (7.4), and using symmetry considerations:

$$U_{tot} = \frac{v_0^2 L}{4}\left[EI\left(\frac{2\pi}{L}\right)^4 + \beta - P\left(\frac{2\pi}{L}\right)^2\right] \tag{7.9}$$

The equilibrium follows from setting the partial derivative $\frac{\partial U_{tot}}{\partial v_0} = 0$, resulting in:

$$P = \frac{4\pi^2 EI}{L^2} + \frac{\beta L^2}{4\pi^2} \tag{7.10}$$

As an example the relation between the buckling load P and the wavelength L is given in Figure 7.5 for EI = 8000 kN/m² and for some values of β in the range 10 to 200 kN/m².

The maximum value of the buckling load is restricted to 2500 kN in Figure 7.5 because higher values are unrealistic. For instance, in the case of UIC 60 rails and an extreme temperature increase of 65 °C, the maximum normal force which can occur in CWR track amounts to 2414 kN (=2*210e^{6}*76.9e^{-4}*1.15e^{-5}*65).

Stability is ensured if $\frac{\partial^2 U_{tot}}{\partial v_0^2} > 0$, hence if:

$$P < \frac{4\pi^2 EI}{L^2} + \frac{\beta L^2}{4\pi^2} \tag{7.11}$$

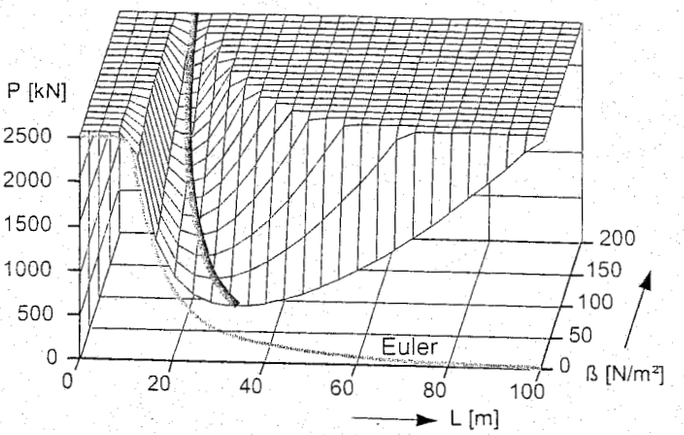


Figure 7.5: Buckling diagram in case of elastic resistance

This shows that each equilibrium point on the surface is indifferent with respect to stability.

From the condition $\frac{\partial P}{\partial L} = 0$ the following critical quantities are determined:

$$L_{crit}^2 = 4\pi^2\sqrt{\frac{EI}{\beta}} \tag{7.12}$$

$$P_{crit} = 2\sqrt{\beta EI} \tag{7.13}$$

This relation between L, β, and P has been indicated in Figure 7.5 with a thick line.

If there is no lateral resistance (β=0), then (7.10) will reduce to the well-known expression for the Euler buckling load for straight free beams, second mode:

$$P = \frac{4\pi^2 EI}{L^2} \tag{7.14}$$

This curve is also drawn in Figure 7.5 and is marked with small circles.

7.1.2 Track with misalignment and constant lateral shear resistance

If applied to rail track situations, the value of P_{crit} in (7.13) is only significant when the lateral ballast resistance is very low and approaches the Euler buckling load. However, this situation is not representative for railway track, for in the elastic model there is always sufficient resistance and according to the above-mentioned example unrealistic high values of the buckling load would be calculated. With respect to railway track two important factors should be taken into account.

First, the lateral resistance, which is caused by the shear resistance between track panel and ballast bed, has a limiting value. Secondly, real track is never perfectly straight, but shows some form of geometrical imperfection or misalignment. Therefore, in this section a more realistic model for rail track will be used.

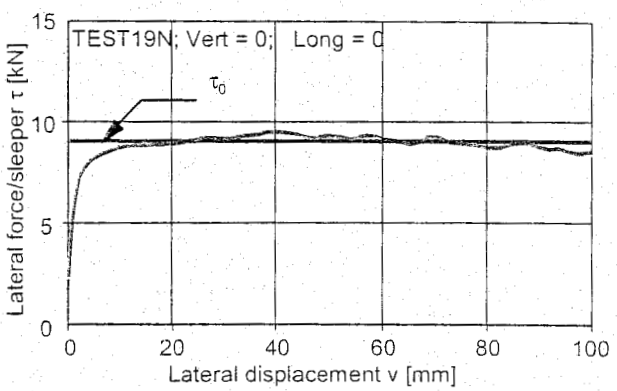


Figure 7.6: Lateral resistance measurement

In Figure 7.6 a recent measurement of the lateral shear resistance characteristic of a track panel in ballast bed is given, expressed in force per sleeper. The curve could, for instance, be approximated by a bi-linear function, but for our purpose here we will approximate the curve by a constant plastic shear resistance which opposes the axial displacement and therefore can be defined as:

$$\tau = \tau_0 sign(v)^{1} \qquad (7.15)$$

The bending stiffness EI is represented here by the horizontal bending stiffness of the track panel. This stiffness incorporates the bending stiffness of the rails, the sleepers, and the rotational resistance of the fastenings. It is assumed that this bending stiffness is constant. The straight track is supposed to possess an initial sinusoidal misalignment in the unloaded situation (Figure 7.7), given by:

$$f = f_0 \sin\frac{2\pi x}{L} \qquad (7.16)$$

It is also assumed that the additional bending curve has the same relationship as the misalignment due to the buckling load P:

$$v = v_0 \sin\frac{2\pi x}{L} \qquad (7.17)$$

Finally, the buckling force P and the wavelength L are assumed to remain constant during the buckling process.

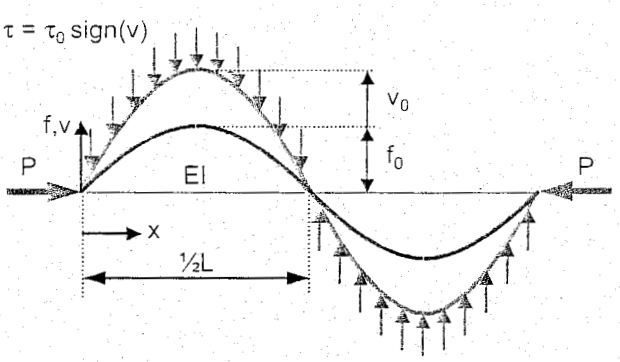


Figure 7.7: Lateral buckling of track panel with plastic lateral shear resistance

Under these conditions the energy parts can be calculated analogous to the scheme in the first example.

1. 'sign' refers to the mathematical function signum meaning here: sign (v) returns 1 if v>0, 0 if v=0, and -1 if v < 0

1. *Bending energy of the beam:*

$$U_{beam} = 2\int_0^{L/4} EIv''^2 dx = \frac{Lv_0^2 EI}{4}\left(\frac{2\pi}{L}\right)^4 \tag{7.18}$$

where v''^2 is d^2v/dx^2.

2. *Deformation energy of lateral shear resistance*, with (7.17):

$$U_{res} = 4\int_0^{L/4} \tau_0 v dx = \frac{2\tau_0 v_0 L}{\pi} \tag{7.19}$$

3. *Work done by the compressive force* (modified to include the effect of f(x)):

$$W_P = 2\int_0^{L/4} P[(v'+f')^2 - f'^2]dx = \frac{PL}{4}[(v_0+f_0)^2 - f_0^2]\left(\frac{2\pi}{L}\right)^2 \tag{7.20}$$

The total energy per wavelength in this case becomes:

$$U_{tot} = \frac{L}{4}\left[v_0^2 EI\left(\frac{2\pi}{L}\right)^4 + \frac{8\tau_0 v_0}{\pi} - P[(v_0+f_0)^2 - f_0^2]\left(\frac{2\pi}{L}\right)^2\right] \tag{7.21}$$

The equilibrium follows from setting the partial derivative $\frac{\partial U_{tot}}{\partial v_0} = 0$, resulting in:

$$P = \frac{v_0}{v_0+f_0}\left[\frac{4\pi^2 EI}{L^2} + \frac{\tau_0 L^2}{\pi^3 v_0}\right] \tag{7.22}$$

From $\frac{\partial P}{\partial L} = 0$ the following critical quantities can be determined:

$$L_{crit}^2 = 2\sqrt{\frac{\pi^5 f_0 EI}{\tau_0}} \tag{7.23}$$

$$P_{crit} = 2\sqrt{\frac{\tau_0 EI}{\pi f_0}} \tag{7.24}$$

$$\left(\frac{v_0}{f_0}\right)_{crit} = 1 \tag{7.25}$$

The critical value of P_{crit} is the smallest load of the compressive force P, with corresponding wavelength, which causes buckling.
The stability criterion of the equilibrium can be determined by checking the inequality $\frac{\partial^2 U_{tot}}{\partial v_0^2} > 0$, which gives the condition for stability:

$$P < \frac{4\pi^2 EI}{L^2} \tag{7.26}$$

This means that for a stable equilibrium the axial load P should be less than the Euler buckling load.

In Figure 7.8 the graphical result is given of an example based on the following parameters:

τ_0 = 10 kN/m;
EI = 8000 kNm²;
f_0 = 0.025 m.

The critical values can be calculated using (7.23) and (7.24):

L_{crit} = 12.51 m;
P_{crit} = 2018 kN.

The critical value of the compressive load lies well within the range of possible values (under rather extreme conditions, see former section).

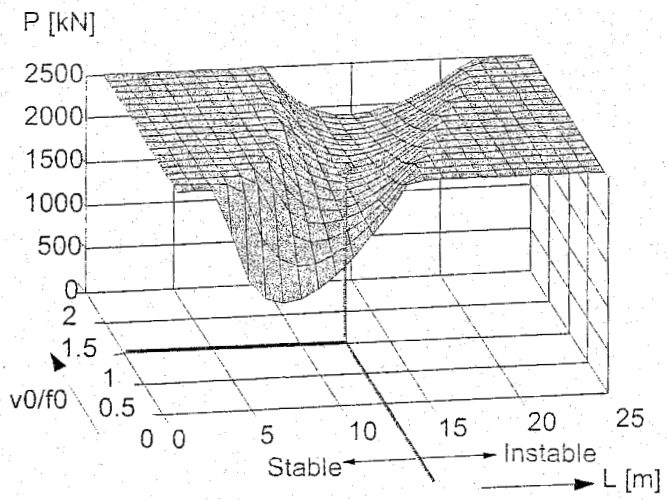


Figure 7.8: Buckling diagram in case of plastic resistance and misalignment

7.2 Track stability: finite element modelling

7.2.1 General considerations

In the 1950's and 1960's ORE Committee D14 investigated the stability of tracks with CWR under the influence of temperature loads on various geometrical imperfections [208], [209]. For this purpose analytical models were developed as described in reference [19]. Although these mathematical models had been verified by many field tests, several limitations and assumptions still had to be accepted in some areas. One of these limitations was that the mathematical model was only allowed to be applied to curve radii greater than 500 m. It is a fact that in tight curves an increase in temperature is followed by a lateral displacement over the total curve length. This decreases the axial compression force in the track. In such a case the longitudinal ballast resistance forces play an important role, especially in transition curves, and this aspect was lacking in the analytical models.

The above-mentioned limitations do not apply to calculations based on the finite element method (FEM). Using the model described in [115], the stability of a number of track structures under the influence of a temperature load was analysed. The non-linearities in this process require an incremental approach in which temperature loads are increased step by step. Moreover, the model allows for any design geometry for the track, including geometrical deviations. Boundary conditions are no longer a limiting factor.

7.2.2 Finite element model

The railway track is considered to be an elastically supported beam with a constant bending stiffness (EI) and a torsional resistance of the fastenings at the sleepers. The ballast resistance is represented by springs in the lateral and longitudinal directions. Displacement occurs in the two-dimensional, horizontal plane.

Each of the three spring types mentioned has a specific bi-linear characteristic according to Figure 7.9. This means that in the first part of the diagram the resistance increases in proportion to the displacement and in the second part the resistance is constant with displacement. This is known as the plastic phase.

The mathematical model permits an elastic spring reaction on removal of the load when the spring has already reached the plastic phase. The beam element length is constant for the overall track construction and corresponds to the sleeper spacing. As indicated in Figure 7.10, the joints possess three degrees of freedom u, v, and ϕ and are connected to two ballast springs and one rotation spring.

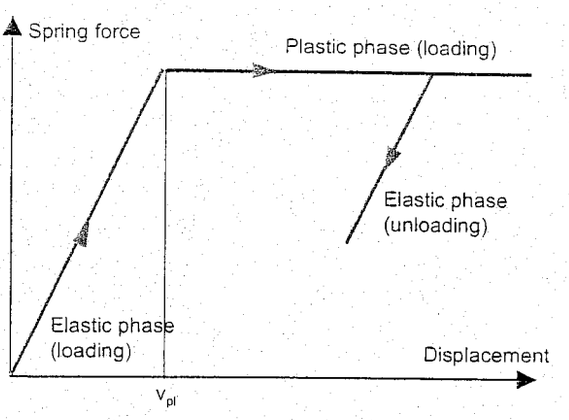


Figure 7.9: bi-linear spring characteristic

The equations of equilibrium for the track construction, composed of the beam elements, provide a system of equations in terms of the unknown displacement increments according to the matrix equation:

$$[S][\Delta v] = [\Delta F] \qquad (7.27)$$

where:
[S] = tangent stiffness matrix [N/mm];
[Δv] = displacement increment vector [mm];
[ΔF] = load increment vector [N].
The tangent stiffness matrix [S] consists of the following components:

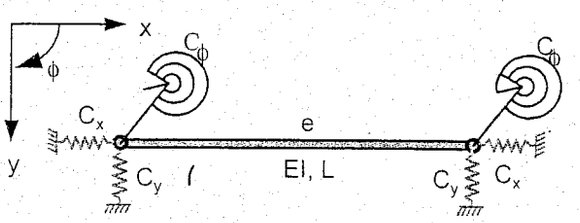


Figure 7.10: Beam element applied in FEM

Element stiffness matrix [S_1]:

The beam element used in the FEM consists of a combined truss-bending element. Therefore, the matrix [S_1] contains the strain and bending stiffness constants. The linear relation between displacement [v^e] and forces [F^e] of a beam element is defined according to the matrix equation:

$$[S_1][v^e] = [F^e] \qquad (7.28)$$

in which the displacement vector [v^e] contains the three degrees of freedom u, v, and ϕ per joint. The displacement field u of the truss element is governed by the differential equation:

$$EA\frac{d^2u}{dx^2} = 0 \qquad (7.29)$$

which has the solution:

$$u(x) = c_1 + c_2 = u_1^e + \frac{x}{L}(u_2^e - u_1^e) \qquad (7.30)$$

The displacement field is presented in Figure 7.11.

The bending part in the displacement field, associated with v and ϕ, is based on what are known as shape functions. These are third degree polynomials in x which satisfy the differential equation:

$$EI\frac{d^4u}{dx^4} = 0 \qquad (7.31)$$

with the solution:

$$v(x) = c_1 + c_2 x + c_3 x^2 + c_4 x^3 \qquad (7.32)$$

The constants c_1 to c_4 are dependent on the joint displacements. This makes it possible to express v(x) in the degrees of freedom according to:

$$v(x) = \alpha_1(x) v_1^e + \alpha_2(x) \phi_1^e + \alpha_3(x) v_2^e + \alpha_4(x) \phi_2^e \qquad (7.33)$$

The shape functions are presented in Figure 7.12 and can be expressed by:

$$\alpha_1 = 2\left(\frac{1}{2} - \frac{x}{L}\right)^2 \left(1 + \frac{x}{L}\right) \qquad (7.34)$$

$$\alpha_2 = \left(\frac{1}{2} + \frac{x}{L}\right)\left(\frac{1}{2} - \frac{x}{L}\right)^2 L \qquad (7.35)$$

$$\alpha_3 = 2\left(\frac{1}{2} + \frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right) \qquad (7.36)$$

$$\alpha_4 = -\left(\frac{1}{2} - \frac{x}{L}\right)\left(\frac{1}{2} - \frac{x}{L}\right)^2 L \qquad (7.37)$$

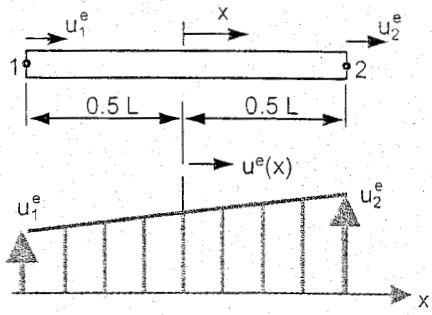


Figure 7.11: Displacement field of truss element

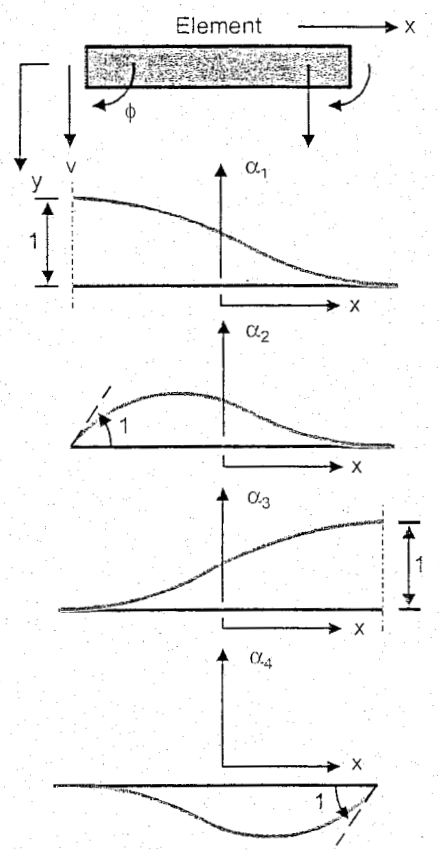


Figure 7.12: Shape functions of bending element

Stiffness matrix [S_2]:

This matrix takes the geometrical non-linearity based on the second order theory for small strains and large displacements into account and is derived by using the strains from the displacements. As depicted in Figure 7.13, the strain e is related to the displacements u and v by the following expression:

$$\varepsilon = \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial v}{\partial x}\right)^2 \qquad (7.38)$$

With this theory, the stability of the equilibrium of the two dimensional beam constructions can be investigated according to the principle of virtual work. There is equilibrium if the variation in potential energy equals zero, or:

$$\delta E_{pot} = \iiint_V \sigma \delta\varepsilon \, dV - \iint_l F \delta v \, dx = 0 \qquad (7.39)$$

Assume that due to the load q_0 the axial forces N_0 and bending moments M_0 are present in the construction. If variations ∂u and ∂v are applied to the displacements u and v, with an unchanged external loading, the following variation in the potential energy of the construction will emerge:

$$\delta E_{pot} = \int \left[N_0 \delta\varepsilon + M_0 \delta\kappa + \frac{1}{2} dN \delta\varepsilon + \frac{1}{2} dM \delta\kappa - q_0 \delta v \right] dx \qquad (7.40)$$
$$= \delta^1 E_{pot} + \delta^2 E_{pot}$$

The first term:

$$\delta^1 E_{pot} = \int \left[N_0 \frac{\partial \delta u}{\partial x} - M_0 \frac{\partial^2 \delta v}{\partial x^2} - q_0 \delta v \right] dx \qquad (7.41)$$

describes the equilibrium of the internal forces with the external load. The second term describes the strain energy and is split up into a linear and a non-linear part:

$$\delta^2 E_{pot} = \delta^2 E_{pot,1} + \delta^2 E_{pot,2} \qquad (7.42)$$

In this expression the linear and non-linear parts are respectively:

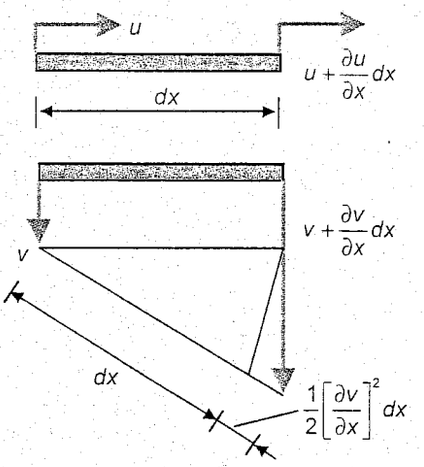


Figure 7.13: Axial strain due to displacements u and v

$$\delta^2 E_{pot,1} = \frac{1}{2} \int EA \left(\frac{\partial \delta u}{\partial x} \right)^2 dx + \frac{1}{2} \int EI \left(\frac{\partial^2 \delta v}{\partial x^2} \right)^2 dx \qquad (7.43)$$

$$= \frac{1}{2} [v^e]^T [S_1^e] [\partial v^e] \qquad (7.44)$$

$$\delta^2 E_{pot,2} = \frac{1}{2} \int N_0 \left(\frac{\partial \delta v}{\partial x} \right)^2 dx \qquad (7.45)$$

$$= \frac{1}{2} [v^e]^T [S_2^e] [\partial v^e] \qquad (7.46)$$

The matrix $[S_1^e]$ is the symmetrical element stiffness matrix. The matrix $[S_2^e]$ contains factors which depend on the normal force. It creates the extra part in the matrix [S] which expresses the geometrical non-linearity.

Spring stiffness matrix [S_v]:

The bi-linear springs at the joints contribute to the total construction stiffness. However, they do not add extra degrees of freedom to the system. As long as a spring is in the elastic phase, the matrix relation for a linear spring applies according to:

$$\begin{bmatrix} C_x & & \\ & C_y & \\ & & C_\phi \end{bmatrix} \begin{bmatrix} u \\ v \\ \phi \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ M_\phi \end{bmatrix} \qquad (7.47)$$

or:

$$[S_v][v] = [F] \qquad (7.48)$$

As shown, the matrix [S_v] contains the stiffness of ballast and fastenings which are placed on the matrix main diagonal. After having reached the plastic phase, the spring force no longer increases but remains constant with displacement. The spring's contribution to the track stiffness equals zero for the next loading increment.

By direct summation of these three matrices the tangent stiffness matrix [S] is formed. The support conditions are taken into consideration when reducing the matrix [S]. The calculating process is as follows:

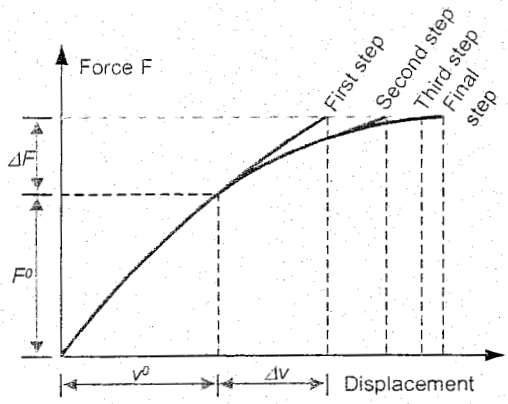


Figure 7.14: Newton-Raphson iteration process

After applying the temperature load and solving the static equilibrium equations, the displacement vector [Δv] is found. After that a new (rotated) tangent stiffness matrix [S] is drawn up with the help of the new total displacement (v°+Δv) and the internal stresses. Back substitution of the same displacement vector [Δv] gives a new load vector [ΔF*]. The difference between the original load vector [ΔF] and the load vector [ΔF*] serves as a new load vector in the new total displacement situation (v°+Δv) until the desired accuracy has been reached. This process, schematically presented in Figure 7.14, is often described as the Newton-Raphson iteration process.

The load on the track construction can consist of two components, namely:

Temperature load

This load is due to a temperature rise which causes compressive forces in the track. One of the consequences can be lateral track instability (buckling). This load is active along the total track. According to Section 7.3.2, the following formula (7.54) for the normal rail force N due to a temperature change ΔT with respect to the neutral or initial temperature is found:

$$N_{max} = -EA\alpha\Delta T \tag{7.49}$$

Hence, the compression force on the track (two rails) is:

$$P = 2EA\alpha|\Delta T| \tag{7.50}$$

in which:

P = compressive force on the track, taken as positive value in subsequent calculations [N];
E = Young's modulus for the rail [N/mm²];
A = total cross-sectional area of one rail [mm²];
α = coefficient of expansion [/°C];
$\Delta T = T_{actual} - T_{neutral}$ [°C].

External joint loads

The possibility of applying individual joint loads in the lateral and longitudinal directions is created. This was done in order to verify the mathematical model by means of data from lateral resistance measurements described in [73].

Both load types can be applied individually or in combination.

7.2.3 Results

After a thorough and comprehensive verification of the mathematical model and computer program, efforts were made to reproduce the results presented in [19]. In order to do this, the necessary input data were derived from the measurement data originally used. The results are presented in Figure 7.15 and show a high level of similarity. The same figure also indicates the existence of a critical wavelength which refers to a minimum buckling force in the relation buckling force/buckling length (= half wavelength L). It also refers to a lower critical compression force (buckling force) when the amplitude y of the misalignment is larger.

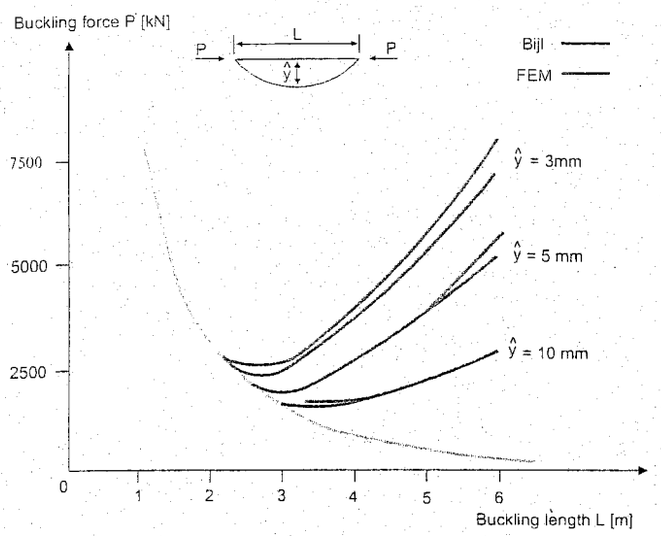


Figure 7.15: Buckling force versus buckling length according to Bijl and calculated with FEM

In the calculations published in [19] a sinusoidal initial displacement, i.e. geometrical imperfection of the track, was assumed. At different wavelengths, with the wavelength being twice the buckling length, the buckling force was determined for different initial conditions. In practice, wavelength and geometrical deviation magnitude are strongly correlated. Starting from some tenths of a millimetre at wavelengths of a few metres and some millimetres at wavelengths of 10 to 20 m, the amplitudes can become much larger than 10 mm for longer waves.

In order to take this correlation into account, two sections of about 30 m length were selected for the calculations on the basis of BMS recordings. Both sections form part of a 200 m section with a standard deviation of 1 mm for alignment, which corresponds approximately to the maintenance intervention level on NS rail. The geometrical deviations in the lateral direction (alignment) are presented in Figure 7.16. For both initial deviations of the track geometry the differential temperature, with respect to the neutral temperature at 25°C at which the track becomes unstable, was calculated. This value amounts to 140 - 150°C [89].

Subsequently, these calculations were repeated by increasing the amplitude of the initial geometrical deviation in proportion to the standard deviation of alignment per 200 m section. Buckling forces and differential temperatures for standard deviations of 0.5, 1, 2, and 4 mm are presented in Figure 7.16. The analysis results show that only in extremely poor tracks differential temperatures are attained which will lead to instability.

The NS have carried out many lateral resistance measurements with an adapted tamping machine [73]. To simulate this loading process with an FEM analysis, two point loads were applied to straight track in the lateral direction. Figure 7.17 shows some results of lateral resistance measurements car-

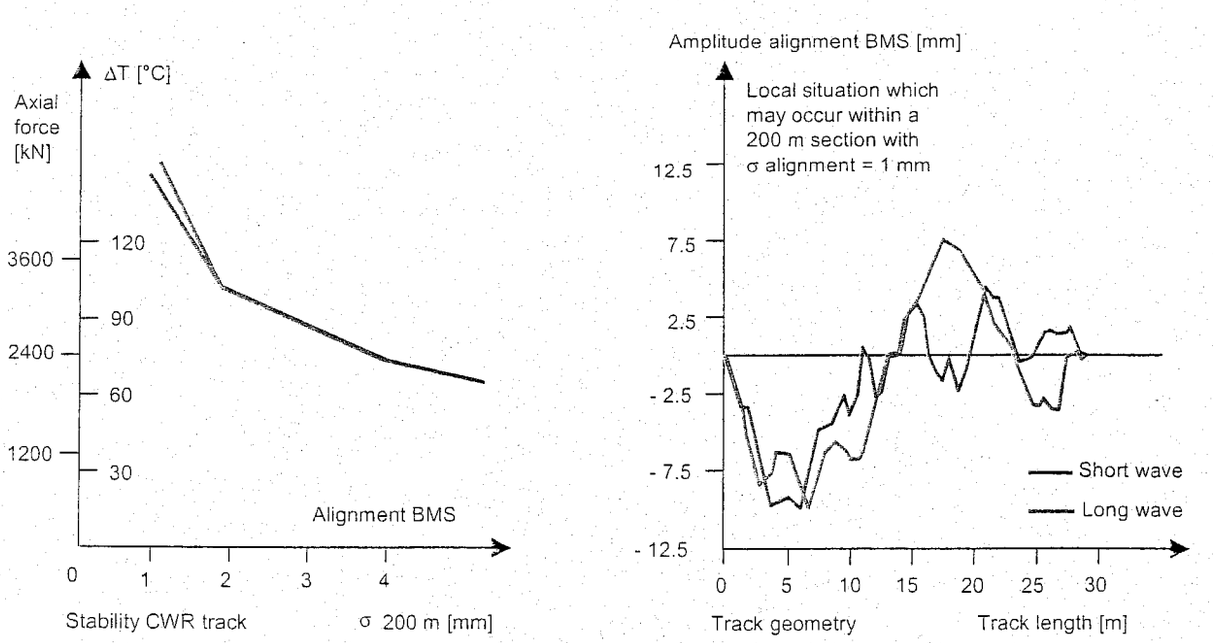


Figure 7.16: Buckling force versus track quality

ried out on NS UIC 54 track and on RET (Rotterdam Metro) NP 46 track. The FEM approximation and relevant parameters are also indicated. The correspondence between the measured and calculated value is quite good.

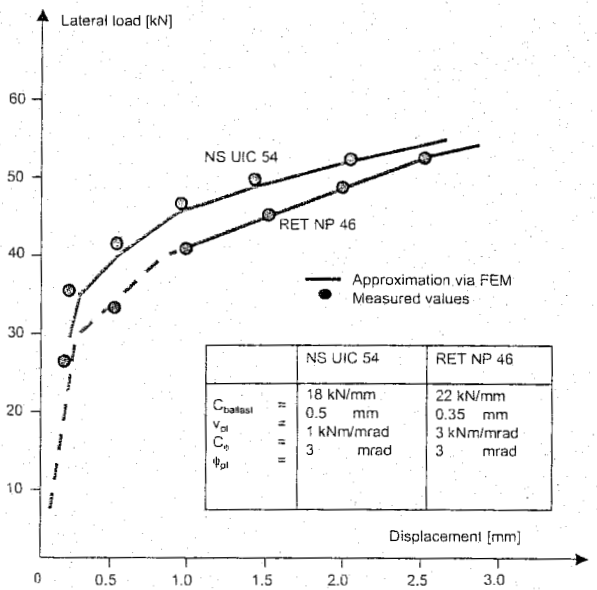


Figure 7.17: Approximating lateral resistance measurements by FEM

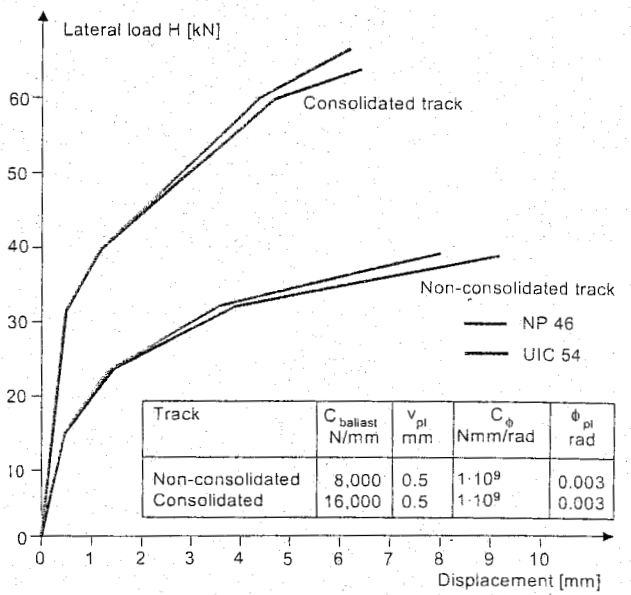


Figure 7.18: Force-displacement characteristics for different railtypes and different stiffness coefficients

To illustrate the influence of different spring characteristics, the load-displacement diagram is plotted in Figure 7.18 for tracks with NP 46 and UIC 54 rail profiles. The influence of the rail profile is relatively low. The consolidated and non-consolidated characteristics are FEM approximations of the lateral resistance measurements mentioned earlier.

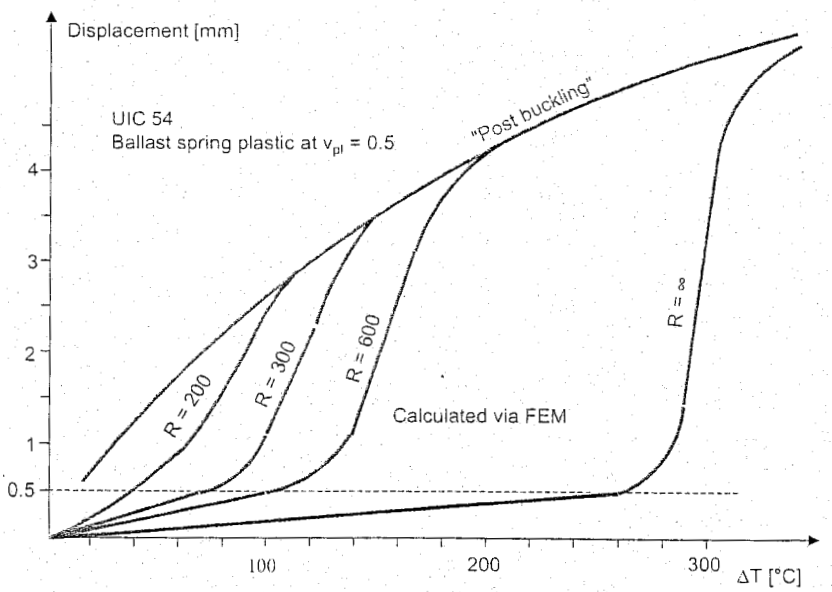


Figure 7.19: Lateral displacement versus differential temperature for different curve radii

On the basis of the non-consolidated characteristic in Figure 7.18, the relationship between lateral displacement and differential temperature has been calculated for different curve radii. The bi-linear ballast spring force has a maximum value of 8.5 kN which is attained at a displacement of v = v_{pl} = 0.5 mm. The FEM analysis was carried out with up to about 1 mm lateral displacement.

The post-buckling behaviour was investigated for only a few cases. The results, presented in Figure 7.19, reveal that the tighter the curve the more gradual the transition to the post-buckling phase develops.

Regarding NP 46 track, the curve radius versus differential temperature, at which the track starts to displace permanently, was also determined for consolidated tracks. For both NP 46 and UIC54 tracks the load-displacement curve for the NP 46 track, shown in Figure 7.17, was applied. The analysis results are shown in Figure 7.20. These values indicate that consolidated tracks in curves with a radius of 200 m will start to develop permanent lateral displacements at differential temperatures above 80°C for UIC54 and 100°C for NP46. Non-consolidated tracks start to displace permanently at about half this value. It should be mentioned that the occurrence of a limited permanent displacement should be considered as permissible.

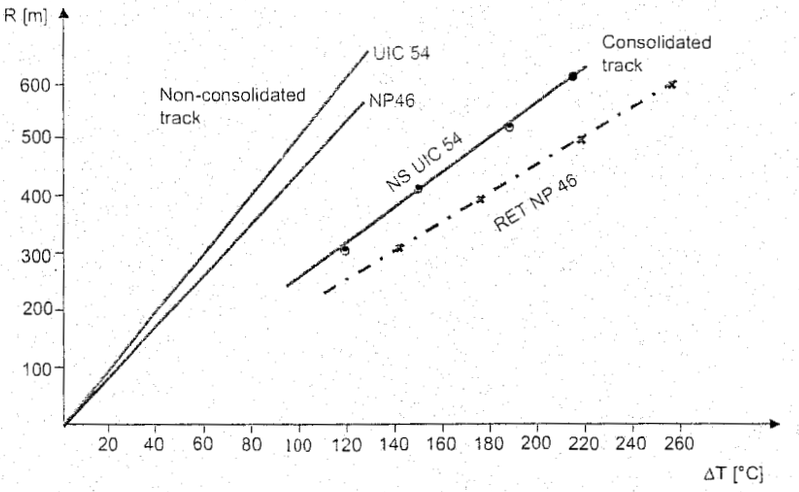


Figure 7.20: Curve radius versus temperature at which permanent track displacements start to occur

7.2.4 Continuous welded switches

In the case of continuous welded switches, the axial forces in the rails connected to the common crossing should be relieved by the sleepers so that they have vanished at the start of the switch blade. Depending on the sleeper stiffness, these forces are partly transferred to both continuous rails and partly to the ballast bed. Sleeper loads and corresponding displacements have been measured by DB for a 1:12 switch. These results, obtained from reference [146], are presented in Figure 7.21.

As a result of the reaction forces exerted on the continuous rails the axial compression force in these rails is reduced in the vicinity of the common crossing. Near the front end of the switch blades the axial compression forces in the continuous rails increase. According to Figure 7.21 this increase may amount to approximately 40%. DB have also carried out measurements at 1:12 crossovers. Owing to the limited transition length, an increase in axial compression force of only 7% was found.

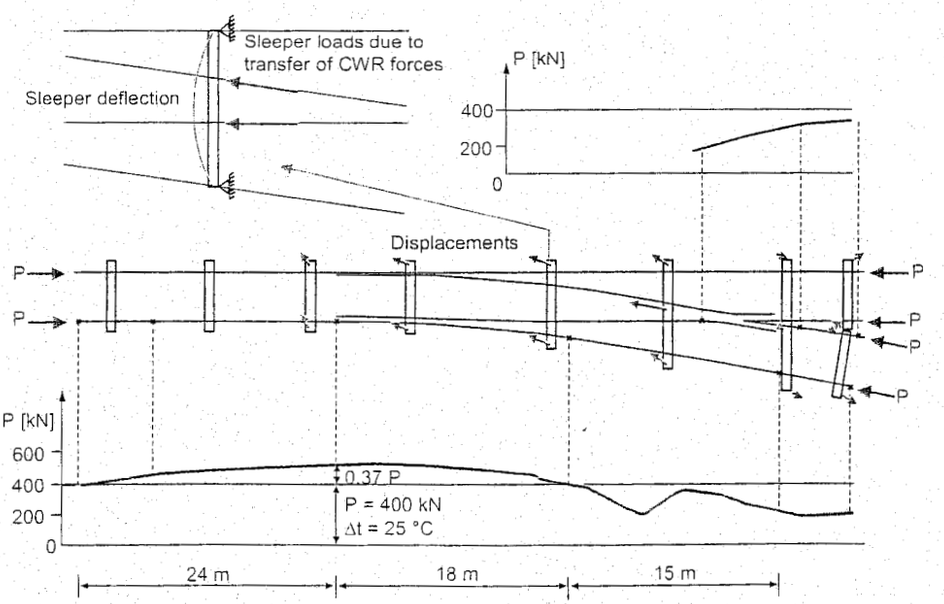


Figure 7.21: Forces in CWR 1:12 switch measured by DB

In general, τ is a function of u, which in turn, is a function of x. Using the formulas (7.51) and (7.52), the following differential equation can be derived:

$$\frac{d^2u}{dx^2} - \frac{\tau(u)}{EA} = 0 \tag{7.53}$$

Once the displacement function u(x) is found, the normal force follows from (7.52):

$$N = EA\left(\frac{du}{dx} - \alpha\Delta T\right) \tag{7.54}$$

In the case of continuous welded rail (CWR), the length of the rail is so great that a plane strain condition exists in the central part of the rail preventing axial displacement of the rail completely. The force caused by a temperature increase is, according to (7.54):

$$N_{max} = -EA\alpha\Delta T \tag{7.55}$$

which means that a temperature increase results in a compressive normal force, as could be expected.

Special case 1: Plastic shear resistance

Sometimes it may be assumed that the shear resistance is constant while the sign of the resistance depends only on the sign of the displacement:

$$\tau = \tau_0 sign(u) \tag{7.56}$$

At the point of transition from CWR to jointed track, according to Figure 7.23, the CWR force is reduced over the so-called breathing length. From (7.53) and (7.56), it follows for $x \geq 0$ that:

$$u(x) = \frac{\tau_0 x^2}{2EA} \tag{7.57}$$

which satisfies the transition conditions at x = 0. With (7.54) the normal force in the breathing length zone is:

$$N = \tau_0 x - EA\alpha\Delta T \tag{7.58}$$

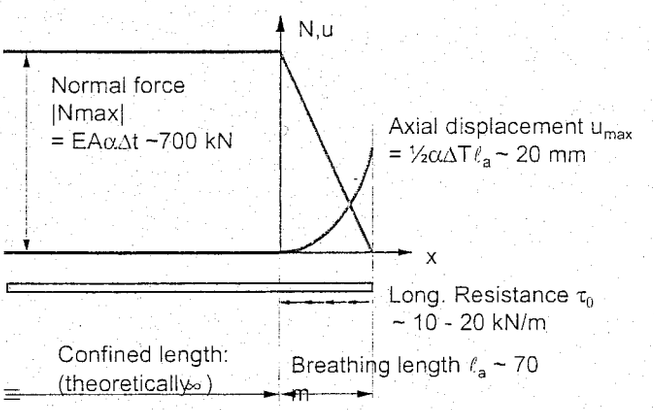


Figure 7.23: Distribution of temperature force and displacement in CWR (plastic shear resistance)

At the free end of the rail ($x = \ell_a$) there is no force, hence the breathing length:

$$\ell_a = \frac{N_{max}}{\tau_0} \tag{7.59}$$

The expressions for the maximum normal force and maximum axial displacement are also indicated in Figure 7.23.

The results in this case are very plausible and may in fact be deduced heuristically. This is the solution often found in text books.

As the plastic resistance characteristic is inherently non-linear, residual stresses remain at the rail ends if the rail system is subjected to more than one temperature cycle, e.g. consecutive warming up/ cooling down periods. In this case the rail end displacement follows a hysteresis loop.

Special case 2: Elastic shear resistance

In the case of elastic rail fastening systems the assumption of a linear shear resistance is more appropriate:

$$\tau = ku \tag{7.60}$$

Substitution in (7.53) results in the differential equation:

$$\frac{d^2u}{dx^2} - \frac{k}{EA}u = 0 \tag{7.61}$$

The solution of (7.61) which satisfies the boundary conditions: x = 0 => N = 0 and $x \rightarrow \infty$ => u = 0 can be written as:

$$u = -\frac{\alpha \Delta T}{\mu} e^{-\mu x} \tag{7.62}$$

in which:

$$\mu = \sqrt{\frac{k}{EA}} \tag{7.63}$$

The normal force according to (7.54) becomes:

$$N = -EA\alpha\Delta T(1 - e^{-\mu x}) \tag{7.64}$$

The expressions for the maximum normal force and maximum axial displacement are indicated in Figure 7.24. Theoretically, there is no distinction here between the breathing length and confined length as was the case with the plastic shear resistance. Practically, though, the large central part of the rail may be regarded as being in a plain strain situation. As this solution is linear elastic, no hysteresis effects are taking place when the system is subjected to consecutive temperature variations. In the next section the more complex case of temperature effects in the combined system track/viaduct will be examined.

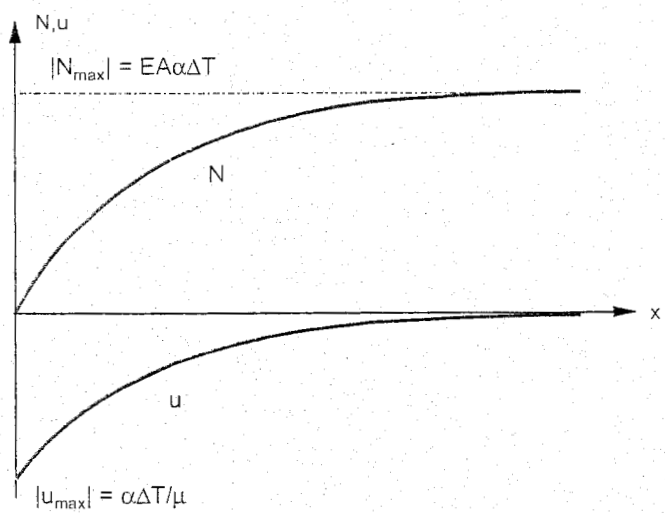


Figure 7.24: Distribution of temperature force and displacement in CWR (elastic shear resistance)

7.3.3 Modelling of the longitudinal interaction problem

In Section 7.3.2 the effect of temperature forces in rails was examined using a simple model (Figure 7.22). To assess the complex temperature effects in the system track on a bridge or a viaduct, we will use a more generalised model, shown in Figure 7.25.

Although there may be more than one track fixed on the bridge, we here only consider one continuous welded rail on a corresponding part of the whole bridge.

In analogy with the theory in Section 7.3.2, we can write down the following mixed equations taking into account the interaction of the longitudinal shear resistance between rail and bridge:

$$\frac{d^2u}{dx^2} + \frac{\tau(u_b - u)}{EA} = 0 \qquad (7.65)$$

$$\frac{d^2u_b}{dx^2} - \frac{\tau(u_b - u)}{(EA)_v} = 0 \qquad (7.66)$$

$$N = EA\left(\frac{du}{dx} - \alpha\Delta T\right) \qquad (7.67)$$

$$N_b = (EA)_b\left(\frac{du_b}{dx} - (\alpha\Delta T)_b\right) \qquad (7.68)$$

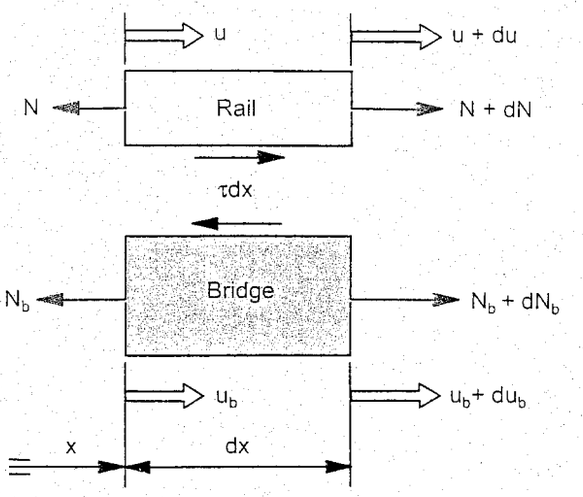


Figure 7.25: Rail and bridge element

in which:

u, u_b = displacement of the rail and bridge, respectively;
N, N_b = normal force in the rail and the bridge, respectively;
$\tau=\tau(u_b-u)$ = axial shear resistance, depending on the difference of the displacements;
EA, $(EA)_b$ = axial normal stiffness of the rail and bridge, respectively;
$\alpha\Delta T$, $(\alpha\Delta T)_b$ = temperature strain of the rail and bridge respectively.

To simplify matters only elastic displacements are assumed:

$$\tau = k(u_b - u) \qquad (7.69)$$

Moreover, it is assumed that the normal stiffness of the rail is much less than the corresponding part of the bridge:

$$EA \ll (EA)_b \qquad (7.70)$$

In this case the bridge exhibits an almost uniform axial expansion or shrinking and the general solution of (7.65) and (7.66) can be written as:

$$\begin{aligned} u &= C_1 \sinh \mu x + C_2 \cosh \mu x + C_3 + C_4 \\ u_b &= C_3 x + C_4 \end{aligned} \qquad (7.71)$$

in which:

$$\mu = \sqrt{\frac{k}{EA}} \qquad (7.72)$$

Consider two special cases:

case 1: $C_3 = 0; C_4 = 0 \quad => \quad u_b = 0; \qquad N_b = N_{max} = -(EA\alpha\Delta T)_b$

In this case the bridge is confined completely and the solution of the rail is, with appropriate boundary conditions, identical to the solution (7.55).

case 2: $C_3 = (\alpha\Delta T)_b; C_4 = 0 => \quad ub = (\alpha\Delta T)_b x; \quad N_b = 0$

The bridge now has a fixed point (support) at x = 0 and can for the rest expand freely.

In this case the complete solution for the rail part of the system follows from (7.71) first equation, and (7.67):

$$u = C_1 \sinh \mu x + C_2 \cosh \mu x + (\alpha\Delta T)_b x \tag{7.73}$$

$$N = EA\mu(C_1 \cosh \mu x + C_2 \sinh \mu x) + EA(\alpha\Delta T)_b - EA\alpha\Delta T \tag{7.74}$$

The last term on the right side in (7.74) represents the normal force in the confined continuous welded rail; the other terms represent the modification of the rail force due to the bridge interaction.

Several bridge configurations have been studied with appropriate boundary conditions using the equations (7.73) and (7.74). As an example we will examine the case of a continuous welded rail fixed on an ever repeating system of short bridges subjected to a temperature decrease.
The boundary conditions are in this case, because of continuity:

$$x = 0 \quad => u(0) = u(\ell); \quad N(0) = N(\ell)$$

The solution of (7.73) and (7.74) is therefore:

$$u = \tfrac{1}{2}(\alpha\Delta T)_b \ell [\cosh \mu x - \coth \tfrac{1}{2}\mu\ell \sinh \mu x] + (\alpha\Delta T)_b x \tag{7.75}$$

$$N = \tfrac{1}{2} EA(\alpha\Delta T)_b \mu\ell [\sinh \mu x - \coth \tfrac{1}{2}\mu\ell \cosh \mu x] + EA(\alpha\Delta T)_b - EA\alpha\Delta T \tag{7.76}$$

In Figure 7.26 a numerical example is given using the following parameters:

E	$= 2.1 \cdot 10^8$	kN/m²;
A	$= 60 \cdot 10^{-4}$	m²;
α	$= 1.15 \cdot 10^{-5}$	1/°C;
ΔT	= 40	°C;
ΔT_b	= 10	°C;
k	= 7636	kN/m²;
L	= 48.7	m.

If no bridge interaction existed, the normal force in the rail would be N_{max} = 580 kN. However, due to the interaction the normal force at the support amounts to 142 kN (24 %) which is quite substantial. The maximum rail displacement is half of that of the maximum bridge displacement.

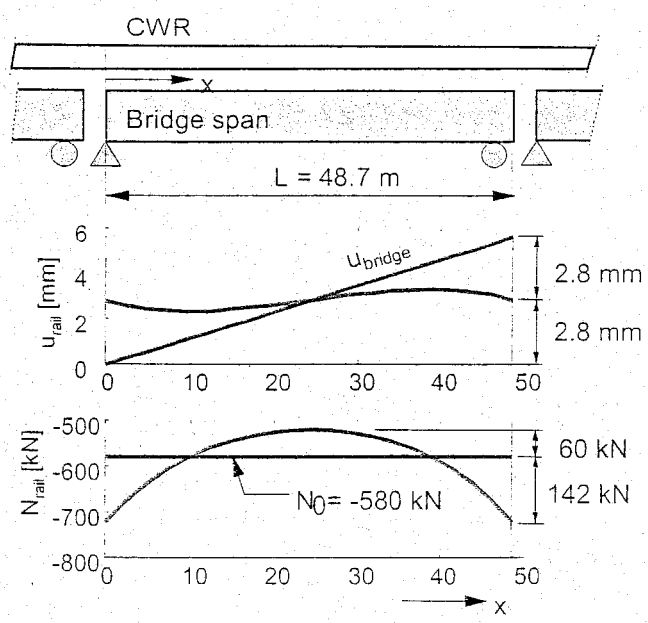


Figure 7.26: Longitudinal displacement and force in CWR track on a repeating bridge configuration

7.4 Longitudinal forces: finite element modelling

7.4.1 General considerations

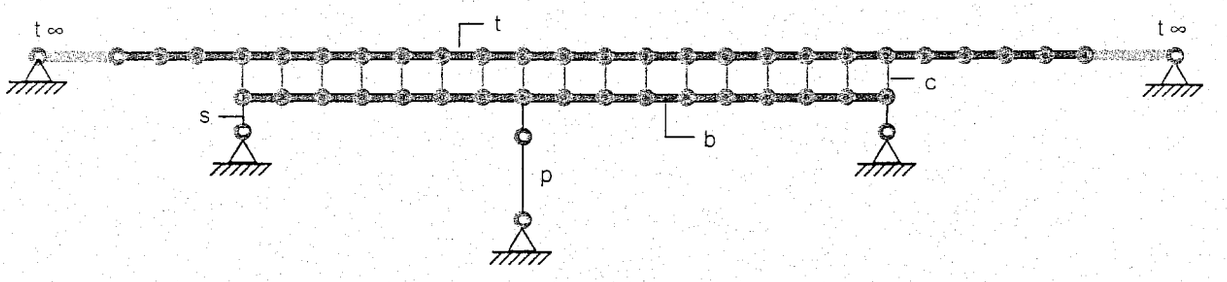
The relatively simple analytical approach, as discussed in Section 7.3, of the longitudinal problem is instructive in order to understand the temperature effects in the track. It should be noted, however, that the modelling used there is based on a number of limitations and assumptions, viz.:

- lateral bending stiffness EI is constant;
- lateral shear resistance is constant;
- compressive force P = constant;
- no vertical loading;
- no longitudinal resistance;
- no axial strain;
- misalignment sinusoidal;
- additional bending sinusoidal;
- no curves.

7.4.2 Finite element model

To obtain a more realistic description of the problem, a finite element model, called PROLIS [281], has been developed to calculate longitudinal track forces in a similar way to the model described earlier regarding track stability. This model comprises track elements, ballast elements, and elements representing the bridge construction including abutments and pillars. The model allows for an arbitrary number of parallel tracks. Figure 7.27 shows the element composition which can be used to model a track/bridge construction.

The ballast spring is, as in the case of the stability program discussed in Section 7.2.2, also modelled as a bi-linear spring according to Figure 7.28. The maximum force, i.e. the force at which yielding starts, depends on the current vertical track load. Two variants have been investigated to describe the plastic behaviour. The first one is sketched in Figure 7.29 and assumes that the elastic limit always coincides with a fixed displacement u_{pl}. This means an increase in ballast stiffness in accordance with a growing vertical track load. This assumption does not exclude discontinuities from occurring in the spring force if vertical loads are added or removed. The second variant, shown in Figure 7.30, consists of a spring with constant stiffness in relation to the vertical track load. In this case, the displacement at which plastic deformations start grows linearly with the track load. From the physical point of view this approach is more consistent. However, simulation tests have shown that there is no significant difference between the results of both methods as the displacements in the areas of interest are often substantially greater than u_{pl}.



○	Joint	c	Ballast element with Coulomb friction
t	Track element (2 rails)	s	Support element
t ∞	Element representing infinitely long track	p	Pillar element

Figure 7.27: Finite element model for calculation of axial forces in tracks and on bridges

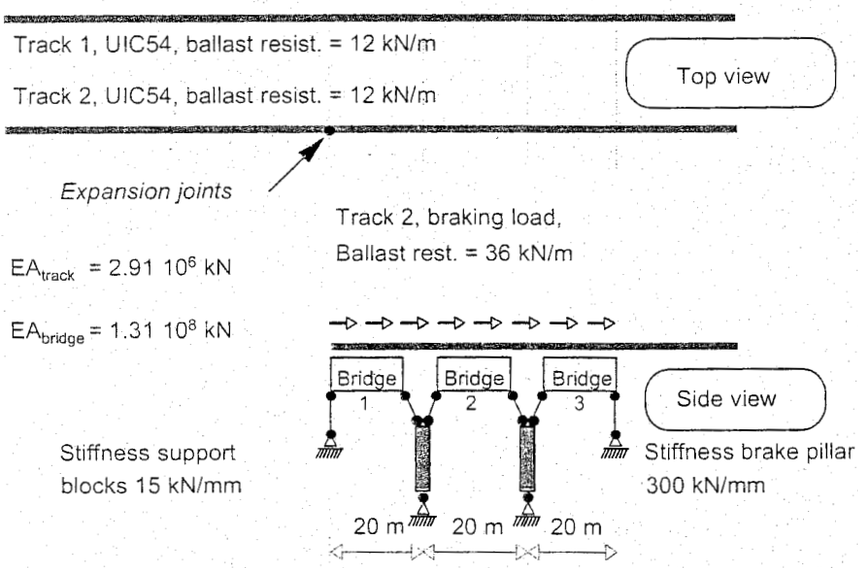
7.4.3 Examples of longitudinal force calculations

Bridge in Amsterdam West Branch

This construction consists of 3 bridges with a length of 20 m each. The bridges carry 2 ballasted tracks. The tracks are continuous welded although track 2 is provided with expansion joints at the beginning of the first bridge. The temperature loads consist of $\Delta T = -45°C$ for the tracks and $\Delta T = -25°C$ for the bridge. A braking force of 8kN/m is applied to track 2. The situation is sketched in Figure 7.32. The maximum ballast force for the non-loaded tracks is taken as 12 kN/m and for the loaded tracks 36 kN/m is used. It is assumed that the temperature loads are first applied during which all tracks have the same ballast yield force. Subsequently, the ballast yield force for the loaded part of track 2 is raised and the braking loads are applied.

The resulting track forces and track displacements are presented in Figure 7.33. Obviously, the largest displacements are achieved at the expansion joints in track 2. At this location, by definition, the longitudinal force is zero in track 2 and has its maximum value in track 1. The longitudinal force of track 2 has been partly transferred to track 1 by means of the ballast and bridge elements. This effect shows great similarities to the force transfer observed in the switch discussed in Section 7.2.4.

The calculation was repeated for two variants. In the first variant the spring characteristic, according to Figure 7.29, was replaced by a ballast spring with constant stiffness as described in Figure 7.30. Comparing the results for the displacements of track 2 and the forces of track 1 revealed that the spring characteristic with respect to the vertical load does not influence the calculated forces. The displacements are 6% higher.



***1st loading step** : temperature load tracks*
Δt = - 45 °C, bridges : Δt = -25 °C

***2nd loading step** : braking load 8 kN/m on track 2 above the bridges, ballast resistance under vertical load = 36 kN/m*

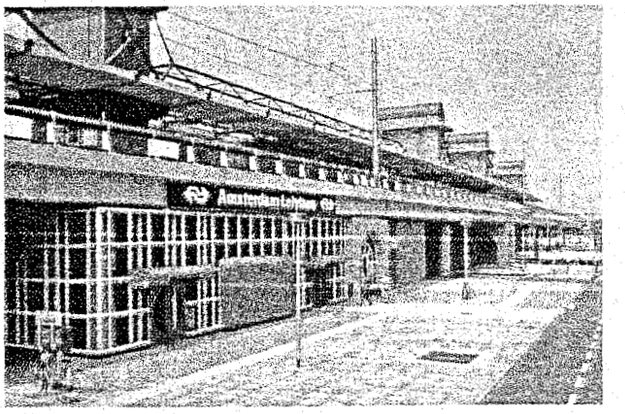


Figure 7.32: Loads on bridges in Amsterdam West Branch

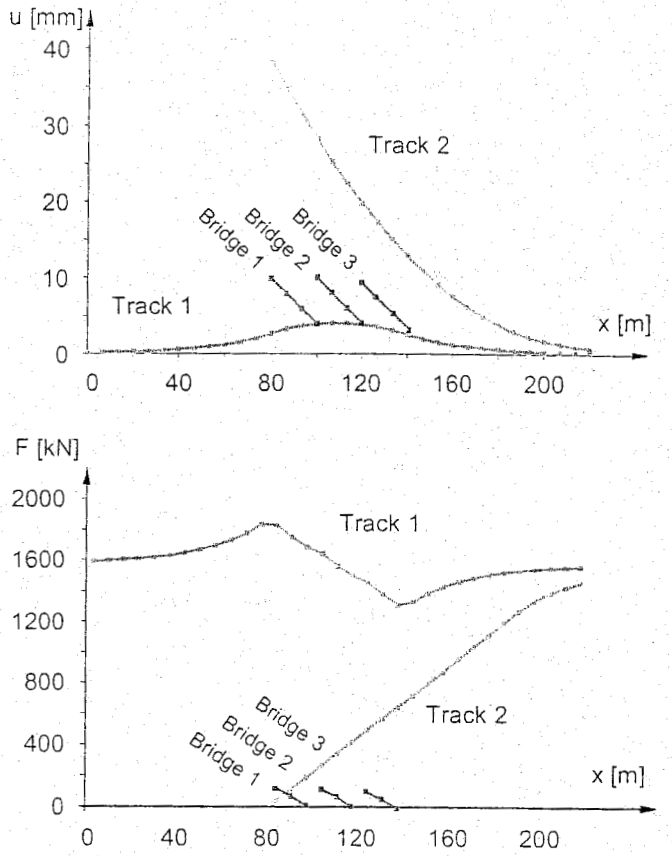


Figure 7.33: Track forces and track displacements resulting from the loads specified in Figure 7.32

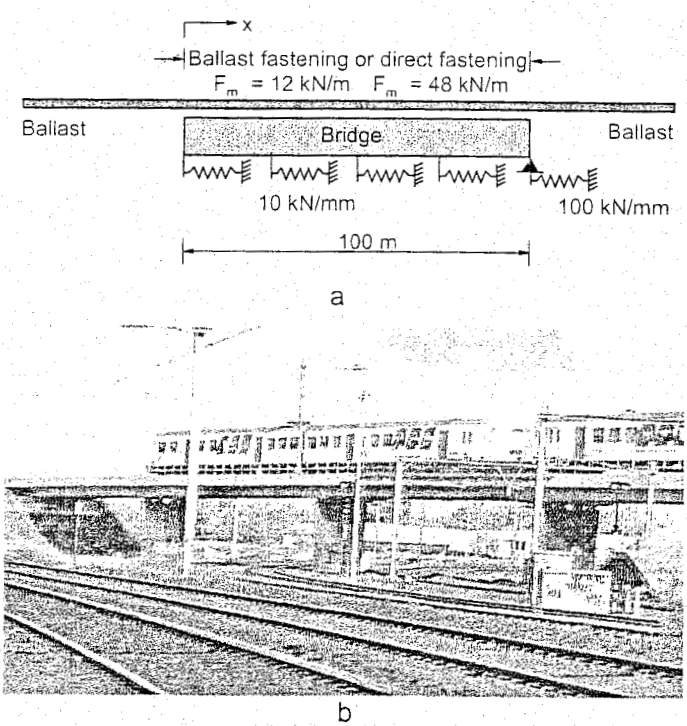


Figure 7.34: Utrecht flyover bridge

In the second variant, the elevated ballast stiffness of track 2 is applied immediately when the temperature load is raised. It was shown that this case leads to an underestimation of the displacements of the order of 37% and an overestimation of the force of the order of 3%.

In order to simulate a brittle rail failure, upon raising the temperature load track 2 was first considered to be long-welded without expansion joints. When applying the braking loads, the rails of track 2 were assumed to be broken, i.e. having expansion joints. The results deviate by less than 1% from the earlier results presented in Figure 7.33. Dynamic effects due to rail fracture have not been taken into account.

Utrecht flyover bridge

This example concerns a 100 m long flyover bridge carrying a single long-welded track. The support conditions of the bridge are presented in Figure 7.34a while Figure 7.34b shows a picture of the bridge. The track is subjected to a temperature load of $\Delta T = -45°C$ and the bridge to $\Delta T = -25°C$.

In this case 6 alternatives were considered which are described in Table 7.1. They consist of CWR without expansion joints, CWR with expansion joints at the left end of the bridge, and the expansion joints replaced by fastenings with teflon pads allowing for a relative displacement between rails and sleeper over a short length of track. For the maximum longitudinal force, referred to as F_m, frozen ballast, normal ballast and direct fastening conditions were considered.

The results of these calculations are presented in Figure 7.35 and Figure 7.36, showing the axial rail forces and the axial rail displacements respectively. The peak stresses and displacements are summarized in Table 7.1. Without expansion joints the frozen ballast, combined with direct fastenings on the bridge, causes the highest rail stresses of the order of 180 N/mm². This means an increase of over 60% compared to the undisturbed temperature stress. In the case of normal ballast conditions and direct fastenings on the bridge, the maximum stress drops 8% compared to the frozen situation. Obviously, a continuous ballast bed smooths the peak stress substantially as is demonstrated in load case 4.

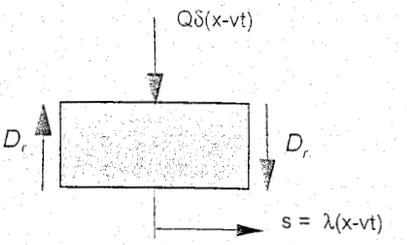


Figure 6.16: Transfer condition at s = 0

The transition conditions at $s = 0$ result in the following equations:

$$\begin{aligned} w_l &= w_r && \rightarrow A_3 + A_4 = A_1 + A_2 \\ \frac{dw_l}{ds} &= \frac{dw_r}{ds} && \rightarrow A_3\gamma_3 + A_4\gamma_4 = A_1\gamma_1 + A_2\gamma_2 \\ M_l &= M_r && \rightarrow A_3\gamma_3^2 + A_4\gamma_4^2 = A_1\gamma_1^2 + A_2\gamma_2^2 \\ Q + D_r &= D_l && \rightarrow \frac{Q}{EI\lambda^3} + A_3\gamma_3^3 + A_4\gamma_4^3 = A_1\gamma_1^3 + A_2\gamma_2^3 \end{aligned} \qquad (6.53)$$

Note that the last condition in (6.53) can be obtained using equations (6.28), (6.29), and Figure 6.16.

Writing this system of equations in matrix notation gives:

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ \gamma_1 & \gamma_2 & -\gamma_3 & -\gamma_4 \\ \gamma_1^2 & \gamma_2^2 & -\gamma_3^2 & -\gamma_4^2 \\ \gamma_1^3 & \gamma_2^3 & -\gamma_3^3 & -\gamma_4^3 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = w_0 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 8 \end{bmatrix} \qquad (6.54)$$

in which the factor $w_0 = Q/(8EI\lambda^3)$ represents the maximum static deflection. Apparently, if this factor is set to one, the solution corresponds with the dimensionless relative displacement $\eta(s)$ in the static case ($v = 0$). This will be useful to compare the dynamic solution ($v \neq 0$) with the static one for various values of α and β, and to determine the amplification factor.

After solving the matrix equation (6.54) the constants A_i are known and the relation between the deflection and the distance can be drawn. Figure 6.17 shows the characteristic wave shapes of the relative displacement of the beam for several values of α and β in the case of undamped ($\beta = 0$), lightly damped ($\beta = 0.1$), over critically damped systems ($\beta = 1.1$), static situations ($\alpha = 0$), subcritical velocities ($\alpha = 0.5$), critical velocities ($\alpha = 1$), and super critical velocities ($\alpha = 2$).

The fourth row in Figure 6.17 shows that the maximum amplitude of the displacements is moving behind the location of the load for super critical velocity. For the critical speed ($\alpha = 1$) and undamped case ($\beta = 0$) the wave amplitudes became infinite.

For a lightly damped system (middle column) a similar behaviour takes place. The wave shape calculated for $\alpha = 1$ shows large amplifications. For an over critically damped case ($\beta = 1.1$) the wave forms are asymmetric with respect to the load and show no amplifications anymore with respect to the static case.

In Figure 6.18 the ratio of the maximum deflection is given as function of the ratio of the load velocity and the critical velocity α for several values if the damping ratio is β. The equivalence to the frequency response function of the simple spring-mass system is striking. For small damping ratios the wave amplitude shows severe amplifications.

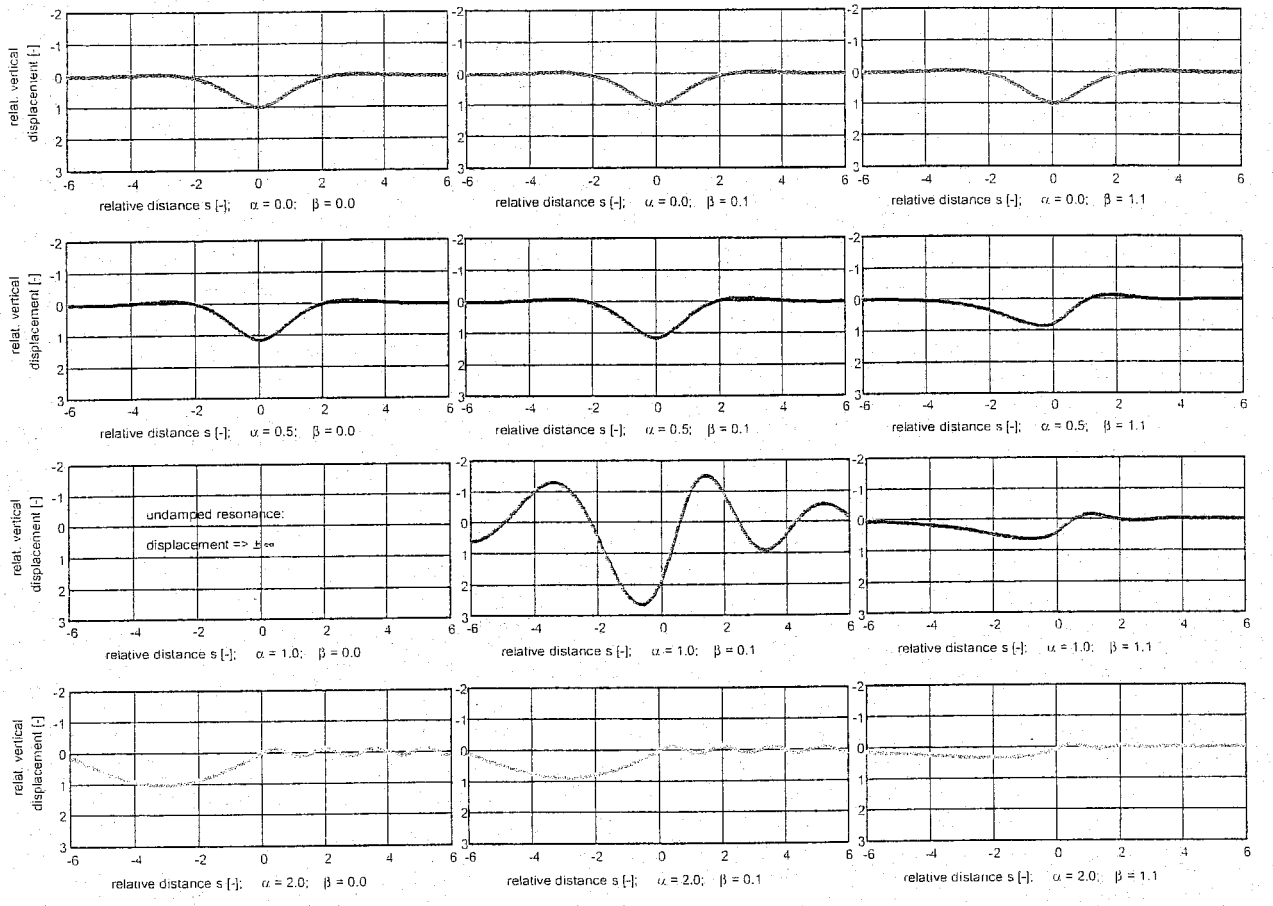


Figure 6.17: Wave shapes versus relative distance

Critical train speed

As can be seen from Figure 6.18, the critical speed or velocity is situated on or near the velocity ratio $\alpha = 1$. According to equation (6.48), it can be derived for the critical speed v_{cr} that:

$$v_{cr}^2 = \frac{2}{m}\sqrt{kEI} \qquad (6.55)$$

in which:
m = rail mass per length;
k = track stiffness;
EI = bending stiffness.

At conventional speeds the influence is negligible as these speeds are much lower than the critical speed v_{cr}. For instance, using the track parameters listed in Table 6.1 , the critical speed amounts to 475 m/s. A speed of 200 km/h thus corresponds with $\alpha = 0.12$. According to Figure 6.18 this would give a very low dynamic amplification and the effect of the load travelling speed can therefore be neglected.

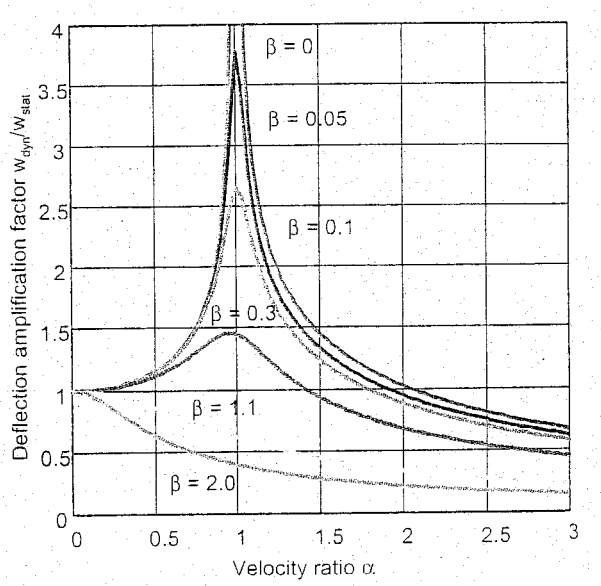


Figure 6.18: Dynamic amplification versus speed due to moving load

For tracks of good quality the critical speed lies far beyond the operating speed, but with poor soil conditions or other mass/spring configurations the critical speed can be so low that special measures are required. In case the train speed approaches the wave propagation speed, the soil may experience a liquefaction type of phenomenon as seen in Figure 6.19. An actual measurement in track on soft soil is shown in Figure 6.20.

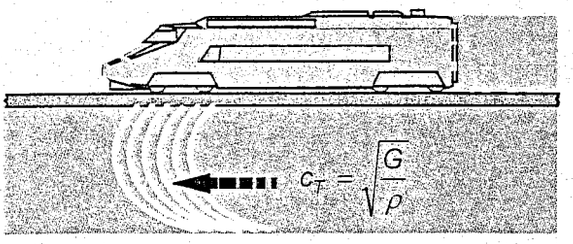


Figure 6.19: Wave propagation at high speed

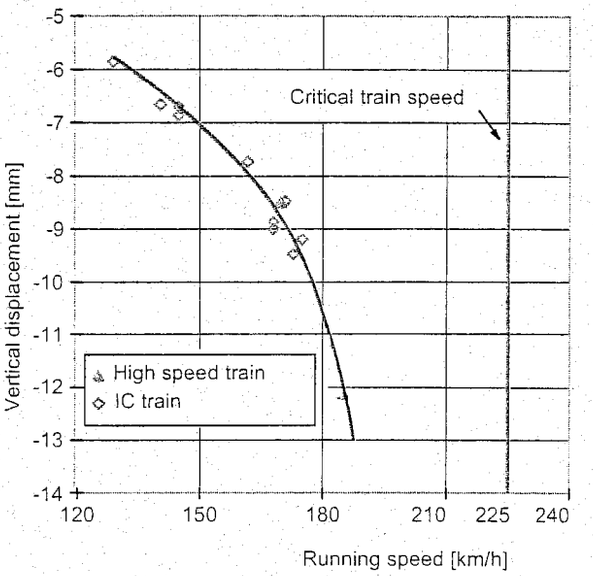


Figure 6.20: Actual measurement on soft soil

For the undamped case (left column of Figure 6.17) a simple formula exists [98] for the dynamic amplification:

$$\frac{w_{dyn}}{w_{stat}} = \frac{1}{\sqrt{\left|1 - \left[\frac{v}{v_{cr}}\right]^2\right|}} \qquad (6.56)$$

6.3.4 Discrete support

The model in Figure 6.10(c), in which the rail is supported in a discrete manner, gives the best approximation. Such an approach also lends itself to the application of standard element programs programs which will be discussed later in Section 6.9. These element method programs give great flexibility as regards load forms and support conditions.

6.4 Vertical wheel response

6.4.1 Hertzian contact spring

During vehicle/track interaction the forces are transmitted by means of the wheel/rail contact area. On account of the geometry of the contact area between the round wheel and the rail, the relationship between force and compression, represented by the Hertzian contact spring, is not linear as has already been discussed in Section 2.7. The relationship between force F and indentation y of the contact surface can be written as:

$$F = c_H y^{3/2} \qquad (6.57)$$

in which c_H [$Nm^{-3/2}$] is a constant depending on the radii and the material properties.

Since a description of the wheel/rail relationship using transfer functions requires that all components are linear, the Hertzian spring must also be linearised. This linearised value of the stiffness can be found by considering the relationship between the force and displacement increments around the static wheel load. The linearised Hertzian spring stiffness k_H is then:

$$k_H = \frac{dF}{dy} = \frac{3}{2} c_H^{2/3} F^{1/3} \tag{6.58}$$

Jenkins e.a. [137] determined the k_H value for old and new wheels as a function of the wheel diameter. For a wheel diameter of 1 m and a static wheel load of 75 kN, a k_H value of 1.4 10^9 N/m is found for new wheels and 1.6 10^9 N/m for old wheels (see also Section 4.11).

6.4.2 Transfer functions between wheel and rail

Figure 6.21 shows the model of a wheel which is connected to the rail by means of a Hertzian spring. From the equilibrium the following is obtained:

$$F_H + M_w \ddot{y}_w = 0 \tag{6.59}$$

with:

$$y_w = \hat{y}_w e^{i2\pi ft} = \hat{y}_w e^{i\omega t} \tag{6.60}$$

The transfer function of the wheel is obtained from (6.59) according to:

$$H_w(f) = \frac{\hat{y}_w}{F_H} = \frac{1}{M_w \omega^2} \tag{6.61}$$

In the following, the relationships between wheel displacement at axle box level and vertical rail geometry, as well as axle box acceleration and vertical rail geometry are examined. These relationships are important when analysing phenomena associated with corrugations and poor quality welds. These transfer functions also formed the basis of the calculations which were carried out when designing the BMS-2 system discussed in Chapter 16.

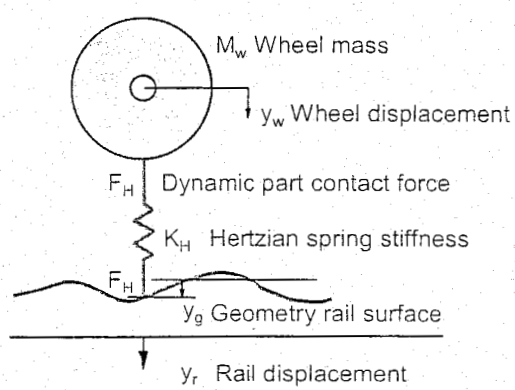


Figure 6.21: Hertzian spring force acting between wheel and rail

The relation between the interaction force F_H and the change in length of the Hertzian spring is determined by:

$$F_H = k_H [y_w - y_r - y_g] \tag{6.62}$$

in which:
y_w = vertical wheel displacement at the level of the axle box;
y_r = vertical rail displacement under the effect of F_H;
y_g = vertical rail geometry;
k_H = linearised stiffness of Hertzian spring;
F_H = dynamic component of wheel/rail force.

If (6.62) is transformed to the frequency domain and the Fourier transformations are indicated in capital letters, the expression can be written as:

$$Y_g(f) = Y_w(f) - Y_r(f) - F_H(f)/k_H \tag{6.63}$$

Using the previously derived transfer functions for the double beam in Section 6.3.3, wheel (6.61), and the rail (6.35), the wheel and rail displacements can be expressed in the wheel/rail force:

$$Y_w(f) = H_w(f)F_H(f) \tag{6.64}$$

$$Y_r(f) = H_r(f)F_H(f) \tag{6.65}$$

After substitution of both in (6.63), this expression becomes:

$$Y_g(f) = [H_w(f) - H_r(f) - 1/k_H]F_H(f) \tag{6.66}$$

$$= H_1(f)F_H(f) \tag{6.67}$$

The relation between wheel displacement $Y_w(f)$ and rail geometry $Y_g(f)$ is now obtained by substitution of (6.64) in (6.67), which results in:

$$Y_g(f) = \frac{H_1}{H_w}Y_w(f) \tag{6.68}$$

Furthermore, by differentiating the wheel displacement twice according to:

$$\ddot{Y}_w(f) = -\omega^2 Y_w(f) \tag{6.69}$$

and substituting this result together with (6.61) in (6.68), the relation between axle box acceleration and rail geometry becomes:

$$Y_g(f) = -M_w H_1(f)\ddot{Y}_w(f) = H_{\ddot{y}}(f)\ddot{Y}_w(f) \tag{6.70}$$

in which:

$$H_{\ddot{y}}(f) = -M_w\left[H_w(f) - H_r(f) - \frac{1}{k_H}\right] \tag{6.71}$$

This transfer function forms the basis of the measuring principle of BMS-2 [268] (see Chapter 16) and is illustrated in Figure 6.22 in which the moduli of the various contributions are plotted, as is the modulus of the resulting transfer function.

The contribution of the rail is calculated using the double beam model based on the appropriate data in Table 6.1 (double beam).

Figure 6.22 shows that the wheel produces by far the greatest contribution in the frequency band up to about 50 Hz. The rail is mainly responsible for the behaviour in the 50 to 1000 Hz band and the Hertzian spring determines the behaviour above 1000 Hz.

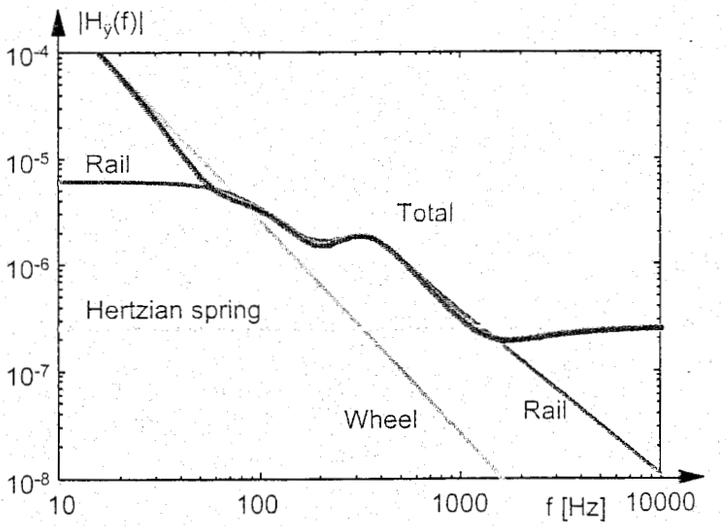


Figure 6.22: Transfer function between rail geometry and axle box acceleration

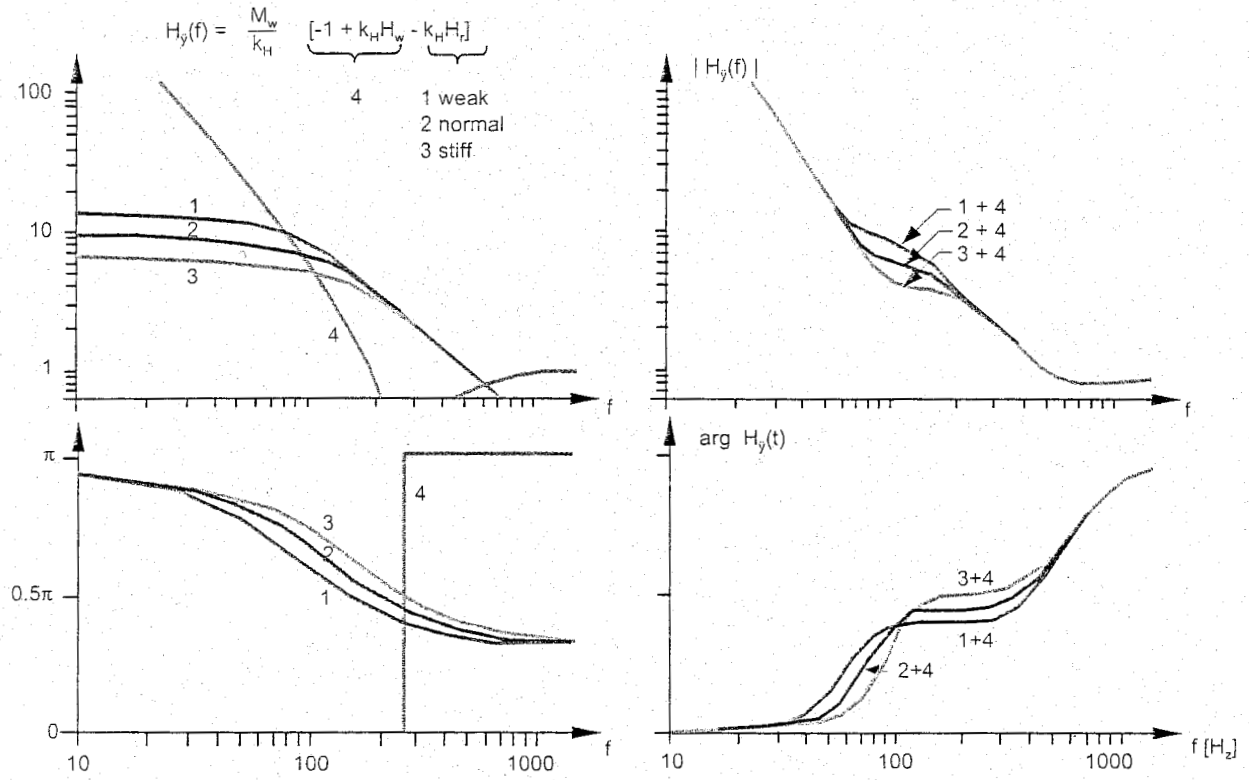


Figure 6.23: Influence of track stiffness on transfer functions $H_{\ddot{y}}(f)$

Since corrugations appear predominantly between 10 and 1500 Hz, it is clear that the track construction in particular has a very great influence. The question is whether, when measuring corrugations by means of axle box accelerations, the variation in track condition can be disregarded. This is examined by varying the track stiffness k_1. Figure 6.23 shows the various contributions made to the transfer function according to formula (6.42) for standard track with a stiffness k_1, for track with a low stiffness of 0.5 k_1, and for track with a high stiffness of 2 k_1. Differences due to the characteristics of the track only show up in the frequency band between 60 and 200 Hz. As a result of system damping due to half-space radiation, for which we refer the reader to [208] and [231], the effect remains limited. No special measures have therefore been taken in the BMS-2 system.

6.5 Linear vehicle model

6.5.1 Schematisation

Transfer functions between track geometry and vehicle reactions can be determined using mathematical models. Figure 6.24 shows a very simple model which describes the main dynamic response of the vehicle. This model can be used to calculate the various relations between track geometry (consisting of cant, level, and alignment) and vehicle response in the form of Q and Y-forces between wheel and rail, but also to calculate horizontal and vertical car body accelerations [269].

In the following the various transfer functions will be derived first. After a discussion of ISO filtering of the vehicle body acceleration signals to take account of human perception, a number of examples are given of transfer functions calculated based on the NS measuring coach.

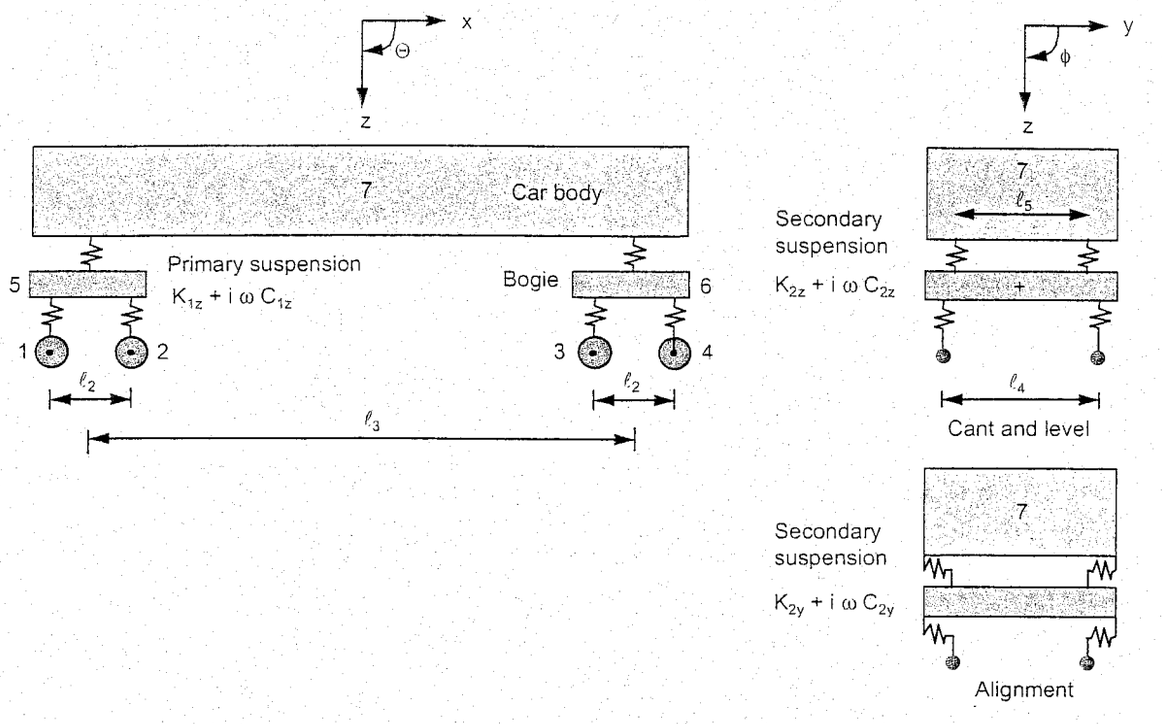


Figure 6.24: Mathematical model for describing vehicle/track interaction

6.5.2 Response to irregularities in level

When examining the response to an irregularity in level the z displacements and θ rotations are of importance. Furthermore, the input will appear four times in succession, i.e. on each of the axles.

The strategy followed here in determining the transfer functions can be summed up as follows:

- determine the response for four symmetrical and asymmetrical wheel displacements;
- combine these so that the response for a single displacement of a wheel is obtained;
- combine these 4 inputs from the separate wheels with their respective time lags.

Since the wheels can be assumed to be infinitely stiff, the inputs act directly on the primary suspension at points 1, 2, 3, and 4 in Figure 6.24. To start with, the mass of the wheelsets is disregarded. Its effect is taken into account later.

Level: case 1

The first mode of vibration is purely a translation characterized by:

$$z_1 = z_2 = z_3 = z_4 = v_1 = \hat{v}_1 e^{i\omega t} \tag{6.72}$$

If the amplitude is normalized to 1, the following amplitude vector is obtained for points 1 to 4:

$$\underline{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \tag{6.73}$$

The equations of motion for this load case can be described in the frequency domain as:

$$-M_5\omega^2 Z_5 = 4F_1 - 2F_2 \tag{6.74}$$

$$-M_7\omega^2 Z_7 = 4F_2 \tag{6.75}$$

$$F_1 = (K_{1z} + i\omega C_{1z})(V_1 - Z_5) \tag{6.76}$$

$$F_2 = (K_{2z} + i\omega C_{2z})(Z_5 - Z_7) \tag{6.77}$$

After substitution of:

$$a = \tfrac{1}{2}M_5\omega^2 \qquad \textit{bogie} \tag{6.78}$$

$$b = \tfrac{1}{2}M_7\omega^2 \qquad \textit{car body} \tag{6.79}$$

$$c = K_{1z} + i\omega C_{1z} \qquad \textit{primary suspension} \tag{6.80}$$

$$d = K_{2z} + i\omega C_{2z} \qquad \textit{secondary suspension} \tag{6.81}$$

the transfer function H^*_{ij} can be calculated between the displacement amplitude v_1 and the response, i.e. the Q-force F_1 and the displacement Z_7. At a later stage the displacement Z_7 is converted to the acceleration $\ddot{Z}_7 = -\omega^2 Z_7$. The transfer function H_{ij} is defined as follows:

$$H^*_{1j} = Q \textit{ force } F_1 \textit{ for } \hat{v}_j = 1 \tag{6.82}$$

$$H^*_{2j} = \textit{car body displacement } Z_7 \textit{ for } \hat{v}_j = 1 \tag{6.83}$$

$$H^*_{3j} = \textit{car body rotation } \Theta_7 \textit{ for } \hat{v}_j = 1 \tag{6.84}$$

$j = 1, 4$, *index of the 4 forced modes of displacement.*

For the first displacement mode the transfer functions are as follows:

$$H^*_{11} = F_1/V_1 = \frac{1}{n_0}\left[-acd + \frac{1}{2}abc - \frac{1}{2}bcd\right] \tag{6.85}$$

$$H^*_{21} = Z_7/V_1 = \frac{1}{n_0}2cd \tag{6.86}$$

$$H^*_{31} = \Theta_7/V_1 = 0 \tag{6.87}$$

where:

$$n_0 = 2cd - bc + \frac{1}{2}ab - ad - \frac{1}{2}bd \tag{6.88}$$

Level: case 2

With this displacement mode both bogies move asymmetrically according to:

$$z_1 = z_2 = -z_3 = -z_4 = v_2 = \hat{v}_2 e^{i\omega t} \tag{6.89}$$

The amplitude vector is:

$$\underline{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \tag{6.90}$$

As a result of the asymmetrical movement the car body only rotates and the z displacement is zero. The equations of motion are:

$$-M_5\omega^2 Z_5 = 4F_1 - 2F_2 \tag{6.91}$$

$$-J_{7y}\omega^2\Theta_7 = 2\ell_3 F_2 \tag{6.92}$$

$$F_1 = (K_{1z} + i\omega C_{1z})(V_2 - Z_5) \tag{6.93}$$

$$F_2 = (K_{2z} + i\omega C_{2z})\left(Z_5 - \frac{1}{2}\ell_3 Z\Theta_7\right) \tag{6.94}$$

With:

$$b_1 = \frac{2}{\ell_3^2} J_{7y}\omega^2 \tag{6.95}$$

the following transfer functions are found:

$$H_{12}^* = F_1/V_2 = \frac{1}{n_1}\left[-acd + \frac{1}{2}ab_1c - \frac{1}{2}b_1cd\right] \tag{6.96}$$

$$H_{22}^* = Z_5/V_2 = \frac{1}{n_1}c(2d - b_1) \tag{6.97}$$

$$H_{32}^* = \Theta_7/V_2 = \frac{4cd}{\ell_3 n_1} \tag{6.98}$$

with:

$$n_1 = 2cd + \frac{1}{2}ab_1 - b_1c - \frac{1}{2}b_1d - ad \tag{6.99}$$

Level: cases 3 and 4

In case 3 the displacement mode and amplitude vector take the following form:

$$z_1 = -z_2 = z_3 = -z_4 = v_3 = \hat{v}_3 e^{i\omega t} \tag{6.100}$$

The amplitude vector is:

$$\underline{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \tag{6.101}$$

For case 4 they are:

$$z_1 = -z_2 = -z_3 = z_4 = v_4 = \hat{v}_4 e^{i\omega t} \tag{6.102}$$

The amplitude vector is:

$$\underline{v}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \tag{6.103}$$

In neither case is the car body loaded. The primary suspensions vibrate independently of each other, symmetrically in case 3 and asymmetrically in case 4. Whenever a wheelset is being considered the following equations of motion apply:

$$-J_{5y}\omega^2\Theta_5 = 2\ell_2 F_1 \tag{6.104}$$

$$F_1 = (K_{1z} + i\omega C_{1z})\left(V_3 - \frac{1}{2}\ell_2\Theta_5\right) \tag{6.105}$$

With:

$$a_1 = \frac{2}{\ell_2^2} J_{5y}\omega^2 \tag{6.106}$$

the following transfer functions are found:

$$H_{13}^* = F_1/V_3 = \frac{-a_1 c}{a_1 - 2c} \tag{6.107}$$

$$H_{14}^* = F_1/V_4 = \frac{-a_1 c}{a_1 - 2c} \tag{6.108}$$

6.5.3 Combination of level results

The response due to the movement of one axle can be obtained by combining the above-mentioned results. For instance, by adding all the results the displacements for axles 1, 3, and 4 equal zero and axle 2 has a displacement of 4. It can clearly be seen that if the k-th axle undergoes a forced movement with an amplitude of 1, the response, i.e. the transfer function H_{ik}^+, can be derived as follows from the above results:

$$H_{ik}^+ = \frac{1}{4}\sum_{j=1}^{4} H_{ij}^* v_{jk} \tag{6.109}$$

In this equation v_{jk} is the k-th element of the amplitude vector v_j. The resulting transfer function between track geometry, in this case level z_t, and the respective response component can now be obtained by adding together the contributions from the 4 axles taking their lag into account. This results in the following expressions:

$$Q = H_1 Z_t \tag{6.110}$$

$$\ddot{Z}_c = (Z_7 - \ell_x\Theta_7)\omega^2 = H_2 Z_t \tag{6.111}$$

Here the transfer functions H_1 and H_2 have the following form:

$$H_1 = -\frac{1}{2}m_1\omega^2 + \sum_{k=1}^{4} H_{1k}^{+} e^{i\omega\tau_k} \tag{6.112}$$

$$H_2 = -\omega^2 \sum_{k=1}^{4} \{H_{2k}^{+} - \ell_x H_{3k}^{+}\} e^{i\omega\tau_k} \tag{6.113}$$

where:

$$\tau_1 = 0 \qquad \textit{reference axle} \tag{6.114}$$

$$\tau_2 = \ell_2 / v \tag{6.115}$$

$$\tau_3 = \ell_3 / v \tag{6.116}$$

$$\tau_4 = \frac{\ell_2 + \ell_3}{v} \tag{6.117}$$

ℓ_x = the distance between the centre of gravity of the vehicle body in the x-direction and the point to which the response relates;
v = running speed.

The first term in (6.112) represents the proportion of the wheelset mass in the Q-force.

6.5.4 Response to irregularities in alignment

The vehicle reactions to an alignment input show a strong resemblance to the reactions to a level input. The degrees of freedom change from z to y and from Q to ϕ. If the ϕ-rotation is disregarded the model used for level fits exactly. This assumption will be used here. The transfer function H_1, according to (6.112), gives the relation between the Y-force between wheel and rail and the alignment, whereas H_2, according to (6.113), represents the relation between horizontal vehicle body acceleration and alignment.

6.5.5 Response to irregularities in cant

In the case of excitation due to cant, only ϕ-rotations occur. As with the response to alignment, where the ϕ-rotations are disregarded, in the case of response to cant lateral spring reactions are not taken into consideration. Calculation of the relationships between cant and vehicle response is similar to that for level input. Here again the same 4 forced displacement modes are selected and applied to the degree of freedom ϕ, from which the effect of each axle can then be derived separately.

Cant: case 1

The mode of displacement is as follows:

$$\phi_1 = \phi_2 = \phi_3 = \phi_4 = v_1 = \hat{v}_1 e^{i\omega t} \tag{6.118}$$

The interaction forces for each axle are equal because of the symmetry. The reaction forces at the left

and right rail are in opposite directions. The equations of motions for this system are:

$$-J_{5x}\omega^2\Phi_5 = 2\ell_4 F_1 - \ell_5 F_2 \tag{6.119}$$

$$-J_{7x}\omega^2\Phi_7 = 2\ell_5 F_2 \tag{6.120}$$

$$F_1 = (K_{1z} + i\omega C_{1z})\left(V_1 - \frac{1}{2}\Phi_5\ell_4\right) \tag{6.121}$$

$$F_2 = (K_{2z} + i\omega C_{2z})(\Phi_5 - \Phi_7)\frac{1}{2}\ell_5 \tag{6.122}$$

With:

$$a = J_{5x}\omega^2 \qquad \textit{bogie} \tag{6.123}$$

$$b = J_{7x}\omega^2 \qquad \textit{car body} \tag{6.124}$$

$$c = (K_{1z} + i\omega C_{1z})\ell_5^2 \qquad \textit{primary suspension} \tag{6.125}$$

$$d = (K_{2z} + i\omega C_{2z})\ell_4^2 \qquad \textit{secondary suspension} \tag{6.126}$$

$$n = 2cd - bc + \frac{1}{2}ab - ad - y\frac{1}{2}bd \tag{6.127}$$

the transfer functions for this load case are:

$$H_{11}^{\bullet} = F_1/V_1 = \frac{1}{\ell_5}\frac{1}{n}\left[-acd + \frac{1}{2}abc - \frac{1}{2}bcd\right] \tag{6.128}$$

$$H_{21}^{\bullet} = \Phi_7/V_1 = \frac{1}{n}2cd \tag{6.129}$$

Cant: case 2

In this case the displacement mode is as follows:

$$\phi_1 = \phi_2 = -\phi_3 = -\phi_4 = v_2 = \hat{v}_2 e^{i\omega t} \tag{6.130}$$

In such a case the vehicle body will not rotate. The solution is obtained from load case 1 by putting $\phi_7 = 0$. The transfer functions for load case 2 are:

$$H_{12}^{\bullet} = F_1/V_2 = \frac{1}{\ell_5}\frac{c(d-a)}{2c+d-a} \tag{6.131}$$

$$H_{22}^{\bullet} = \Phi_7/V_2 = 0 \tag{6.132}$$

Cant: cases 3 and 4

$$\phi_1 = -\phi_2 = \phi_3 = -\phi_4 = v_3 = \hat{v}_3 e^{i\omega t} \tag{6.133}$$

$$\phi_1 = -\phi_2 = -\phi_3 = \phi_4 = v_4 = \hat{v}_4 e^{i\omega t} \tag{6.134}$$

In these cases both car body and bogie are not moving. The transfer functions are now:

$$H_{13}^{*} = \frac{F_1}{V_3} = -\frac{c}{\ell_5} \tag{6.135}$$

$$H_{14}^{*} = \frac{F_1}{V_4} = -\frac{c}{\ell_5} \tag{6.136}$$

6.5.6 Combination of cant results

Excitation due to cant produces two response components, i.e. F_1 and ϕ_7. Two outputs can be derived from ϕ_7 i.e.:

$$\ddot{Y}_b = \omega^2 \ell_z \Phi_7 \tag{6.137}$$

$$\ddot{Z}_b = -\omega^2 \ell_y \Phi_7 \tag{6.138}$$

Here ℓ_z and ℓ_y are the distances between the centre of gravity of the car body and the point to which the response relates in z and y directions.

Combination of the 4 load cases in order to obtain the input for one axle is achieved using the formula (6.109) in which:

$$H_{1j}^{*} = Q \text{ force} \qquad \text{for } \hat{v}_1 = 1 \tag{6.139}$$

$$H_{2j}^{*} = \Phi_7 \qquad \text{for } \hat{v}_1 = 1 \tag{6.140}$$

The response due to cant:

$$\Delta Z = Z_l - Z_r = -1.435\Phi_t \tag{6.141}$$

can now be determined as follows using (6.109):

$$Q = \left\{ J_{1x} \frac{\omega^2}{1.435} + \sum_{k=1}^{4} H_{1k}^{+} e^{i\omega\tau_k} \right\} \phi_t \tag{6.142}$$

The first term inside the brackets is the contribution of the unsprung mass of the wheelset. Bearing in mind (6.141), the transfer function H_1 defined as:

$$Q = H_1 \Delta Z \tag{6.143}$$

follows from the relation:

$$H_1 = -J_{1x}\left[\frac{\omega}{1.435}\right]^2 - \frac{1}{1.435}\sum_{k=1}^{4} H_{1k}^{+} e^{i\omega\tau_k} \tag{6.144}$$

For the car body accelerations according to (6.137) and (6.138) the relationships regarding cant input are:

$$\ddot{Y}_b = H_2 \Delta Z \tag{6.145}$$

where:

$$H_2 = -\frac{1}{1.435}\omega^2 \ell_z \sum_{k=1}^{4} H_{2k}^{+} e^{i\omega\tau_k} \tag{6.146}$$

$$\ddot{Z}_b = H_3 \Delta Z \tag{6.147}$$

in which:

$$H_3 = -\frac{1}{1.435}\omega^2 \ell_y \sum_{k=1}^{4} H_{2k}^{+} e^{i\omega\tau_k} \tag{6.148}$$

6.5.7 ISO weighting of car body accelerations

In order to apply the transfer functions to the calculation of vehicle reactions from measured track geometry, it is necessary to filter the transfer functions. On the one hand, the purpose is to confine unlimited growth in the higher frequencies. On the other hand, the transfer functions for calculating car body accelerations are weighted according to the ISO characteristics which are also incorporated in the ride index meters.

The Q and Y-forces are filtered with a 6th order low-pass filter with the modulus of a Butterworth filter and a zero phase according to:

$$H(f) = \sqrt{\frac{1}{1+\left[\frac{f}{f_c}\right]^{12}}} \tag{6.149}$$

in which the 3 dB point is at f_c. The value $1/f_c$ is set at 3 m. Figure 6.25 shows the modulus of this filter function.

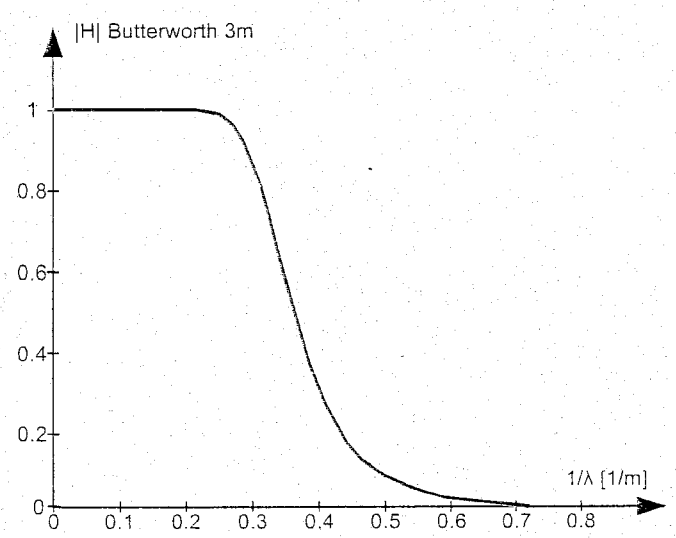


Figure 6.25: Butterworth low-pass filter applied to Q and Y forces

The ISO weighting is implemented using two filters for vertical and horizontal accelerations respectively. These transfer functions are as follows:

vertical:

$$H_v(s) = \frac{0.248\,s\,(0.065\,s+1)\,(0.01457\,s+1)}{(3.5\cdot 10^{-3}\,s^2+0.184\,s+1)\,(5.96\cdot 10^{-4}\,s^2+0.04096\,s+1)} \tag{6.150}$$

horizontal:

$$H_h(s) = \frac{0.6\,s\,(0.0135\,s+1)\,(0.051\,s+1)}{(1.67\cdot 10^{-3}\,s^2+0.151\,s+1)\,(7.19\cdot 10^{-3}\,s^2+0.141\,s+1)} \tag{6.151}$$

Here s is the complex frequency which in this application can be set at:

$$s = i\omega = i2\pi f = i\frac{2\pi}{\lambda} \tag{6.152}$$

where $\lambda = v/f$ = wavelength.

The moduli of the ISO filters are shown in Figure 6.26 for vertical weighting and in Figure 6.27 for horizontal weighting. The functions are calculated for speeds of 90, 120, 140, 160, and 200 km/h.

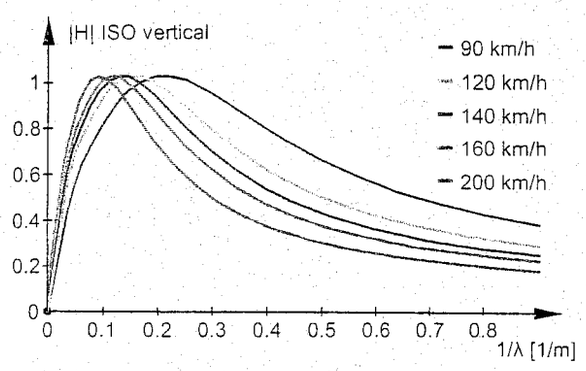


Figure 6.26: ISO characteristic applied to horizontal car body acceleration

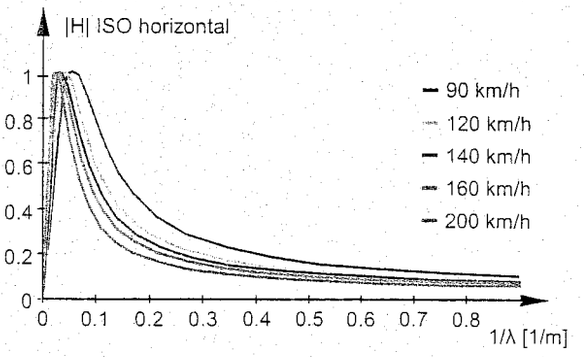


Figure 6.27: ISO characteristic applied to vertical car body acceleration

6.5.8 Calculated transfer functions for the NS measuring coach

The transfer functions derived above were calculated for an NS vehicle, i.e. the measuring coach in which the BMS track recording system is installed.

Table 6.2 summarizes the relevant parameters. The results are given in the form of the modulus of the transfer function and the argument. The unit impulse response function is also calculated.

Figure 6.28 shows the relationship between Q-force and level, Figure 6.29 between vertical car body acceleration and level, Figure 6.30 between Y-force and alignment, Figure 6.31 between lateral car body acceleration and alignment, Figure 6.32 between increase in Q-force and cant, and Figure 6.33 between lateral car body acceleration and cant

Primary suspension		**Secondary suspension**	
K_{1y}	= 4.75 10^6 N/m	K_{2y}	= 0.18 10^6 N/m
K_{1z}	= 0.70 10^6 N/m	K_{2z}	= 0.41 10^6 N/m
C_{1y}	= 3.99 10^4 Ns/m	C_{2y}	= 1.50 10^4 Ns/m
C_{1z}	= 5.88 10^3 Ns/m	C_{2z}	= 2.20 10^3 Ns/m
Unsprung mass			
m	= 1500 kg		
J_{1x}	= 730 kgm²		
Bogie frame		**Car body**	
M_5	= 3.15 10^3 kg	M_7	= 3.37 10^4 kg
J_{5x}	= 2.02 10^3 kgm²	J_{7x}	= 5.24 10^4 kgm²
J_{5y}	= 2.02 10^3 kgm²	J_{7y}	= 7.67 10^5 kgm²
J_{5z}	= 3.56 10^3 kgm²	J_{7z}	= 7.36 10^5 kgm²
Dimensions			
ℓ_2	= 2.56 m	ℓ_x	= -6.95 m
ℓ_3	= 15.66 m	ℓ_y	= 0.00 m
ℓ_4	= 2.78 m	ℓ_z	= 1.45 m
ℓ_5	= 2.00 m		

Table 6.2: Relevant parameters for the NS vehicle containing BMS

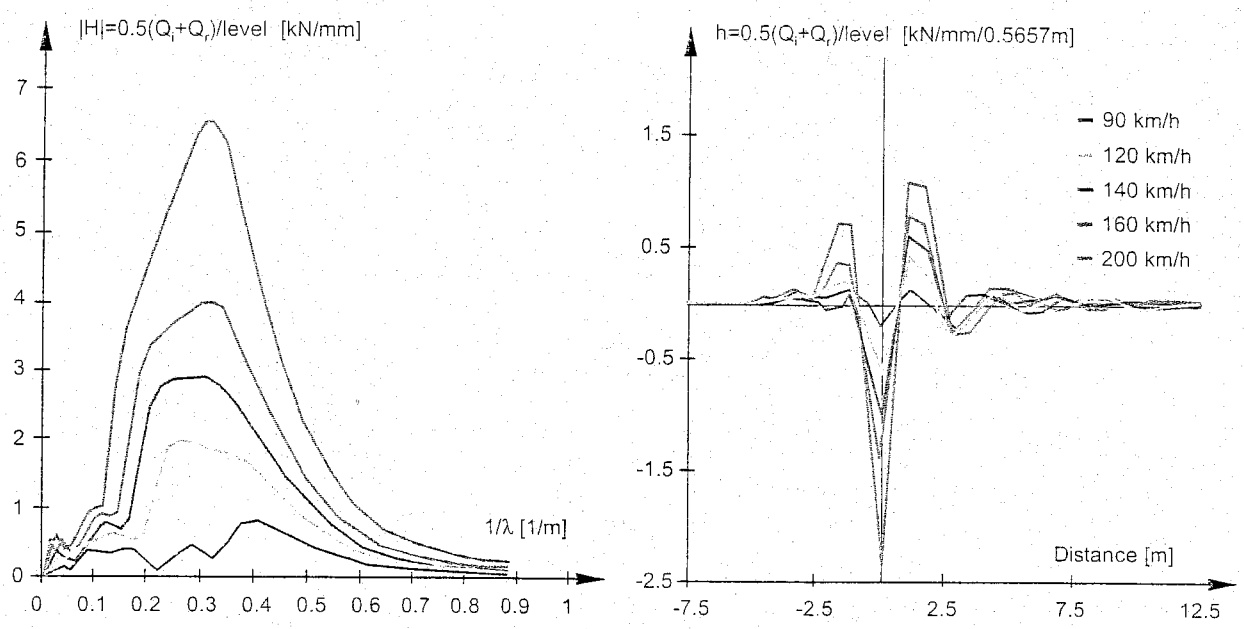


Figure 6.28: Relationships between Q-force and level

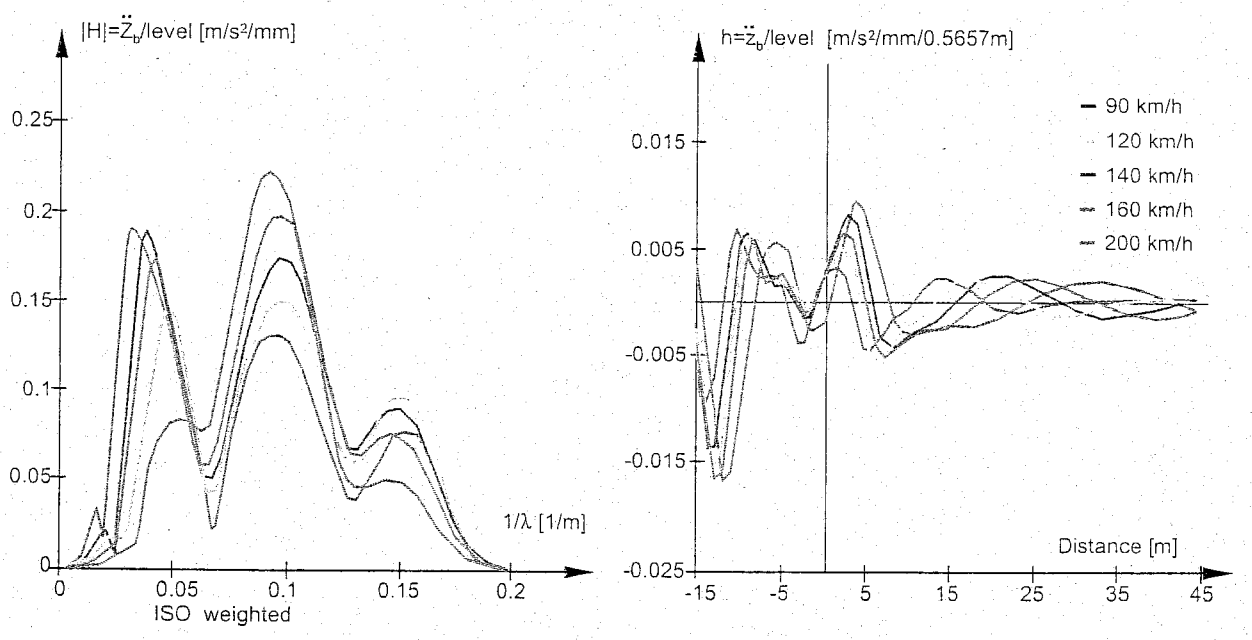


Figure 6.29: Relationships between vertical car body acceleration and level

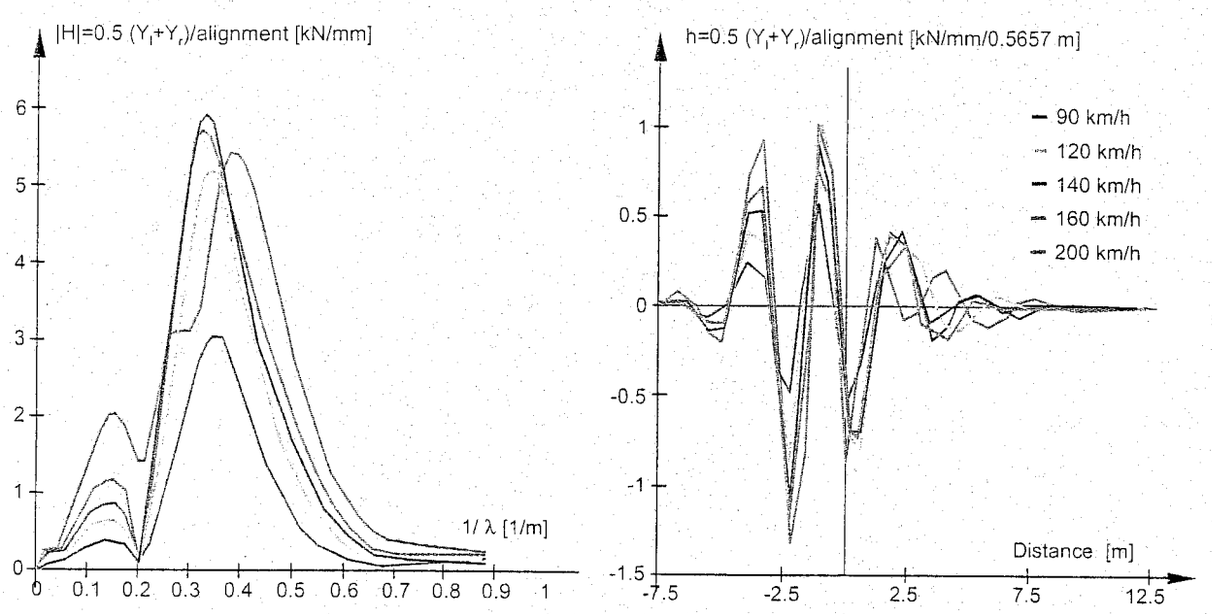


Figure 6.30: Relationships between Y-force and alignment

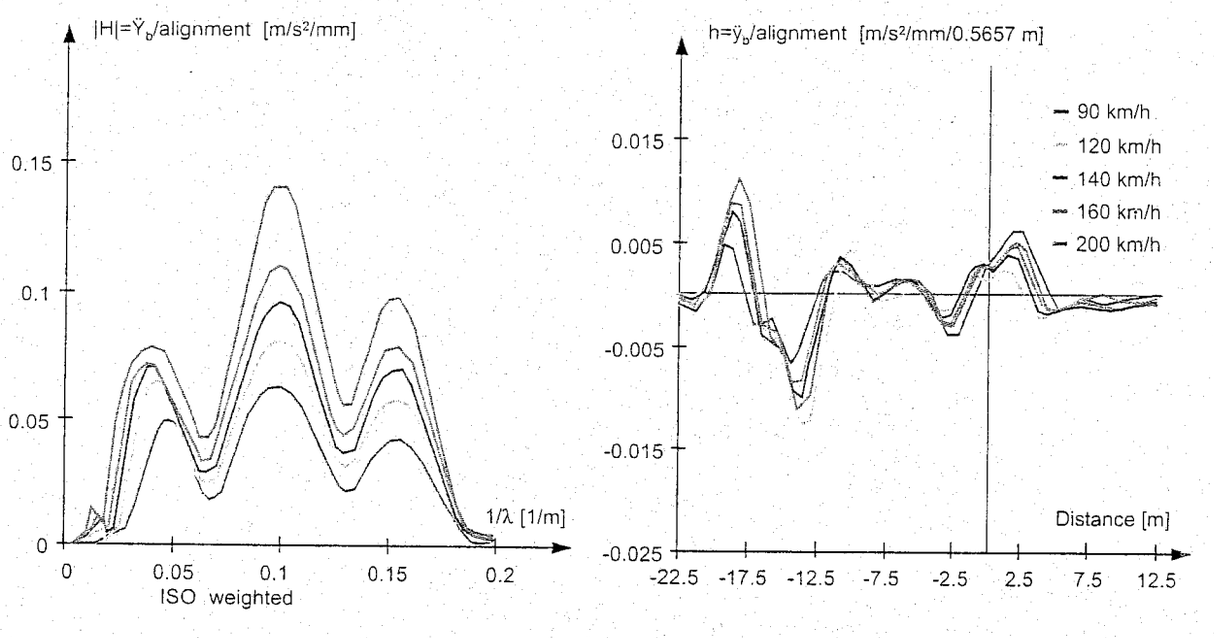


Figure 6.31: Relationships between lateral car body acceleration and alignment

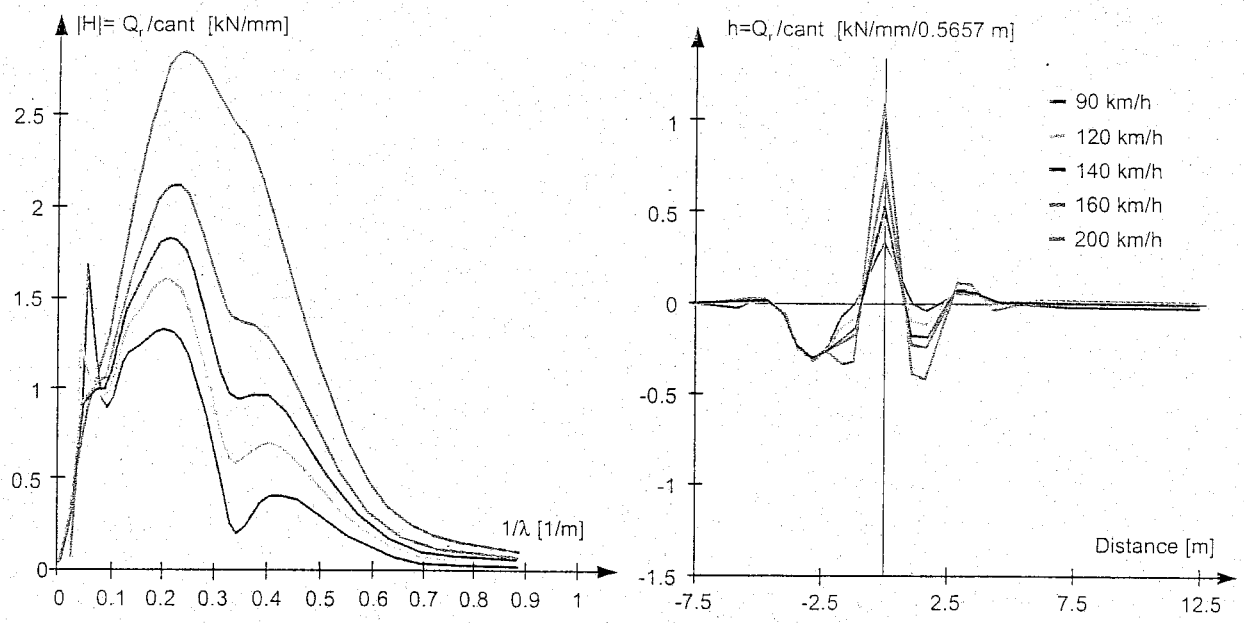


Figure 6.32: Relationships between increase in Q-force and cant

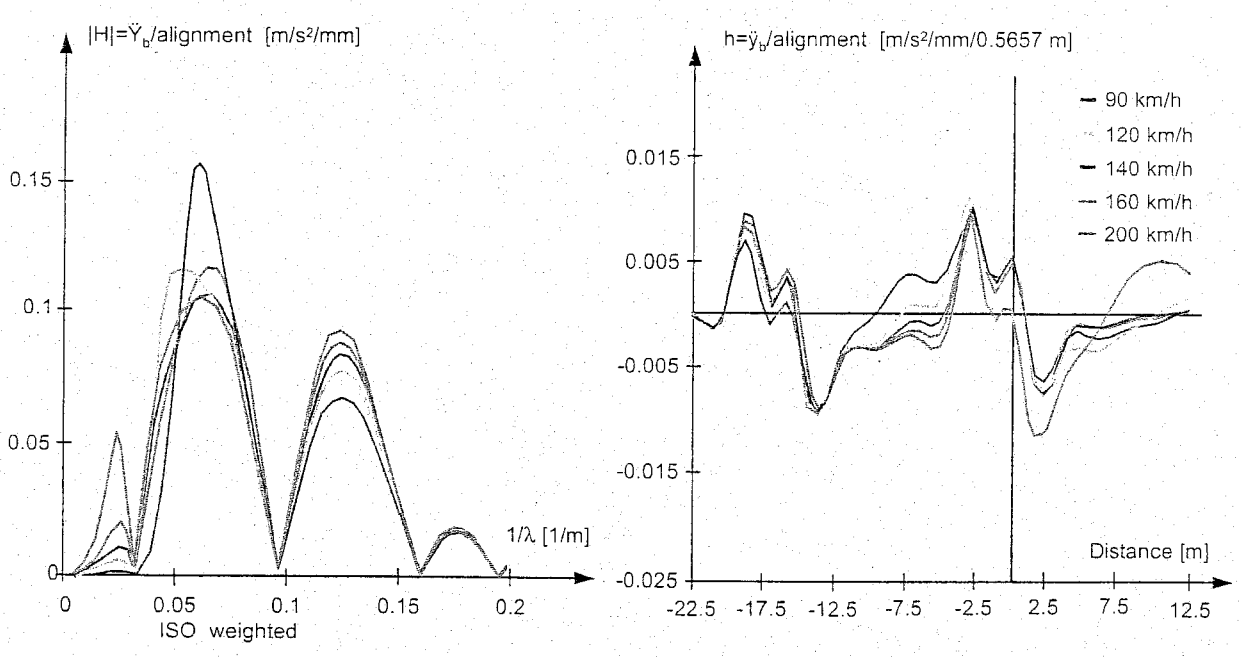


Figure 6.33: Relationships between lateral car body acceleration and cant

6.6 Estimate of transfer functions using measured data

6.6.1 General concept

The relationships between vehicle and track can be described in terms of transfer functions (see Section 6.5), for instance indicating how the various track geometry components contribute to a given response component of the vehicle. In the previous part the mathematical model approach based on a schematisation according to masses, springs, and dampers was discussed.

The method described here uses measured geometry signals as inputs and a corresponding response signals of the vehicle as output to establish transfer functions with the aid of the MISO method (Multiple Input Single Output) based on the theory of random signal analysis. The ORE Committee C 152, set up in 1979, has also dealt with this method and has meanwhile published reports [204], [205], and [74].

A number of examples are discussed concerning the estimate of transfer functions using recorded data, with special emphasis on the reliability aspect. In addition, a new concept for the dynamic measurement of Q and Y forces using measuring wheelsets is discussed.

6.6.2 Basic principles for 1 input and 1 output

Since the primary objective of the theory presented here consists of giving a survey of the main trends of the theory without entering into all sorts of minor details, no derivations will be discussed. As far as details and more basic considerations are concerned, reference is made, in the first instance, to the standard work by Bendat and Piersol [15] and to references [68], [69], [70], [74], and [78] which, in addition to practical implementations, include information on rail applications.

The theory of random signal analysis distinguishes between the time domain for dynamic processes (or the spatial domain for geometrical processes) and the frequency domain. The frequency is composed of the reciprocal time or the reciprocal distance for dynamic or geometrical processes, respectively. Although in the following text the magnitude t is used as time variable, this may be replaced by distance (x or s) too. Likewise, the frequency f may represent both the reciprocal time and the reciprocal distance. As a matter of fact, the variables time and distance are interlinked by the running speed.

If the signal x(t) denotes a magnitude in the time domain, the representation in the frequency domain is obtained by means of the so-called Fourier transformation. Provided that $\int_{-\infty}^{\infty}|x(t)|dt<\infty$ and consequently also $\int_{-\infty}^{\infty}|X(f)|df<\infty$, both transformations from and to the time domain read as follows:

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft}dt \qquad (6.153)$$

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{-i2\pi ft}df \qquad (6.154)$$

If these transformations are made digitally, this is done with the aid of the Fast Fourier Transformation (FFT) which is at present readily available in hardware.

If the 1-input-1-output model shown in Figure 6.34 is composed of a linear physically realisable system, the transfer function H(f) can be explicitly determined on the basis of the system parameters. For a measured input x(t), with the corresponding X(f), an output value Y(f) can be calculated for any f as follows:

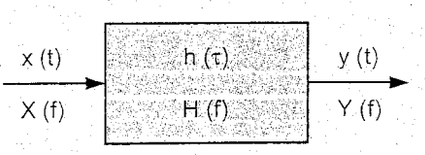


Figure 6.34: One-degree-of-freedom system

$$Y(f) = H(f)X(f) \qquad (6.155)$$

If, however, both input and output are measured, it might be wrongly inferred from (6.155) that the transfer function would follow from the quotient of output and input. However:

$$H(f) \neq \frac{Y(f)}{X(f)} \tag{6.156}$$

Both the real and the imaginary parts of the complex Fourier transformations have, in general, a rather irregular shape. Therefore, it is necessary to use quadratic spectral density functions, which must first be subjected to an averaging procedure, so as to obtain an acceptable statistical degree of reliability. Only after this may an estimate of H(f) be made. This will be discussed later on.

The relationship between input and output is described in the time domain as the convolution product of h and the input x according to (6.157). In the frequency domain this complicated procedure is reduced to a simple multiplication according to (6.155).

$$y(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)d\tau \qquad \textit{convolution} \tag{6.157}$$

$$Y(f) = H(f)X(f) \qquad \textit{multiplication} \tag{6.158}$$

In these expressions $h(\tau)$ represents the unit impulse response and H(f) the transfer function, interrelated as follows:

$$H(f) = \int_{-\infty}^{\infty} h(\tau)e^{-i2\pi f\tau}d\tau \tag{6.159}$$

$$h(\tau) = \int_{-\infty}^{\infty} H(f)e^{-i2\pi f\tau}df \tag{6.160}$$

From the Fourier transformations, spectral density functions may be deduced by multiplying the two moduli with each other and by subsequently dividing them by the record length T. This leads to the complex cross-spectrum $S_{xy}(f)$. If y is replaced by x, a real valued auto-spectrum $S_{xx}(f)$ is obtained. In (6.161), $\overline{X}$ stands for the complex conjugate of X.

$$S_{xy}(f) = \lim_{T\to\infty} \frac{1}{T}\overline{X}(f)Y(f) \tag{6.161}$$

$$R_{xy}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_0^T x(t)y(t+\tau)d\tau \tag{6.162}$$

Here, too, an equivalent operation in the time domain exists and, thus, leads to the cross correlation function $R_{xy}(t)$ shown in (6.162). From the point of view of the calculation technique, this expression is very similar to the convolution process discussed before. By replacing y by x, the autocorrelation function R_{xx} is calculated.

In an absolutely analogue way, as in (6.159) and (6.160), $S_{xy}(f)$ and $R_{xy}(t)$ are interrelated by means of a Fourier transformation. These expressions are known in literature as the Wiener-Kintchine relationships which read as follows:

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau)e^{-i2\pi f\tau}d\tau \tag{6.163}$$

$$R_{xy}(\tau) = \int_{-\infty}^{\infty} S_{xy}(f)e^{-i2\pi f\tau}df \tag{6.164}$$

An important feature of the auto-spectra is that they are symmetrical with respect to the line f=0, which is illustrated in Figure 6.35. Moreover, the area equals the variance according to:

$$\sigma_x^2 = \int_{-\infty}^{\infty} S_{xx}(f)df = 2\int_0^{\infty} S_{xx}(f)df \qquad (6.165)$$

The cross-correlation function is particularly useful in determining the shift between two signals. This displacement corresponds, in fact, to the place where maximum correlation occurs. Figure 6.36 illustrates this approach.

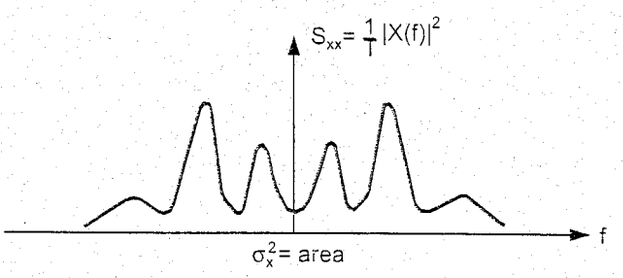


Figure 6.35: Symmetrical auto-spectrum

In quantifying the correlation between two signals, a consideration in the frequency domain is once more applicable, designated as the coherence $\gamma_{xy}^2(f)$, as well as a consideration in the time domain resulting in the correlation function $\rho_{xy}^2(\tau)$.
These expressions read as follows:

$$\gamma_{xy}^2(f) = \frac{|S_{xy}(f)|^2}{S_{xx}(f)S_{yy}(f)} \qquad 0 \le \gamma_{xy}^2(f) \le 1 \qquad (6.166)$$

$$\rho_{xy}^2(\tau) = \frac{R_{xy}^2(\tau)}{R_{xx}(0)R_{yy}(0)} \qquad 0 \le \rho_{xy}^2(\tau) \le 1 \qquad (6.167)$$

In (6.167) $R_{xx}(0) = \sigma_x^2$ and $R_{yy}(0) = \sigma_y^2$. In addition, it should already be observed here that the coherence according to (15.136) only furnishes useful information if the spectra S_{xx}, S_{yy}, and S_{xy} have been averaged according to the rules to be discussed in section 6.6.4. This also applies to the formulae for estimating the transfer function H(f) which will be discussed now.

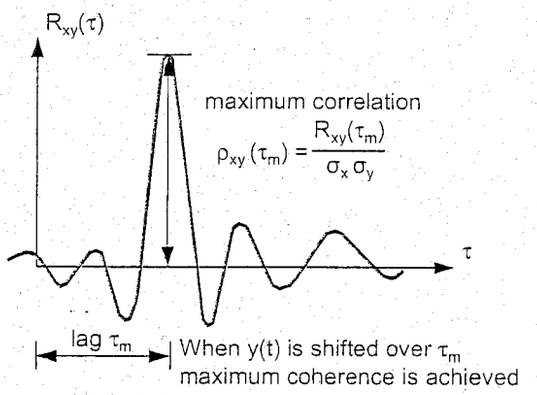


Figure 6.36: Lag to achieve maximum correlation

From the relationship between input and output, the following relation on a spectral level may be deduced:

$$S_{xy}(f) = H(f)S_{xx}(f) \qquad (6.168)$$

From this it follows for the transfer function H(f) that:

$$H(f) = \frac{S_{xy}(f)}{S_{xx}(f)} \qquad (6.169)$$

The coherence function $\gamma_{xy}^2(f)$ is obtained from:

$$\gamma_{xy}^2(f) = \frac{H(f)S_{xy}(f)}{S_{yy}(f)} \qquad (6.170)$$

A simulation study of the influence of non-correlated contributions to the output on the error in estimating the transfer function, published in [205], has shown that about 10% of non-correlated data in the output leads to an error of about 10% in the transfer function estimate, the coherence being reduced to about 0.8. Seen from this angle, whilst also allowing for other possible causes, only estimates of H(f) for which $\gamma^2_{xy}(f) > 0.85$ should be accepted.

6.6.3 Multiple input single output (MISO)

The model depicted in Figure 6.37 shows how the q inputs $x_i(t)$ produce, by means of q linear systems, q outputs $y_i(t)$ which together constitute the overall output according to:

$$y(t) = \sum_{i=1}^{q} y_i(t) \qquad (6.171)$$

The outputs $y_i(t)$ follow from the inputs by means of the convolution products:

$$y_i(t) = \int_{-\infty}^{\infty} h_i(\tau) x(t-\tau) d\tau \qquad (6.172)$$

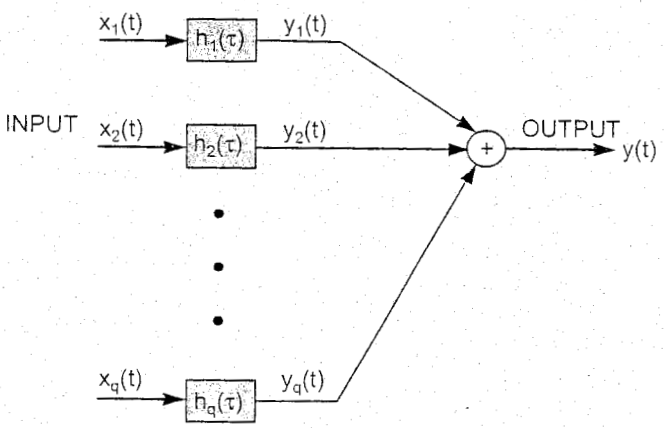


Figure 6.37: MISO model

Assuming the process to be stationary, auto- and cross-correlation functions may be deduced furnishing, by means of Fourier transformation, the following set of equations in the frequency domain:

$$S_{iy}(f) = \sum_{j=1}^{q} S_{ij}(f) H_j(f) \qquad (6.173)$$

in which according to (6.161):

$$S_{iy}(f) = \lim_{T \to \infty} \frac{1}{T} \overline{X}_i(f) Y(f) \qquad (6.174)$$

$$S_{ij}(f) = \lim_{T \to \infty} \frac{1}{T} \overline{X}_i(f) X_j(f) \qquad (6.175)$$

In matrix notation (6.173) reads as follows:

$$\{S_{xy}\} = [S_{xx}]\{H\} \qquad (6.176)$$

The generation and solution of these q complex equations is discussed in section 6.6.5. However, the formal solution can be written provisionally as follows:

$$\{H\} = [S_{xx}]^{-1}\{S_{xy}\} \qquad (6.177)$$

The reliability of the transfer functions thus estimated follows from the multiple coherence function $\gamma^2_{y \bullet x}(f)$ which depicts the ratio between the output spectrum calculated on the basis of (6.177) and the measured output spectrum according to:

$$\gamma^2_{y \bullet x}(f) = \frac{S_{yy\ calculated}}{S_{yy\ measured}} \qquad 0 \le \gamma^2_{y \bullet x}(f) \le 1 \qquad (6.178)$$

After substitution this leads to:

$$\gamma^2_{y\bullet x}(f) = \frac{\sum_{i=1}^{q} \bar{H}_i(f) S_{iy}(f)}{S_{yy}(f)} = \frac{\{\bar{H}\}^T \{S_{xy}\}}{S_{yy}(f)} \tag{6.179}$$

Summation in the numerator indicates the contribution of the various inputs to the output spectrum. In this case, too, it is necessary that for a reliable estimate $\gamma^2_{y\bullet x}(f)$ must lie in the interval

$$0.85 \leq \gamma^2_{y\bullet x}(f) \leq 1 \tag{6.180}$$

6.6.4 Statistical reliability

In this section some attention is paid to random errors ε_r and systematic errors (bias errors) ε_b. Systematic errors can be compensated for by correction or calibration according to:

$$X = \tilde{X}(1 - \varepsilon_b) \qquad -1 \leq \varepsilon_b \leq 1 \tag{6.181}$$

where X denotes the true value and $\tilde{X}$ the estimator. If the systematic error is negligibly small (ε_b < 0.02) and the random error is small (ε_r < 0.10), the 95% confidence interval for X is given approximately by:

$$\tilde{X}(1 - 2\varepsilon_r) \leq X \leq \tilde{X}(1 + 2\varepsilon_r) \tag{6.182}$$

In this section only some main trends are given. For a more detailed discussion see references [15], [16], [17], and [106].

Random errors

As stated before, spectral density functions must be averaged so as to keep the random error at an acceptably low value. This averaging procedure can be carried out in two different ways: either by averaging out the records (ensemble averaging), this number amounts to NSEC, or by combining a number of frequency components known as frequency smoothing, this number amounts to NA. The overall number of averaging operations n is thus:

$$n = NA * NSEC \tag{6.183}$$

The random error ε_r in spectral density functions is inversely proportional to $\sqrt{n}$ and follows from:

$$S_{xx} \rightarrow \varepsilon_r = \frac{1}{\sqrt{n}} \tag{6.184}$$

$$S_{xy} \rightarrow \varepsilon_r = \frac{1}{|\gamma^2_{y\bullet x}|\sqrt{n}} \tag{6.185}$$

The random error ε_r in the modulus of the transfer function H_i, determined by means of MISO, can be approximated according to:

$$|H_i| \rightarrow \varepsilon_{r|H_i|} = \sqrt{\frac{q}{n-q} F_{n_1, n_2, \alpha} \frac{1 - \gamma^2_{y\bullet x}}{1 - \gamma^2_{i\bullet x}} \frac{S_{yy}}{S_{xx}}} \tag{6.186}$$

where:
q = number of inputs;
n = number of averaging operations;
F = 100 α percentage point of the F-distributions, with $n_1 = 2q$ and $n_2 = 2(n-q)$;
$\gamma^2_{i \bullet x}$ = multiple coherence function between input x_i and the other inputs.

From the first term under the square root sign it becomes apparent that the number of averaging operations n must at least be equal to q + 1. In practice, the value of n will certainly have to be one order of magnitude higher. The random error in the argument of H_i follows from:

$$\arg H_i \rightarrow \varepsilon_r = \arcsin \frac{\varepsilon_{r|H_i|}}{|H_i|} \tag{6.187}$$

Bias errors

On the subject of estimating bias errors relatively little is known. For auto-spectra, a Taylor-series expansion for approximating the bias error may be applied. This leads to the following formula:

$$\varepsilon_b \cong \frac{B_e^2}{24} \frac{S''_{xx}}{S_{xx}} \cong -\frac{1}{3} \frac{B_e^2}{B_r^2} \tag{6.188}$$

where:
B_e = resolution band width = Δf = NA/L;
L = record length;
NA = number of frequency smoothing operations;
B_r = "half power point" band width ≈ 0.03 m^{-1} for track geometry spectra.

Summary of random and bias errors

Obviously, the requirements to keep the random error and the bias error small are sometimes incompatible since:

$$\varepsilon_b \therefore \frac{(NA)^2}{L^2} \tag{6.189}$$

$$\varepsilon_r \therefore \frac{1}{\sqrt{NA * NSEC}} \tag{6.190}$$

Averaging over frequencies soon leads to large bias errors whereas long records and a high value of NSEC are attractive, but soon lead to measuring problems. The following maximum values of random and bias errors could be adhered to:

$$\begin{aligned} \varepsilon_r &< 0.10 \\ \varepsilon_b &< 0.025 \end{aligned} \tag{6.191}$$

With $B_r \approx 0.03$ m^{-1}, this leads to the following choice of parameters:

$$\begin{aligned} NA &= 4 \\ NSEC &= 25 \\ L &= 500\ m \end{aligned} \tag{6.192}$$

The distance to be measured in this case is therefore 12.5 km. In [205] a minimum length of 10 km was applied.

$$[\tilde{G}xx] = R + iQ \tag{6.197}$$

can be converted into the following system of 2q real equations:

$$\begin{bmatrix} \underline{a} \\ \underline{b} \end{bmatrix} = \begin{bmatrix} R & -Q \\ Q & R \end{bmatrix} \begin{bmatrix} \underline{c} \\ \underline{d} \end{bmatrix} \tag{6.198}$$

The sub-matrices have the following properties:

$$R^T = R \tag{6.199}$$

$$Q^T = -Q \tag{6.200}$$

$$Q_{ii} = 0 \tag{6.201}$$

The set (6.198) may now be written as:

$$\underline{y} = A\underline{x} \tag{6.202}$$

in which, by virtue of (6.199), (6.200), and (6.201), matrix A is symmetrical. This set of equations can be solved using the decomposition method of Crout-Cholesky briefly described in the following section.

Solving equations

In set (6.202) matrix A describes the inputs while the output only occurs in the $\underline{y}$ vector. In practice, a full series of output signals is usually involved, implying that the (6.202) system would then have to be solved just as many times. In the solution technique of Crout-Cholesky, the arithmetical operations are therefore split up into two parts, i.e. into one part which is independent of the output, thus only bearing on matrix A, and into one part affecting the whole system. The single-time operation on matrix A is designated to factorize and splits up this matrix into two triangular matrices and one diagonal matrix, according to:

$$A = U^T D\, U \tag{6.203}$$

where:
U = upper triangular matrix;
U_{ij} = 0 for i < j; U_{ii} = 1;
D = diagonal matrix; D_{ij} = 0 for i = j.

This is executed in the subroutine FACBAN. The solution proper is now carried out in two steps by means of the subroutine SYMBAN:

$$U^T\underline{z} = \underline{y} \tag{6.204}$$

$$D\, U\, \underline{x} = \underline{z} \tag{6.205}$$

In equation (6.204) vector $\underline{z}$ is solved from top to bottom, after which vector $\underline{x}$, the solution vector of set (6.202), is found by going through (6.205) from bottom to top.

6.6.6 Applications

The MISO applications described here are primarily confined to the field of interaction between vehicle and track. The method is aimed at determining the relationships between track geometry compo-

nents, serving as inputs, and a vehicle response magnitude, representing the output. The model describing the approach by means of MISO is depicted in Figure 6.39. The geometry components cant, level, alignment, and gauge constitute the input whereas so far only car body accelerations have been considered as output.

Within the scope of the ORE C 152 work program a great number of measurements were carried out, during which the track geometry and the vehicle response were recorded on magnetic tape. During evaluation of these measurements the coherence turned out to be much too low in various cases. In most cases this could be ascribed to problems in the signal-to-noise ratio. For example, on some measured line sections the quality was so high that car body accelerations were barely measurable. If, in addition, the signals were not fully amplified before being recorded onto tape, little else but noise remains for analysis.

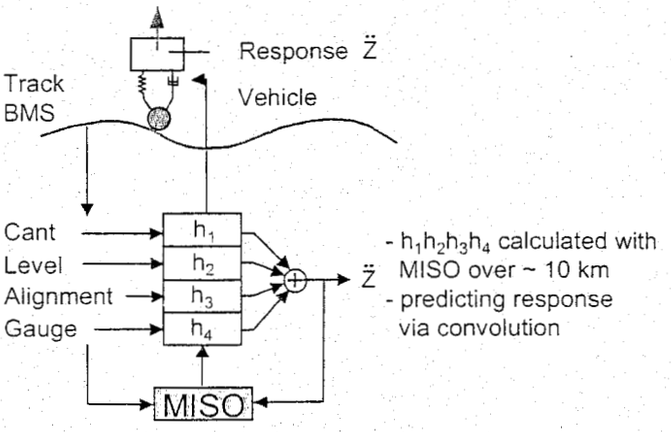


Figure 6.39: MISO model for estimating transfer functions between track geometry and vehicle reactions

Another important factor is the frequency range within which the measured signals fall. If this range is different for input and output, the correlation between the two cannot be expected to be good either. Such problems occur, for example, if low frequency car body accelerations which have a wavelength of 30 - 40 m in the higher speed ranges are compared with track geometry measured with a conventional track recording system capable only of measuring wavelengths up to 20 - 25 m.

Conversely, however, problems will also arise if an attempt is made to relate vehicle reactions of high frequency, such as axle box accelerations and dynamic Q and Y forces, to the track geometry measured in the waveband between 0.5 and 25 m. In this case the geometry will have to be high-pass filtered to remove the long waves with relatively high energy.

Some examples

The NS track recording car

The NS recording car, in which the BMS system is installed, is fitted with Y-32 bogies which have a very linear spring characteristic.

Within the scope of the C 152 studies, a MISO analysis was applied to this recording car with the vertical body acceleration furnished by BMS which is considered to be the vehicle reaction.

The results are given in Figure 6.40 to Figure 6.45.

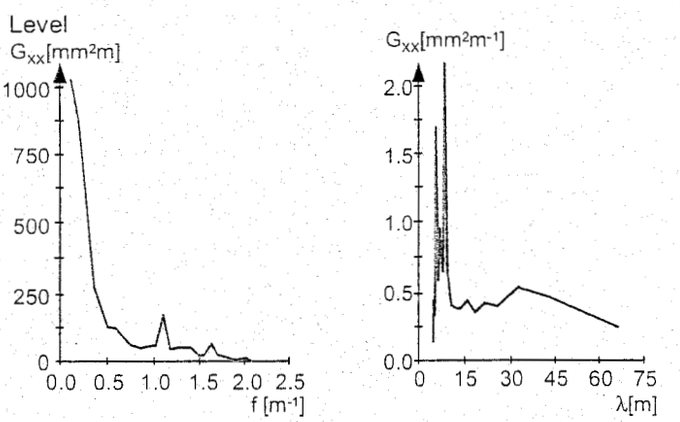


Figure 6.40: Geometry spectra recorded by BMS

Figure 6.40 shows the power spectrum of the track geometry component "level" as a function of the spatial frequency f [m^{-1}] and the wavelength λ [m]. The response spectra at 80 and 120 km/h are shown in Figure 6.41. These spectra were calculated using 20 records of 500 m length and a frequency smoothing factor NA = 4, so that the bias and random errors remain within the limits mentioned in (6.191).

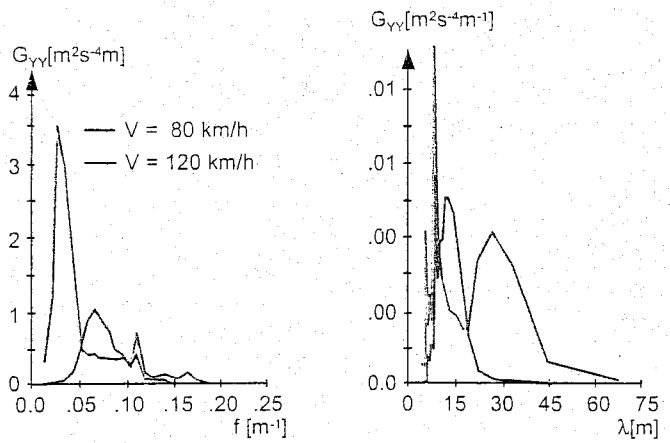


Figure 6.41: Vertical car body acceleration spectra measured on the NS recording car

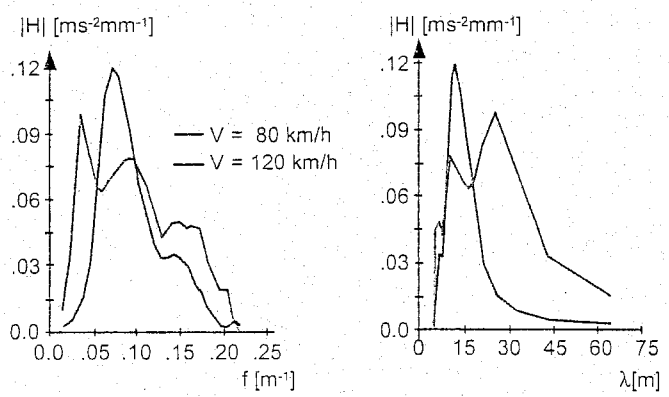


Figure 6.42: Transfer functions between level and vertical car body acceleration estimated for the NS recording car

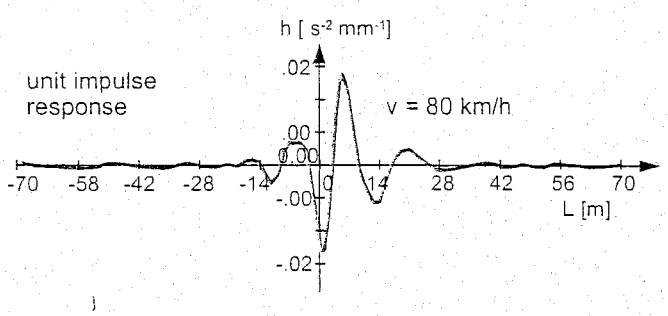


Figure 6.43: Unit impulse response function based on the data in Figure 6.42

As an example Figure 6.42 shows the transfer functions H_2 between level and vertical car body response; the other transfer functions are negligibly small.

These are, in fact, the primary results of the MISO analysis according to (6.177). In agreement with formula (6.160), these functions have been Fourier transformed so as to obtain the unit impulse response functions h_i of which h_2 is illustrated in Figure 6.43. In this case, too, all other h-functions can be neglected, implying that only the level contributes to the vertical acceleration.

The degree to which the transfer function values are reliable is shown by the multiple coherence function $\gamma^2_{y\bullet x}(f)$ depicted in Figure 6.44. As stated before, the $\gamma^2_{y\bullet x}(f)$ value should be higher than 0.85 if practical applications are to be made possible. On further analysis of the results, it appears that the $\gamma^2_{y\bullet x}(f)$ value only meets this requirement in the frequency bands where the measured signals contain enough energy. This is rather obvious and it also explains, perhaps in a different way, why long measuring sections should be chosen, preferably with maximum possible variation in the geometry spectra.

To complete the sequence of computations, the response is once more calculated as a function of the distance covered by means of the convolution principle according to (6.157), using the unit impulse response functions previously obtained, and is compared with the response signal originally measured. Figure 6.45 contains a graphical representation of the calculated and measured signals; the similarity between the measured and the calculated response is remarkable.

To quantify the deviation between the two signals, the RDS value (relative difference between standard deviations) is determined for each 200 m sub-section. The RSD value is defined as:

$$RDS = \frac{\sigma_{measured} - \sigma_{calculated}}{\sigma_{measured}} \qquad (6.206)$$

This value roughly conforms to the value $1-|\gamma^2_{y\bullet x}(f)|$, with respect to which the mean value of $|\gamma^2_{y\bullet x}(f)|$ must be imagined over the area in which the energy in the signal is concentrated. In fact, RSD denotes the error in standard deviation if the latter has been calculated using the transfer functions obtained from MISO.

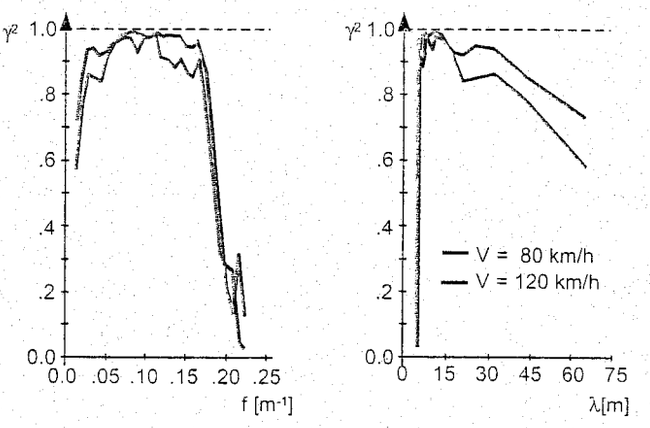


Figure 6.44: Multiple coherence for the estimates in Figure 6.42

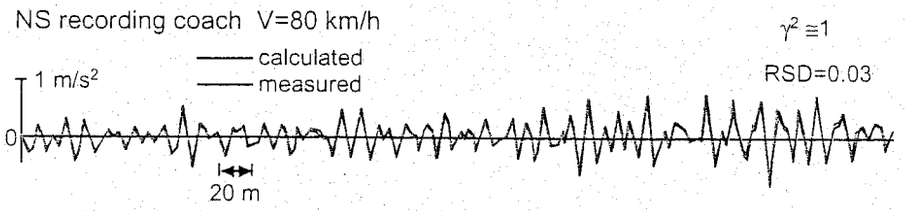


Figure 6.45: Comparison between measured and calculated response based on the estimated transfer function in Figure 6.42

Results of other measurements

As part of the C 152 study mentioned above, various other measurements were analysed using MISO. Some characteristic results will now be highlighted.

The data contained in Figure 6.46 and Figure 6.47 refer to SNCF measurements of vertical car body accelerations of a CORAIL coach. The relationship between these measurements and the track geometry recordings made by the Mauzin car have been established. Figure 6.48 shows the coherence to be very poor with a maximum value of 0.6.

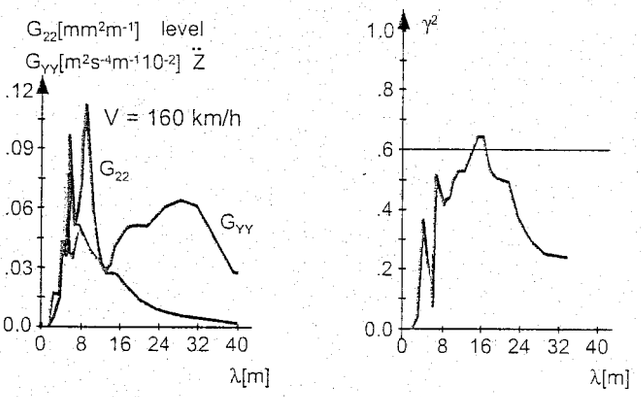


Figure 6.46: Example of bad coherence for transfer estimate regarding SNCF Corail coach

The factors mentioned earlier concerning compatibility in frequency range and proper signal amplification are, in all probability, responsible for this. The calculated and measured responses according to Figure 6.49 show, in a different way, that in such a case the results are useless for practical application.

The last example concerns measurements on two-axle wagons of KS design taken by the Hungarian State Railways (MAV). Despite the rather non-linear spring characteristic of this type of rolling stock, the coherence turns out to be fairly high. Figure 6.50 shows the coherence vertically with a peak value of 0.95. The response calculated according to Figure 6.49 consequently conforms very well to the measured vehicle response. Considered laterally, the coherence shown in Figure 6.50 turns out to be somewhat lower with a peak value of about 0.8. As can be seen from Figure 6.51, the agreement between the measured and calculated response figures is nevertheless quite reasonable.

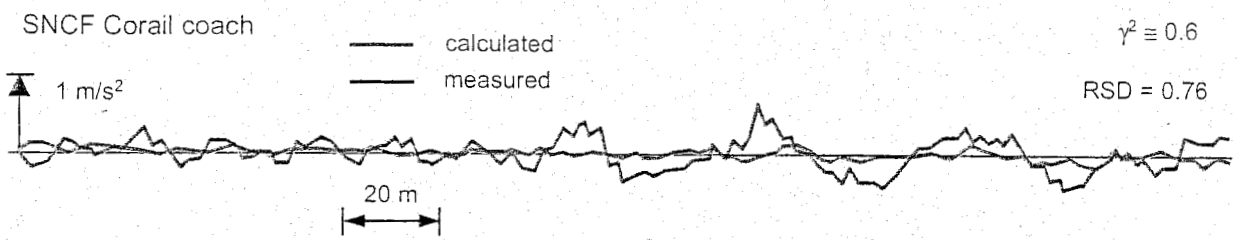


Figure 6.47: Comparison between measured and calculated response based on the data in Figure 6.46

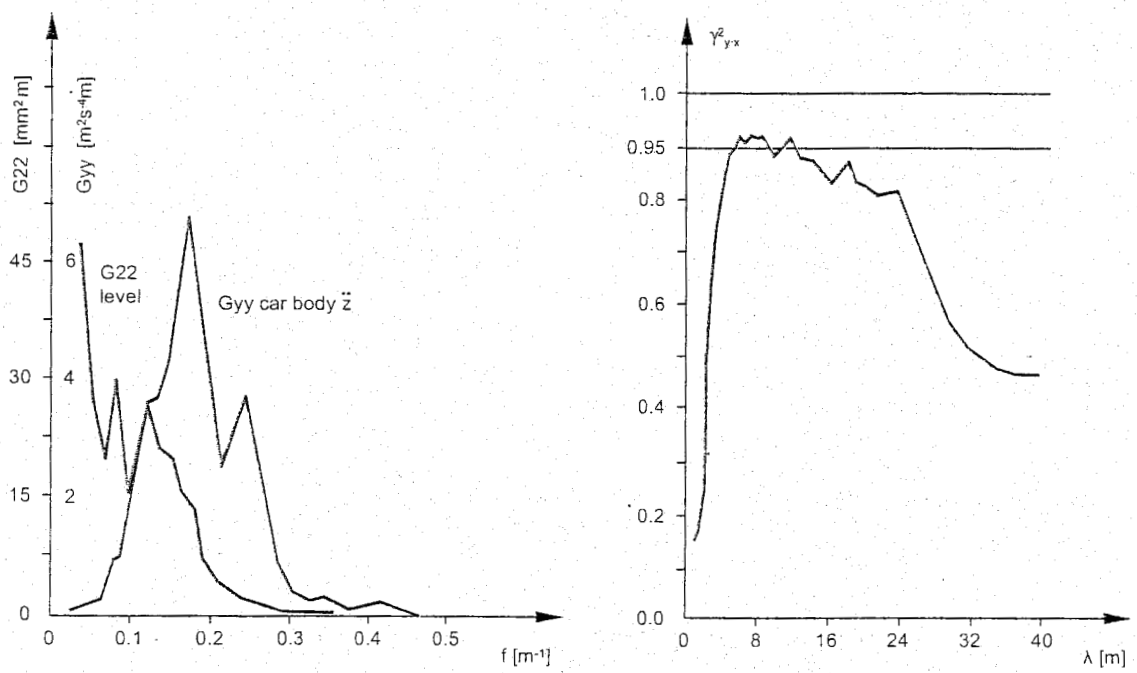


Figure 6.48: Example of vertical response for MAV freight car

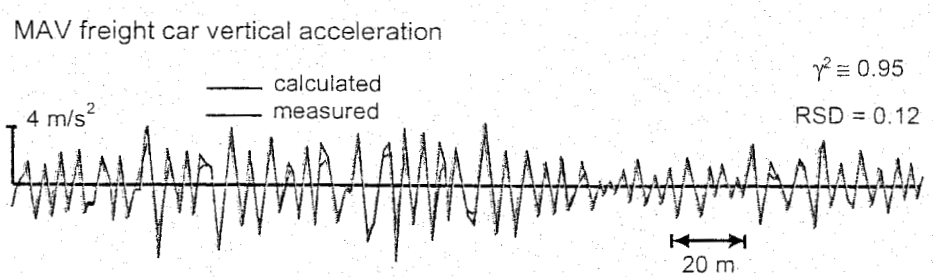


Figure 6.49: Comparison between measured and calculated response based on the data in Figure 6.48

In 1986 a short BMS measuring car campaign was carried out in Sweden to provide SJ with geometry data for vehicle model input. On this occasion SJ also recorded Q and Y forces on a locomotive with the aid of measuring wheelsets. With these data and the geometry recorded by BMS some MISO analyses were carried out. These results are summarized in Figure 6.52. In the vertical direction the multiple coherence is sufficiently high for wavelengths shorter than 8 meter. The transfer function increases progressively with frequency. This reflects the strong influence of the unsprung mass which contributes to the transfer function in proportion to the square of the frequency. In the lateral direction the coherence is somewhat lower. The shape of this estimate has a peak at about 16 m wavelength, remains more or less constant between 13 and 14.5 m and then increases progressively for shorter waves. Although the theoretical transfer functions presented in Figure 6.28 and Figure 6.30 refer to much lower unsprung masses, the shapes resemble the estimated transfer functions quite well.

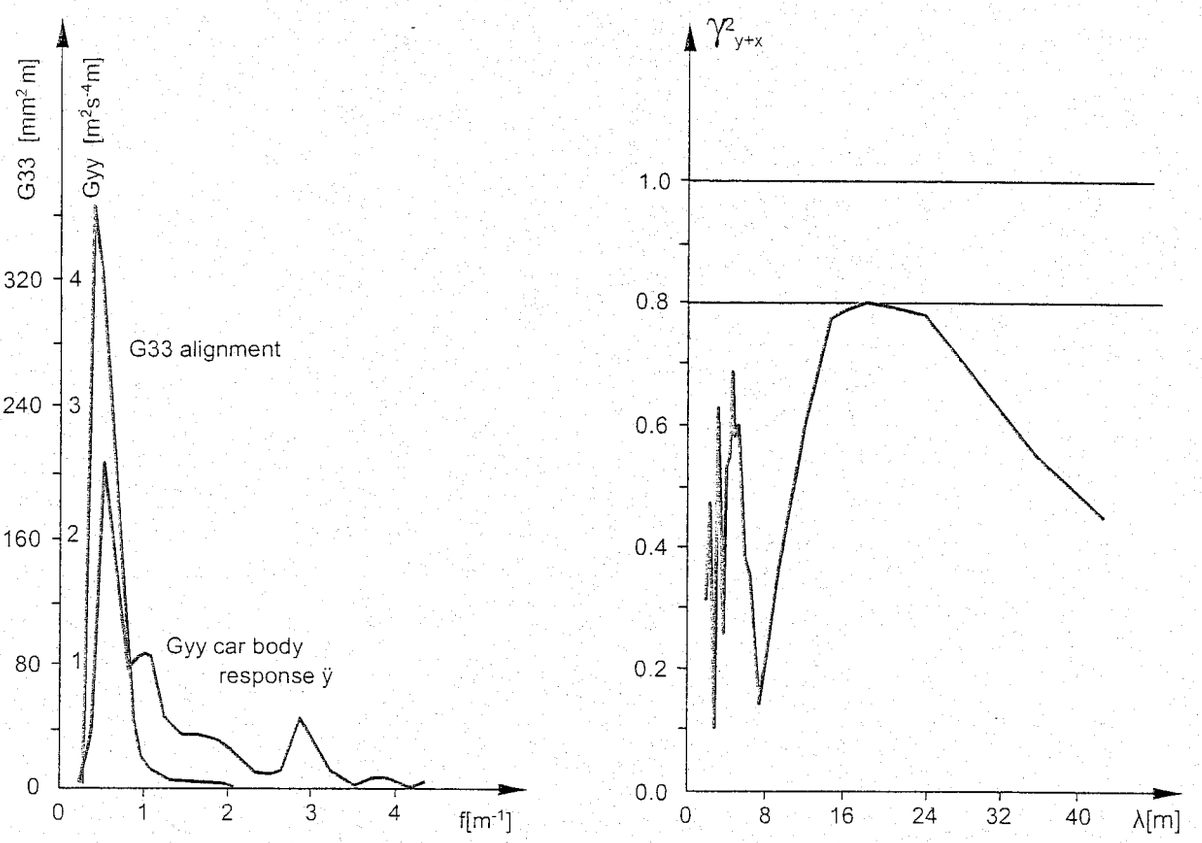


Figure 6.50: Example of lateral response for MAV freight car

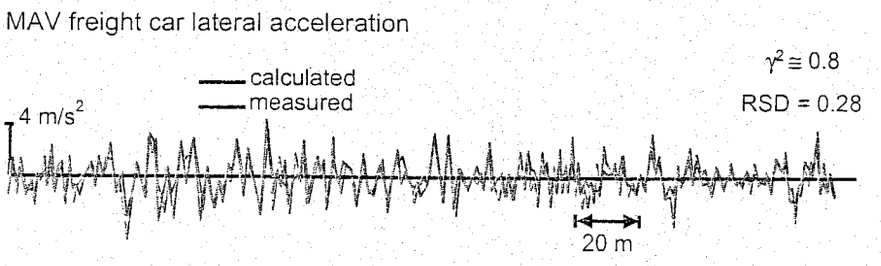


Figure 6.51: Comparison between measured and calculated response based on the data in Figure 6.50

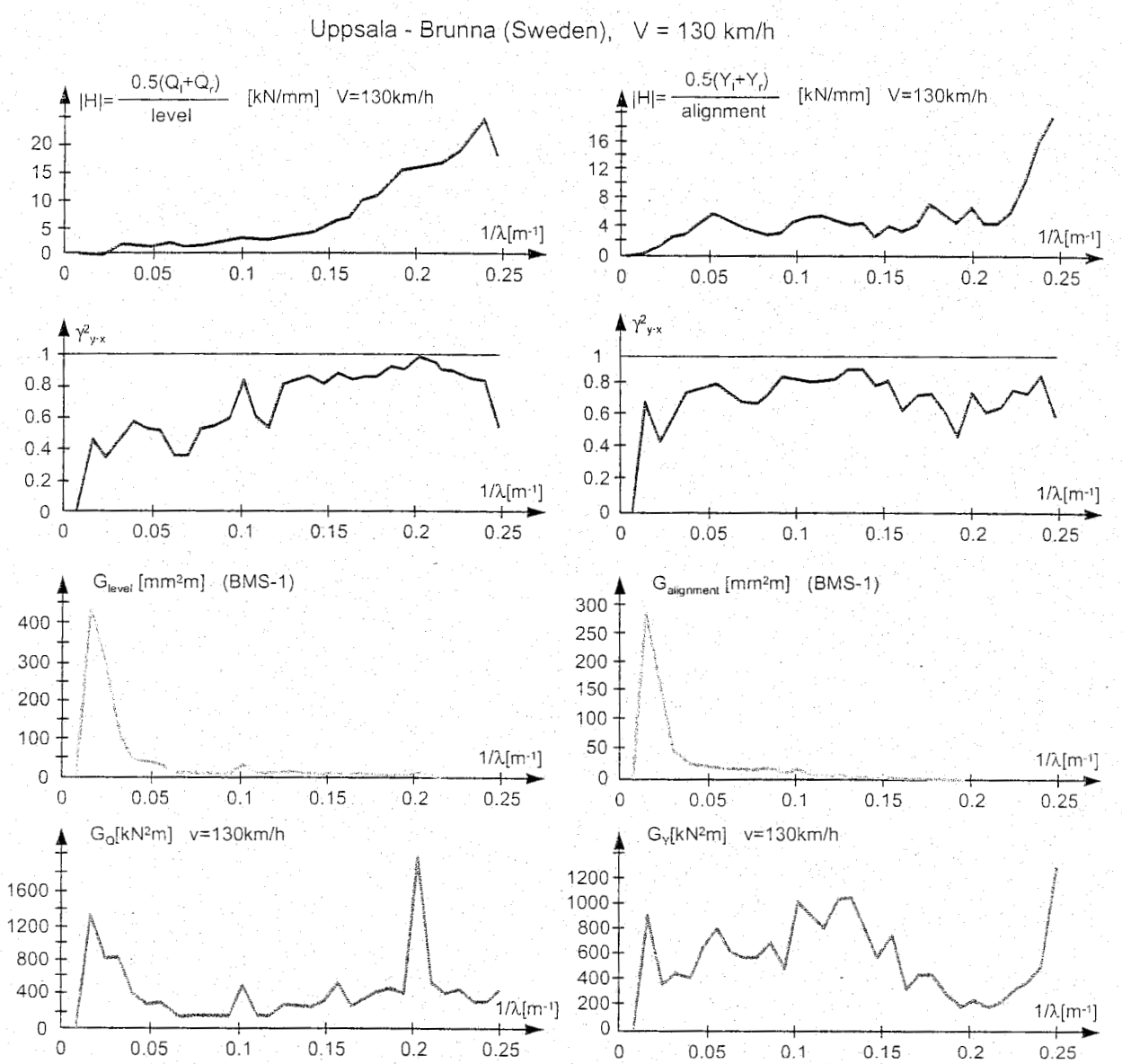


Figure 6.52: Estimated transfer function for the Q and Y forces of an SJ locomotive

Measuring wheelsets in order to determine Q and Y forces

When measuring Q and Y forces with the aid of measuring wheelsets, use is made of strain gauges from which the respective forces are deduced. The required relationships between strains and forces are determined experimentally on a roller test rig. Here, the variation in strain versus distance covered is calculated for one force component which is then kept constant. This is in fact a static calibration which restricts Q and Y force measurements to quasi-static phenomena [228]. Such a system gives incorrect values for real dynamic phenomena at frequencies between 2 and 150 Hz. As a matter of fact, strain gauges do furnish a signal in this frequency band, but the transfer functions between strains and forces are unknown. These transfer functions could be estimated using MISO during calibration on the roller test rig.

Static calibration requires that the number of strain measuring points are equal to the number of forces to be measured, since the experimental test arrangement only enables the relation between force and strain to be determined; the relationship sought in the reverse direction then follows from this by matrix inversion. The dynamic relationships in the frequency domain offer a similar picture, though all the magnitudes are now frequency-dependent.

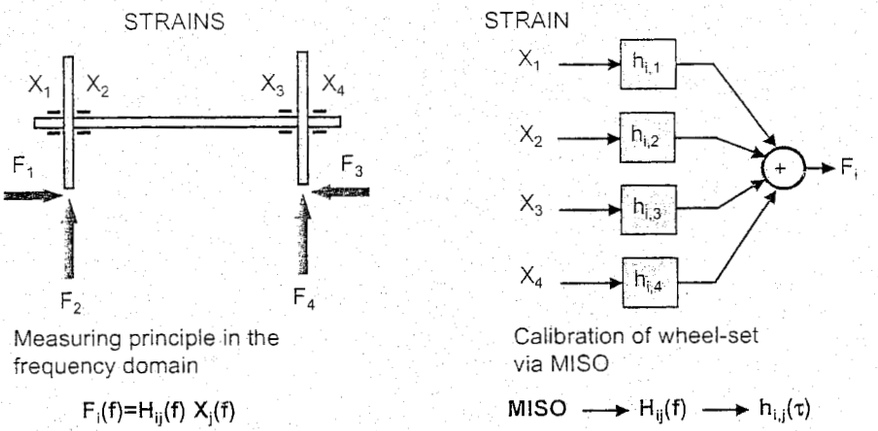


Figure 6.53: Principle of force measuring by means of a wheelset

Figure 6.53 shows the arrangement of the strain gauges and the forces acting on the measuring wheelset. The conventional quasi-static approach remains invariably valid and must be considered separately from the dynamic calibration.

In order to estimate the transfer functions between the strain gauges and each force component, the forces exerted by the test rig on the wheels must be varied randomly in the frequency range of application. If sampling is carried out as a function of distance or revolution, these realisations should be repeated at different running speeds. Using these measurement series, the different transfer functions can be estimated for each speed by means of MISO by considering the strains as input and each force as output, as indicated in Figure 6.53.

During actual measurements the Q and Y forces can be derived from the strains in real time by applying the estimated transfer functions. These should be continuously adapted to the actual running speed by interpolation.

6.6.7 Comparison between transfer functions estimated by MISO and calculated with models

In section 6.5 a simple vehicle model was discussed which was used to derive transfer functions, amongst other things, between level and vertical car body acceleration. These results, presented in Figure 6.29, have been compared with the estimated transfer functions using MISO, depicted in Figure 6.42. Both results for the same speed and running direction are displayed in Figure 6.54. Clearly there is quite good agreement between both functions.

The practical advantage of the model approach is that additional estimates for different speeds can easily be made, whereas in the case of MISO a test run should be carried out for each speed. MISO has of course the big advantage that all important natural frequencies are automatically taken into account.

In practice, a hybrid solution may be feasible in which a MISO estimate is made for one or two speeds. After adapting the model parameters in such a way that the MISO estimates are approximated as well as possible, the transfer values at other speeds can be obtained by means of calculation.

6.7 Vehicle response analysis in real time

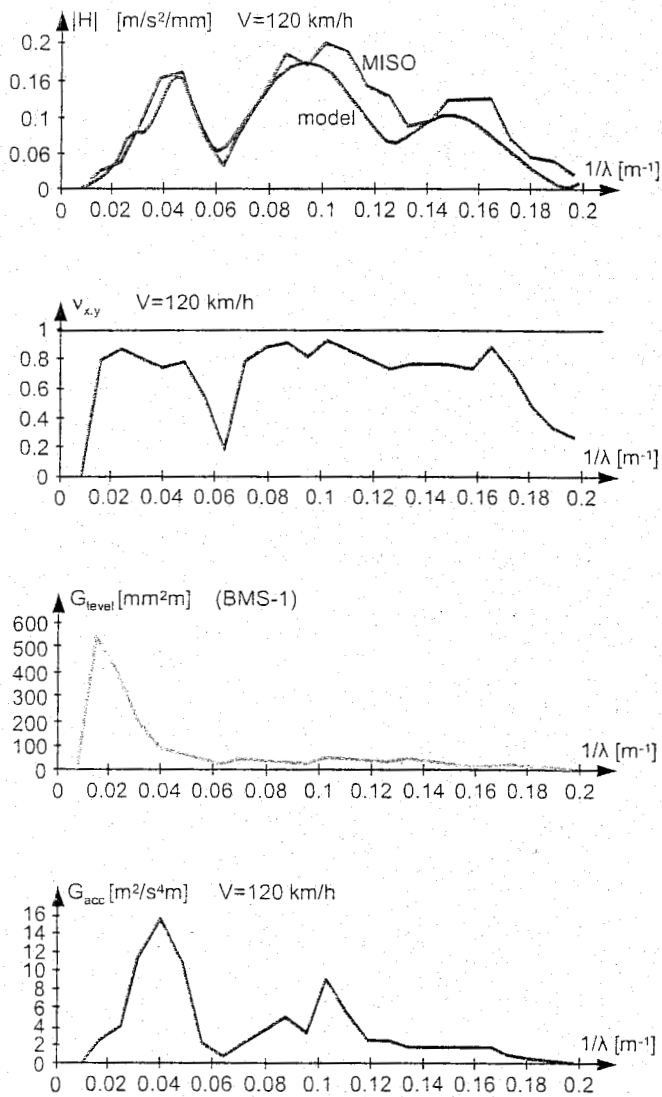


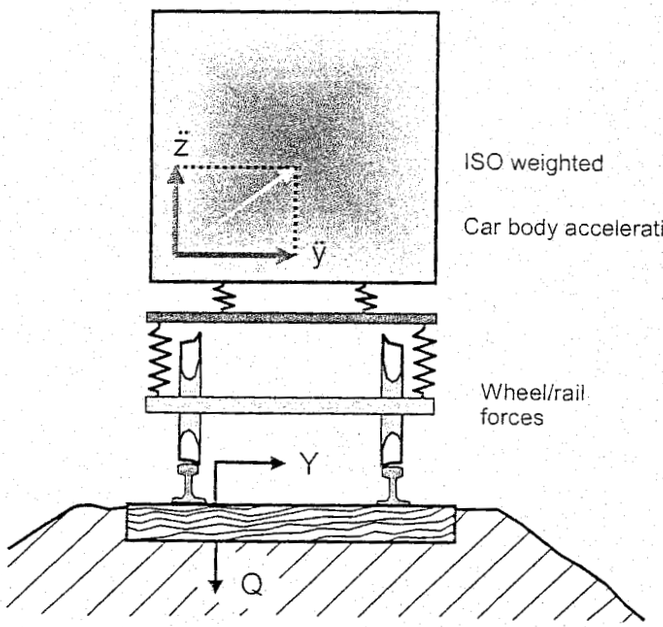
Figure 6.54: Transfer function between level and car body acceleration for NS recording car estimated by means of MISO and calculated with models

In 1988 the computers of the NS track recording system BMS were replaced by one powerful computer system. Progress in both information technology and railway vehicle dynamics, as described earlier in this chapter, allowed the development of a Vehicle Response Analysis (VRA) system using a state-of-the-art Array Processor in combination with a central minicomputer system.

The central minicomputer is a Digital PDP 11/53 minicomputer connected to an Analogue Devices AP 500 Array Processor. The software for the PDP 11/53 was developed under Digital's RT11 real time operating system. The combination of a modern PDP 11/53 under RT11 with the AP 500 Array Processor gives a state-of-the-art real time signal-processing system.

The sampling interval of the 5 analogue BMS-1 signals is 0.25 m. At the maximum speed of the measuring vehicle of 45 m/s (= 162 km/h) the sampling rate is 180 Hz. For BMS-2 the 8 channels are sampled at a constant rate of 250 Hz. While the on-line sampling is performed continuously for both systems, all 200 m block calculations have to be done by the system within 4.4 seconds.

From the BMS-1 signals cant, level, alignment, and gauge in the 0 - 25 m waveband and alignment in the 0 - 70 m waveband VRA calculates the vehicle reactions indicated in Figure 6.55. For 3 types of rolling stock and 5 speeds the horizontal and vertical car body acceleration, weighted according to ISO 2631, and the vertical and horizontal track load are computed.



Calculation of vehicle reactions in real time for:
- 3 vehicles: locomotive
  passenger coach
  freight car
- 5 speeds: 40-160 m/h

Calculations via array processor:
- Fourier transform of 4 geometry signals
- Computation of vehicle reactions via 240 transfer functions
- Determination of largest variance per speed range: these decisive 12 values are printed
- Back transformation of selected 12 responses
- Vectorial summation of car body accelerations
- Recording resulting in 9 response signals
- Determination of exceedences

Figure 6.55: Vehicle reactions calculated by VRA

The calculation flow in VRA is shown in Figure 6.56. The track geometry data of a 600 m block, represented by a 1024 point complex array, is transformed to the frequency domain using the FFT (= Fast Fourier Transformation) capabilities of the Array Processor. All communication between the computer and the Array Processor is performed using a high-speed DMA (= Direct Memory Access) channel.

The 600 m record is composed of the current 200 m section preceded by a section of 350 m and followed by a section of 50 m. The section of 350 m should cover the attenuation of the vehicle, i.e. the length of the unit impulse response. As far as BMS-1 is concerned, the longest wavelength in the analysis is 70 m and consequently the preceding section should be a number of times this wavelength. The section of 50 m takes the part of the vehicle ahead of the measuring point into account.

At the systems power-up, the transfer functions of three representative railway vehicles at five different speeds are loaded into the Array Processor. The frequency domain responses are calculated by multiplying the transformed signals with the corresponding transfer functions.

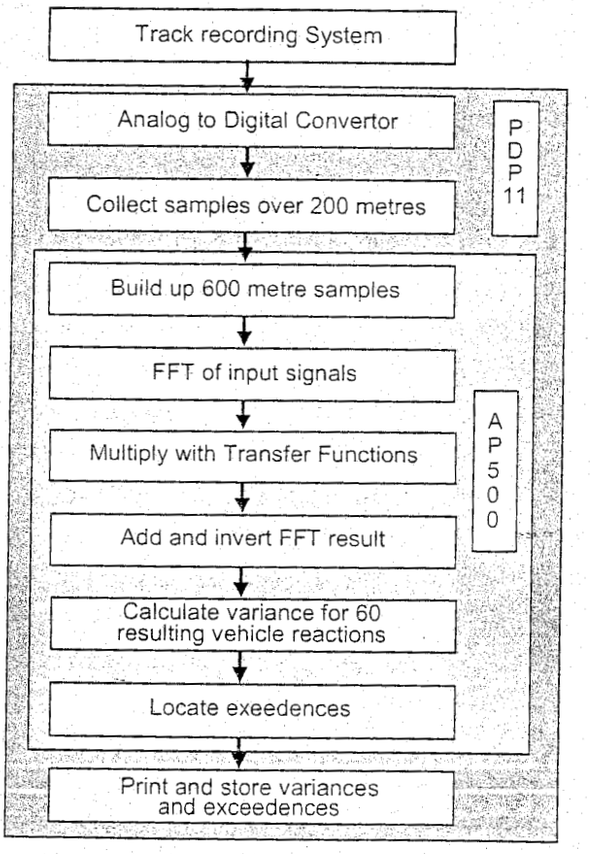


Figure 6.56: Calculation flow VRA

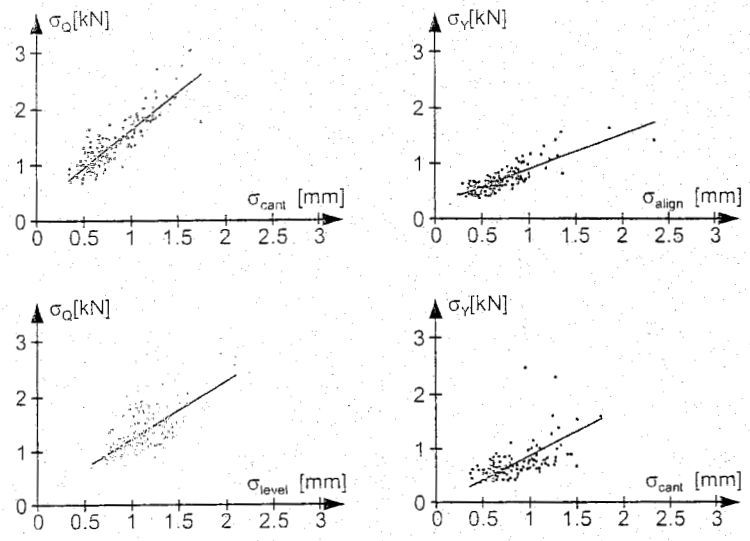


Figure 6.57: VRA-calculated forces versus track geometry per 200 m

These results are then transformed back into the time domain, i.e. per vehicle and response component 5 signals for 5 different speeds. Next the variances for the current 200 m section, and also the force and acceleration levels which are exceeded are calculated. Per vehicle and response component the maximum is taken for the 5 speeds. Both variances are added for the car body accelerations.
By taking the square root, standard deviations are obtained and these are printed, if desired, after having been normalized.

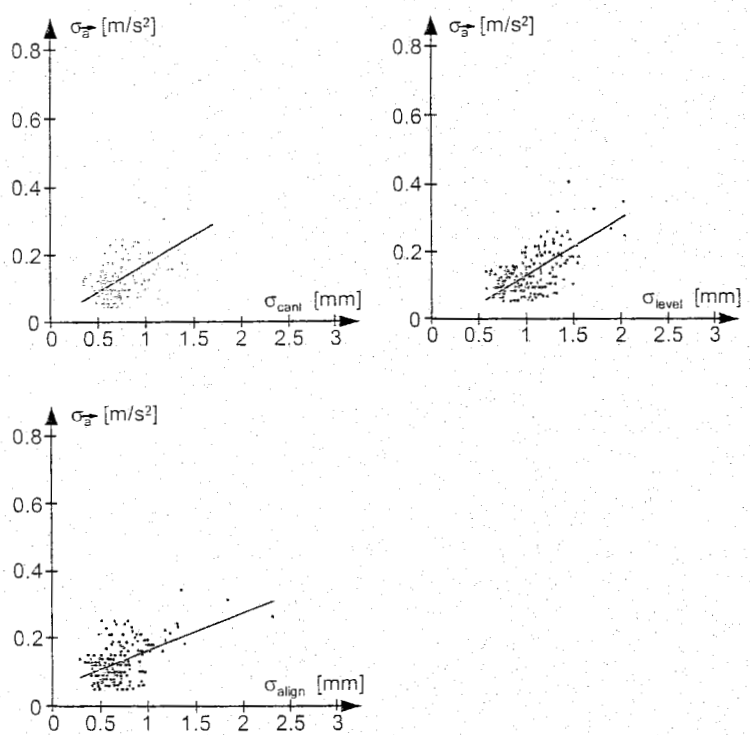


Figure 6.58: VRA-calculated accelerations versus track geometry per 200 m

At the time VRA was completed not all the required transfer functions were available. Although the final aim is to provide the Array Processor with transfer functions estimated by means of MISO, it was decided that the transfer functions calculated for the model of section 6.5. would first be implemented. Regarding this implementation Figure 6.57 shows a comparison between track geometry and vehicle reactions produced by VRA.

Per 200 m the standard deviations of Q-force and vertical car body acceleration are plotted against the standard deviation of level. Figure 6.58 presents the same information in the lateral direction.

Obviously, large variations in vehicle reactions are found for the same geometry values. It will be evident that the dynamic responses of railway vehicles are by far a better criterion for the actual maintenance status of a railway track than the pure track geometry. From this point of view it is expected that VRA will contribute to significantly better control of the track maintenance process.

6.8 Relation between Sperling's Ride Index Wz and ISO-weighted accelerations

The Wz Ride Index introduced by Sperling is the classical way of evaluating vehicle ride quality and passenger comfort [100]. Using modern analysing tools, the Wz factors are evaluated by multiplying the power spectral density function of the acceleration $G_a(f)$ with the squared modulus of the weighting function H(f) and integrating this result over the frequency range of interest to obtain the variance in that band according to:

$$\sigma^2 = \int_0^F G_a(f)|H(f)|^2 df \tag{6.207}$$

Finally, the Wz factor follows from:

$$Wz = [\sigma^2]^{1/6.67} \tag{6.208}$$

In this expression σ has the units cm/s^2. For passenger comfort in the vertical direction:

$$H(f) = H_{CV} = 0.588\left[\frac{1.911\,f^2 + (0.25\,f^2)^2}{(1 - 0.277\,f^2) + (1.563\,f - 0.0368\,f^3)^2}\right] \tag{6.209}$$

In lateral direction:

$$H(f) = H_{CL} = 1.25\,H_{CV} \tag{6.210}$$

For vehicle ride quality, the weighting in vertical and lateral directions is the same and reads:

$$H(f) = H_{RV} = H_{RL} = 1.14\left[\frac{[(1 - 0.056\,f^2)^2 + (0.645\,f^2)^2](3.55\,f^2)}{[(1 - 0.252\,f^2)^2 + (1.547\,f - 0.00444\,f^3)^2](1 + 3.55\,f^2)}\right]^{1/2} \tag{6.211}$$

The transfer functions are represented as a function of frequency in Figure 6.59 and Figure 6.60. The evaluation scales for the Wz factors were constructed based on vibration tests on people and were supplemented by other test results.

The assessment can be summarized as follows:

Wz passenger comfort		Wz vehicle ride comfort	
1	Just noticeable	1	Very good
2	Clearly noticeable	2	Good
2.5	More pronounced but not Unpleasant	3	Satisfactory
3	Strong, irregular, but still tolerable	4	Acceptable for running
3.25	Very irregular	4.5	Not acceptable for running
3.5	Extremely irregular, unpleasant, Annoying; prolonged exposure Intolerable	5	Dangerous
4	Extremely unpleasant; prolonged Exposure harmful		

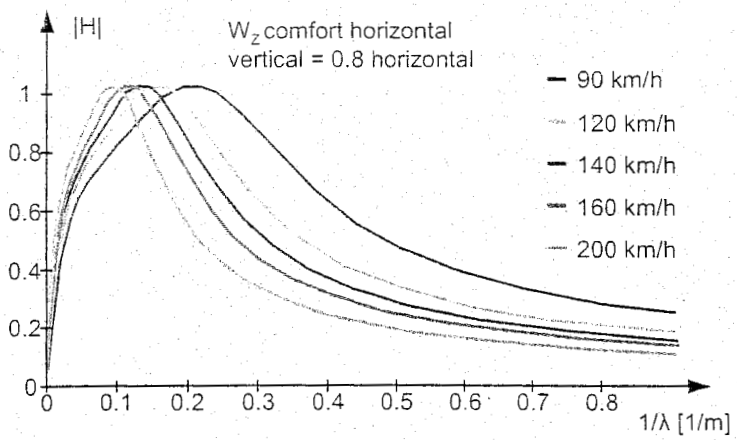


Figure 6.59: Transfer functions used to determine Sperling's ride index Wz for comfort

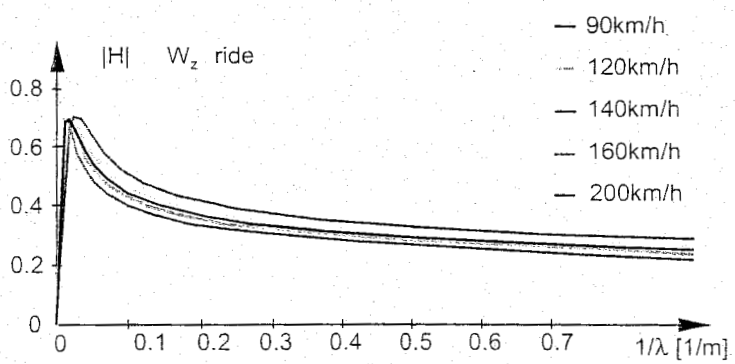


Figure 6.60: Transfer functions used to determine Sperling's ride index Wz for ride

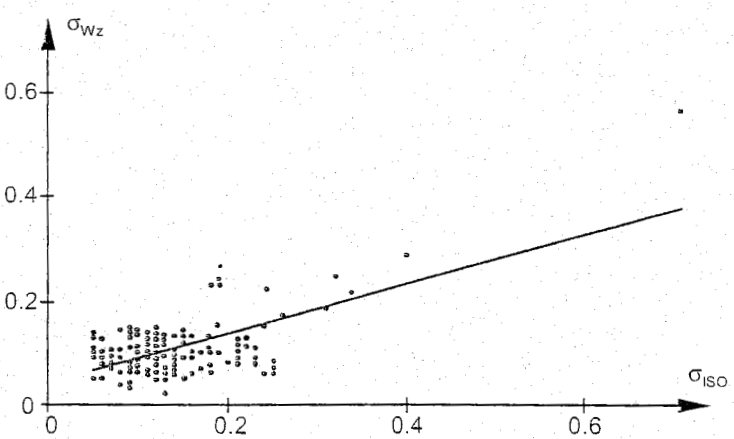


Figure 6.61: Relationship between Wz and ISO obtained from VRA simulations

To get a better understanding of how the ISO-weighted car body accelerations should be interpreted, the Sperling weighting functions were also stored temporarily in VRA. In this way it was possible to produce standard deviations per 200 m section for the ISO-weighted accelerations and the accelerations weighted according to the Sperling transfer functions for comfort. The latter have subsequently been treated according to (6.208).

Figure 6.61 shows the results obtained per 200 m section with the corresponding regression lines. It should be stated here that in the vertical direction the ISO and Wz weighting functions are quite similar. In the lateral direction Wz weighting deviates strongly from the ISO characteristic. In fact, the Wz weighting is not appropriate for evaluating passenger comfort in the lateral direction. For this purpose the ISO transfer function is now commonly used.

6.9 Applications of advanced dynamic models

6.9.1 Introduction

In the preceding chapters we discussed how to model and analyse dynamic track problems. Many methods were related to single degree of freedom (SDOF) systems, sometimes more complicated systems were discussed. These methods are valuable in that, with the limited support of some computer power, they are easy to handle in a design process. The disadvantage is that the structure and loadings are sometimes oversimplified and much experience is required to interpret, if still possible, the results in a meaningful way.

The application of much computer power makes it possible to apply 3D finite element method (f.e.m.) models. However, the use of 3D f.e.m. models is very time consuming, both of computer power and manpower. The time necessary to prepare, to execute, and to interpret one single problem is prohibitive in a design process. Unknown material properties and lack of homogeneity limit the meaningfulness of a 3D analysis.

On the other hand, the support of computer power makes it possible to apply more sophisticated models than were used before. Here an example is given of such a model, called RAIL, in addition to the models discussed in section 6.3. The position of RAIL in the field of track models is indicated in Figure 6.62.

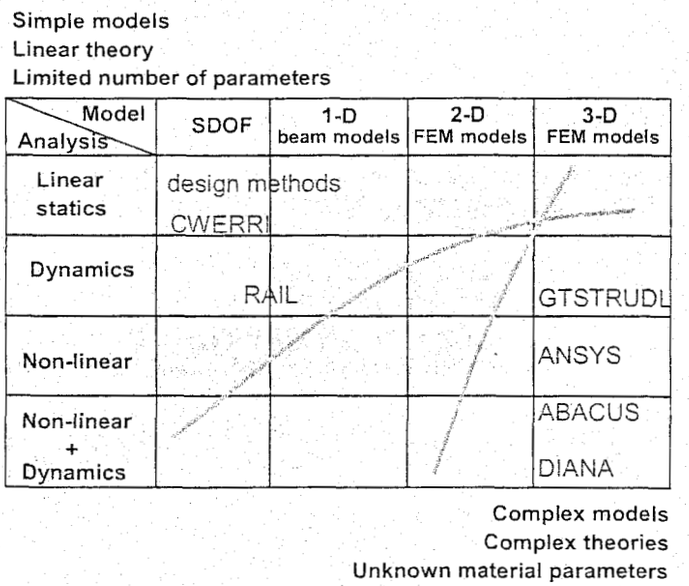


Figure 6.62: Overview of available software and railway track models for static and dynamic calculations

6.9.2 The RAIL-model

To assess the vertical dynamic behaviour of railway track due to moving railway vehicles, an integrated model called RAIL was developed at Delft University of Technology. The model was built up of two structures, namely the moving train and the railway track.

The coaches were schematised by rigid bodies, pivoted at the bogies by springs and dampers. Bogies and wheelsets were also modelled as rigid masses connected with springs and dampers. The vehicle model and the track model are schematically depicted in Figure 6.63. Only vertical forces and displacements etc. are considered.

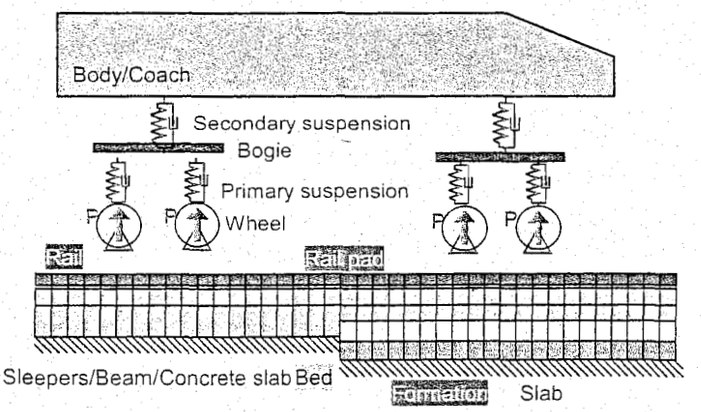


Figure 6.63: Model of moving vehicle on track structure

Rail and track irregularities are one of the most important sources of dynamic loads generated by a moving train. In the RAIL model the loads were introduced by means of Hertzian springs travelling along a sine-shaped or irregularly shaped rail and track surface as explained in Figure 6.64.

The equations of motion were formulated on the basis of the discussed structural models and were evaluated according to a direct integration procedure, based upon the concept of time-space elements. During a time step Δt the train moves over a distance Δx, which is dependent on the train velocity v, and 'crosses' the time-space elements as indicated in Figure 6.65.

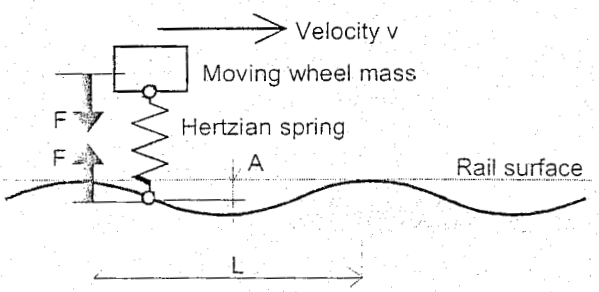


Figure 6.64: Modelling of wheel/rail loads

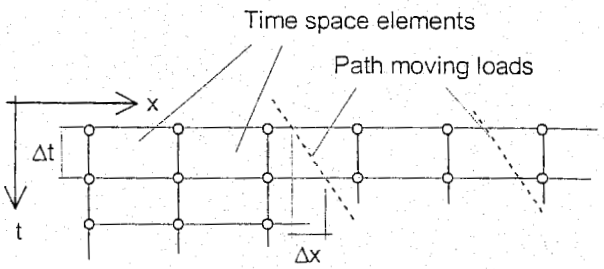


Figure 6.65: Calculation principle of moving trains

6.9.3 A comparison of several different track types

In order to study the effect of several different track types maximum elastic rail deflection, wheel rail contact force, and vertical body accelerations were investigated. In the underlying example it is shown how the properties of both vehicle and track are integrated into one complete model that takes the full interaction between the two structures into account.

The investigated structures consist of a Thalys train (special type TGV) which travels at different speeds on either a classic ballast track or an embedded rail structure (ERS). The ERS track is supported either by a rigid slab or a flexible slab which is itself discretely supported by a rigid pile foundation. The loads between track and vehicle are introduced by a sine-shaped surface deformation of the rail. Different wavelengths are investigated for their impact on the results. Three subjects are investigated: the elastic displacement of the rail head under the wheels, the contact forces between wheel and rail, and, finally, the vertical accelerations of the body, referring to the passenger coach or the locomotive.

The entire model is built up of two structures, namely, the moving train and the railway track. Each structure is modelled separately. During the dynamic analysis, that is the numerical integration procedure, the interaction of the two structures will be taken into account.

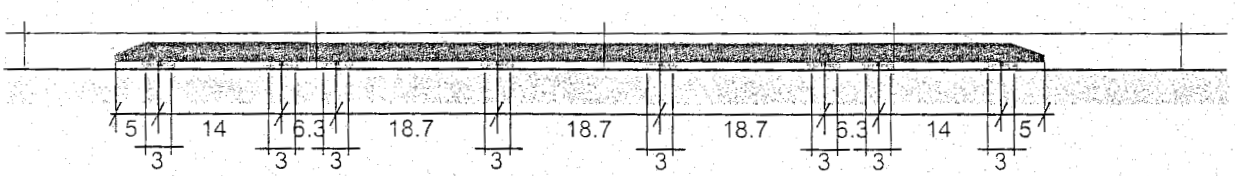


Figure 6.66: Dimensions of a (short) Thalys train as applied in the model

The train modelled here (short Thalys, see Figure 6.66) consists of five vehicles. It is typical for Thalys that bogies between two coaches are shared. The total length between the first wheel and the last wheel is 99.7 m. Rigid bodies, supported at the bogies by springs and dampers, schematise the coaches. Springs, dampers, and rigid bodies also model bogies and wheels. Only vertical forces and displacements are considered.

Regarding the Thalys we applied the following data:

- Mass (per item): coach 27140 kg, locomotive 54280 kg, bogie 2791 kg, wheel 1013.5 kg;
- Primary suspension spring (per wheel): K=1150 kN/m, C=2.5 kNs/m;

- Secondary suspension spring (per bogie): K=600 kN/m, C=4 kNs/m;
- Wheel radius: 0.42 m;
- To the coach next to the locomotive an additional lumped mass of 13570 kg. This equals all static wheel loads to 85 kN.

For our analyses we investigated the following track structures (Figure 6.67):

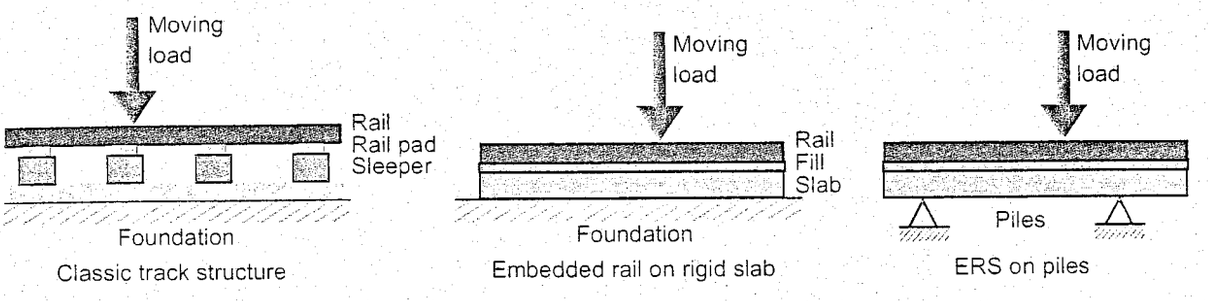


Figure 6.67: Analysed track structures

a A classic ballast track, using sleepers, rail pads, and an elastic ballast bed.

b An embedded rail structure (ERS) in which the rail, embedded in a fill material, has been put into a through in the rigid supporting concrete slab.

c A comparable embedded rail structure (ERS) with a more flexible concrete slab and supported by concrete piles (Figure 6.68).

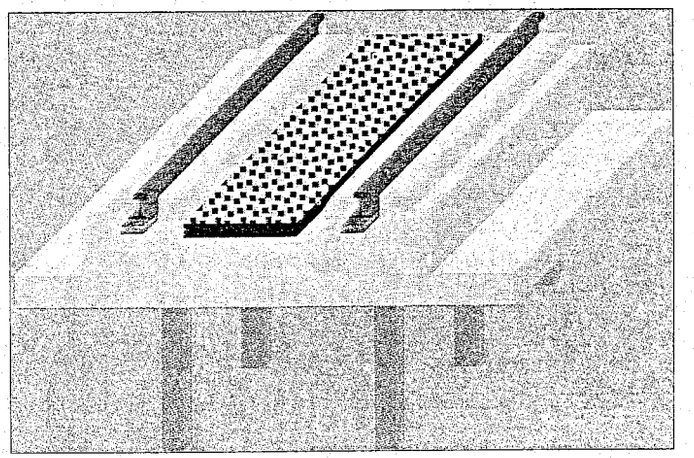


Figure 6.68: ERS with piles as foundation

For the analysis of the track we only took one rail into consideration assuming both rails and both sides of the train would show equal and synchronous behaviour.

The following data have been used for the classic ballast track:

- Rail: UIC-60;
- Rail pads (per rail): K=100 10^3 kN/m, C=15 kNs/m;
- Sleeper (per rail): mass 150 kg, spacing 0.60 m, width 0.15 m;
- Ballast bed (distributed per rail):
  k=180 10^3 kN/m^2, c=82 kNs/m^2 or per sleeper: K=27 10^3 kN/m, C=12.3 kNs/m.

For the ERS we applied:

- Rail: UIC-54;
- Fill material: k=52.5 10^3 kN/ m^2, c=4.98 kNs/ m^2;

- Slab rigidly supported (structure b);
- Slab supported on piles (structure c): A=0.765 m^2, I_{zz}=0.02746 m^4;
- Concrete properties: E=3.1 10^7 kN/m^2, ν=0.30, ρ=2500 kg/m^3;
- Pile distances: 3.0 and 6.0 m.

We considered a track length of 150 m for the analysis.

The surface geometry of the rail is one of the most important sources of load applications to a moving train. The extremes occur with short surface waves and high speeds. We considered three surface shapes $z_i(x)$ given by the algebraic formula:

$$z_i(x) = A_i\left(1 - \cos\frac{2\pi x}{L_i}\right) \tag{6.212}$$

where we substituted the following data:

- shape 1: L_i =3.0 m, A_i =1.5 mm;
- shape 2: L_i =12.5 m, A_i =3.0 mm;
- shape 3: L_i =50.0 m, A_i =6.0 mm.

The moving train has been analysed with regards to the speeds v=30 m/s, 60 m/s, and 90 m/s. The analysis has been performed with 1000 rail track elements of 0.15 m each and time steps Δt = 0.0015s. Dependent on the rail surface wavelengths, we applied 1600 time steps (2.40 s) for shape 1 and 4000 time steps (6.00 s) for shape 3.

The railway dynamics problem has been formulated by the equations of motion and the initial conditions. This means that, starting at t=0, a steady state solution will be obtained after damping of the initial disturbances. This requires some time.

This can be seen in a comparable calculation of a moving Thalys on an ERS at 3 different speeds. In Figure 6.69 the vertical body accelerations of the first Thalys vehicle are depicted for shape 1 rail surface geometry. The time-invariant period appears after almost 1 second of simulation. Moreover, the eigen period of the vehicle body is clearly present and lasts almost 1.5s.

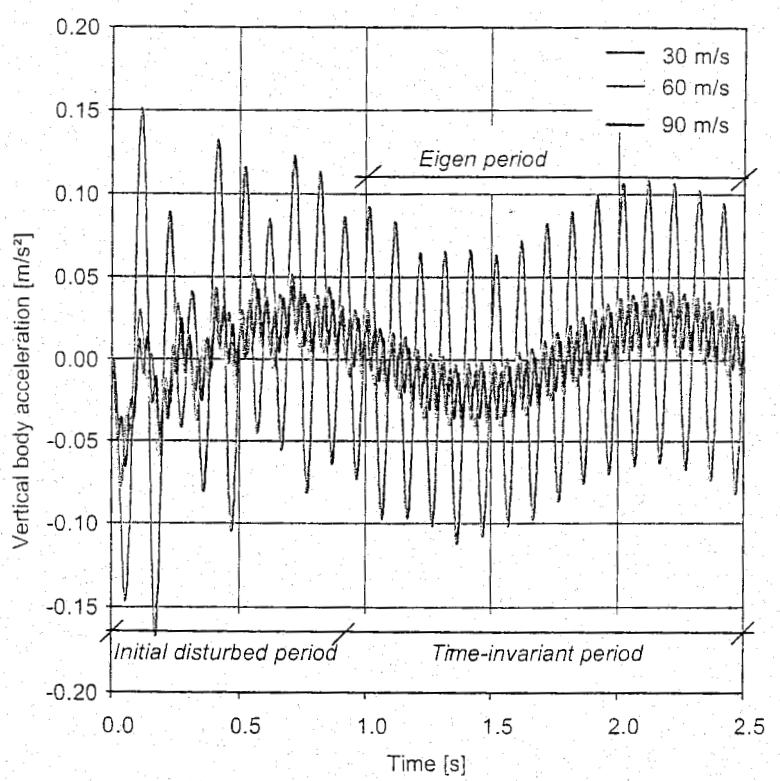


Figure 6.69: Vertical body accelerations of the first vehicle of Thalys at 3 different speeds

To study the velocity parameter and the dynamic response following our numerical process, the results for displacements and contact forces between t=1.8 s and t=2.4 s regarding the Thalys travelling at three different speeds on an ERS are shown in Figure 6.70 and Figure 6.71. As the speed increases, the variation in displacements and in contact forces also increases.

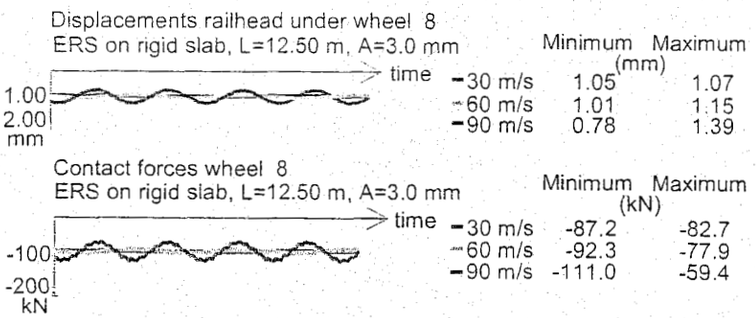


Figure 6.70: Displacements and contact forces between t=1.8s and t=2.4 s

When comparing different track structures, ERS on a continuously supported slab or a slab with small pile distance (3.00 m) performs much better than the classic ballast track or an ERS track on a slab with large pile distance. Continuously supported ERS shows a factor 2 between static and maximum dynamic contact forces, while the classic ballast track and an ERS on a widely spaced pile foundation (6.00 m) show a factor 3. For the elastic displacements of the rail head under the wheels, the size of the variations is even more distinct: here the classic ballast track shows the poorest performance as well.

The performed analyses confirm, of course, many observations that are made by intuition or by simple hand calculations. The analyses make it possible to compare the performance of a classic ballast track with respect to the ERS track, and the effect of pile spacing on an ERS track support. Certainly with respect to the contact forces and the rail head displacements, the ERS track performs better than the classic ballast track.

Besides full train simulations, a simpler approach is also available within RAIL. Moving loads or moving lumped mass models can be a suitable simplification of a train while running over predefined surfaces and structures of rail and track (see Figure 6.72). This gives e.g. vertical body accelerations in Figure 6.73, which are comparable to those already shown earlier in Figure 6.69. Please note that the track structure is different from that previously mentioned.

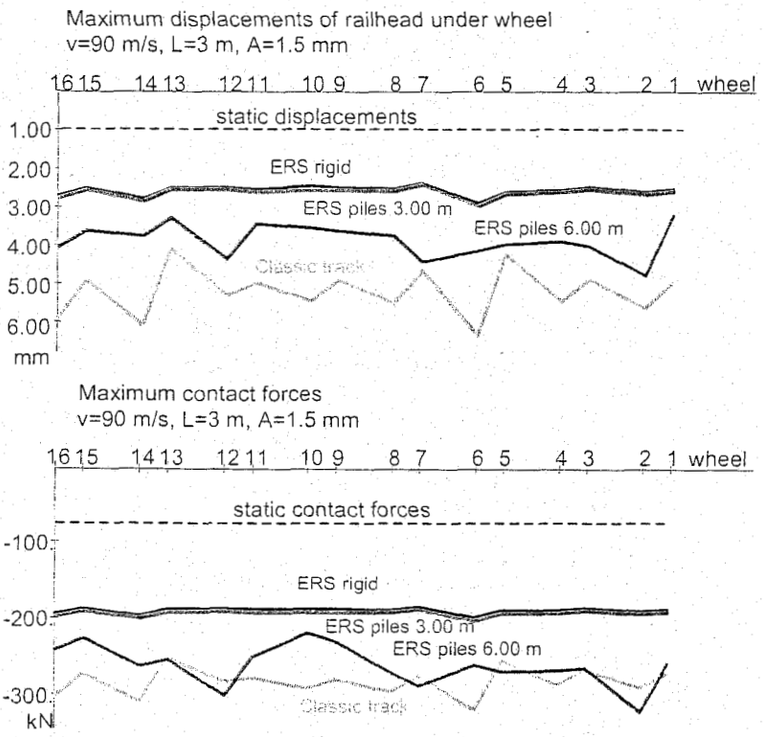


Figure 6.71: Maximum displacements and contact forces

It is obvious that not all of the possible applications of the integrated model of track and vehicle have been explored completely. It is also obvious that the integrated model offers many new possibilities for an integrated investigation of track and vehicle properties. For example, further applications of the RAIL program can be found in Chapter 13 and Chapter 14.

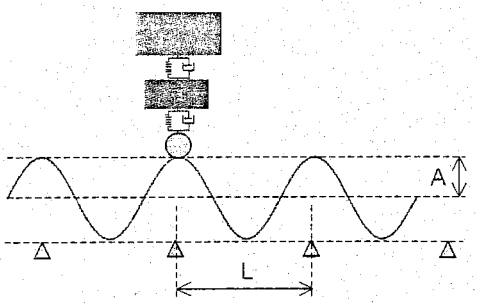


Figure 6.72: Simplified lumped mass vehicle model in RAIL

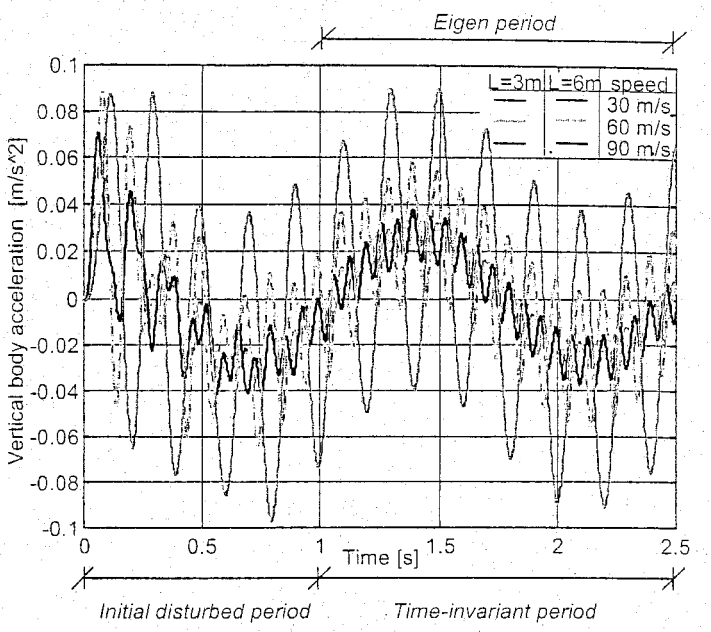


Figure 6.73: Vertical body (upper lumped mass) accelerations at different speeds on L = 3.0 m and L = 6.0 m and A = 3.0 mm rail surface geometry

6.9.4 Transitions in railway track on embankments and bridges

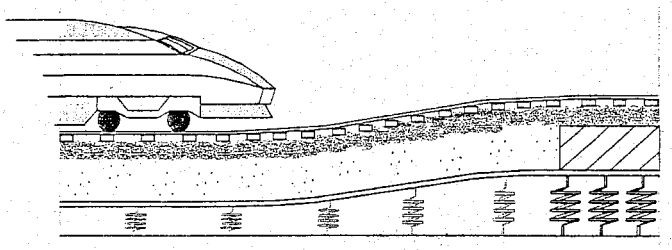


Figure 6.74: Principle of train response at a transition

Experience has shown that the transition between bridge and plain track (see Figure 6.74) often causes problems. Immediately after the track is laid, in particular, the plain track is subject to relatively high subsidence, at a rate different to that of the bridge. These transitions have been shown to cause problems, both in theory ([254], [127]) and in practice.

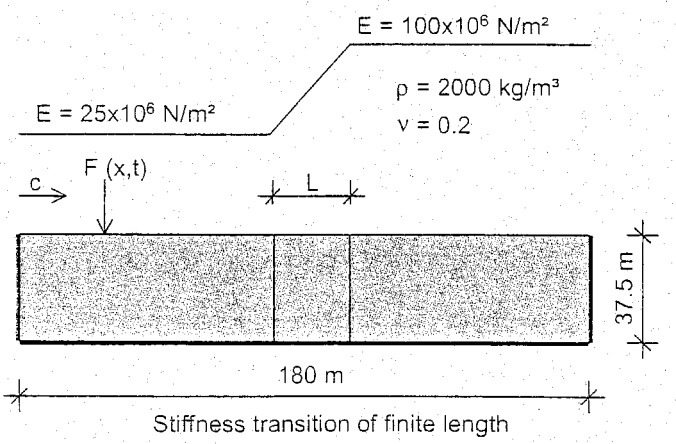


Figure 6.75: Stiffness transition of finite length

For the theoretical solution, finite element routines are available since a certain time, enabling to model and calculate the behaviour of trains at a transition.

The change in stiffness causes increased dynamic forces, the extent of which is determined by speed, stiffness ratio, damping and the length of the transition. In [254] a study is described into the effects of changes in vertical stiffness on the dynamic response.

Figure 6.75 defines the stiffness transition between a soft soil (clay) and a stiff soil (sand), with a linear stiffness transition of length L. Figure 6.76 shows the dynamic amplification factor for two situations. If the ratio between train speed and Rayleigh speed is low, an asymptotic D.A.F. value of 1.21 is reached after a relatively short length of about 8 m. If the train speed is close to the critical speed, a much larger asymptotic value of 2.51 is found, which is reached at a transition length beyond 25 m. So it obvious to stay far from the critical speed. The change in track geometry resulting from subsidence thus causes increased dynamic forces.

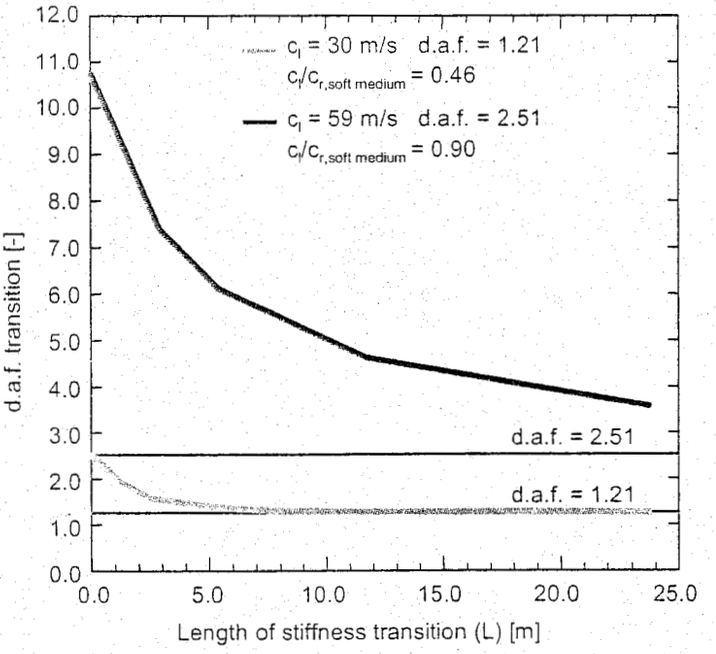


Figure 6.76: Dynamic amplification within the stiffness transition

Like large changes in subgrade stiffness, changes in the vertical alignment of the track lead to increasingly pronounced vertical accelerations in the vehicle, which may mean that the criteria for passenger comfort or maximum dynamic track force are no longer met. A marked increase in forces leads to accelerated deterioration of the track geometry and hence to additional maintenance.

The dynamic effects at the transition between bridge and plain track is a typical part of research into the problems of high speed rail systems. TU Delft has carried out a study [243], which have revealed that the effects of a change in height are generally more significant than those of a change in stiffness. Figure 6.77 shows an example of the car body accelerations calculated for a discrete event (change in height) of 30 mm, over a length of 30 m. Parametric studies have been carried out on the basis of the model described in [127], enabling accelerations and vertical wheel loads to be derived for various lengths of event and changes in height, at 300 km/h. Figure 6.78 shows the differences in height that are acceptable for HSL rolling stock for a given length of event, on the basis of a maximum permissible dynamic wheel force Q_{dyn} of 170 kN and a vertical car body acceleration a_v of 1.0 m/s^2.

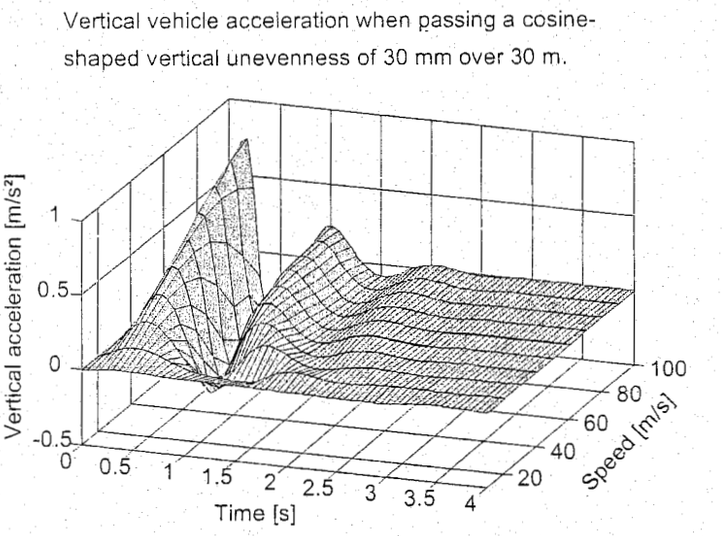


Figure 6.77: Calculated car body acceleration at transition

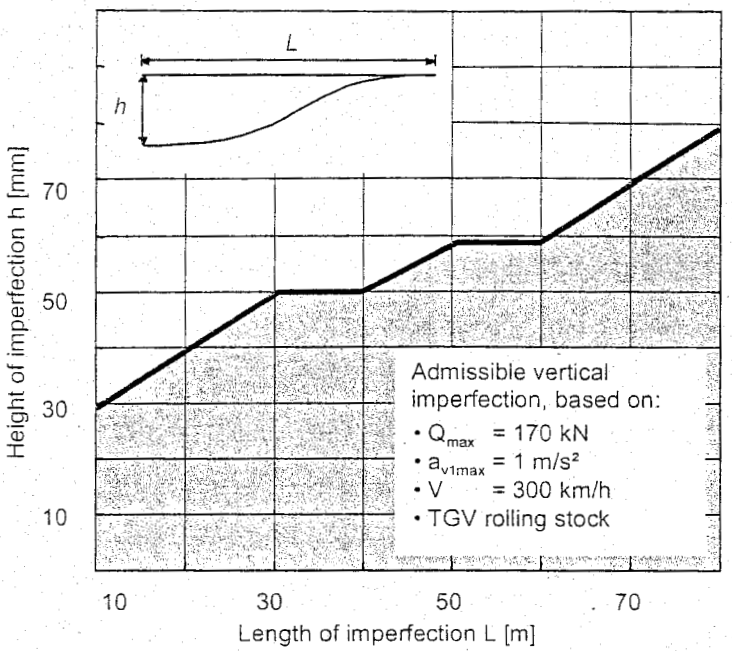


Figure 6.78: Design graph for a transition structure

6.10 Track response to a moving axle load

In railway transport, the continuous demand for capacity increase as well as for more extensive mobility networks automatically requires critical assessment of the effectiveness of currently operating and newly built railway lines. In order to meet the high standards regarding effectiveness, in many countries train cruising speeds of 200 km/h and more have been employed. As a result of such high train velocities, the track response in these high-speed lines has a typically dynamic character.

In order to reliably predict track safety and track deterioration, it is necessary to develop and utilise appropriate track models that cover the relevant vibration effects and wave propagation phenomena. In this section, the influence of the train velocity on the dynamic response of a railway track is considered. Accordingly, some models are discussed that consider the elasto-dynamic wave propagation under a moving load. These models provide further insight into the dynamic behaviour of a railway track, revealing typical effects that can only be captured by means of advanced analyses.

6.10.1 Track response at the critical train velocity

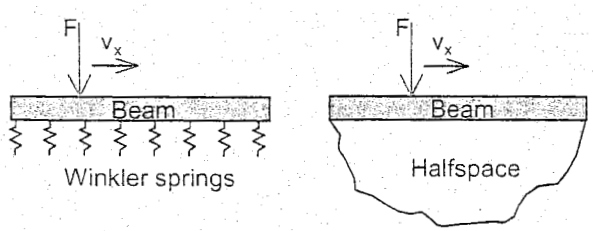


Figure 6.79: Beam-spring configuration and beam-halfspace configuration subjected to load F that moves with velocity v_x.

The analysis of the instantaneous response of a homogeneous railway track to a moving train axle commonly occurs by employing moving load models. In these models the train axle is modelled by a force of a constant (or a harmonically varying) amplitude F that moves with a velocity v_x over a configuration representing the railway track. The compound system of rails, sleepers, and ballast, known as the superstructure, is thereby often modelled by a beam, while the supporting subgrade, known as the substructure, is modelled by either discrete Winkler springs ([263], [52], [98]) or by a continuous halfspace ([95], [154], [50], [254], [163], [148]), see Figure 6.79.

In the current section it is illustrated what kind of results can be obtained from a moving load model and how these results relate to railway practice. First, the model presented in [255] will be discussed which considers a Timoshenko beam supported by a two-dimensional halfspace and subjected to a moving load with a constant force magnitude. Here, the Timoshenko beam simulates the elastic bending and shearing behaviour of the superstructure, while the halfspace simulates the elastic response of the subgrade. It is reasonable to assume that the structure behaves elastically, since the response of a railway track during the passage of an individual train axle is almost completely reversible (see for instance [247]).

In order to analyse the wave propagation in the beam-halfspace system, in [254] the corresponding boundary value problem has been elaborated on. This elaboration was based on combining the equations of motion of the system with the boundary conditions. By searching for harmonic solutions, the characteristic waves in the system have been computed where two body waves (compression wave, shear wave) can be distinguished and a wave that propagates along the surface of the track system (surface wave).

The surface wave commonly conveys the largest part of the energy generated by the train. When the velocity of the surface wave is of a similar magnitude as the velocity of the train, the generated energy remains close to the train. This leads to accumulation of energy under the train as time progresses. This phenomenon can be identified as resonance. Because this resonance may result in large track amplifications, the train's velocity is often designated as 'critical' when resonance occurs.

Although several critical velocities may be identified for a specific track configuration (see for instance, [50], [254]), the lowest critical velocity is the most important since this is the first one to be met by an accelerating train vehicle. As confirmed by 'in-situ' measurements, rail deflections can increase to more than three times the static deflection when a train reaches the lowest critical velocity ([126], [164]). It is easily understood that for reasons of safety and limitation of train/track deterioration such track amplifications can not be tolerated.

The main features of a train that reaches the critical regime will be illustrated by assuming the halfspace in Figure 6.79 to be softer than the Timoshenko beam. Accordingly, the railway superstructure is considered to be supported by a relatively soft formation of clay or peat. In the model [254], the load starts to move from zero velocity and accelerates up to a velocity larger than the lowest critical velocity.

The analysis of the response has been performed by means of a finite element model with dimensions b x h = 180 m x 37.5 m, see Figure 6.80. The moving character of the load is simulated by means of a set of discrete pulses which act successively on the element nodes at the surface along which the load is supposed to propagate. Furthermore, the half-infinite character of the halfspace is simulated by means of viscous damping elements that are connected to the artificial model boundaries (energy-absorbing boundary). By providing the damping elements with dynamic impedances that are similar to that of the adjacent continuous medium, the energy of the waves which encounters the artificial model boundaries is (almost completely) absorbed. For more details on the finite element model, see [254].

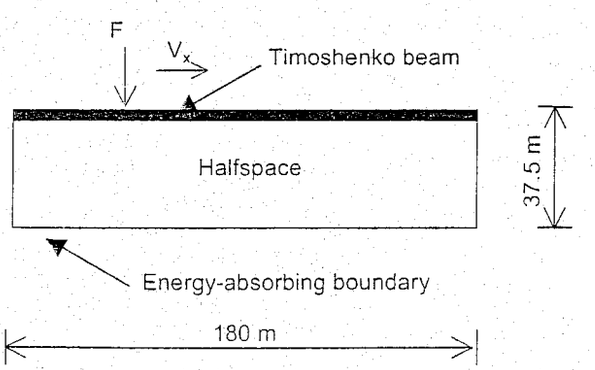


Figure 6.80: Moving load F subjected to a Timoshenko beam-halfspace configuration. Waves arriving at artificial model boundaries are absorbed.

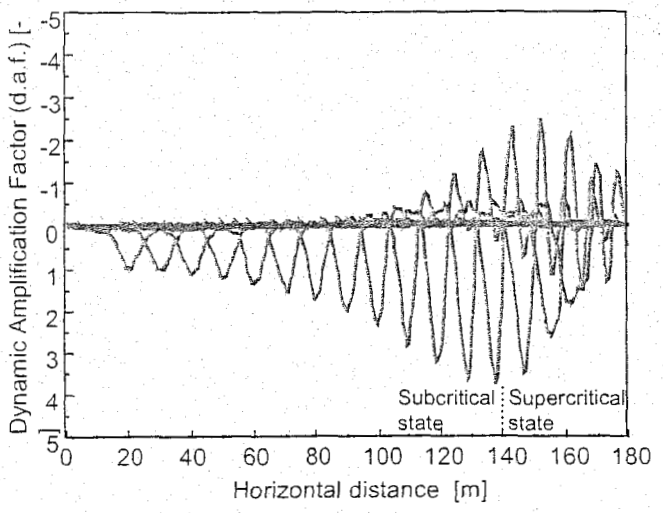


Figure 6.81: Stroboscopic development of the dynamic amplification factor (d.a.f.) under an accelerating load. Timoshenko beam is supported by a relatively soft half-space.

In Figure 6.81, the stroboscopic development of the dynamic amplification of the normal stress in vertical direction, σ_{zz}, is depicted. The dynamic amplification relates to a point at 5.68 m below the surface, and has been computed by dividing the dynamic stress by its static counterpart. The total horizontal distance of 180 m plotted on the horizontal axis reflects the total width of the finite element model. The load starts to move at a horizontal distance of 20 m from the left-side model boundary at which the dynamic amplification equals 1.0, thus corresponding to a static response.

Due to the fact that the load velocity increases with increasing horizontal distance, the dynamic amplification grows in a monotonic manner. At a horizontal distance of 140 m from the left-side model boundary, a maximum of about 4 times the static response is reached. At this stage, the load velocity has approached the shear wave velocity c^s of the halfspace and the system behaves critically. After this critical stage has passed the response becomes super critical.

Obviously, in the supercritical range the system response decreases under increasing load velocity. This is because the load propagates faster than the energy transmitted by the radiated surface waves. Correspondingly, the energy can no longer accumulate under the load.

When the load velocity is strongly super critical, the amplitude of the response appears to be of the same order of magnitude as that of the static response. The characteristic behaviour sketched in Figure 6.81 is, actually, similar to that of an aeroplane passing the sound barrier. This results from the fact that the phenomenon of an aeroplane catching up with a sound wave is completely analogue to that of a train catching up with surface waves.

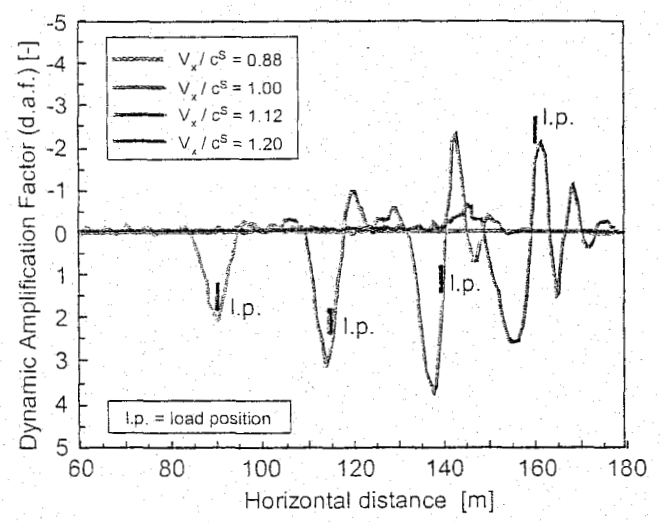


Figure 6.82: The dynamic amplification factor (d.a.f.) under an accelerating load in the velocity range 0.88 $c^S < v_x < 1.20$ c^S. Timoshenko beam is supported by a relatively soft half-space.

In order to study the response at the critical stage in more detail, in Figure 6.82 the dynamic amplification has been plotted for a load velocity (v_x) in between 0.88 and 1.20 times the shear wave velocity (c^s) of the halfspace. Obviously, at v_x=0.88c^s the response is still approximately symmetric. For larger load velocities, however, the response becomes increasingly asymmetric. This is the result of the appearance of oscillating Mach waves upon reaching the critical regime. These Mach waves have also been monitored during 'in-situ' track deflection measurements [164], thereby revealing a maximum upward deflection of 9 mm in front of the front train axle and a maximum amplitude of 12 mm behind the front train axle.

For obvious reasons, the generation of Mach waves may have a detrimental effect on both the track and the train. It may, in addition, cause the train to derail. This can be motivated from the supercritical response at $v_x = 1.20c^s$ which shows that the response directly below the load position (l.p.) acts in a direction opposite to the loading direction. In other words, the train axle is lifted up by the surface waves.

The analysis above has illustrated that in the case of a railway track built on a relatively soft subgrade, the critical regime is reached when the train velocity is near the shear wave velocity of the subgrade. The magnitude of the shear wave velocity is determined by material properties, as computed from

$$c^s = \sqrt{\mu/\rho} \tag{6.213}$$

where μ is the shear modulus and ρ is the density of the material.

For a subgrade of soft clay or peat, the shear wave velocity commonly lies in between 150 and 250 km/h. 'In-situ' track deflection measurements in Great-Britain [126] and Sweden [164]) confirmed that the track response may become critical in this velocity range. For a railway track resting on a subgrade of relatively soft clay, the measurements demonstrated that the track response is amplified to more than three times the static response when the train reaches a velocity of about 200 km/h.

Track amplifications are decreased when the soft subgrade is replaced by a stiffer structure such as a sand embankment of considerable thickness. Accordingly, the shear wave velocity of the supporting subgrade can be increased to a value that lies outside the velocity domain of the high-speed train vehicle. An alternative solution is to disconnect the railway superstructure from the soft subgrade by means of a pile foundation. The soft subgrade then will not affect the track response anymore, since the piles convey the generated waves into the stiff soil layer that supports the piles. Another option is to leave the soft subgrade as it is and limit the train velocity to a level at which the track amplifications are acceptable. However, this option is certainly not preferred as it may considerably increase the transportation time on the specific railway line.

As already pointed out in Figure 6.79, moving load models may be divided into two categories: beam-spring models and beam-halfspace models. Because the analysis of a beam-halfspace model is considerably more complicated than the analysis of a beam-spring model, there may be a strong preference to use beam-spring models when examining wave propagation phenomena in a railway track. Nonetheless, it should be realised that a discrete spring support is not able to transmit waves. As a consequence, incorporation of the commonly accepted spring properties for modelling the track subgrade results in an inadequate description of the track dynamics. In fact, the dynamic amplification predicted by a beam-spring model then becomes significant only for velocities far beyond the range of operational train speeds. This may lead to the misconception that the dynamic track amplification caused by the train is generally negligible.

6.10.2 Dynamic response of a ballast layer

For a railway track supported by a soft subgrade the wavelength of the surface waves is usually relatively long, typically ranging from 5 to 20 meters for a peat or clay formation. These long wavelengths are caused by the soft nature of the subgrade. When the superstructure is supported by a stiff substratum, such as a rock formation, a concrete bridge, or a concrete tunnel, the waves which propagate at the surface of the track are not necessarily restricted to the domain of long wavelengths.

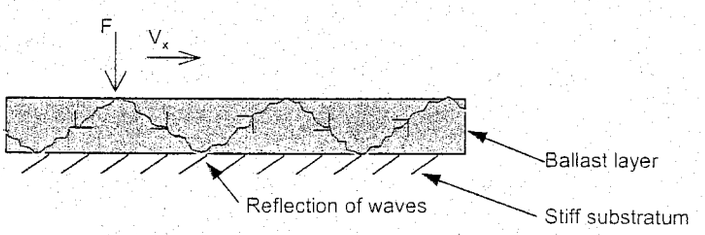


Figure 6.83: Ballast layer supported by a stiff substratum. Multiple reflection of waves.

This is because a stiff substratum reflects body waves of all wavelengths back into the superstructure, see Figure 6.83. Consequently, a ballast layer on a concrete bridge or tunnel acts as a waveguide, conveying waves of both long and short wavelengths.

The shorter waves in the ballast layer can have a wavelength of the order of magnitude of the ballast particle size which may perturb the individual particles. In order to model relative motions by the ballast particles, it is necessary to incorporate the particle size into a mechanical track model. This can be done either by employing continuum models that are derived from the micro-mechanical particle behaviour ([256], [257]), or by using discrete particle models ([258], [259]). In these models the particles are assumed to have an ideal spherical shape. The interaction with neighbouring particles is prescribed at particle contact points by means of a contact law. Moving load analyses carried out with these models have demonstrated that a ballast material consisting of large particles increases the intensity of the wave radiation, especially when the damping capacity of the ballast is low. Hence, to suppress such effects, the damping capacity of a ballast material should be sufficiently high.

The damping capacity of ballast generally depends on two effects: the inter-particle friction and the distribution of particle sizes. As far as the first effect is concerned, when the friction between the particles is high (i.e. a coarse-grained ballast) a relatively large amount of energy is dissipated at the particle contacts, thus causing the ballast to have a high damping capacity. The second effect stems from the fact that a wide particle size distribution excludes a dominant appearance by large particles that may act as a resonator. Accordingly, a ballast material with a random distribution of various particle sizes has a better damping capacity than a ballast material which consists of identical particles of a relatively large size.

In general, dynamic amplifications are not only caused by the train velocity, but also by impact loads such as those generated by the sleeper distance effect or by other track irregularities. In the case of a ballast layer supported by a stiff substratum, the energy transmitted by the track vibrations remains for a large part inside the ballast layer as a the result of (multiple) wave reflections at the stiff substratum, see Figure 6.83.
Consequently, reflected waves may interfere with other (reflected or non-reflected) waves. This causes the amplitude of the response either to increase or to decrease depending on the motion characteristics of the interfering waves. When superimposing the dynamic amplification by load impacts on that by the train speed, it may appear that the track response becomes 'critical' at a considerably lower velocity than the critical velocity which relates to the effect of train speed only ([98]; [257], [259]).

The extent to which this occurs depends on the characteristics of the impact loading (e.g. frequencies and duration of the impact loading) as well as on the geometry and material properties of the track structure (e.g. thickness and stiffness of the ballast layer). An adequate way of reducing the dynamic amplification by load impacts is to apply ballast mats between the ballast layer and the stiff substratum. In fact, the damping characteristics of the ballast mats will reduce the wave reflection depicted in Figure 6.83.

6.10.3 Stiffness transitions

Stiffness transitions emerge when a relatively soft substratum of clay or peat changes into a relatively stiff substratum of sand or concrete (or vice versa). Such transitions appear nearby concrete railway bridges or railway tunnels, but also when the characteristics of the natural soil formation (i.e. the substructure) change abruptly. Stiffness transitions form the basis for the emergence of differential settlements, which, once initiated, may grow considerably as time progresses.

The growth rate of differential settlements is governed by the dynamic properties of the train, i.e. the mass and velocity of the train in combination with its spring and damping characteristics. The mechanism of growing differential settlements can be explained as follows. A stiffness difference increases the dynamic loading on the track. As a consequence of the increased dynamic loading, differential settlements may emerge. Differential settlements activate additional train vibrations which cause the dynamic loading to become larger the next time the train passes the stiffness transition. Accordingly,

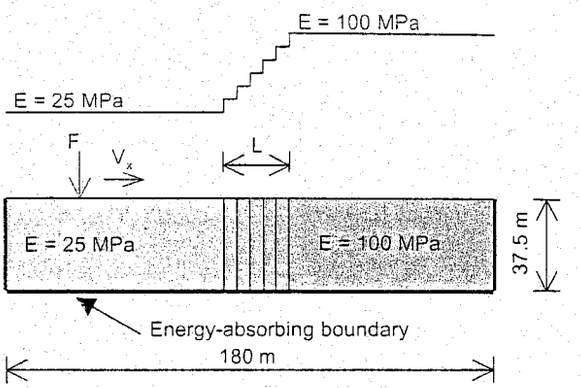


Figure 6.84: Stiffness transition L in an elastic halfspace comprising a relatively soft medium (E = 25 MPa) and a relatively stiff medium (E = 100 MPa).

the differential settlements grow further and the above mechanism repeats itself.

Since differential settlements may cause excessive track deterioration, they should be controlled as much as possible. Therefore, it is of practical importance to understand how the dynamic response of a stiffness transition, which forms the basis for the emergence of settlement differences, is influenced by a passing train axle. This is demonstrated by means of a finite element model, see Figure 6.84, that has been discussed previously in [253].

The model consists of an elastic halfspace comprising a relatively soft material of E=25 MPa (e.g. clay) and a relatively stiff material of E=100 MPa (e.g. sand), in which E represents the Young's modulus of the materials.

The stiffness transition L between the soft and the stiff material occurs in a stepwise manner, with a step-size equal to the transition length divided by 0.75 m (= element size). The modelling of the moving load and the absorption of the wave energy at the artificial model boundaries occurs in the same manner as explained for the model in Figure 6.80.

For an ideal homogeneous halfspace, the critical velocity is equal to the Rayleigh wave velocity c^r which has a magnitude close to that of the shear wave velocity [1]).

The influence of the load velocity v_x on the dynamic stress amplification at the transition L has been examined by considering a relatively low load velocity and a relatively high load velocity. Here, the low load velocity v_x=30.0 m/s is about half the critical velocity of the soft medium, while $v_x/c^{r,\ \text{soft medium}} = 0.46$.The high load velocity v_x=59.0 m/s is very near the critical velocity of the soft medium, $v_x/c^{r,\ \text{soft medium}} = 0.90$.

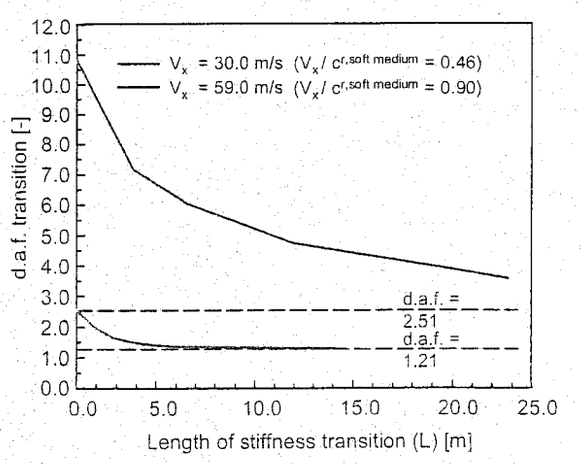


Figure 6.85: Dynamic amplification factor (d.a.f.) versus the length of the stiffness transition. Moving load with a relatively low velocity ($v_x/c^{r,\ \text{soft medium}} = 0.46$) and a relatively high velocity ($v_x/c^{r,\ \text{soft medium}} = 0.90$).

The effect of the transition length L on the dynamic stress amplification has been depicted in Figure 6.85. Clearly, for both load velocities the dynamic amplification factor (d.a.f.) at the stiffness transition decreases with an increasing transition length. As for the previous analysis, the dynamic amplification factor has been computed by dividing the dynamic stress σ_{zz} at 5.68 m below the surface by its static counterpart.

For the high load velocity v_x=59.0 m/s, the dynamic amplification appears to be significantly larger than for the low load velocity v_x=30.0 m/s The depicted horizontal lines d.a.f. = 1.21 and d.a.f. = 2.51 represent the dynamic amplification for v_x=30.0 m/s and v_x=59.0 m/s in case the halfspace would be homogeneous, with E=25 MPa. Hence, these lines act as asymptotes to which the dynamic amplification factor approaches when the transition length increases. Obviously, for the low load velocity v_x=30.0 m/s the asymptote is reached much faster than for the high load velocity v_x=59.0 m/s.

From the current analysis two important conclusions can be drawn. Firstly, in order to adequately reduce the dynamic amplification at a stiffness transition, the length of the transition should be sufficiently large. At a certain stage, a further increase of the transition length will no further reduce the dynamic amplification. Secondly, railway lines constructed for fast(er) trains require a large(r) transition length to reduce the dynamic track amplification. Apart from these two effects there is another effect that influences the dynamic amplification: the magnitude of the stiffness difference. Trivially, an increasing stiffness difference increases the dynamic amplification at the stiffness transition so that a large stiffness difference should be bridged by large transition length L. The quantitative effect of the magnitude of the stiffness difference on the dynamic amplification is exemplified in [253] by means of finite element analyses and energy considerations.

6.10.4 Brief discussion

The models presented in this section have provided insight into some important features of the dynamic response of a railroad track to a moving axle load. It is possible to extend the models in several ways in order to provide a more realistic representation of an 'in-situ' railway system. Here, one may think of replacing the moving load by a mass-dashpot-spring system that accounts for the train vehicle dynamics, or of adding the influence of ground water and frictional sliding to the ballast and subgrade behaviour.

7 TRACK STABILITY AND LONGITUDINAL FORCES

7.1 Introduction

In conventional non-welded tracks the rails are connected by means of joints to allow for length changes caused by temperature fluctuations. Using joints prevents the development of axial forces and the consequent risk of track buckling at high temperatures. However, the penalty for this is the care for maintenance-intensive joints which generate high dynamic loads during train passage. These loads are responsible for many problems like rapid deterioration of vertical track geometry, plastic deformation of the rail head, dangerous rail cracks as well as damage to sleepers and fastenings. These problems increase progressively as speed increases. As a rule, joints have a very considerable negative effect on the service life of all track components.

Tracks with continuous welded rails (CWR) do not possess the above drawbacks. Owing to the absence of joints the quality of the track geometry is better by an order and this results in a substantial decrease in the total life cycle cost. CWR does not, however, only have advantages. As was pointed out in Chapter 5, the stresses resulting from the plane strain situation may be of the order of 100 N/mm² and should be added to the residual rail stresses and bending stresses caused by train loads which are of the same order of magnitude. Temperature stresses especially are responsible for failure of welds with small imperfections at low temperatures. On the other hand, lateral stability should be sufficiently great to resist compression forces developing at temperatures above the neutral temperature of 25°C, as buckling may otherwise occur as, for example, illustrated in Figure 7.1. The principle of this phenomenon is sketched in Figure 7.2 showing the compressive forces and the resistance forces on the track and the resulting typical buckling shape.

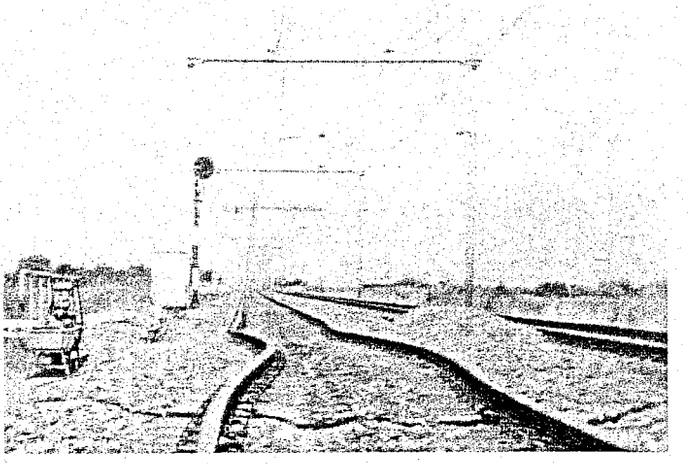


Figure 7.1: Example of track buckling

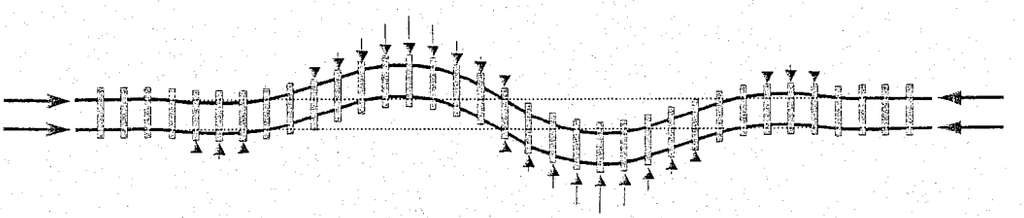


Figure 7.2: Typical buckling shape

On bridges and viaducts the deformation regime deviates from the plain track situation. The rails follow the construction which can undergo large displacements with respect to the adjacent track. Without adequate measures this would result in high rail stresses. To avoid these stresses expansion joints are applied.

This chapter is devoted to track stability and track longitudinal problems which, in the case of compression forces, are strongly interrelated. For both fields analytical and finite element modelling approaches are presented with examples. The last section discusses recently developed advanced models which describe safety considerations about track buckling or deal with more general or complicated track systems.

7.1.1 Straight track and elastic lateral resistance

In order to analyse this basic problem the track is modelled as a straight bending beam with a length L[1]. The beam is loaded by axial compressive forces and is elastically supported by a laterally distributed resistance which is proportional to the lateral deflection:

$$\tau = \beta v \qquad (7.1)$$

where ß is a constant and v(x) = lateral deflection.

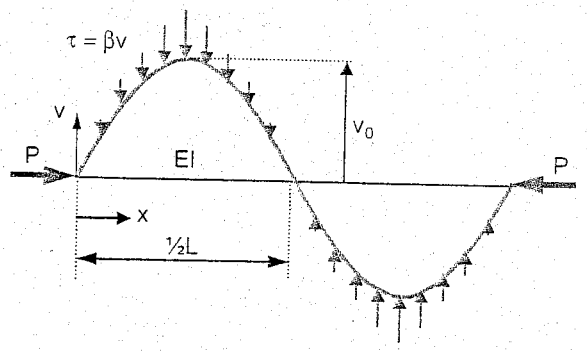


Figure 7.3: Buckling of straight beam with elastic lateral resistance

In view of the treatment of the rail buckling problem, a continuous sinusoidal shape will be assumed. Only one wavelength need to be considered in Figure 7.3. The energy method will be used to solve the problem because it not only gives the equilibrium equation, but also the nature (stable or unstable) of the equilibrium. The total energy consists of three parts:

1. Bending energy of the beam:

$$U_{beam} = \tfrac{1}{2}\int_0^L EIv''^2 dx \qquad (7.2)$$

where v"2 is d2v/dx2.

2. Deformation energy lateral pressure, with (1-1):

$$U_{press} = \tfrac{1}{2}\int_0^L \beta v^2 dx \qquad (7.3)$$

3. Work done by the compressive force:

$$W_P = \tfrac{1}{2}\int_0^L Pv'^2 dx \qquad (7.4)$$

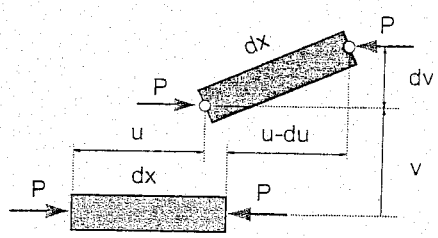


Figure 7.4: Rail element in buckling

While the first two energies are obvious, the last one may need some explanation. Referring to Figure 7.4 the elementary energy can be written as:

$$dW_P = Pdu = P(dx - \sqrt{dx^2 - dv^2}) = Pdx(1 - \sqrt{1 - v'^2}) \qquad (7.5)$$

where P is assumed to be constant and axial deformation is neglected. Linearisation gives:

$$dW_P = \tfrac{1}{2}Pv'^2 dx \qquad (7.6)$$

which explains equation (7.4).

The total energy is then:

$$U_{tot} = U_{beam} + U_{press} - W_P \qquad (7.7)$$

1. Not to be confused with the same symbol L used for the characteristic length of the track

Giving as solution:

$$v(x) = v_0 \sin\frac{2\pi x}{L} \tag{7.8}$$

the total energy per wavelength becomes, after inserting (7.8) in (7.2), (7.3), (7.4), and using symmetry considerations:

$$U_{tot} = \frac{v_0^2 L}{4}\left[EI\left(\frac{2\pi}{L}\right)^4 + \beta - P\left(\frac{2\pi}{L}\right)^2\right] \tag{7.9}$$

The equilibrium follows from setting the partial derivative $\frac{\partial U_{tot}}{\partial v_0} = 0$, resulting in:

$$P = \frac{4\pi^2 EI}{L^2} + \frac{\beta L^2}{4\pi^2} \tag{7.10}$$

As an example the relation between the buckling load P and the wavelength L is given in Figure 7.5 for EI = 8000 kN/m² and for some values of β in the range 10 to 200 kN/m².

The maximum value of the buckling load is restricted to 2500 kN in Figure 7.5 because higher values are unrealistic. For instance, in the case of UIC 60 rails and an extreme temperature increase of 65 °C, the maximum normal force which can occur in CWR track amounts to 2414 kN (=2*210e^{6}*76.9e^{-4}*1.15e^{-5}*65).

Stability is ensured if $\frac{\partial^2 U_{tot}}{\partial v_0^2} > 0$, hence if:

$$P < \frac{4\pi^2 EI}{L^2} + \frac{\beta L^2}{4\pi^2} \tag{7.11}$$

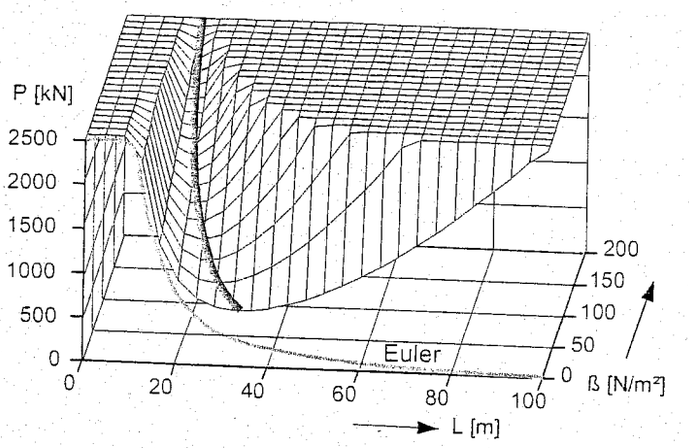


Figure 7.5: Buckling diagram in case of elastic resistance

This shows that each equilibrium point on the surface is indifferent with respect to stability.

From the condition $\frac{\partial P}{\partial L} = 0$ the following critical quantities are determined:

$$L_{crit}^2 = 4\pi^2\sqrt{\frac{EI}{\beta}} \tag{7.12}$$

$$P_{crit} = 2\sqrt{\beta EI} \tag{7.13}$$

T
his relation between L, β, and P has been indicated in Figure 7.5 with a thick line.

If there is no lateral resistance (β=0), then (7.10) will reduce to the well-known expression for the Euler buckling load for straight free beams, second mode:

$$P = \frac{4\pi^2 EI}{L^2} \tag{7.14}$$

This curve is also drawn in Figure 7.5 and is marked with small circles.

7.1.2 Track with misalignment and constant lateral shear resistance

If applied to rail track situations, the value of P_{crit} in (7.13) is only significant when the lateral ballast resistance is very low and approaches the Euler buckling load. However, this situation is not representative for railway track, for in the elastic model there is always sufficient resistance and according to the above-mentioned example unrealistic high values of the buckling load would be calculated. With respect to railway track two important factors should be taken into account.

First, the lateral resistance, which is caused by the shear resistance between track panel and ballast bed, has a limiting value. Secondly, real track is never perfectly straight, but shows some form of geometrical imperfection or misalignment. Therefore, in this section a more realistic model for rail track will be used.

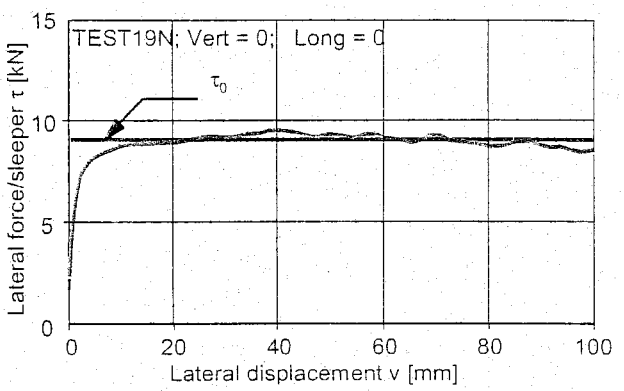


Figure 7.6: Lateral resistance measurement

In Figure 7.6 a recent measurement of the lateral shear resistance characteristic of a track panel in ballast bed is given, expressed in force per sleeper. The curve could, for instance, be approximated by a bi-linear function, but for our purpose here we will approximate the curve by a constant plastic shear resistance which opposes the axial displacement and therefore can be defined as:

$$\tau = \tau_0 sign(v)^{1} \qquad (7.15)$$

The bending stiffness EI is represented here by the horizontal bending stiffness of the track panel. This stiffness incorporates the bending stiffness of the rails, the sleepers, and the rotational resistance of the fastenings. It is assumed that this bending stiffness is constant. The straight track is supposed to possess an initial sinusoidal misalignment in the unloaded situation (Figure 7.7), given by:

$$f = f_0 \sin\frac{2\pi x}{L} \qquad (7.16)$$

It is also assumed that the additional bending curve has the same relationship as the misalignment due to the buckling load P:

$$v = v_0 \sin\frac{2\pi x}{L} \qquad (7.17)$$

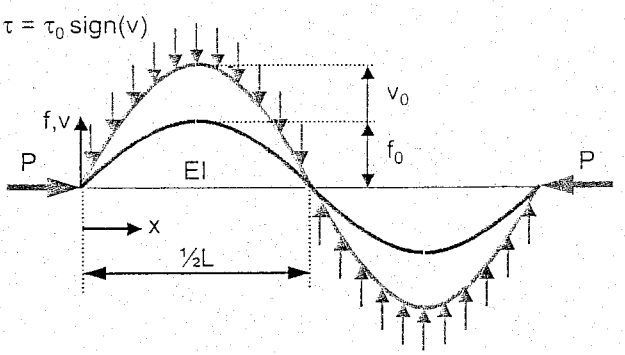


Figure 7.7: Lateral buckling of track panel with plastic lateral shear resistance

Finally, the buckling force P and the wavelength L are assumed to remain constant during the buckling process.

Under these conditions the energy parts can be calculated analogous to the scheme in the first example.

1. 'sign' refers to the mathematical function signum meaning here: sign (v) returns 1 if v>0, 0 if v=0, and -1 if v < 0

1. *Bending energy of the beam:*

$$U_{beam} = 2\int_0^{L/4} EIv''^2 dx = \frac{Lv_0^2 EI}{4}\left(\frac{2\pi}{L}\right)^4 \tag{7.18}$$

where v''^2 is d^2v/dx^2.

2. *Deformation energy of lateral shear resistance,* with (7.17):

$$U_{res} = 4\int_0^{L/4} \tau_0 v dx = \frac{2\tau_0 v_0 L}{\pi} \tag{7.19}$$

3. *Work done by the compressive force* (modified to include the effect of f(x)):

$$W_P = 2\int_0^{L/4} P[(v'+f')^2 - f'^2]dx = \frac{PL}{4}[(v_0+f_0)^2 - f_0^2]\left(\frac{2\pi}{L}\right)^2 \tag{7.20}$$

The total energy per wavelength in this case becomes:

$$U_{tot} = \frac{L}{4}\left[v_0^2 EI\left(\frac{2\pi}{L}\right)^4 + \frac{8\tau_0 v_0}{\pi} - P[(v_0+f_0)^2 - f_0^2]\left(\frac{2\pi}{L}\right)^2\right] \tag{7.21}$$

The equilibrium follows from setting the partial derivative $\frac{\partial U_{tot}}{\partial v_0} = 0$, resulting in:

$$P = \frac{v_0}{v_0+f_0}\left[\frac{4\pi^2 EI}{L^2} + \frac{\tau_0 L^2}{\pi^3 v_0}\right] \tag{7.22}$$

From $\frac{\partial P}{\partial L} = 0$ the following critical quantities can be determined:

$$L_{crit}^2 = 2\sqrt{\frac{\pi^5 f_0 EI}{\tau_0}} \tag{7.23}$$

$$P_{crit} = 2\sqrt{\frac{\tau_0 EI}{\pi f_0}} \tag{7.24}$$

$$\left(\frac{v_0}{f_0}\right)_{crit} = 1 \tag{7.25}$$

The critical value of P_{crit} is the smallest load of the compressive force P, with corresponding wavelength, which causes buckling.
The stability criterion of the equilibrium can be determined by checking the inequality $\frac{\partial^2 U_{tot}}{\partial v_0^2} > 0$, which gives the condition for stability:

$$P < \frac{4\pi^2 EI}{L^2} \tag{7.26}$$

This means that for a stable equilibrium the axial load P should be less than the Euler buckling load.

In Figure 7.8 the graphical result is given of an example based on the following parameters:

τ_0 = 10 kN/m;
EI = 8000 kNm²;
f_0 = 0.025 m.

The critical values can be calculated using (7.23) and (7.24):

L_{crit} = 12.51 m;
P_{crit} = 2018 kN.

The critical value of the compressive load lies well within the range of possible values (under rather extreme conditions, see former section).

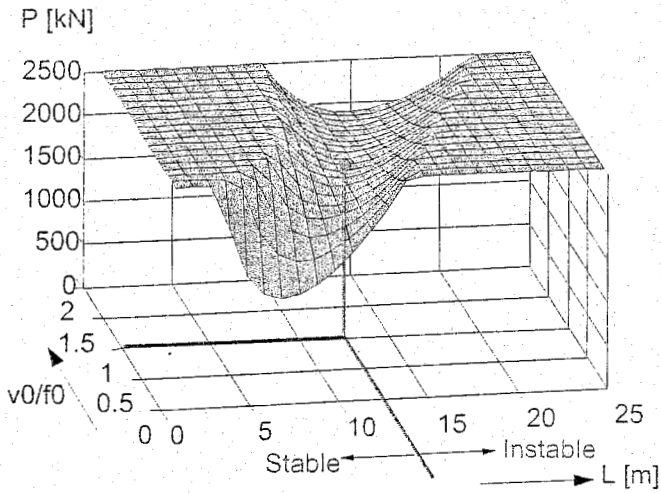


Figure 7.8: Buckling diagram in case of plastic resistance and misalignment

7.2 Track stability: finite element modelling

7.2.1 General considerations

In the 1950's and 1960's ORE Committee D14 investigated the stability of tracks with CWR under the influence of temperature loads on various geometrical imperfections [208], [209]. For this purpose analytical models were developed as described in reference [19]. Although these mathematical models had been verified by many field tests, several limitations and assumptions still had to be accepted in some areas. One of these limitations was that the mathematical model was only allowed to be applied to curve radii greater than 500 m. It is a fact that in tight curves an increase in temperature is followed by a lateral displacement over the total curve length. This decreases the axial compression force in the track. In such a case the longitudinal ballast resistance forces play an important role, especially in transition curves, and this aspect was lacking in the analytical models.

The above-mentioned limitations do not apply to calculations based on the finite element method (FEM). Using the model described in [115], the stability of a number of track structures under the influence of a temperature load was analysed. The non-linearities in this process require an incremental approach in which temperature loads are increased step by step. Moreover, the model allows for any design geometry for the track, including geometrical deviations. Boundary conditions are no longer a limiting factor.

7.2.2 Finite element model

The railway track is considered to be an elastically supported beam with a constant bending stiffness (EI) and a torsional resistance of the fastenings at the sleepers. The ballast resistance is represented by springs in the lateral and longitudinal directions. Displacement occurs in the two-dimensional, horizontal plane.

Each of the three spring types mentioned has a specific bi-linear characteristic according to Figure 7.9. This means that in the first part of the diagram the resistance increases in proportion to the displacement and in the second part the resistance is constant with displacement. This is known as the plastic phase.

The mathematical model permits an elastic spring reaction on removal of the load when the spring has already reached the plastic phase. The beam element length is constant for the overall track construction and corresponds to the sleeper spacing. As indicated in Figure 7.10, the joints possess three degrees of freedom u, v, and ϕ and are connected to two ballast springs and one rotation spring.

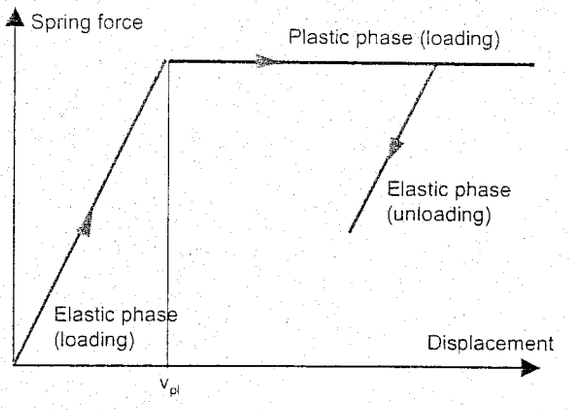


Figure 7.9: bi-linear spring characteristic

The equations of equilibrium for the track construction, composed of the beam elements, provide a system of equations in terms of the unknown displacement increments according to the matrix equation:

$$[S][\Delta v] = [\Delta F] \qquad (7.27)$$

where:
[S] = tangent stiffness matrix [N/mm];
[Δv] = displacement increment vector [mm];
[ΔF] = load increment vector [N].
The tangent stiffness matrix [S] consists of the following components:

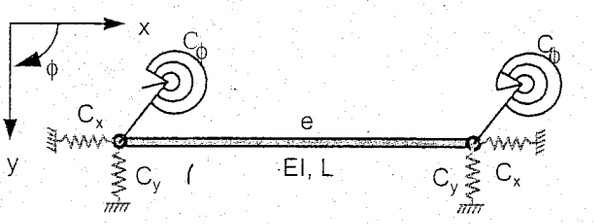


Figure 7.10: Beam element applied in FEM

Element stiffness matrix [S_1]:

The beam element used in the FEM consists of a combined truss-bending element. Therefore, the matrix [S_1] contains the strain and bending stiffness constants. The linear relation between displacement [v^e] and forces [F^e] of a beam element is defined according to the matrix equation:

$$[S_1][v^e] = [F^e] \qquad (7.28)$$

in which the displacement vector [v^e] contains the three degrees of freedom u, v, and ϕ per joint. The displacement field u of the truss element is governed by the differential equation:

$$EA\frac{d^2u}{dx^2} = 0 \qquad (7.29)$$

which has the solution:

$$u(x) = c_1 + c_2 = u_1^e + \frac{x}{L}(u_2^e - u_1^e) \qquad (7.30)$$

The displacement field is presented in Figure 7.11.

The bending part in the displacement field, associated with v and ϕ, is based on what are known as shape functions. These are third degree polynomials in x which satisfy the differential equation:

$$EI\frac{d^4u}{dx^4} = 0 \qquad (7.31)$$

with the solution:

$$v(x) = c_1 + c_2 x + c_3 x^2 + c_4 x^3 \qquad (7.32)$$

The constants c_1 to c_4 are dependent on the joint displacements. This makes it possible to express v(x) in the degrees of freedom according to:

$$v(x) = \alpha_1(x) v_1^e + \alpha_2(x) \phi_1^e + \alpha_3(x) v_2^e + \alpha_4(x) \phi_2^e \qquad (7.33)$$

The shape functions are presented in Figure 7.12 and can be expressed by:

$$\alpha_1 = 2\left(\frac{1}{2} - \frac{x}{L}\right)^2 \left(1 + \frac{x}{L}\right) \qquad (7.34)$$

$$\alpha_2 = \left(\frac{1}{2} + \frac{x}{L}\right)\left(\frac{1}{2} - \frac{x}{L}\right)^2 L \qquad (7.35)$$

$$\alpha_3 = 2\left(\frac{1}{2} + \frac{x}{L}\right)^2 \left(1 - \frac{x}{L}\right) \qquad (7.36)$$

$$\alpha_4 = -\left(\frac{1}{2} - \frac{x}{L}\right)\left(\frac{1}{2} - \frac{x}{L}\right)^2 L \qquad (7.37)$$

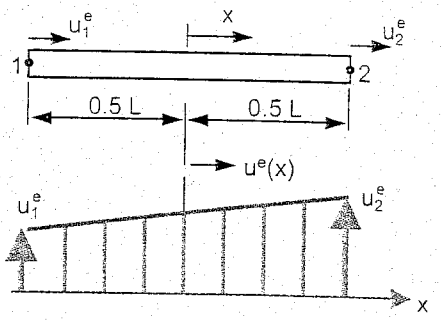


Figure 7.11: Displacement field of truss element

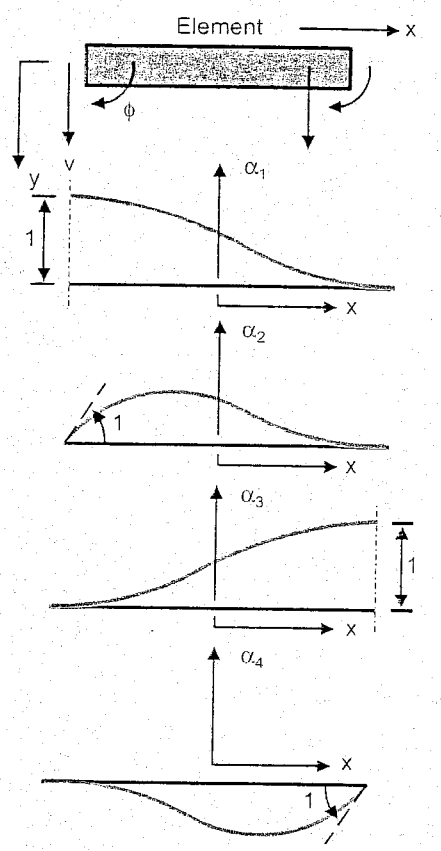


Figure 7.12: Shape functions of bending element

Stiffness matrix $[S_2]$:

This matrix takes the geometrical non-linearity based on the second order theory for small strains and large displacements into account and is derived by using the strains from the displacements. As depicted in Figure 7.13, the strain e is related to the displacements u and v by the following expression:

$$\varepsilon = \frac{\partial u}{\partial x} + \frac{1}{2}\left(\frac{\partial v}{\partial x}\right)^2 \qquad (7.38)$$

With this theory, the stability of the equilibrium of the two dimensional beam constructions can be investigated according to the principle of virtual work. There is equilibrium if the variation in potential energy equals zero, or:

$$\delta E_{pot} = \iiint_V \sigma \delta\varepsilon \, dV - \iint_l F \delta v \, dx = 0 \qquad (7.39)$$

Assume that due to the load q_0 the axial forces N_0 and bending moments M_0 are present in the construction. If variations ∂u and ∂v are applied to the displacements u and v, with an unchanged external loading, the following variation in the potential energy of the construction will emerge:

$$\delta E_{pot} = \int \left[N_0 \delta\varepsilon + M_0 \delta\kappa + \frac{1}{2} dN \delta\varepsilon + \frac{1}{2} dM \delta\kappa - q_0 \delta v \right] dx$$
$$= \delta^1 E_{pot} + \delta^2 E_{pot} \qquad (7.40)$$

The first term:

$$\delta^1 E_{pot} = \int \left[N_0 \frac{\partial \delta u}{\partial x} - M_0 \frac{\partial^2 \delta v}{\partial x^2} - q_0 \delta v \right] dx \qquad (7.41)$$

describes the equilibrium of the internal forces with the external load. The second term describes the strain energy and is split up into a linear and a non-linear part:

$$\delta^2 E_{pot} = \delta^2 E_{pot,1} + \delta^2 E_{pot,2} \qquad (7.42)$$

In this expression the linear and non-linear parts are respectively:

$$\delta^2 E_{pot,1} = \frac{1}{2} \int EA \left(\frac{\partial \delta u}{\partial x} \right)^2 dx + \frac{1}{2} \int EI \left(\frac{\partial^2 \delta v}{\partial x^2} \right)^2 dx \qquad (7.43)$$

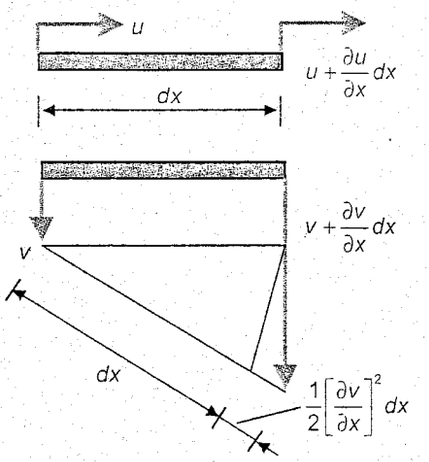


Figure 7.13: Axial strain due to displacements u and v

$$= \frac{1}{2} [v^e]^T [S_1^e] [\partial v^e] \qquad (7.44)$$

$$\delta^2 E_{pot,2} = \frac{1}{2} \int N_0 \left(\frac{\partial \delta v}{\partial x} \right)^2 dx \qquad (7.45)$$

$$= \frac{1}{2} [v^e]^T [S_2^e] [\partial v^e] \qquad (7.46)$$

The matrix $[S_1^e]$ is the symmetrical element stiffness matrix. The matrix $[S_2^e]$ contains factors which depend on the normal force. It creates the extra part in the matrix [S] which expresses the geometrical non-linearity.

Spring stiffness matrix [S_v]:

The bi-linear springs at the joints contribute to the total construction stiffness. However, they do not add extra degrees of freedom to the system. As long as a spring is in the elastic phase, the matrix relation for a linear spring applies according to:

$$\begin{bmatrix} C_x & & \\ & C_y & \\ & & C_\phi \end{bmatrix} \begin{bmatrix} u \\ v \\ \phi \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ M_\phi \end{bmatrix} \qquad (7.47)$$

or:

$$[S_v][v] = [F] \qquad (7.48)$$

As shown, the matrix [S_v] contains the stiffness of ballast and fastenings which are placed on the matrix main diagonal. After having reached the plastic phase, the spring force no longer increases but remains constant with displacement. The spring's contribution to the track stiffness equals zero for the next loading increment.

By direct summation of these three matrices the tangent stiffness matrix [S] is formed. The support conditions are taken into consideration when reducing the matrix [S]. The calculating process is as follows:

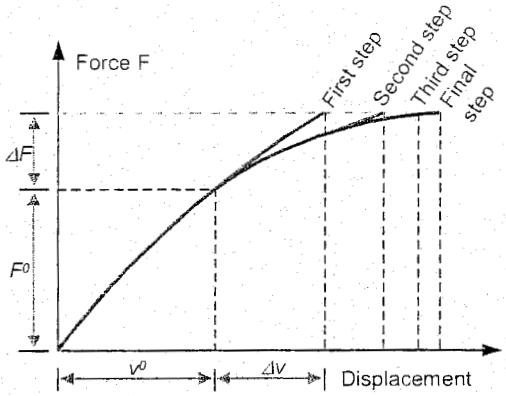


Figure 7.14: Newton-Raphson iteration process

After applying the temperature load and solving the static equilibrium equations, the displacement vector [Δv] is found. After that a new (rotated) tangent stiffness matrix [S] is drawn up with the help of the new total displacement ($v^\circ+\Delta v$) and the internal stresses. Back substitution of the same displacement vector [Δv] gives a new load vector [ΔF*]. The difference between the original load vector [ΔF] and the load vector [ΔF*] serves as a new load vector in the new total displacement situation ($v^\circ+\Delta v$) until the desired accuracy has been reached. This process, schematically presented in Figure 7.14, is often described as the Newton-Raphson iteration process.

The load on the track construction can consist of two components, namely:

Temperature load

This load is due to a temperature rise which causes compressive forces in the track. One of the consequences can be lateral track instability (buckling). This load is active along the total track. According to Section 7.3.2, the following formula (7.54) for the normal rail force N due to a temperature change ΔT with respect to the neutral or initial temperature is found:

$$N_{max} = -EA\alpha\Delta T \tag{7.49}$$

Hence, the compression force on the track (two rails) is:

$$P = 2EA\alpha|\Delta T| \tag{7.50}$$

in which:

P = compressive force on the track, taken as positive value in subsequent calculations [N];
E = Young's modulus for the rail [N/mm²];
A = total cross-sectional area of one rail [mm²];
α = coefficient of expansion [/°C];
$\Delta T = T_{actual} - T_{neutral}$ [°C].

External joint loads

The possibility of applying individual joint loads in the lateral and longitudinal directions is created. This was done in order to verify the mathematical model by means of data from lateral resistance measurements described in [73].

Both load types can be applied individually or in combination.

7.2.3 Results

After a thorough and comprehensive verification of the mathematical model and computer program, efforts were made to reproduce the results presented in [19]. In order to do this, the necessary input data were derived from the measurement data originally used. The results are presented in Figure 7.15 and show a high level of similarity. The same figure also indicates the existence of a critical wavelength which refers to a minimum buckling force in the relation buckling force/buckling length (= half wavelength L). It also refers to a lower critical compression force (buckling force) when the amplitude y of the misalignment is larger.

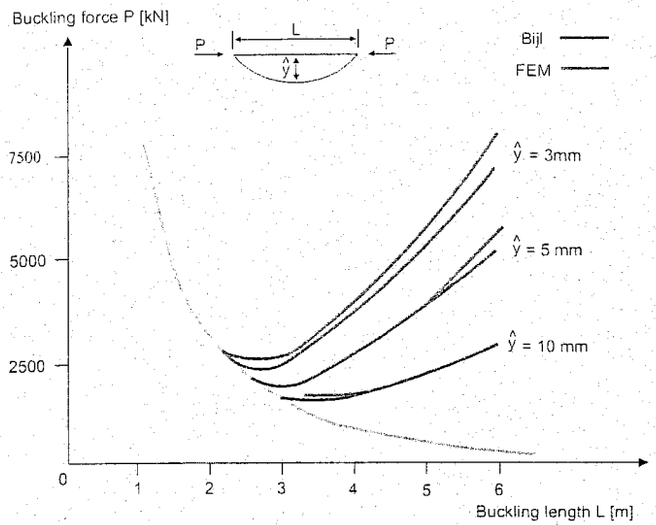


Figure 7.15: Buckling force versus buckling length according to Bijl and calculated with FEM

In the calculations published in [19] a sinusoidal initial displacement, i.e. geometrical imperfection of the track, was assumed. At different wavelengths, with the wavelength being twice the buckling length, the buckling force was determined for different initial conditions. In practice, wavelength and geometrical deviation magnitude are strongly correlated. Starting from some tenths of a millimetre at wavelengths of a few metres and some millimetres at wavelengths of 10 to 20 m, the amplitudes can become much larger than 10 mm for longer waves.

In order to take this correlation into account, two sections of about 30 m length were selected for the calculations on the basis of BMS recordings. Both sections form part of a 200 m section with a standard deviation of 1 mm for alignment, which corresponds approximately to the maintenance intervention level on NS rail. The geometrical deviations in the lateral direction (alignment) are presented in Figure 7.16. For both initial deviations of the track geometry the differential temperature, with respect to the neutral temperature at 25°C at which the track becomes unstable, was calculated. This value amounts to 140 - 150°C [89].

Subsequently, these calculations were repeated by increasing the amplitude of the initial geometrical deviation in proportion to the standard deviation of alignment per 200 m section. Buckling forces and differential temperatures for standard deviations of 0.5, 1, 2, and 4 mm are presented in Figure 7.16. The analysis results show that only in extremely poor tracks differential temperatures are attained which will lead to instability.

The NS have carried out many lateral resistance measurements with an adapted tamping machine [73]. To simulate this loading process with an FEM analysis, two point loads were applied to straight track in the lateral direction. Figure 7.17 shows some results of lateral resistance measurements car-

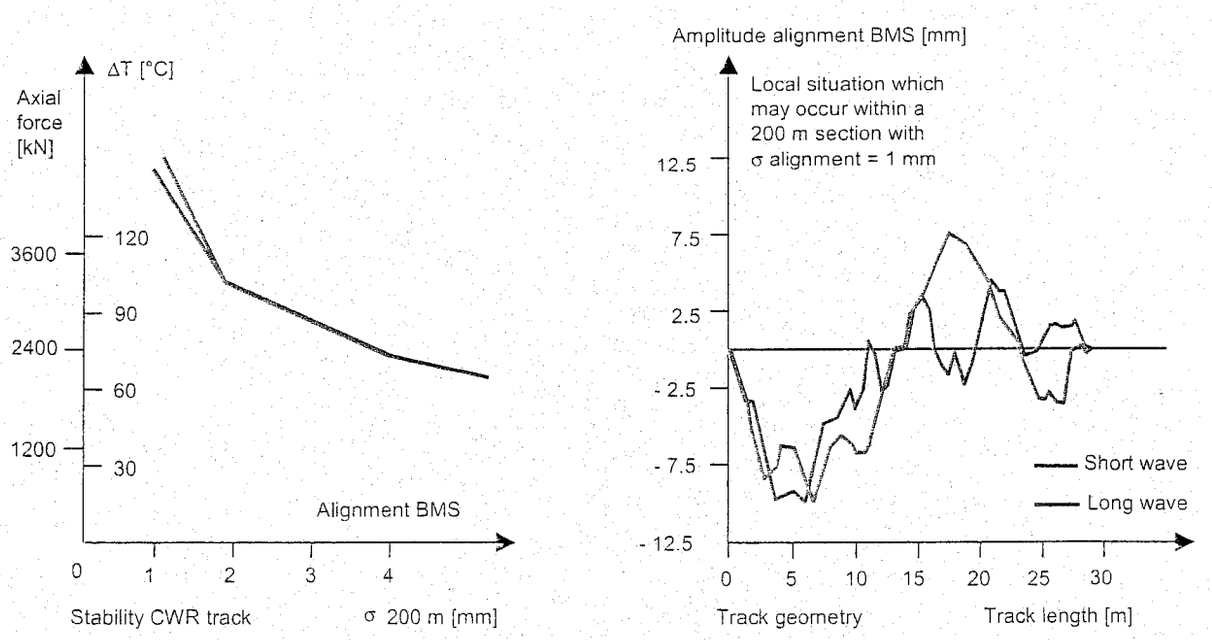


Figure 7.16: Buckling force versus track quality

ried out on NS UIC 54 track and on RET (Rotterdam Metro) NP 46 track. The FEM approximation and relevant parameters are also indicated. The correspondence between the measured and calculated value is quite good.

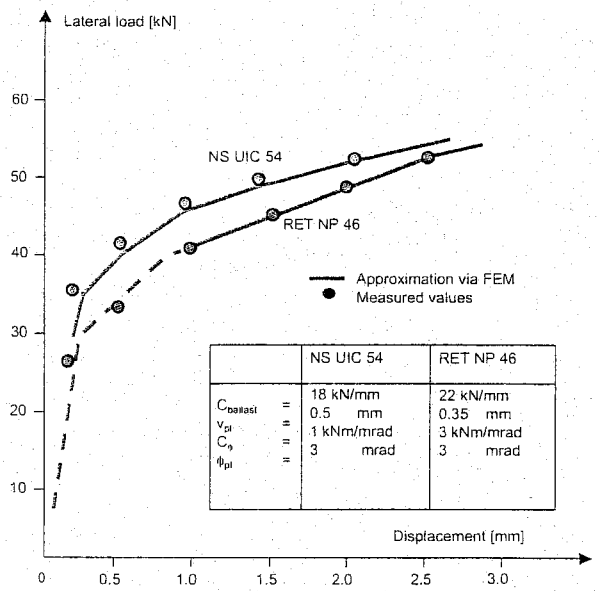


Figure 7.17: Approximating lateral resistance measurements by FEM

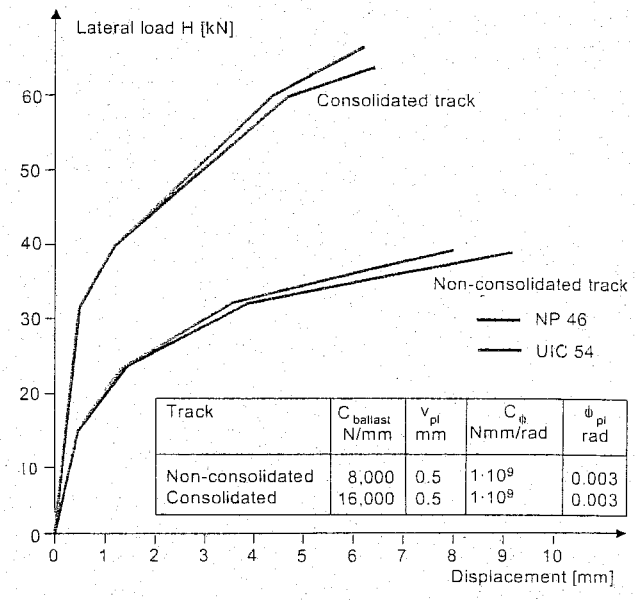


Figure 7.18: Force-displacement characteristics for different railtypes and different stiffness coefficients

To illustrate the influence of different spring characteristics, the load-displacement diagram is plotted in Figure 7.18 for tracks with NP 46 and UIC 54 rail profiles. The influence of the rail profile is relatively low. The consolidated and non-consolidated characteristics are FEM approximations of the lateral resistance measurements mentioned earlier.

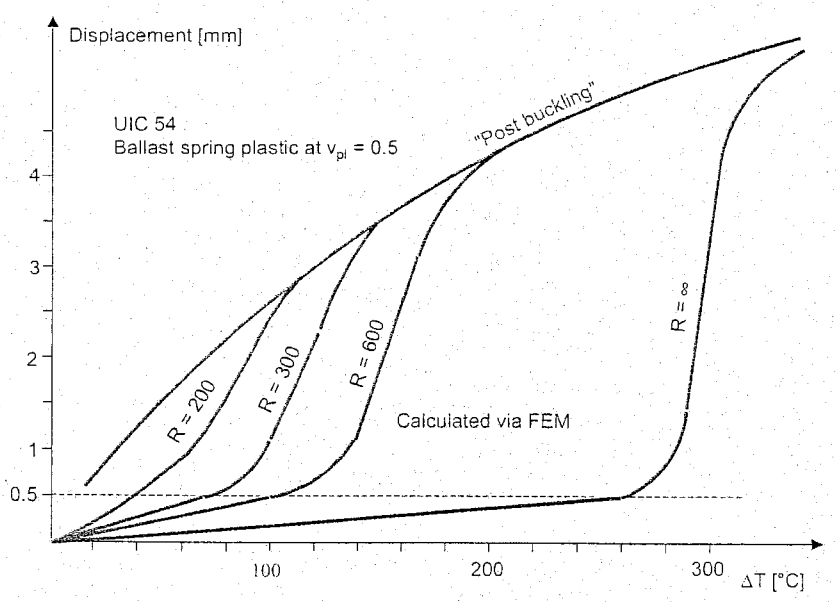


Figure 7.19: Lateral displacement versus differential temperature for different curve radii

On the basis of the non-consolidated characteristic in Figure 7.18, the relationship between lateral displacement and differential temperature has been calculated for different curve radii. The bi-linear ballast spring force has a maximum value of 8.5 kN which is attained at a displacement of v = v_{pl} = 0.5 mm. The FEM analysis was carried out with up to about 1 mm lateral displacement.

The post-buckling behaviour was investigated for only a few cases. The results, presented in Figure 7.19, reveal that the tighter the curve the more gradual the transition to the post-buckling phase develops.

Regarding NP 46 track, the curve radius versus differential temperature, at which the track starts to displace permanently, was also determined for consolidated tracks. For both NP 46 and UIC54 tracks the load-displacement curve for the NP 46 track, shown in Figure 7.17, was applied. The analysis results are shown in Figure 7.20. These values indicate that consolidated tracks in curves with a radius of 200 m will start to develop permanent lateral displacements at differential temperatures above 80°C for UIC54 and 100°C for NP46. Non-consolidated tracks start to displace permanently at about half this value. It should be mentioned that the occurrence of a limited permanent displacement should be considered as permissible.

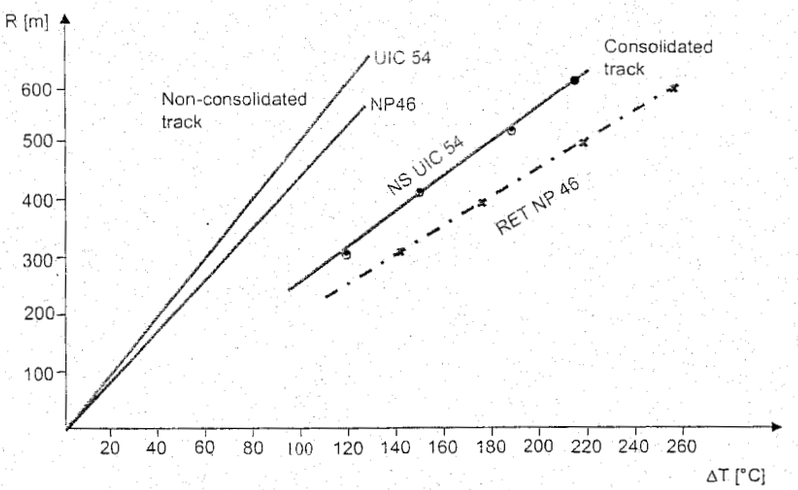


Figure 7.20: Curve radius versus temperature at which permanent track displacements start to occur

7.2.4 Continuous welded switches

In the case of continuous welded switches, the axial forces in the rails connected to the common crossing should be relieved by the sleepers so that they have vanished at the start of the switch blade. Depending on the sleeper stiffness, these forces are partly transferred to both continuous rails and partly to the ballast bed. Sleeper loads and corresponding displacements have been measured by DB for a 1:12 switch. These results, obtained from reference [146], are presented in Figure 7.21.

As a result of the reaction forces exerted on the continuous rails the axial compression force in these rails is reduced in the vicinity of the common crossing. Near the front end of the switch blades the axial compression forces in the continuous rails increase. According to Figure 7.21 this increase may amount to approximately 40%. DB have also carried out measurements at 1:12 crossovers. Owing to the limited transition length, an increase in axial compression force of only 7% was found.

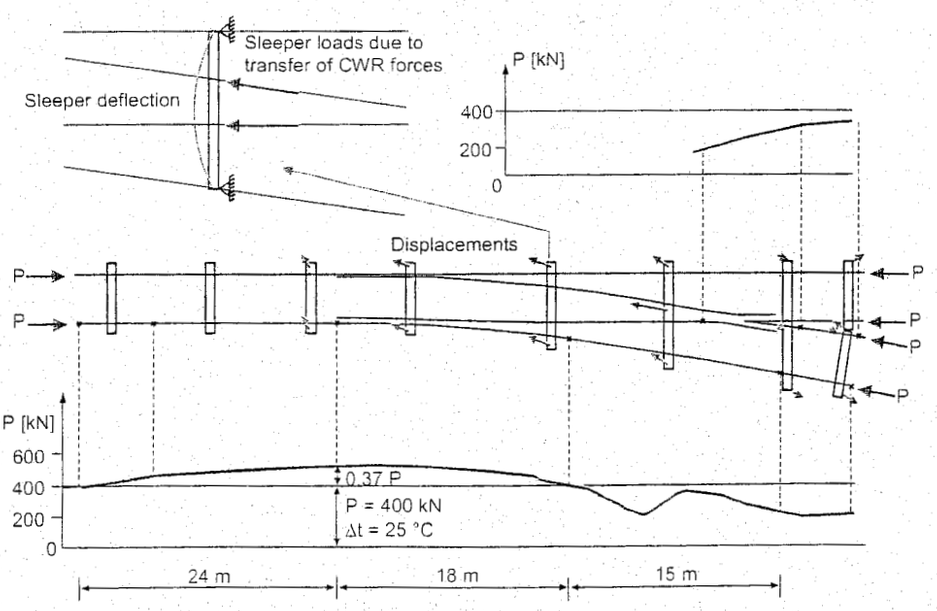


Figure 7.21: Forces in CWR 1:12 switch measured by DB

7.3 Longitudinal forces: analytical modelling

7.3.1 General considerations

In the previous section lateral track stability under the influence of longitudinal compressive forces was discussed. These forces consisted of temperature loads. There are, however, also other types of track load in the longitudinal direction. These loads originate from braking and acceleration forces. On bridges and viaducts differences in temperature between track and fixed installation may also occur. Obviously, rails respond much faster to temperature changes than solid concrete and steel constructions do causing relative movements between the parts.

From the point of view of track design, two load classes may be distinguished: longitudinal compression forces at high temperatures combined with traction forces which can introduce track buckling; longitudinal tensile forces, associated with low temperatures, in combination with traction forces which may easily lead to rail failure. In the design of bridges and viaducts the tensile load cases are normally the determining factor.

7.3.2 Axial rail model

As was already remarked in Chapter 4, free expansion or contraction of the rail due to temperature differences does not occur. A longitudinal resistance force distribution emanating from the rail fixation limits the rail displacements, but may cause high forces.

To investigate this problem, a simple yet instructive model is presented to determine the axial displacement of the rail under temperature load as well as a shear force distribution representing the longitudinal resistance.

In Figure 7.22 a small rail element is given which is subjected to a temperature increase ΔT with respect to the neutral or initial temperature, as well as a shear force $\tau(u)dx$ which opposes the displacement $u(x)$. The longitudinal force in the rail is $N(x)$.

Equilibrium demands:

$$dN = \tau(x)dx \tag{7.51}$$

The total strain is the sum of temperature strain and strain according to Hooke's law:

$$\frac{du}{dx} = \alpha\Delta T + \frac{N}{EA} \tag{7.52}$$

in which:

E = Young's modulus for the rail [N/mm²];
A = total cross-sectional area of the rail [mm²];
α = linear expansion coefficient of rail steel [/°C];
ΔT = change in temperature (defined as $\Delta T = T_{actual} - T_{initial}$).

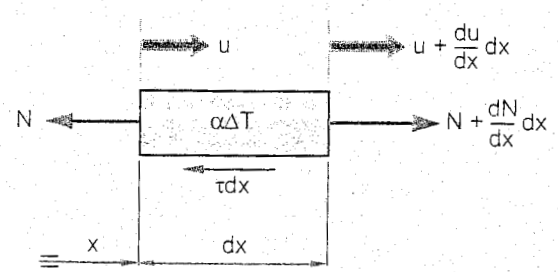


Figure 7.22: Differential rail element meant to study the temperature effect

In general, τ is a function of u, which in turn, is a function of x. Using the formulas (7.51) and (7.52), the following differential equation can be derived:

$$\frac{d^2u}{dx^2} - \frac{\tau(u)}{EA} = 0 \tag{7.53}$$

Once the displacement function u(x) is found, the normal force follows from (7.52):

$$N = EA\left(\frac{du}{dx} - \alpha\Delta T\right) \tag{7.54}$$

In the case of continuous welded rail (CWR), the length of the rail is so great that a plane strain condition exists in the central part of the rail preventing axial displacement of the rail completely. The force caused by a temperature increase is, according to (7.54):

$$N_{max} = -EA\alpha\Delta T \tag{7.55}$$

which means that a temperature increase results in a compressive normal force, as could be expected.

Special case 1: Plastic shear resistance

Sometimes it may be assumed that the shear resistance is constant while the sign of the resistance depends only on the sign of the displacement:

$$\tau = \tau_0 sign(u) \tag{7.56}$$

At the point of transition from CWR to jointed track, according to Figure 7.23, the CWR force is reduced over the so-called breathing length. From (7.53) and (7.56), it follows for $x \geq 0$ that:

$$u(x) = \frac{\tau_0 x^2}{2EA} \tag{7.57}$$

which satisfies the transition conditions at x = 0. With (7.54) the normal force in the breathing length zone is:

$$N = \tau_0 x - EA\alpha\Delta T \tag{7.58}$$

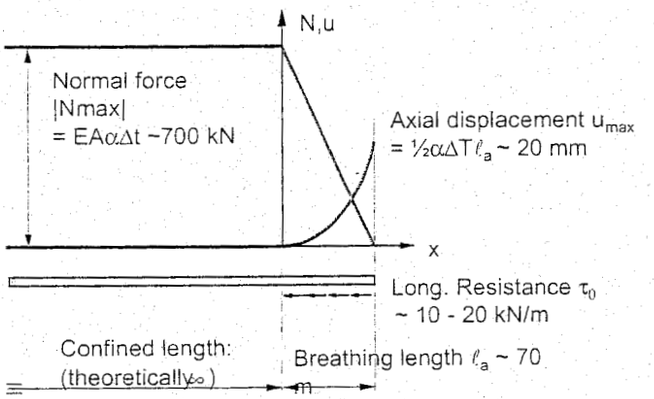


Figure 7.23: Distribution of temperature force and displacement in CWR (plastic shear resistance)

At the free end of the rail ($x = \ell_a$) there is no force, hence the breathing length:

$$\ell_a = \frac{N_{max}}{\tau_0} \tag{7.59}$$

The expressions for the maximum normal force and maximum axial displacement are also indicated in Figure 7.23.

The results in this case are very plausible and may in fact be deduced heuristically. This is the solution often found in text books.

As the plastic resistance characteristic is inherently non-linear, residual stresses remain at the rail ends if the rail system is subjected to more than one temperature cycle, e.g. consecutive warming up/ cooling down periods. In this case the rail end displacement follows a hysteresis loop.

Special case 2: Elastic shear resistance

In the case of elastic rail fastening systems the assumption of a linear shear resistance is more appropriate:

$$\tau = ku \tag{7.60}$$

Substitution in (7.53) results in the differential equation:

$$\frac{d^2u}{dx^2} - \frac{k}{EA}u = 0 \tag{7.61}$$

The solution of (7.61) which satisfies the boundary conditions: x = 0 => N = 0 and x $\rightarrow \infty$ => u = 0 can be written as:

$$u = -\frac{\alpha \Delta T}{\mu} e^{-\mu x} \tag{7.62}$$

in which:

$$\mu = \sqrt{\frac{k}{EA}} \tag{7.63}$$

The normal force according to (7.54) becomes:

$$N = -EA\alpha\Delta T(1 - e^{-\mu x}) \tag{7.64}$$

The expressions for the maximum normal force and maximum axial displacement are indicated in Figure 7.24. Theoretically, there is no distinction here between the breathing length and confined length as was the case with the plastic shear resistance. Practically, though, the large central part of the rail may be regarded as being in a plain strain situation. As this solution is linear elastic, no hysteresis effects are taking place when the system is subjected to consecutive temperature variations. In the next section the more complex case of temperature effects in the combined system track/viaduct will be examined.

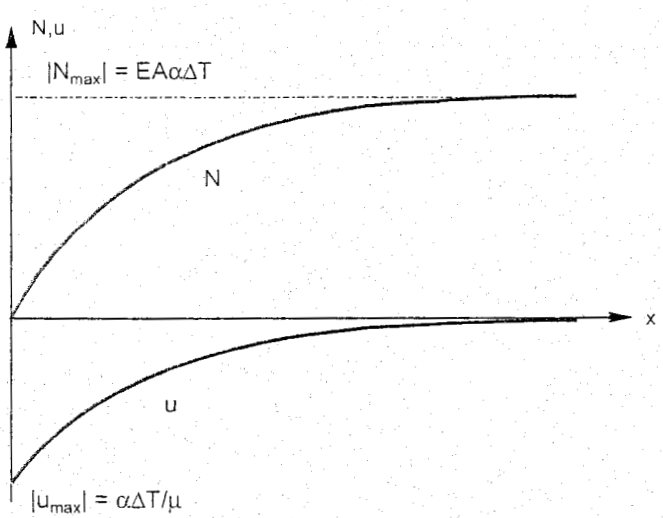


Figure 7.24: Distribution of temperature force and displacement in CWR (elastic shear resistance)

7.3.3 Modelling of the longitudinal interaction problem

In Section 7.3.2 the effect of temperature forces in rails was examined using a simple model (Figure 7.22). To assess the complex temperature effects in the system track on a bridge or a viaduct, we will use a more generalised model, shown in Figure 7.25.

Although there may be more than one track fixed on the bridge, we here only consider one continuous welded rail on a corresponding part of the whole bridge.

In analogy with the theory in Section 7.3.2, we can write down the following mixed equations taking into account the interaction of the longitudinal shear resistance between rail and bridge:

$$\frac{d^2u}{dx^2} + \frac{\tau(u_b - u)}{EA} = 0 \qquad (7.65)$$

$$\frac{d^2u_b}{dx^2} - \frac{\tau(u_b - u)}{(EA)_v} = 0 \qquad (7.66)$$

$$N = EA\left(\frac{du}{dx} - \alpha\Delta T\right) \qquad (7.67)$$

$$N_b = (EA)_b\left(\frac{du_b}{dx} - (\alpha\Delta T)_b\right) \qquad (7.68)$$

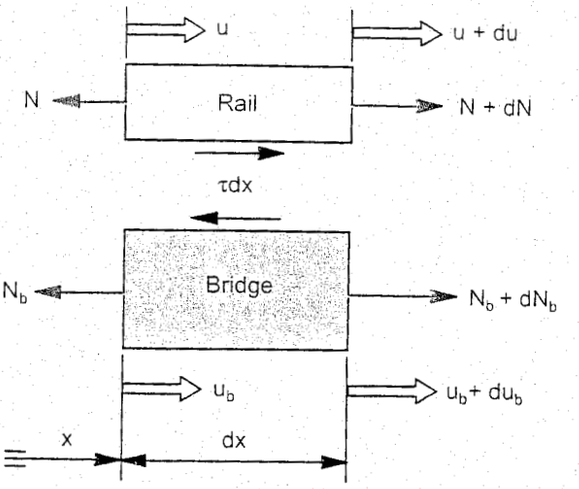


Figure 7.25: Rail and bridge element

in which:

u, u_b = displacement of the rail and bridge, respectively;
N, N_b = normal force in the rail and the bridge, respectively;
$\tau=\tau(u_b-u)$ = axial shear resistance, depending on the difference of the displacements;
EA, $(EA)_b$ = axial normal stiffness of the rail and bridge, respectively;
$\alpha\Delta T$, $(\alpha\Delta T)_b$ = temperature strain of the rail and bridge respectively.

To simplify matters only elastic displacements are assumed:

$$\tau = k(u_b - u) \qquad (7.69)$$

Moreover, it is assumed that the normal stiffness of the rail is much less than the corresponding part of the bridge:

$$EA \ll (EA)_b \qquad (7.70)$$

In this case the bridge exhibits an almost uniform axial expansion or shrinking and the general solution of (7.65) and (7.66) can be written as:

$$u = C_1 \sinh \mu x + C_2 \cosh \mu x + C_3 + C_4$$
$$u_b = C_3 x + C_4 \qquad (7.71)$$

in which:

$$\mu = \sqrt{\frac{k}{EA}} \qquad (7.72)$$

Consider two special cases:

case 1: $C_3 = 0$; $C_4 = 0$ => $u_b = 0$; $N_b = N_{max} = -(EA\alpha\Delta T)_b$

In this case the bridge is confined completely and the solution of the rail is, with appropriate boundary conditions, identical to the solution (7.55).

case 2: $C_3 = (\alpha\Delta T)_b$; $C_4 = 0$=> $ub = (\alpha\Delta T)_b x$; $N_b = 0$

The bridge now has a fixed point (support) at x = 0 and can for the rest expand freely.

In this case the complete solution for the rail part of the system follows from (7.71) first equation, and (7.67):

$$u = C_1 \sinh \mu x + C_2 \cosh \mu x + (\alpha\Delta T)_b x \tag{7.73}$$

$$N = EA\mu(C_1 \cosh \mu x + C_2 \sinh \mu x) + EA(\alpha\Delta T)_b - EA\alpha\Delta T \tag{7.74}$$

The last term on the right side in (7.74) represents the normal force in the confined continuous welded rail; the other terms represent the modification of the rail force due to the bridge interaction.

Several bridge configurations have been studied with appropriate boundary conditions using the equations (7.73) and (7.74). As an example we will examine the case of a continuous welded rail fixed on an ever repeating system of short bridges subjected to a temperature decrease.
The boundary conditions are in this case, because of continuity:

$$x = 0 \Rightarrow u(0) = u(\ell); \quad N(0) = N(\ell)$$

The solution of (7.73) and (7.74) is therefore:

$$u = \tfrac{1}{2}(\alpha\Delta T)_b \ell [\cosh \mu x - \coth \tfrac{1}{2}\mu\ell \sinh \mu x] + (\alpha\Delta T)_b x \tag{7.75}$$

$$N = \tfrac{1}{2} EA(\alpha\Delta T)_b \mu\ell [\sinh \mu x - \coth \tfrac{1}{2}\mu\ell \cosh \mu x] + EA(\alpha\Delta T)_b - EA\alpha\Delta T \tag{7.76}$$

In Figure 7.26 a numerical example is given using the following parameters:

E	$= 2.1 \cdot 10^8$	kN/m²;
A	$= 60 \cdot 10^{-4}$	m²;
α	$= 1.15 \cdot 10^{-5}$	1/°C;
ΔT	= 40	°C;
ΔT_b	= 10	°C;
k	= 7636	kN/m²;
L	= 48.7	m.

If no bridge interaction existed, the normal force in the rail would be N_{max} = 580 kN. However, due to the interaction the normal force at the support amounts to 142 kN (24 %) which is quite substantial. The maximum rail displacement is half of that of the maximum bridge displacement.

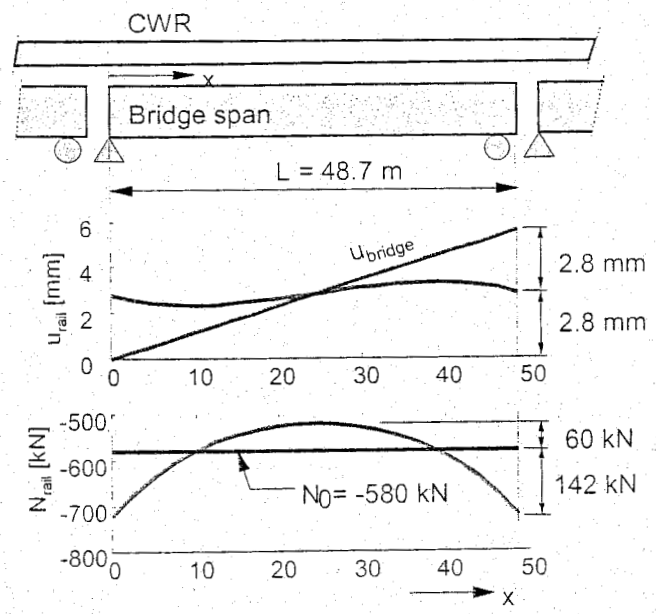


Figure 7.26: Longitudinal displacement and force in CWR track on a repeating bridge configuration

7.4 Longitudinal forces: finite element modelling

7.4.1 General considerations

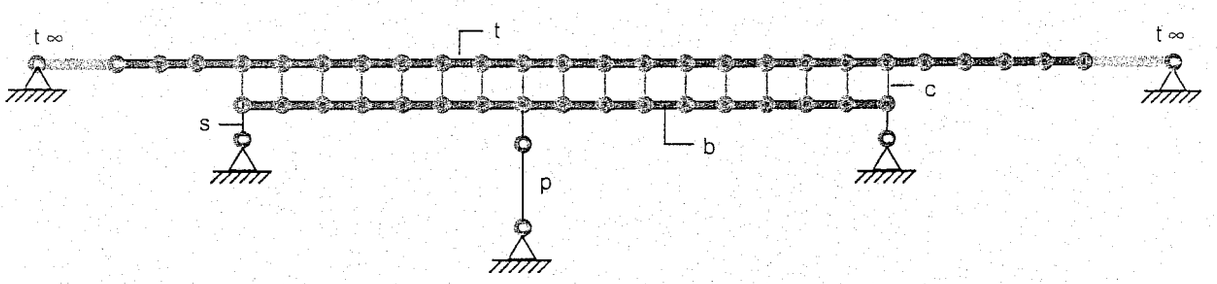
The relatively simple analytical approach, as discussed in Section 7.3, of the longitudinal problem is instructive in order to understand the temperature effects in the track. It should be noted, however, that the modelling used there is based on a number of limitations and assumptions, viz.:

- lateral bending stiffness EI is constant;
- lateral shear resistance is constant;
- compressive force P = constant;
- no vertical loading;
- no longitudinal resistance;
- no axial strain;
- misalignment sinusoidal;
- additional bending sinusoidal;
- no curves.

7.4.2 Finite element model

To obtain a more realistic description of the problem, a finite element model, called PROLIS [281], has been developed to calculate longitudinal track forces in a similar way to the model described earlier regarding track stability. This model comprises track elements, ballast elements, and elements representing the bridge construction including abutments and pillars. The model allows for an arbitrary number of parallel tracks. Figure 7.27 shows the element composition which can be used to model a track/bridge construction.

The ballast spring is, as in the case of the stability program discussed in Section 7.2.2, also modelled as a bi-linear spring according to Figure 7.28. The maximum force, i.e. the force at which yielding starts, depends on the current vertical track load. Two variants have been investigated to describe the plastic behaviour. The first one is sketched in Figure 7.29 and assumes that the elastic limit always coincides with a fixed displacement u_{pl}. This means an increase in ballast stiffness in accordance with a growing vertical track load. This assumption does not exclude discontinuities from occurring in the spring force if vertical loads are added or removed. The second variant, shown in Figure 7.30, consists of a spring with constant stiffness in relation to the vertical track load. In this case, the displacement at which plastic deformations start grows linearly with the track load. From the physical point of view this approach is more consistent. However, simulation tests have shown that there is no significant difference between the results of both methods as the displacements in the areas of interest are often substantially greater than u_{pl}.



○	Joint	c	Ballast element with Coulomb friction
t	Track element (2 rails)	s	Support element
t ∞	Element representing infinitely long track	p	Pillar element

Figure 7.27: Finite element model for calculation of axial forces in tracks and on bridges

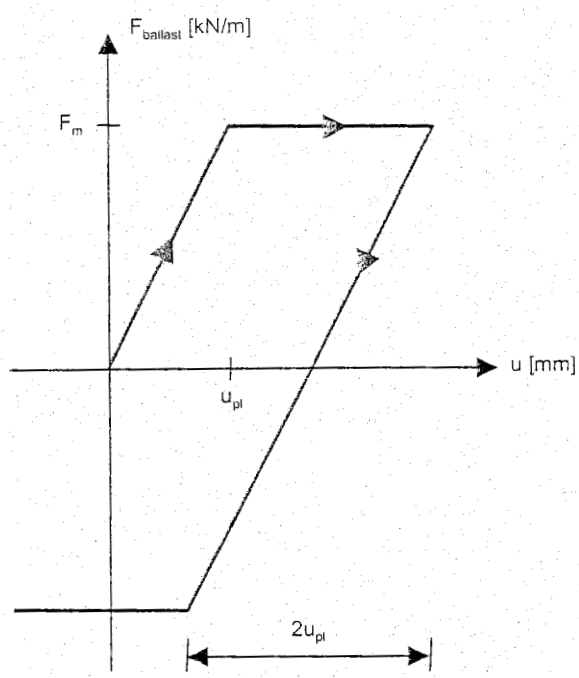


Figure 7.28: Bi-linear ballast element

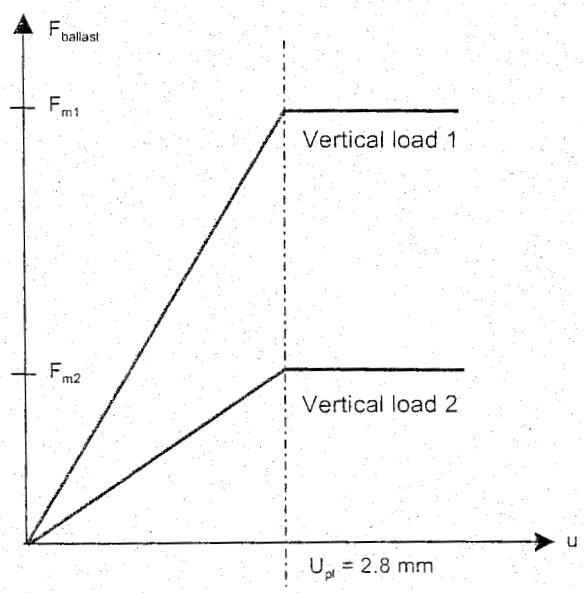


Figure 7.29: Ballast spring with stiffness proportional to the vertical track load

Both the maximum spring force and the relationship with vertical track load, of course, depend on ballast and sleeper type for ballasted tracks and on the type of fastening if a direct fastening system is used. The values used in the NS calculations are summarized in Figure 7.31 for ballasted track and direct fastenings. The loads are given in kN per metre track. Additionally, some DB values published in [116] are presented.

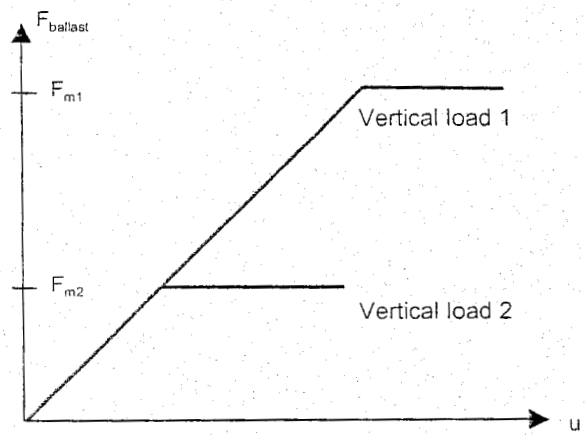


Figure 7.30: Ballast spring with constant stiffness relative to the vertical track load

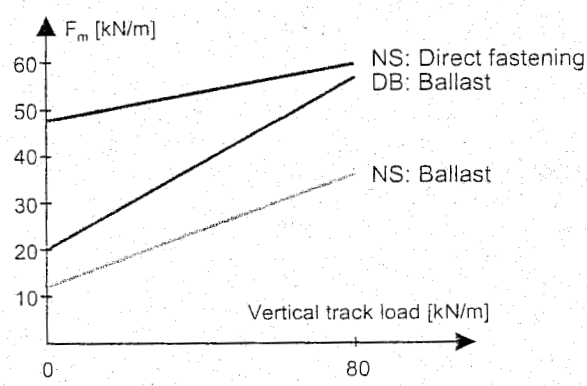


Figure 7.31: Maximum longitudinal force versus vertical track load

The loads consist of temperature loads which can be specified separately per element for track and bridge as the differential temperature relative to the neutral temperature and as braking and acceleration forces. The NS apply a braking load of 25% of the vertical track load. For D4 traffic this means 20kN/m. The acceleration force is 1000 kN over 30 m, i.e. 33 kN/m over 30 m.

The finite element calculation is basically the same as described previously for the stability analysis and, therefore, any further discussion here is superfluous.

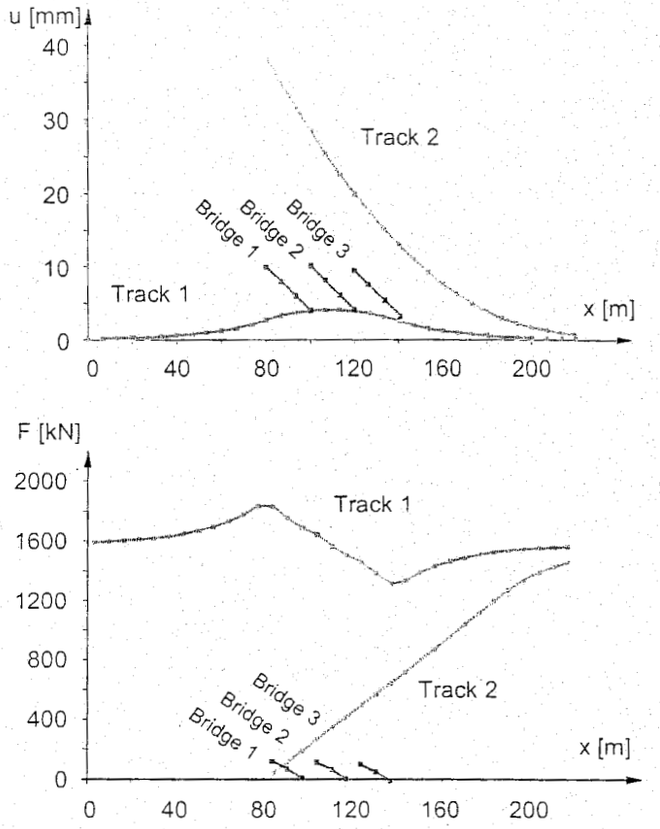


Figure 7.33: Track forces and track displacements resulting from the loads specified in Figure 7.32

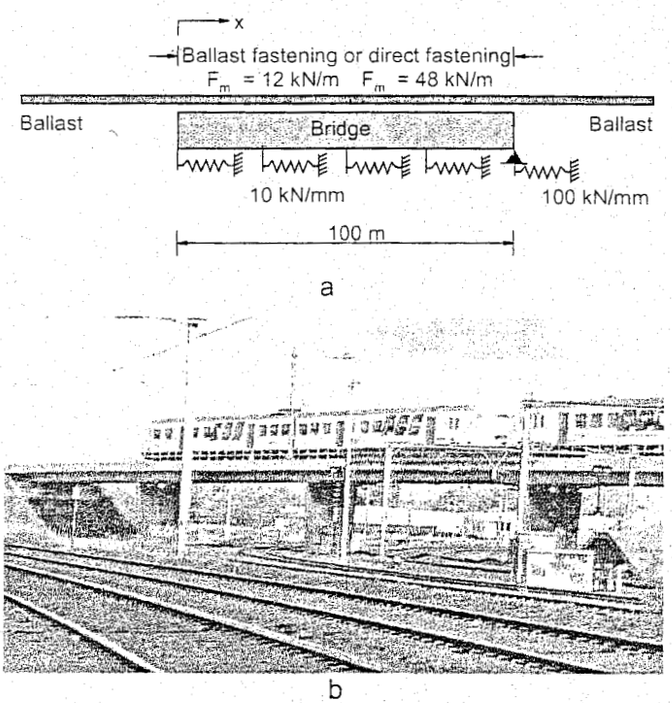


Figure 7.34: Utrecht flyover bridge

In the second variant, the elevated ballast stiffness of track 2 is applied immediately when the temperature load is raised. It was shown that this case leads to an underestimation of the displacements of the order of 37% and an overestimation of the force of the order of 3%.

In order to simulate a brittle rail failure, upon raising the temperature load track 2 was first considered to be long-welded without expansion joints. When applying the braking loads, the rails of track 2 were assumed to be broken, i.e. having expansion joints. The results deviate by less than 1% from the earlier results presented in Figure 7.33. Dynamic effects due to rail fracture have not been taken into account.

Utrecht flyover bridge

This example concerns a 100 m long flyover bridge carrying a single long-welded track. The support conditions of the bridge are presented in Figure 7.34a while Figure 7.34b shows a picture of the bridge. The track is subjected to a temperature load of $\Delta T = -45°C$ and the bridge to $\Delta T = -25°C$.

In this case 6 alternatives were considered which are described in Table 7.1. They consist of CWR without expansion joints, CWR with expansion joints at the left end of the bridge, and the expansion joints replaced by fastenings with teflon pads allowing for a relative displacement between rails and sleeper over a short length of track. For the maximum longitudinal force, referred to as F_m, frozen ballast, normal ballast and direct fastening conditions were considered.

The results of these calculations are presented in Figure 7.35 and Figure 7.36, showing the axial rail forces and the axial rail displacements respectively. The peak stresses and displacements are summarized in Table 7.1. Without expansion joints the frozen ballast, combined with direct fastenings on the bridge, causes the highest rail stresses of the order of 180 N/mm². This means an increase of over 60% compared to the undisturbed temperature stress. In the case of normal ballast conditions and direct fastenings on the bridge, the maximum stress drops 8% compared to the frozen situation. Obviously, a continuous ballast bed smooths the peak stress substantially as is demonstrated in load case 4.

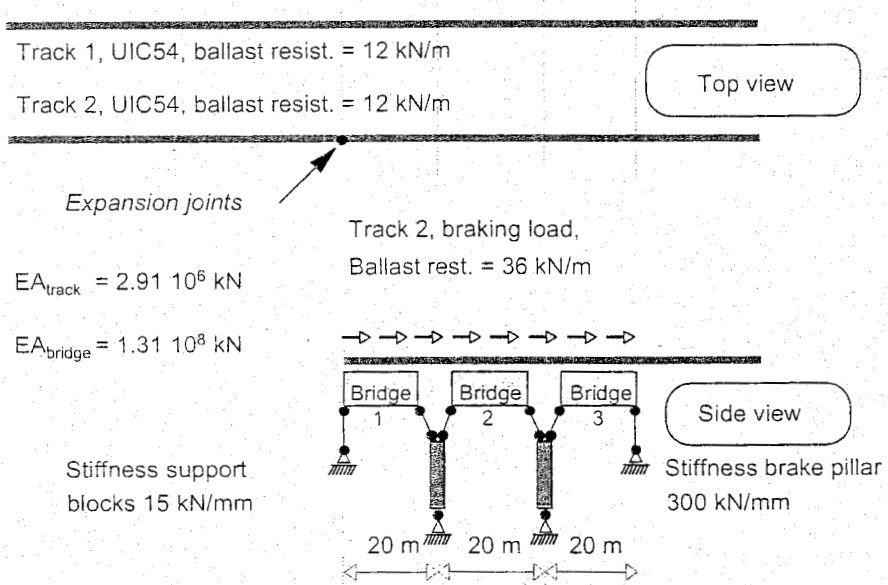
7.4.3 Examples of longitudinal force calculations

Bridge in Amsterdam West Branch

This construction consists of 3 bridges with a length of 20 m each. The bridges carry 2 ballasted tracks. The tracks are continuous welded although track 2 is provided with expansion joints at the beginning of the first bridge. The temperature loads consist of $\Delta T = -45°C$ for the tracks and $\Delta T = -25°C$ for the bridge. A braking force of 8kN/m is applied to track 2. The situation is sketched in Figure 7.32. The maximum ballast force for the non-loaded tracks is taken as 12 kN/m and for the loaded tracks 36 kN/m is used. It is assumed that the temperature loads are first applied during which all tracks have the same ballast yield force. Subsequently, the ballast yield force for the loaded part of track 2 is raised and the braking loads are applied.

The resulting track forces and track displacements are presented in Figure 7.33. Obviously, the largest displacements are achieved at the expansion joints in track 2. At this location, by definition, the longitudinal force is zero in track 2 and has its maximum value in track 1. The longitudinal force of track 2 has been partly transferred to track 1 by means of the ballast and bridge elements. This effect shows great similarities to the force transfer observed in the switch discussed in Section 7.2.4.

The calculation was repeated for two variants. In the first variant the spring characteristic, according to Figure 7.29, was replaced by a ballast spring with constant stiffness as described in Figure 7.30. Comparing the results for the displacements of track 2 and the forces of track 1 revealed that the spring characteristic with respect to the vertical load does not influence the calculated forces. The displacements are 6% higher.



***1st loading step** : temperature load tracks*
Δt = - 45 °C, bridges : Δt = -25 °C

***2nd loading step** : braking load 8 kN/m on track 2 above the bridges, ballast resistance under vertical load = 36 kN/m*

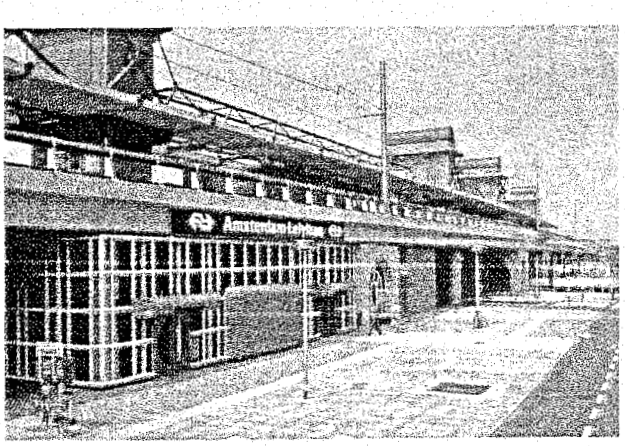


Figure 7.32: Loads on bridges in Amsterdam West Branch

loading: track: ΔT = -45 °C bridge: ΔT = -25 °C					
load case	axial yield force		computational results		
	plain track	bridge	maximum rail stress [N/mm²]		maximum differential displacement [mm]
1	24 kN/m (frozen ballast)	48 kN/m	181	–	
2	12 kN/m	48 kN/m	167	8%	
3	same as 2, but with teflon sliding facilities over 7.5 m at x = 0		163	10%	
4	12 kN/m	12 kN	138	24%	
5	as 4, but with expansion joints at x = 0		113	37%	74
6	as 2, but with expansion joints at x = 0		94	48%	69

Table 7.1: Different alternatives and results for Utrecht flyover bridge.

The stress overshoot is reduced by about 50% compared to the direct fastening system. The penalty, of course, is a greater construction height and increased dead weight.

In the case of expansion joints, according to load cases 5 and 6, the axial forces by definition drop to zero at the joints. However, the displacements are quite considerable, namely of the order of 70 mm. The rails on the continuous ballast bed displace by about 7% more than in the direct fastening solution. One of the major drawbacks of the use of expansion joints is the introduction of high detrimental impact loads. As an alternative, sliding facilities between rail and sleeper were created over 7.5 m. From load case 3, which refers to this situation, it can be deduced that the peak stress is reduced by about 1% per metre of released track.

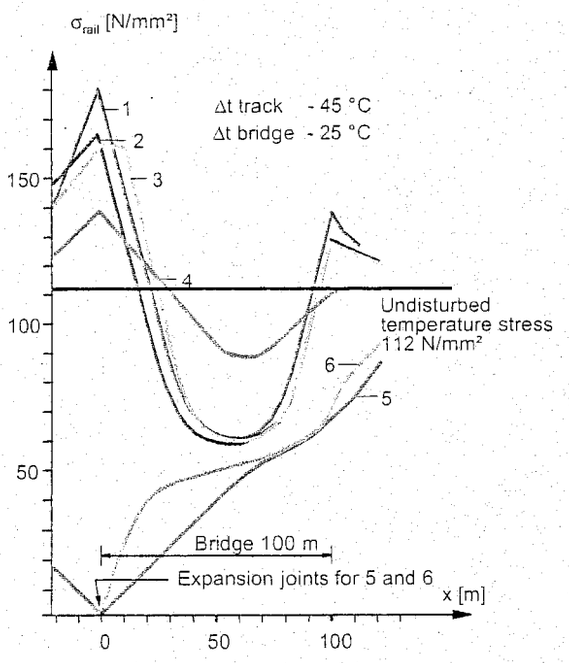


Figure 7.35: Rail stresses due to temperature load in the case of continuous ballast bed, direct fastening, and expansion joints

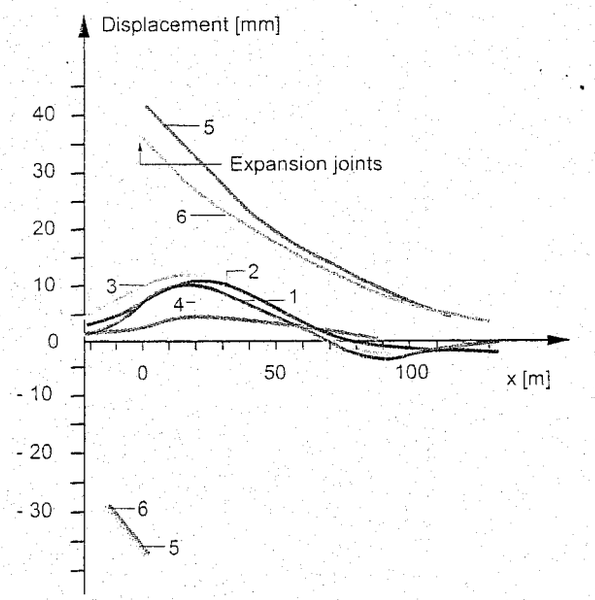


Figure 7.36: Displacements due to temperature load in the case of continuous ballast bed, direct fastening, and expansion joints

The major parameter influencing the stress increase at the bridge end is the length of the bridge. In the previous example the length was 100 m. The analyses were repeated for 50 m and 200 m and the results are presented in Figure 7.37. The rail stresses increase linearly with the logarithm of the bridge length. To achieve the total stress, the values in Figure 7.37 should be increased by about 6% to take account of the braking load: about 100N/mm² for the vertical track load and 150 to 300N/mm² for the residual rail stresses. On the basis of these assumptions, the permissible stresses for several rail grades are also indicated.

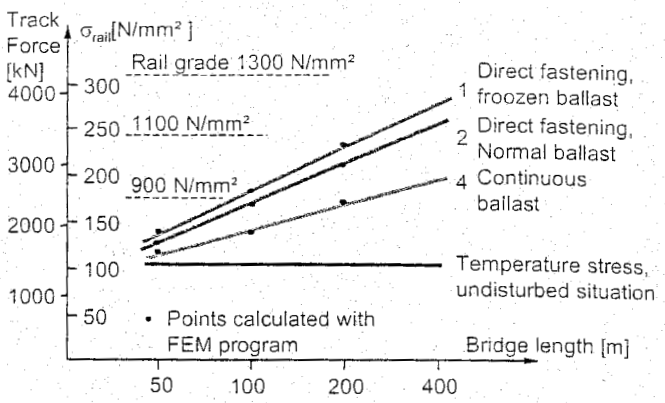


Figure 7.37: Peak stress versus bridge length

To conclude this section on longitudinal forces, it is worthwhile to mention that, in the case of compression loads, both finite element programs discussed here can be used in cascade. The longitudinal force program can determine the load distribution along the track or in a switch. This distribution can be used as input in a subsequent analysis using the stability program to verify whether or not the lateral track resistance is exceeded.

7.5 Advanced numerical models of track buckling

7.5.1 Introduction

The above-mentioned examples help to understand the mechanism of track buckling using simple mathematical means. It should be noted, however, that the theory is based on a number of limitations and assumptions, viz.:

- lateral bending stiffness EI is constant;
- lateral shear resistance is constant;
- compressive force P = constant;
- no vertical loading;
- no longitudinal resistance;
- no axial strain;
- misalignment sinusoidal;
- additional bending sinusoidal;
- no curves.

In order to get a more accurate assessment of the safety limits of CWR track, a lot of research has been carried out recently by the ERRI Committee D202, 'Improved knowledge of forces in CWR track'. The theoretical part of this research consisted of a finite element method called CWERRI (acronym for continuous welded (Rail); European Rail Research Institute). This program was developed at the TU Delft and describes the mechanical behaviour of the railway track much more realistically and can cope with complex situations. The program was based on a discrete element program

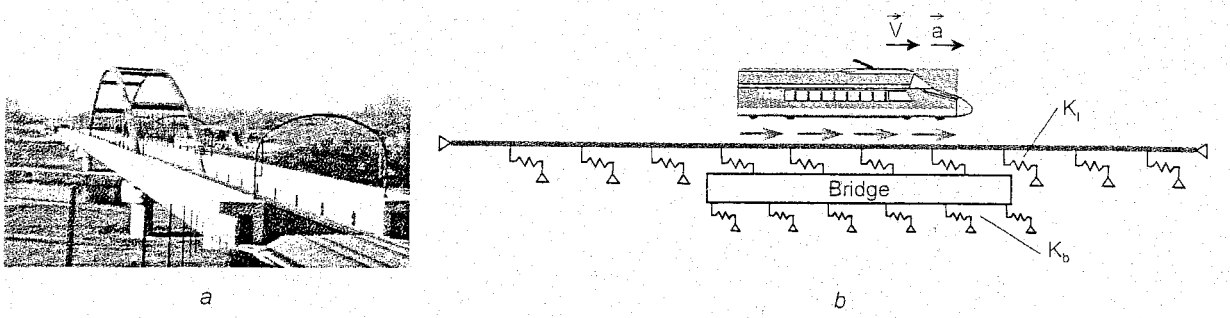


Figure 7.40: Classical track on bridge (a) and its FE model built using CWERR (b). Loading cases: temperature variation and braking of train

The rails and bridge are modelled by beam elements, whereas the longitudinal behaviour of ballast and fasteners is described by spring finite elements (stiffness). Since the bridge itself can also move in the longitudinal direction, spring elements under the bridge (stiffness K_b), representing the longitudinal stiffness of bridge supports, have been introduced. It should be noted that other track structures such as slab track with direct fastening or embedded rail structure on a bridge can be modelled as well by adjusting the spring stiffness K_l.

The effect of the train's braking/accelerating is modelled by distributed loads which are applied to the rails over the length of a train as shown in Figure 7.40. In order to simulate the behaviour of the structure caused by a temperature variation, thermal loads are applied to the rails and bridge. Another feature of the 'track on bridge' model is the possibility to take the effect of eccentricity in the top and centre of the bridge deck into account as shown in Figure 7.41.

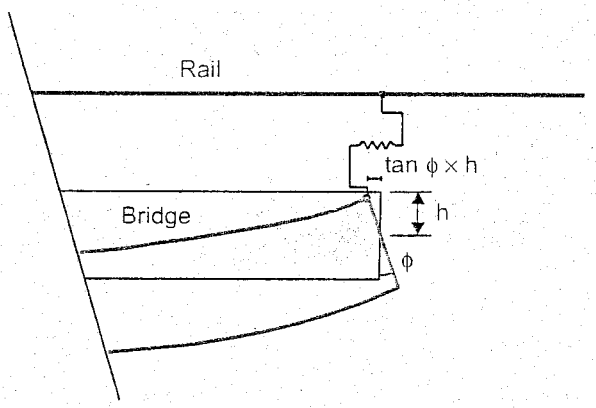


Figure 7.41: Eccentricity of top and centre of bridge deck

Figure 7.42 shows numerical results of a study case regarding a classical track on a concrete bridge (Figure 7.40).The bridge has a length of 125 meters and is subjected to the vertical loads due to braking of a train (along the track length of 20 meters) and temperature variation (35° C for rails and 20° C for a bridge). The results in Figure 7.42 contain longitudinal displacements of rails and bridge as well as normal forces occurring in rails. By analysing the resulting displacements and stresses, i.e. checking whether the maximum allowable values of displacements and stresses of rails and other components of a track have been exceeded, the quality and acceptance of a track design can be estimated. Practical applications of such a model include analyses of flyover bridges in Utrecht and Maartensdijk in the Netherlands [282].

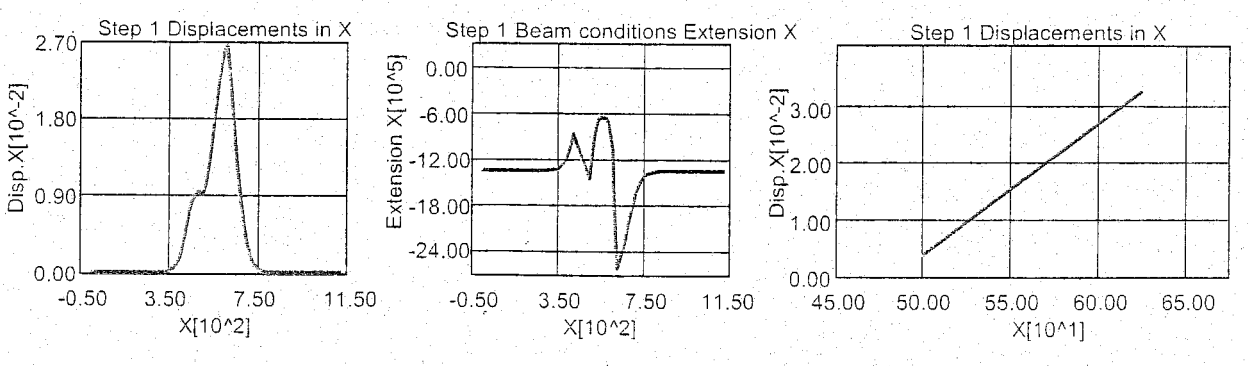


Figure 7.42: Longitudinal displacement (a) and force (b) in rail, and longitudinal displacement of bridge (c) caused by temperature variation and braking of train

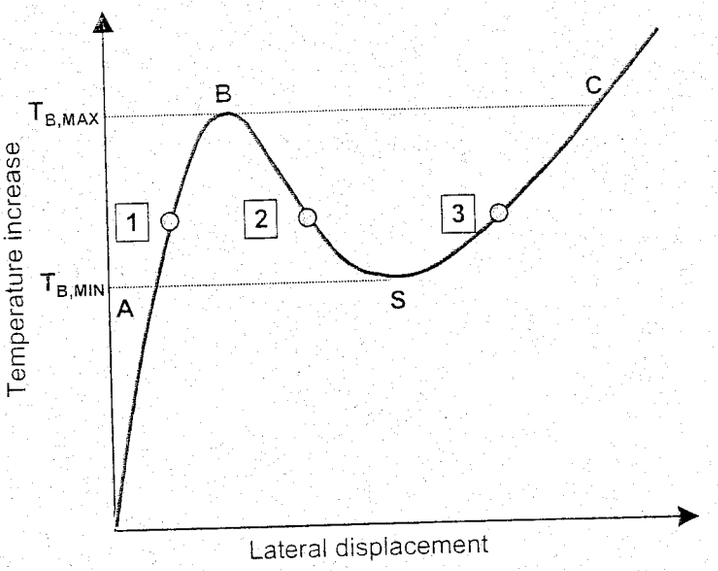


Figure 7.46: Buckling response characteristic

All the equilibrium configurations corresponding to points along BS (branch 2) can be shown to be unstable, and along SC (branch 3) as stable. The path BC indicates a possible snapthrough from the pre-buckling to the final post-buckling position.

The temperature corresponding to the minimum point S will be defined as the lower buckling temperature increase $T_{B,MIN}$. In contrast, the temperature at the maximum point B will be called the upper buckling temperature $T_{B,MAX}$. The significance of the point S is that it is the lowest buckling temperature at which the track can buckle out from the point A if it is given sufficient external energy or disturbance at the point A. Likewise, the track can buckle at any point between A and B. This requires less energy than the temperature increase from A to B. The track will buckle out at B with no external energy, as the theory and experiments have shown. Such a degree of stability is of no use in engineering, and, therefore, it is not possible to operate safely on CWR tracks at $T_{B,MAX}$.

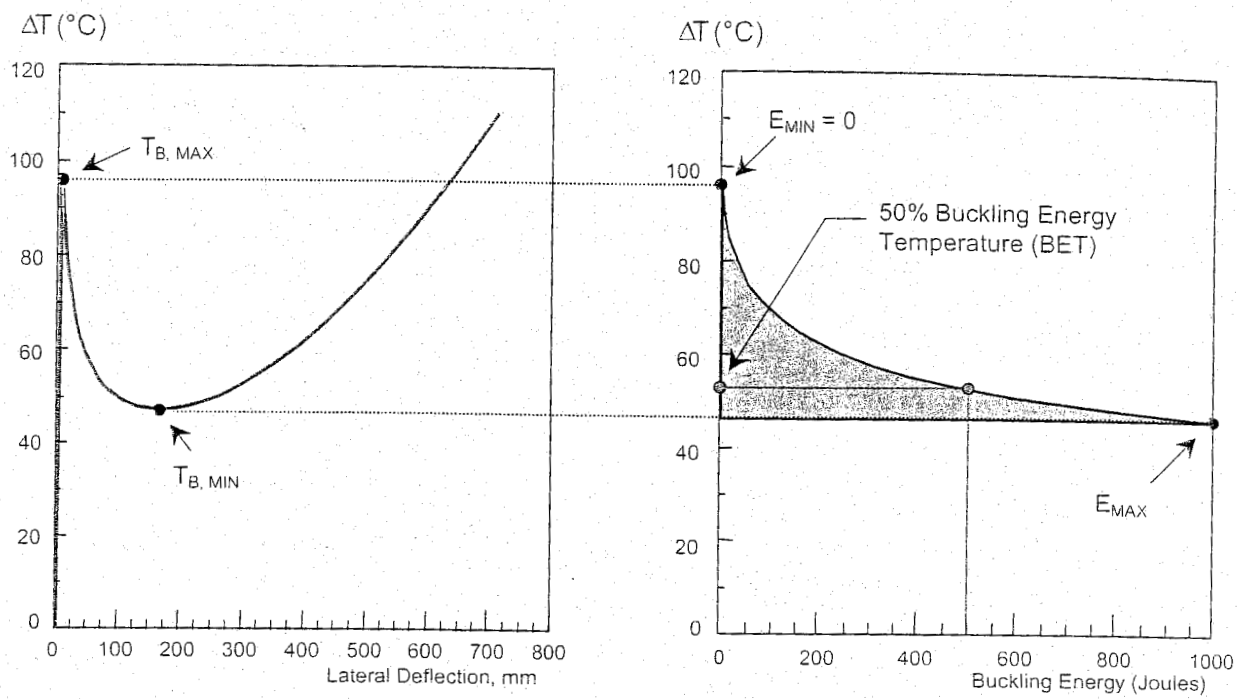
An important feature of sudden explosive buckles is that the rail force drops in the buckled zone compared to the pre-buckled value. This is caused by the large lateral displacement contributing to the rail extension that releases some of the load.

7.5.7 Approach in order to determine the allowable temperature T_{ALL}

The buckling response curves provide the basis for the allowable temperature. The allowable temperature T_{ALL} can be based on either the Lower Critical Temperature, $T_{B,MIN}$, or on a temperature above $T_{B,MIN}$ (but less than $T_{B,MAX}$) depending on the "Levels of Safety desired", i.e.:

- Level 1 Safety: $T_{ALL} = T_{B,MIN}$;
- Level 2 Safety: $T_{ALL} = T_{B,MIN} + \Delta T$.

Level1 tends to be more conservative, i.e. "safer", than Level 2. The choice or determination of the ΔT value is determined based on safety considerations and is not a trivial matter, for buckling potential increases rapidly with temperature above the $T_{B,MIN}$ value. This is based on research which shows that the buckling energy sharply decreases from a maximum value at $T_{B,MIN}$ to zero at $T_{B,MAX}$. Figure 7.47 illustrates the buckling energy decrease as a function of temperature above $T_{B,MIN}$.



Parameters: Rail: UIC60; Sleeper type: Concrete; Torsional resistance: Medium; Longitudinal resistance: Medium; Misalignment: 12 mm in 8 m; R = 300 m; Axle loads: UIC/D4.

Figure 7.47: Buckling energy concept illustration

Approach 1 for determining ΔT

In this approach, the buckling energy versus temperature increase relationship is used as a criterion for choosing ΔT, e.g. Level 2. Safety is based on an allowable temperature which corresponds to a temperature at which a finite buckling energy exists that is larger than zero but less than the maximum value at $T_{B,MIN}$. Determination of the buckling energy is based on a program called CWR-BUCKLE from US DOT (Department Of Transportation). Research to date suggests using the 50% Buckling Energy Level (BEL):

$T_{ALL} = T_{50\%BEL}$

Approach 2 for determining ΔT

If the CWR-BUCKLE model is not available for determining the buckling energy, an alternative definition of ΔT may be based on the model prediction of $T_{B,MAX}$ and $T_{B,MIN}$. The program CWERRI can be used to determine these levels. This safety concept was recently incorporated into UIC Leaflet 720 through ERRI D202. The results can be summarized as follows:
For all CWERRI calculations: first calculate $\Delta T = T_{B,MAX} - T_{B,MIN}$:

if $\Delta T > 20$ °C: $T_{ALL} = T_{B,MIN}$ + 25% of ΔT;
if 5 °C < $\Delta T < 20$ °C: $T_{ALL} = T_{B,MIN}$;
if 0 °C < $\Delta T <$ 5 °C: $T_{ALL} = T_{B,MIN}$ - 5 °C;
if $\Delta T <$ 0 °C: Not allowable in main lines.

In the last case progressive buckling (PB) occurs which means that elastic and plastic lateral deformation easily fade into each other. PB is common in low ballast quality structures.

Figure 7.48: Study case: stability of tram curve track

7.5.8 Study case: Stability of tram track

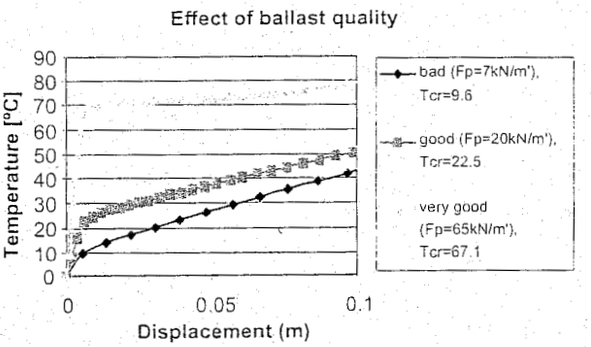
In the following example the stability of a tram track on curve is investigated (see Figure 7.48). A tram track turn is characterised by a relatively small radius of curvature. Therefore, in case of loss of stability, a progressive buckling of a track occurs. To analyse the track structure, the model given in Figure 7.43 has been used. The parameters of the model are given in Figure 7.49. The effect of some of a track structure's properties, such as quality of ballast (Fp), radius of curvature (R), and amplitude of misalignment (h), on its stability is shown in Figure 7.49. From these results it can be deduced that quality of ballast, which is defined by the degree of compactness of ballast particles, is crucial for track stability. As expected, the chance of buckling is higher on sharper (smaller radius of curvature) curve tracks for ballast of the same quality. Misalignment of the rail also has a negative effect on the track stability.

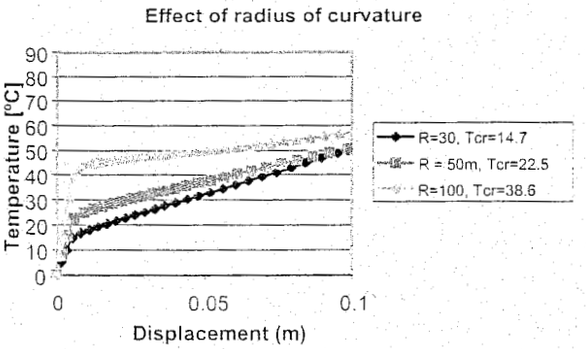
Track geometry

R = 50 m λ = 8 m h = 0.01 m

Ballast properties

Fp = Fl = 20 kN/m' Kl = 5 MN/m/m'
Wp = 0.005 m Kv = 70 MN/m/m'
Wl = 0.02 m Kt = 141 kN/m/rad/m'
φ = 0.86





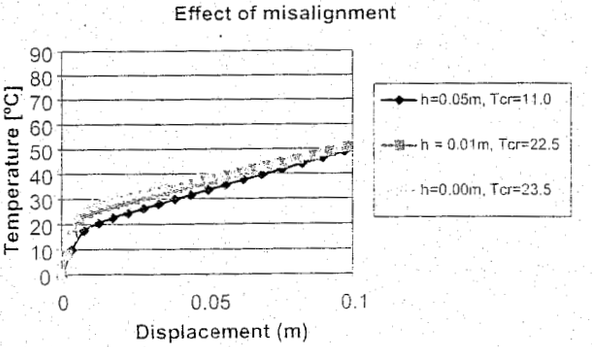


Figure 7.49: Effect of track parameters on its stability

8 BALLASTED TRACK

8.1 Introduction

This chapter deals with the principles according to which ballasted track, also called 'classical track' or 'conventional track', is constructed. A detailed discussion of every type of track structure and its variants is beyond the scope of this book. Only a few examples will be given with the main intention of illustrating the principles.

The classical railway track basically consists of a flat framework made up of rails and sleepers which is supported on ballast. The ballast bed rests on a sub-ballast layer which forms the transition layer to the formation. Figure 8.1 and Figure 8.2 show the construction principle of the classical track structure. The rails and sleepers are connected by fastenings. These components and other structures such as switches and crossings are all considered as part of the track. The particulars of switches and crossings are discussed in Chapter 11.

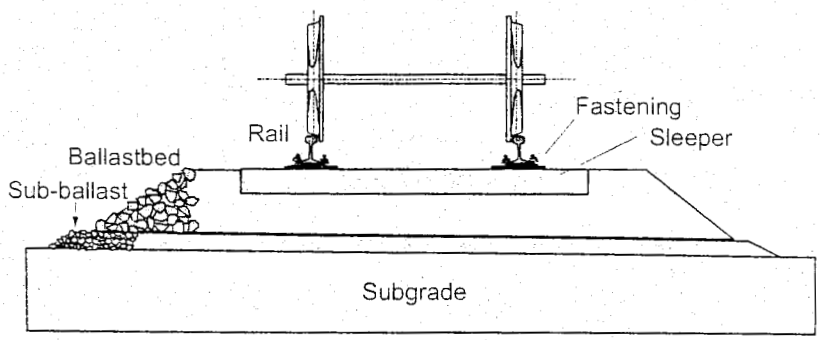


Figure 8.1: Principle of track structure: cross section

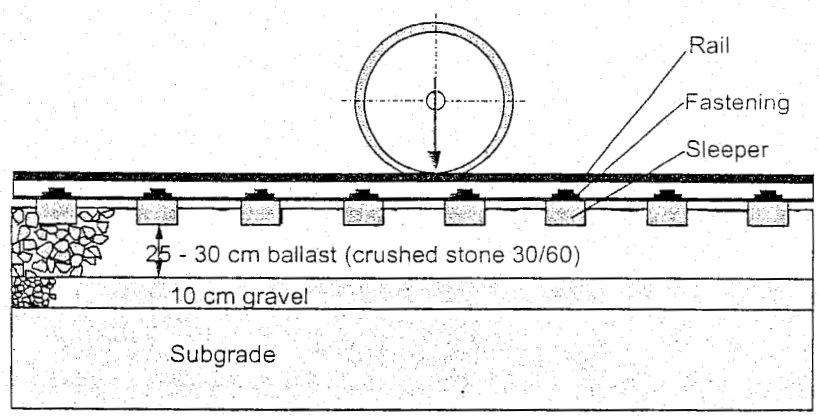


Figure 8.2: Principle of track structure: longitudinal section

Since the beginning of the railways, the principle of the ballasted track structure has not changed substantially. Important developments after the Second World War include: introduction of continuous welded rail, use of concrete sleepers, heavier rail-profiles, innovative elastic fastenings, mechanisation of maintenance, and introduction of advanced measuring equipment and maintenance management systems. As a result, the traditional ballasted superstructure can still satisfy the high demands, as demonstrated by the TGV-tracks in France.

The main advantages of ballasted track are:

- proven technology;
- relatively low construction costs;
- simple replacement of track components;
- relatively simple correction of track geometry (maintenance);
- small adjustments of track lay-out (curves) possible;
- good drainage properties;
- good elasticity;
- good damping of noise.

The type of structure chosen depends not only on the expected axle loads and speeds, but also on the required service life, the type and amount of maintenance, local conditions, and availability of basic materials. This means that the choice of track system is a technical and economical question which has to be answered according to each individual case. Of great importance here is the relation between costs for laying and costs for renewing the track, not forgetting the continual maintenance costs. The eventual aim is to arrive at minimum costs for total maintenance and renewal, or in other words, minimum total life cycle costs.

Due to certain disadvantages of ballasted track which are becoming more and more prominent in modern railway operation, slab track design is evolving into a very competitive alternative. The discussion whether slab track should prevail over ballasted track is discussed separately in Chapter 9, as well as the technological details of the slab track design.

The important track component "rail' is also discussed separately at some length in Chapter 10.

In the following sections the respective track components are discussed.

8.2 Formation

The substructure or subgrade consists of the formation which includes slopes, verges, ditches, and any structures within them.

The formation must have sufficient bearing strength and stability, must show reasonable settlement behaviour, and must provide good drainage of rain and melted snow from the ballast bed. If the existing subgrade cannot meet these requirements properly, the soil can be improved by either digging a trench, consolidating the ground by mechanical means, or stabilizing the ground by chemical means. In Chapter 5 a classification system is presented for the quality of soils.

The formation must be well consolidated and must have adequate bearing strength. Furthermore, the profile must not differ too greatly from the design profile.

Globally the following requirements apply:

- minimum bearing strength CBR > 5% (E_{V2} > 35 MN/m²);
- compaction 97% Proctor;
- deviation from design subgrade profile less than 10 mm.

To promote good filtering action an intermediate layer is placed between ballast bed and formation. This consists of a layer of gravel ballast about 10 cm thick with a grain size of 5/40 mm. The function of the intermediate layer is to separate the coarse-grain ballast from the fine sand. This layer also contributes to better load distribution and provides protection against frost. The global requirements for the sub-ballast layer are:

- bearing strength CBR > 25% (EV2 > 100 MN/m²);
- compaction 100% Proctor;
- maximum deviation from design subgrade less than 10 mm.

In conjunction with an intermediate layer use can also be made of a fabric made of a synthetic material (geotextile). Its function is primarily to prevent finer particles from getting through. In order to prevent damage, it is necessary to insert a fine-grain protection layer below and above the fabric.

8.3 Ballast bed

The ballast bed consists of a layer of loose, coarse grained material which, as a result of internal friction between the grains, can absorb considerable compressive stresses, but not tensile stresses. The bearing strength of the ballast bed in the vertical direction is considerable, but in the lateral direction it is clearly reduced.

The thickness of the ballast bed should be such that the subgrade is loaded as uniformly as possible. The optimum thickness is usually 25 to 30 cm measured from the lower side of the sleeper.

In addition to its load distribution function and provision of lateral resistance, the ballast bed's draining effect is important as its storage capacity during downpours is an aspect which should not be underestimated.

When finishing the subgrade and installing the layers of ballast, great attention must be paid to detail in order to prevent differential settlements with deviations which are limited to 10 mm.

Contamination of the ballast bed may have external or internal causes, such as attrition and weathering of the ballast material or upwards penetration of fine particles in the form of a clay (loam) mixture referred to as slurry.

A contaminated ballast bed hinders water drainage which results in reduced shear resistance and freezing during frost. The most important requirements to be met by the ballast material are hardness, wear resistance, and good particle size distribution. The particles themselves must be cubic and have sharp edges.

Some of the more commonly used types of ballast are:

- *crushed stone*: broken, solid, or sedimentary rock such as porphyry, basalt, granite, gneiss, limestone, sandstone, etc. Grading is 30/60 mm for main lines and 20/40 mm for switch and crossing work and level crossings. Generally speaking, crushed stone has many favourable properties. Some properties are, however, susceptible to weathering, giving rise to the possibility of mud formation.
- *gravel*: obtained from rivers; 20/50 mm grading. Gravel is very hard, but is made up of round grains which means that a gravel ballast bed has a low internal friction level.
- *crushed gravel*: obtained by breaking up large pieces of gravel; 20/40 mm grading. The shear resistance of crushed gravel is greater than that of normal gravel.

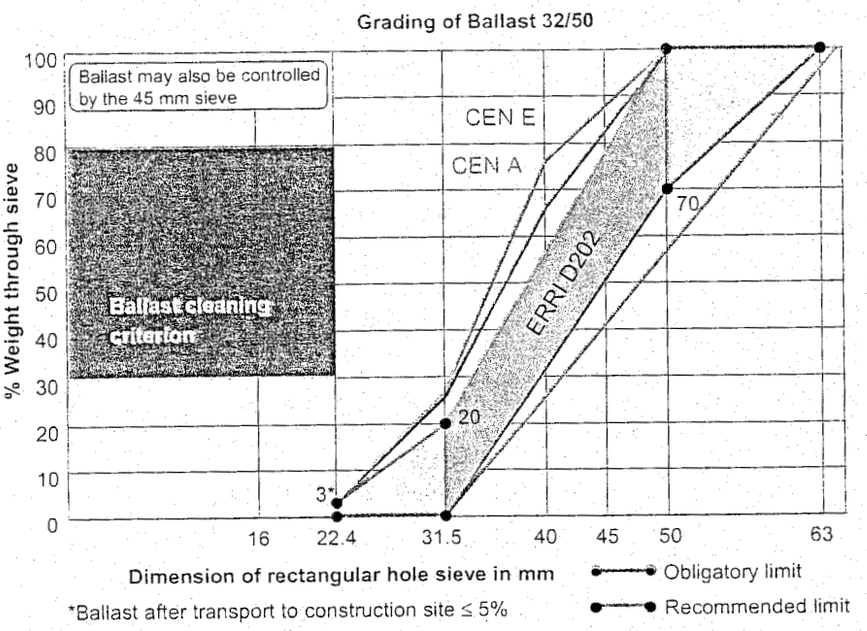


Figure 8.3: Ballast specification

Figure 8.3 shows the result of a comprehensive study into ballast specification carried out by ERRI (European Rail Research Institute), Committee D182, and issued by the CEN.

According to the ERRI findings, ballast cleaning is required if the contamination, expressed as the weight which passes the rectangular hole sieve 22.4, is greater than 30%.

8.4 Rails

8.4.1 Functions

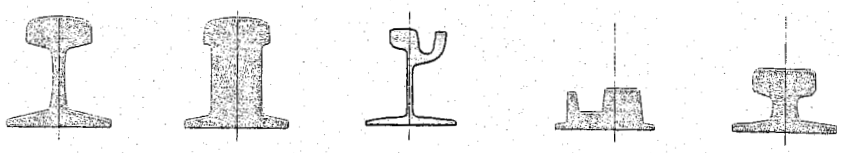
The rail can be seen as the most important component of the track structure and has the following functions:

- it accommodates the wheel loads and distributes these loads over the sleepers or supports;
- it guides the wheel in a lateral direction, any horizontal transverse forces on the rail head being transferred to and distributed over the sleepers and supports;
- it provides a smooth running surface and distributes accelerating and braking forces by means of adhesion;
- it acts as an electrical conductor on an electrified line;
- it conducts signal currents.

8.4.2 Profile types

Figure 8.4 shows some of the different types of rail profile:

- **flat-bottom rail**. This is the standard profile used as a general rule in conventional track;
- **non-standard profile**. This type differs from the flat-bottom rail in that the web thickness is greater in order to cater for switch and crossing components, expansion devices, etc;
- **grooved rail**. This type is used in enclosed track structures such as roadways, yards, etc.



flat-bottom rail	*construction rail*	*grooved rail*	*block rail*	*crane rail*
standard rail track	manufacturing of switch parts	railway track in pavement	railway track in concrete slab	heavy load hoisting cranes

Figure 8.4: Rail profile types

The block rail is part of the Nikex-structure. The crane rail is designed to support hoisting cranes with high wheel loads. In order to reduce the noise emitted by rails, new rail designs are nowadays proposed which are very compact and are used as embedded rails in slab track.

8.4.3 Geometry of flat-bottom rail

The Vignole rail (or flat-bottom rail) is derived from the I-profile in which, for the purposes of support and guidance, the upper flange is converted to form a rail head. Rail profiles which are used extensively in Europe include the 54 E1 (UIC 54) and the 60 E1 (UIC 60) which are shown in Figure 8.5 and Figure 8.6. The numbers refer to the rounded weight in kg per meter.

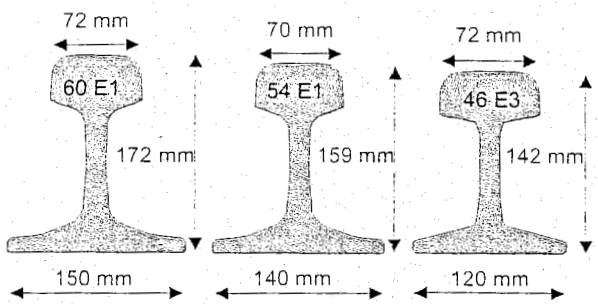


Figure 8.5: Vignole type rail profiles

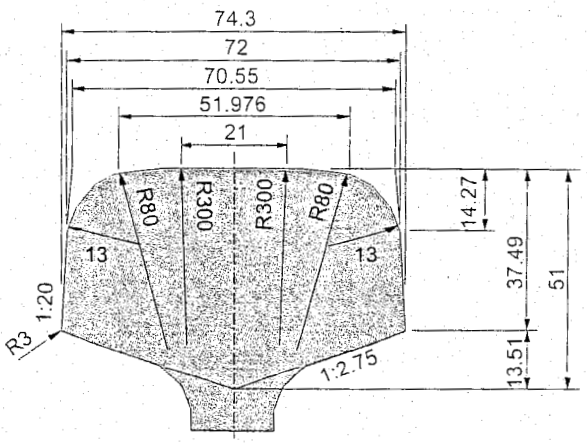


Figure 8.6: 60 E1 (UIC 60) profile

From the point of view of function, flat-bottom rail can be divided into three parts:

- **rail head**: the form must be such that good contact with the wheel tire profile is ensured. The dimensions of the rail head must be sufficient to provide a high wear margin;
- **rail web**: the thickness of the web is determined by the requirement for adequate stiffness against bending and buckling, even when affected by corrosion. In jointed rails the holes for the joints are made in the web. To allow fish-plating, the areas of transition to rail head and rail foot, known as the fishing surfaces, are inclined. The fillet radii must be greater than 6 mm in order to prevent stress concentrations;
- **rail foot**: the width of the rail foot must be large to provide for the stability of the rail profile, load distribution to the sleeper, and the required moment of inertia in the lateral direction.

The rail foot also serves to fasten the rail either directly or indirectly to the sleeper. When replacing the rail with another profile, it must not be forgotten that different foot widths cannot always simply be accommodated by the existing fastening.

Shape and dimension deviations in rail profiles, which arise during the manufacturing process, must not exceed certain limits as these may cause problems during processing and can result in high dynamic forces.

When building the track it is also recommended to use rails from the same batch and to assemble these rails in the same rolling direction. For this purpose, the rails are marked with a coding system.

The metallurgical, wear, and failure aspects of rails are dealt with in the special Chapter 10.

8.5 Rail joints and welds

8.5.1 Introduction

Joints and welds are used to form a connection between separate lengths of rail. Geometrical deviations must be small enough to limit dynamic effects. Strength and stiffness must be approximately that of the complete rail. The following options exist:

- *with* the possibility of expansion:
  - fishplated joints in non-continuous track;
  - joints and expansion devices;
  - bridge transition structures.
- *without* the possibility of expansion:
  - metallurgical welds, such as flash butt welds, thermit welds, and electric arc welds;
  - glued insulated joints.

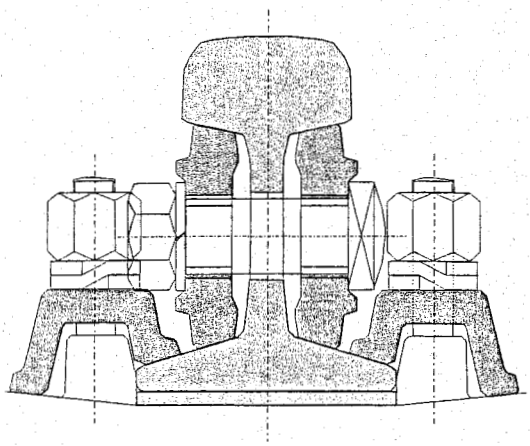
In order to be distinguished from metallurgical welds, the other types are described as fishplated joints.

8.5.2 Fishplated joints

Fishplated joints are used to connect rails in jointed track. The joint is made by means of fish-plates and fish bolts as shown in Figure 8.7. Axial displacements as a result of changes in temperature must be accommodated.

On account of the high impact effect, fishplated joints are often fitted to double sleepers, i.e. two coupled sleepers. Although the load is distributed over two sleepers, the joint still requires a great deal of maintenance. Moreover, the different sleeper pattern demands extra attention during mechanical tamping.

In view of the large amount of extra maintenance required, fishplated joints are used as little as possible.



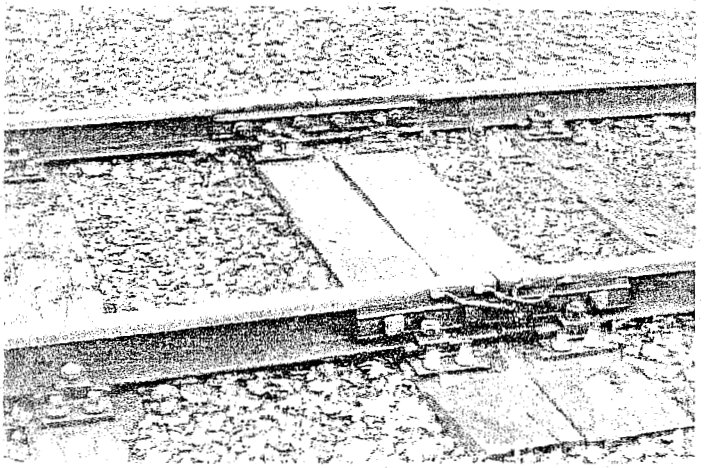


Figure 8.7: Fishplated joint on double sleepers

8.5.3 Expansion joints and expansion devices

These features serve to keep certain structures stress-free if the track on the structure is subject to large displacements as a result of temperature differences or creep.

On some railways, like BR, expansion joints are made from suitably machined standard rails. The expansion joint, shown in Figure 8.8, is constructed from a non-standard rail and allows a maximum axial displacement of 120 mm. Expansion joints are expensive and are only used in jointed track near the moveable bearings of fairly large structures, at the end of CWR track, or at changes in structure such as switches, crossings, and bridges.

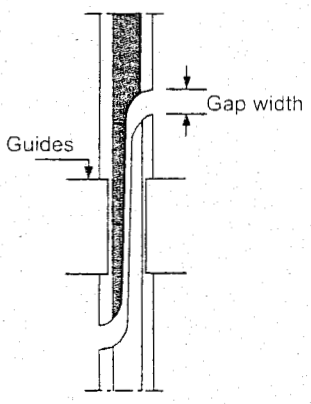


Figure 8.8: Expansion joint

The expansion device consists of a fixed stock rail and a blade, as illustrated in Figure 8.9, which can move parallel to the stock rail over a maximum distance of 220 mm. This device is used on CWR track on structures with a large expansion length.

8.5.4 Bridge transition structures

These structures take account of the lifting or turning movement of the bridge. A distinction is made between two types. The transition structure with a skew joint is used in bridges which move vertically and is similar to the expansion joint. The filler-type transition structure, depicted in Figure 8.10, is used in bridges which cannot be raised such as bascule bridges. In this case, the wheel rests partly on the filler which can be pushed back before the bridge is opened.

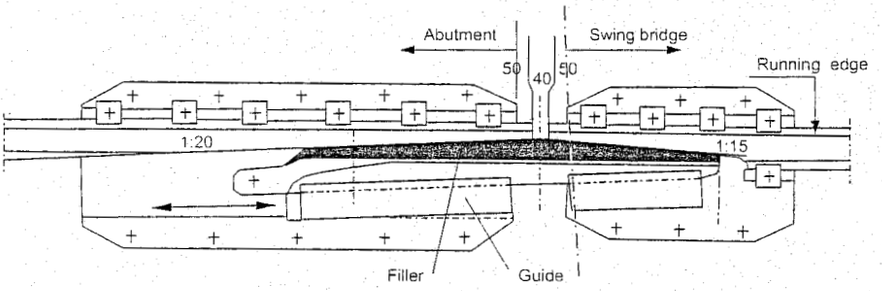


Figure 8.10: Bridge transition with filler

8.5.5 Insulated joint

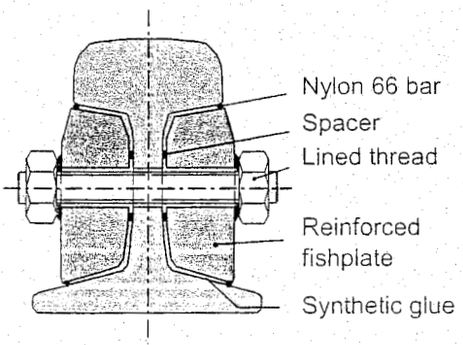
In order to separate track circuits when operating the automatic block system, insulated joints are used. Such rail joints must be used exclusively to insulate each adjoining rail, for the mechanical properties of the plain rail must be maintained as much as possible. The joints are made of fish-plates, bolts, and insulating material.
For the most part a distinction is made between two kinds of insulated joint:

- **non-glued insulated joint**

  This joint includes a 6mm nylon end post in the gap between the rails and a synthetic lining between the fish-plates and rails. The fish bolts and the fish-plates are insulated too. Since large longitudinal forces cannot be withstood, this structure is not suitable for CWR track.

- **glued insulated joint**

This joint is made in a similar way to the non-glued joint except for the linings. The insulating effect in this case, as shown in Figure 8.11, is obtained by using synthetic adhesive. This joint is capable of transferring forces in CWR track of up to about 1000 kN. These joints can either be made in the work-shop or can be prefabricated in a piece of rail which is then later thermit welded into the track. On NS rail the electrical resistance (impedance) must be at least 10 Ω at 100 kHz.



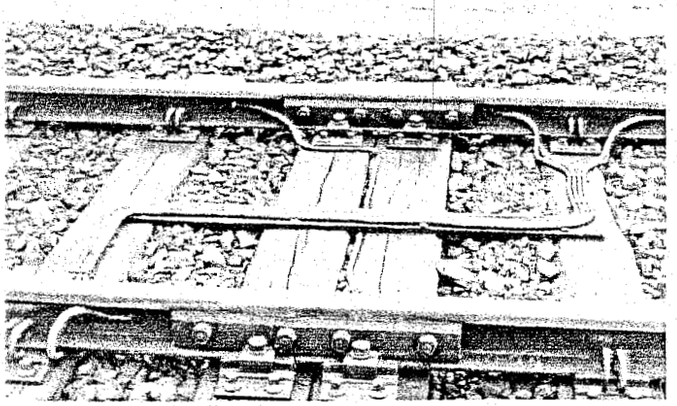


Figure 8.11: Glued insulated joint

Flash butt weld

Flash butt welds are generally produced in welding depots with stationary welding machines. In terms of quality, flash butt welds are better than thermit welds which will be discussed below. Mobile flash butt welding machines also exist, so that these welds can be made on the track.
The metallurgical weld quality, as represented by the fatigue strength, produced by the mobile machine is equivalent to that produced by the stationary machine. The difficulty is, however, to achieve good geometrical properties.

In brief, the procedure is as follows:

- rail ends are shot-peened;
- rails are positioned;
- electric voltage is applied;
- electric arc brings rail ends to forging temperature;
- rail ends are upset under high pressure;
- weld collar is stripped and ground;
- welds are cooled, straightened, and finish-ground.

Further details are given in Section 10.4.

Thermit weld

Thermit welds were developed to enable a good metallurgical weld to be made in the track using relatively simple equipment. The method makes use of a mixture of aluminium powder and iron oxide which at a high temperature is converted to alumina and steel. The heat released causes the temperature to rise to about 2500 °C [185]. The process, which for as far as the specific welding part is concerned lasts about 15 minutes, is basically as follows:

- rail ends are aligned;
- mould is installed around the joint;
- crucible is installed;
- rails are pre-heated to about 900 °C using propane burners;
- mixture is ignited; reaction: $2Al + Fe_2O_3 \rightarrow Al_2O_3 + 2Fe + 850$ kJ;
- crucible and mould are removed;
- weld collar is stripped and ground.

Further details are given in Section 10.4.

Electric arc weld

In the electric arc welding process the weld is built up from the rail foot inside a copper "casing" using electrodes. The procedure lasts 15 to 20 minutes and is basically as follows:

- place casings in position;

- pre-heat up to 200 to 250 °C using propane burners;
- weld gap from the bottom to the top;
- remove casings.

On account of its poor quality, this type of weld is only used if it is not possible to make a thermit weld.

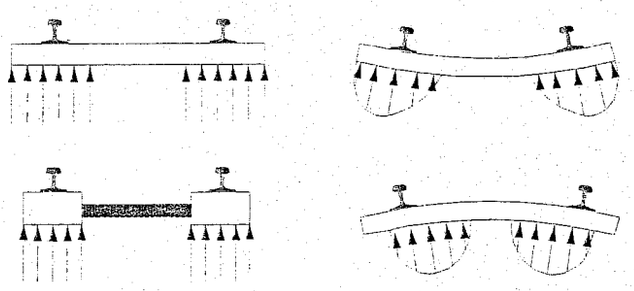
8.6 Sleepers

8.6.1 Introduction

In ballasted track the rails rest on sleepers and together form the built-up portion of the superstructure. Timber and concrete sleepers, and to a limited extent steel sleepers, are used. The advantage of the concrete sleeper is that climatic influences have little effect. Its service life is expected, under certain conditions, to be significantly higher than in the case of timber sleepers. These conditions entail that formation and ballast bed are of good quality as are also the rail and weld geometry. Concrete sleepers are susceptible to impact loads, especially in the 25-300 Hz frequency range.

The general functions and requirements of sleepers are:

- to provide support and fixing possibilities for the rail foot and fastenings;
- to sustain rail forces and transfer them as uniformly as possible to the ballast bed;
- to preserve track gauge and rail inclination;
- to provide adequate electrical insulation between both rails;
- to be resistant to mechanical influences and weathering over a long time period.



Correctly packed sleepers Uncorrectly packed sleepers

Figure 8.12: Sleeper packing

To ensure stability, it is desirable that the sleeper is only supported in the areas beneath the rails. In the case of prismatic sleepers such as the timber sleeper and the monoblock concrete sleeper, this is achieved by only packing this area and leaving the centre portion free as depicted in Figure 8.12. In the case of twin-block concrete sleepers, this is achieved by the actual structure.

Furthermore, it must be ensured that the sleeper does not rotate under the rails as a result of the vertical load, for this can cause gauge to become narrower or to widen and it changes rail inclination. This happens if the sleeper is supported too close to the inside or outside as a result of incorrect tamping.

To ensure that the available ballast resistance in both longitudinal and lateral directions is put to optimal use, the ends and sides of the sleeper must be fully embedded in the ballast. The distance between sleeper centres is usually 60 cm. In lightly loaded CWR track, this value may be increased to 75 cm.

8.6.2 Timber sleepers

The timber sleeper is prismatic in shape, 15 cm high and 25 cm wide. The length is 2.60 - 2.70 m and it weighs about 100 kg, which means that the timber sleeper can be replaced by hand.

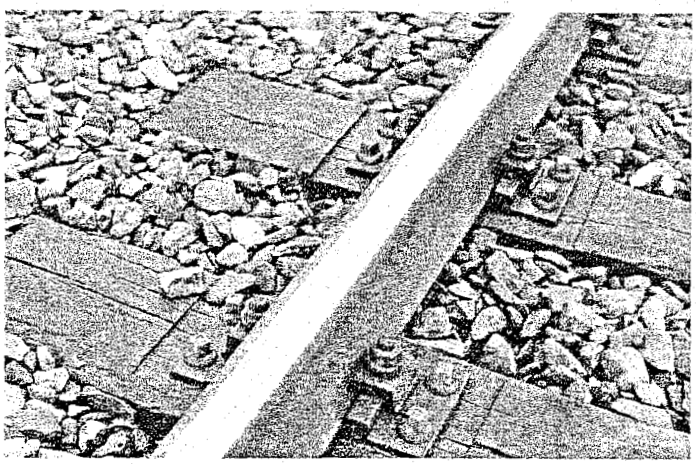


Figure 8.13: Wooden sleeper + baseplate

As an example, Figure 8.13 shows a type of timber sleeper reinforced with a baseplate.

Use is made of the following varieties:

- **softwood sleepers** (pine-wood). As the compressive strength perpendicular to the wood grain in the case of softwood is not great, a steel baseplate has to be placed between sleeper and rail to spread the load over a greater area. In the long term, however, the baseplate cuts into the sleeper wood, causing a gap into which water can penetrate (pumping effect) resulting in rapid deterioration of the quality of the fastening. This phenomenon is slowed down by treating the bearing surface of the sleeper with a synthetic material. This procedure, which is followed by the NS, increases the service life of the sleepers, generally speaking, by 50%;

- **hardwood sleepers** (beech, oak, tropical varieties). This type is stronger and has a longer service life. Hardwood sleepers are used, for instance, in switches and crossings and where fastenings are applied without baseplates.

Switch and crossing sleepers, called bearers, have a normal cross-section yet are longer, up to 5.50 m, and should be completely flat to accommodate switches. Bridge sleepers are specially dimensioned timber sleepers for use on steel bridges.
Once delivered, timber sleepers have to undergo a number of treatments, as described below, before they can be laid in the track:

- *drying period* of at least 9 months until the moisture level has fallen to 20 - 25 %, relative to the dry weight of the wood;

- *processing* such as:

  - notching (milling) of the bearing surfaces;

  - drilling of holes to accommodate the fastenings;

  - binding using a steel band to limit cracks.

- *creosoting* (conserving) timber sleepers against biological attack (fungi, insects etc.). The creosote is impregnated into the wood under high pressure after which some of the oil is recovered by applying a vacuum;

- *fixing* the fastening system.

The total service life in years of some types of timber sleepers is: pinewood 20 - 25, beech 30 - 40, and oak 40 -60.
In contrast to other types of wood, beech-wood can, and must, be creosoted through and through. This accounts for the long service life. If left untreated beech-wood is very quickly affected by fungi and will decay.

On the whole, the service life of treated sleepers does not depend on weathering but on mechanical effects.

8.6.3 Concrete sleepers

The development and use of concrete sleepers became significant after the Second World War owing to the scarcity of wood, the introduction of CWR track, and the improvements in concrete technology and pre-stressing techniques.

Specific advantages and drawbacks of concrete sleepers are:

Advantages:

- heavy weight (200-300 kg), useful in connection with stability of CWR track;
- long service life provided fastenings are good or can be replaced easily;
- great freedom of design and construction;
- relatively simple to manufacture.

Drawbacks:

- less elastic than wood. On poor formation, pumping may occur;
- susceptible to corrugations and poor quality welds;
- risk of damage from impacts (derailment, loading/unloading, tamping tines);
- dynamic loads and ballast stresses can be as much as 25% higher;
- residual value is negative.

There are two basic types of concrete sleeper:

- **twin-block sleeper**. This type consists of two blocks of reinforced concrete connected by a coupling rod or pipe. Until about 1990, NS made use of the UIC 54 concrete sleeper, shown in Figure 8.15 with DE fastenings. The blocks are connected by means of a synthetic pipe filled with reinforced concrete. The upper surface of the blocks has a 1:40 inclination which corresponds with the rail inclination. The weight is approximately 210 kg. Figure 8.17 shows a TGV track on SNCF with concrete twin-block sleepers provided with Nabla rail fasteners (Figure 8.26);

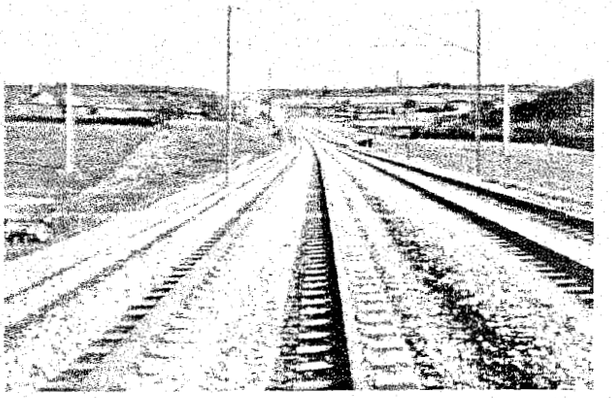


Figure 8.14: Paris-Est TGV track with twin-block concrete sleepers and Nabla rail fasteners

– **monoblock sleeper**. This is based on the shape of a beam and has roughly the same dimensions as a timber sleeper. Since 1990, the NS has used this type of sleeper in new lines and renewals because it is considered to endure the higher and intensive loadings better than the twin-block sleeper. This pre-stressed sleeper, coded NS 90, is shown in Figure 8.16. The rail is fastened using Vossloh clips. Figure 8.17 shows an example of a track fitted with pre-stressed concrete sleepers. The rail is fastened using Pandrol clips.

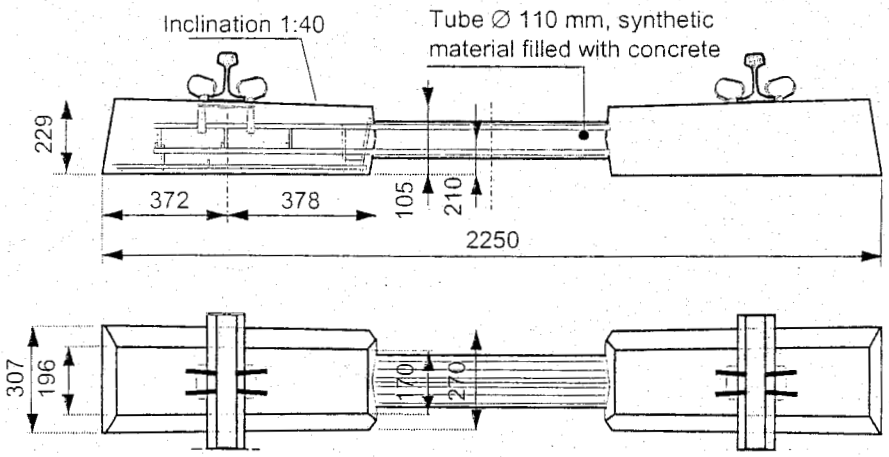


Figure 8.15: Reinforced twin-block sleeper

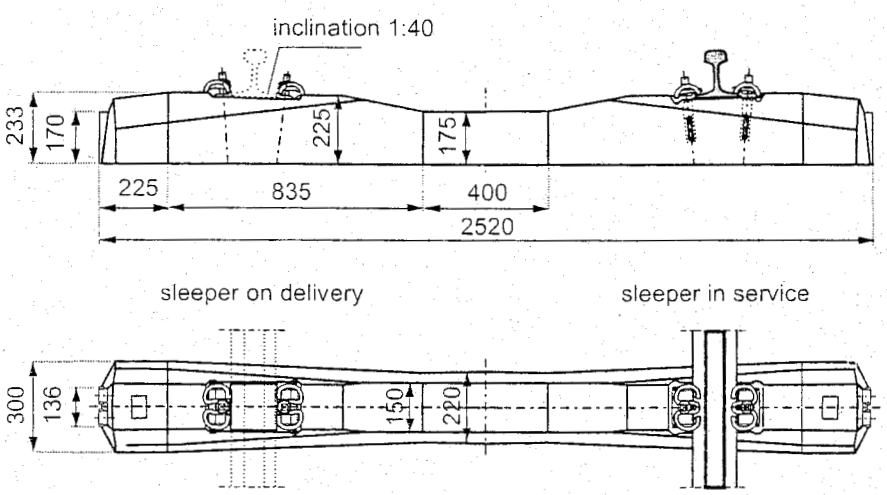


Figure 8.16: Prestressed monoblock sleeper

The advantages of the twin-block sleeper over the monoblock sleeper are:

– well-defined bearing surfaces in the ballast bed;

– high lateral resistance in the ballast bed because of the double number of surface areas.

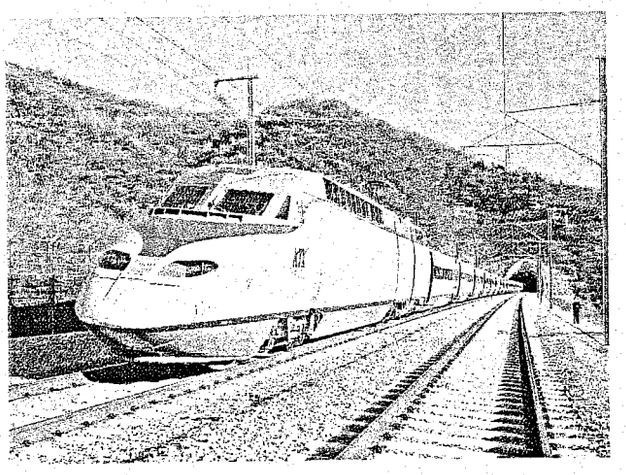


Figure 8.17: KTX running on concrete sleepered track

The advantages of monoblock sleepers are:

- lower price;
- little susceptibility to cracking;
- can be pre-stressed.

8.6.4 Steel sleepers

Steel sleepers are now only used on a very small scale because of problems such as:

- insulation;
- maintenance using tampers;
- relatively high price.

On the other hand, steel sleepers also have a few strong points such as:

- long service life;
- great dimensional accuracy;
- positive residual value.

8.7 Improvements in ballasted tracks

8.7.1 Introduction

New developments have found their way into the design of traditional ballasted super-structures. Ballast spreads the load from underneath the sleepers to the foundation. This loading causes pressure between the contact-points of the grains which causes the ballast roadbed to deteriorate: the grains wear and wander, the ballast becomes polluted causing a diminished drainage-function, and the track quality decreases. Maintenance work is necessary at a regularly basis, depending on the amount of tonnage which has passed.

There is a direct link between the level of pressure and the deterioration of the track. Decreasing the average pressures on the ballast will lead to a lower rate of deterioration. To achieve this, one can use a heavier rail-profile which spreads the load over more sleepers, decrease the distance between the sleepers, or increase the bearing surface of the sleepers themselves by making them longer or broader. Several attempts have been made with broader sleepers and slab sleepers, but these were not successful.

8.7.2 Wide sleeper

Currently, in Germany the "wide sleeper" has been developed as a sequel to these former attempts. The main problem then was the inability to provide specific tamping-techniques. This has now been solved. This wide sleeper is 2.40 m long and 57 cm wide. The distance between them is 60 cm leaving a 3 cm gap which is covered with a rubber cover (Figure 8.18).

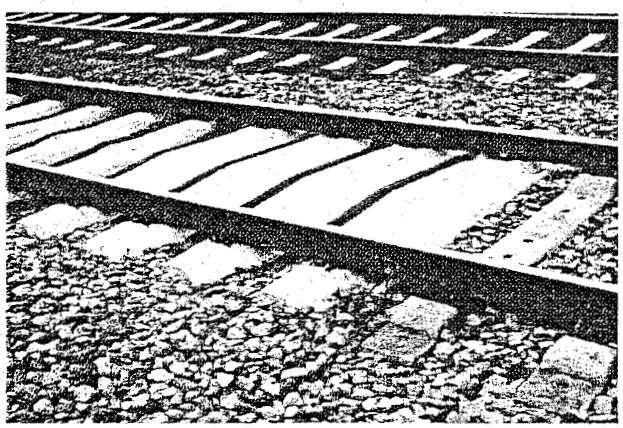


Figure 8.18: wide sleeper track

In the case of the wide sleeper, the sleeper-weight is doubled to 560kg, but with an axle load of 22.5 tonnes this results in an average surface pressure of 2 kg/cm² compared to approximately 3.7 kg/m² in the case of traditional sleepers. The track-bed remains free of rain and vegetation as a result of which the rate of pollution is reduced.

Because the track can now be considered as a continuous slab of discrete elements, tamping can only take place at the end of the sleepers or the edge of the "slab". The tamping device must be turned 90 degrees.

The first trials started in Waghäusel Germany in 1996 with 450m. After 50 MGT, the Q-value was determined at 28 meaning a 'good' to 'very good' track geometry. Afterwards, a track of 6-km between Homburg/Saar and Bexbach was reconstructed and completed in October 1997 next to a traditional reference track. The maximum allowed speed was 120 km/h with traffic loading of 18.500-21.000 tonnes/day, curve-radii up to 350 m with maximum cant of 135 mm.

The continuous measurements show the following results:

- High track bed stability;
- Increase of 15 % in the sideways-stability;
- Increase of the airborne noise emission by 2 dB (A) which can be halved by applying noise-reduction measures;
- Low body noise emissions;
- Favourable deformation behaviour (compared to the reference track less than 50 %);
- During 3.5 years, hardly any maintenance was necessary.

The construction costs are approximately 10 to 20 % higher than for normal ballasted track. These additional costs will be compensated in the medium-term through the decreased amount of maintenance.

8.7.3 Frame sleeper

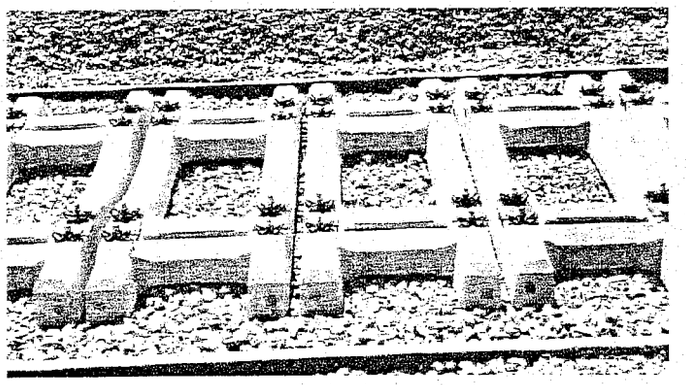


Figure 8.19: The double H-sleeper from SLL, Austria

The frame-sleeper intends to replace the load transmitting structure of a traditional ballast track, i.e. cross sleepers in a regular distance, by a girder-grid, thus combining a continuous longitudinal beam with cross members. The "longitudinal beam" is formed by sections in the frames which are connected and held in place by the rail and the fastenings which are situated at either end of the frame-sleeper. Hence, the running wheel-load is transmitted in a continuous manner onto the ballast bed reducing the pressure under the sleepers substantially and avoiding most of the stress gradients.

The double H-shaped concrete sleeper (Figure 8.19) was tested in Austria. These elements can be considered as two sleepers which are connected at the fastening systems with 'bridges'. The same principal as with the wide sleeper is used to increase the bearing surface of the sleeper and, thus, decrease the pressure on the ballast. The elements are 2.4 m long and 0.95 m wide with two sets of rail fastenings per element. Between the fastening system, the rail itself is supported by the concrete bridge which provides a quasi-continuous rail support. Underneath the sleeper elements a 12 mm thick polymer layer will provide a better spreading of the loading and additional damping. Test results showed a reduction in settlements of two-thirds compared to a normal sleeper.

A parallel Austrian development is the frame sleeper depicted in Figure 8.20 [237]. The arrangement of 4 fasteners per frame-sleeper creates a very high stiffness in the horizontal plane, supporting the stability of the alignment and increasing the buckling strength beyond the limits presently in existence. An elastic pad is incorporated into the concrete structure at the frame-sleeper-bottom to improve the ballast-concrete interface.

Very high lateral resistance and frame stiffness offer best provisions for a durable alignment. The reduced settlements and settlement-differences indicate a much more durable track geometry and thus a higher availability of the infrastructure.

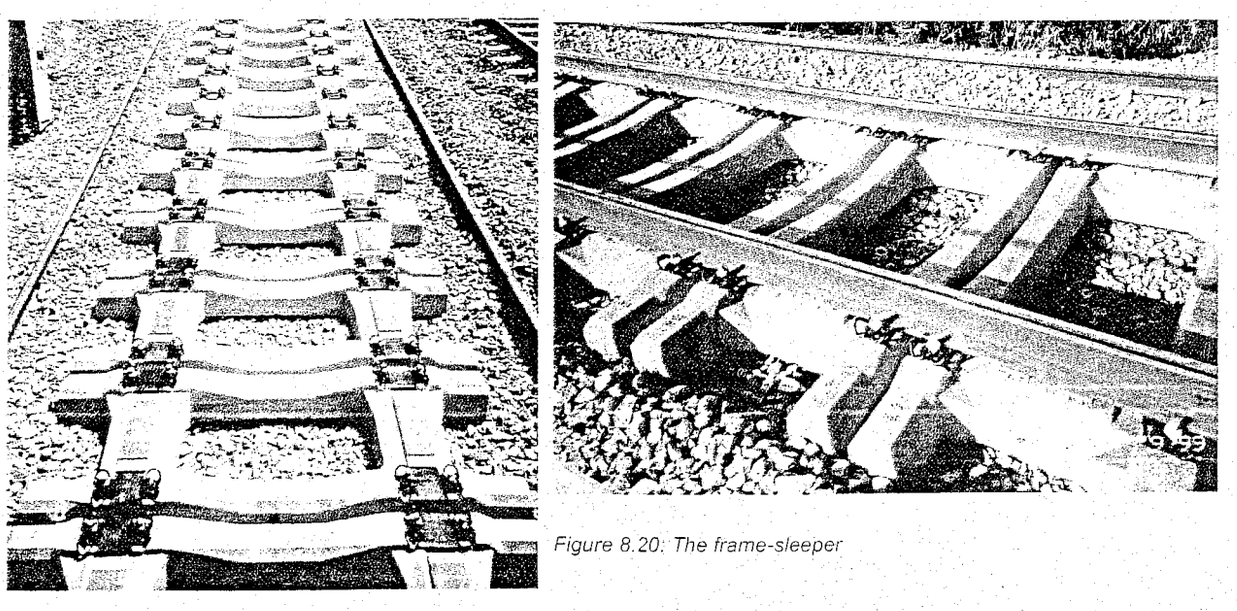


Figure 8.20: The frame-sleeper

8.7.4 Local ballast stabilisation by means of a chemical binder

Using a mixture of liquid resins and hardeners, ballast is glued together. Glueing occurs only at the points of particle contact, transforming loose material into a high-resistance compound material and preventing if from loosening and subsiding from the sleeper front.
Ballast binding is applied in order to prevent damaging of the ballast bed caused by ice blocks falling off trains. In winter ice forms on the underside of railway carriages which means that extremely heavy blocks of ice can build up due to the considerable clearance below the carriage. When those blocks become detached from the carriages at high speeds, they can cause considerable damage. When the blocks of ice strike the ballast they cause explosion-like reactions on the ballast stones.

There is an increased risk of blocks of ice coming loose in long tunnels due to temperature changes or abrupt changes in air pressure. Therefore, in long tunnels the track ballast in front of the sleeper ends is stabilised through an inert, non-toxic binding agent. This agent allows the danger caused by depressions in the ballast bed around the sleepers due to the dispersal of ballast stones to be prevented to a large degree.

This kind of ballast stabilisation is also being applied at transition zones between slab track and ballast superstructures, road crossings, track stops, switches, in order to prevent ballast whirling on high-speed lines, as general ballast stabilisation, to improve the cleaning properties of tracks at stations and to improve drainage properties at shunting emplacements.

8.8 Fastening systems

8.8.1 Introduction

The term "fastening systems" or in short "fastenings" is considered to include all the components which together form the structural connection between rail and sleeper. All over the world a great variety of fastenings exists to which new types are regularly added in order to keep up with changes in requirements and opinions or due to the availability of new materials. The choice of fastening also greatly depends on the properties and structure of the sleeper.

The general functions and requirements of the fastenings are:

- To absorb the rail forces elastically and transfer them to the sleeper. The vertical clamping force of the rail on the sleeper must be sufficient in all load situations, even in case of wear, in order to provide the necessary longitudinal resistance to limit the breathing length in CWR rail, to limit gaps in the case of rail fractures, and to resist creep;

- To damp vibrations and impacts caused by traffic as much as possible;

- To retain the track gauge and rail inclination within certain tolerances;

- To provide electrical insulation between the rails and sleepers, especially in the case of concrete and steel sleepers.

The longitudinal resistance between rail and sleeper should be at least 15 kN so that it is much greater than the longitudinal shear resistance of the sleeper in the ballast bed. This would amount to 5 kN for half a sleeper. This means that in the case of large axial rail forces the fastening remains intact and the sleepers move in the ballast bed.

In the case of CWR track with direct fastening on viaducts, it is necessary to ensure that the longitudinal force caused by differences in temperature, which the fixed support of the viaduct has to transfer, is not too great. This longitudinal force is, at the most, the sum of the sliding forces over the length of the viaduct. By using special nylon rail pads the longitudinal resistance can be reduced.

8.8.2 Subdivision of fastenings

Fastening systems can be categorized as follows:

- **direct fastenings** entail that the rail and, if necessary, the baseplate are fixed to the sleeper using the same fasteners. Direct fastenings also include the fastening of track on structures without ballast bed and sleepers;

- **indirect fastenings** entail that the rail is connected to an intermediate component, such as the baseplate, by other fasteners than those used to fix the intermediate component to the sleeper. The advantages of indirect fastenings are that the rail can be removed without having to undo the fastening to the sleeper and the intermediate component can be placed on the sleeper in advance.

8.8.3 Baseplates

The rails can be fastened to the sleepers with or without the use of steel baseplates. Baseplates are provided with sloping upper surfaces and upright ribs between which the rail is locked. The force components are depicted in Figure 8.21.
Using baseplates has the following advantages:

- The vertical load is distributed over a larger area of the sleeper. This lengthens the service life of the sleeper;

- The horizontal load is absorbed better because of friction and because it is distributed over all the fastenings anchored in the sleeper. Baseplates are excellent for sustaining large lateral forces if large cant deficiencies are provided;

- The overturning moment causes less force on the fastenings in the sleeper;

- Baseplates have a high bending stiffness and grooves in the ribs provide good fastening locations for the rail;

- Baseplates give extra weight to the sleeper.

The disadvantage, on the other hand, is the relatively high price. The baseplate with clip bolts, shown in Figure 8.21, is the classic indirect fastening. German-speaking regions refer to this type as the K-fastening. The baseplate is fixed to the timber or concrete sleeper using coach screws. The rail is fixed by means of clip bolts, rigid sleeper clips, spring washers, and nuts. The clip bolt is inserted into the holes in the plate using a hammer head (to stop it turning). The pad between the rail and baseplate is of masonite. This fastening derives its vertical elasticity from the pad and the spring washers.

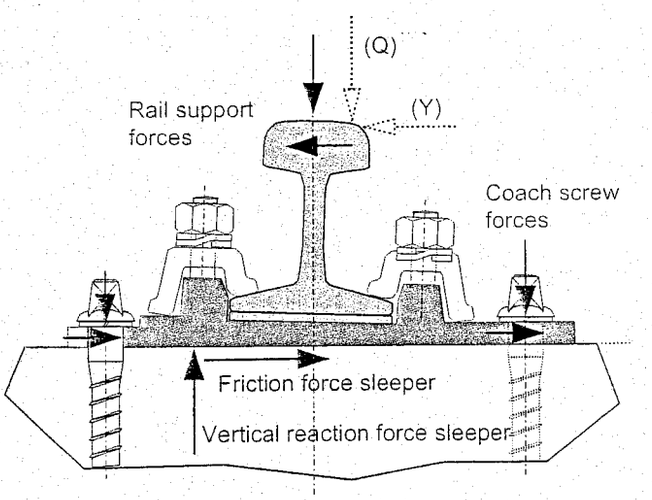


Figure 8.21: Forces on baseplate

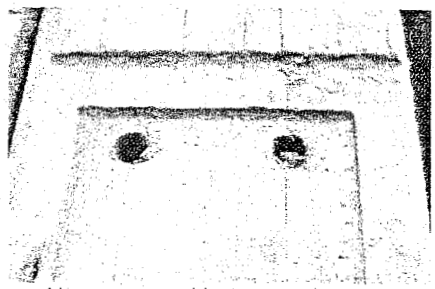
Sleeper life in combination with baseplates can be dramatically extended by applying an araldite compound between baseplate and sleeper. The NS experience was that sleeper life could be doubled to tripled through this measure. Figure 8.22 clearly shows the difference between treated and non-treated sleepers which came from the same test section.

8.8.4 Elastic fastenings

The introduction of CWR track gave rise to the need for fastenings with greater elasticity. Certainly in the case of concrete sleepers, which are susceptible to impacts, this is an absolute necessity. Since the end of the fifties the NS has used the DE clip (Deenik, Eisses). This component is fitted to both timber sleepers and concrete sleepers as shown for instance in Figure 8.15. The DE clip, which can also be used in combination with baseplates, is usually fitted in a holder. The clip holder is fixed to the sleeper by means of pins which are cast into concrete sleepers or, as in the case of timber sleepers, are pushed into pre-drilled holes.The DE clip is installed using special equipment.

As there is no threaded screw connection, in principle no maintenance or adjustment is theoretically required. But this so-called 'fit-and-forget' principle also implies a drawback. If manufacturing tolerances are not met or if excessive wear occurs, there is no means of adjusting the fastening.

Araldite compound between sleeper surface and baseplate



No compound between sleeper surface and baseplate

Figure 8.22: Effect of sleeper treatment with araldite

Other examples of elastic fastenings are the Pandrol standard clip, shown in Figure 8.23, the Pandrol Fastclip, shown in Figure 8.24, the Vossloh fastening, shown in Figure 8.25, and the Nabla clip, depicted in Figure 8.26.

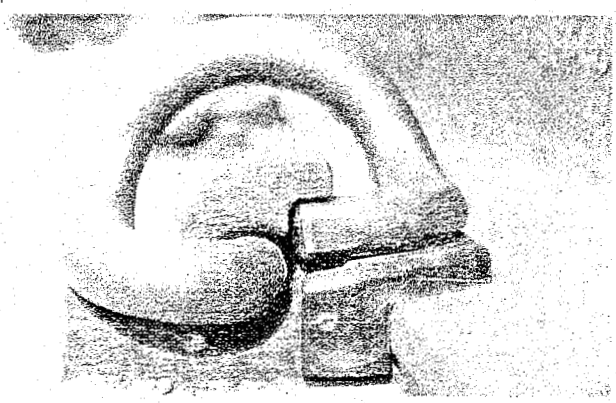


Figure 8.23: Pandrol fastening system

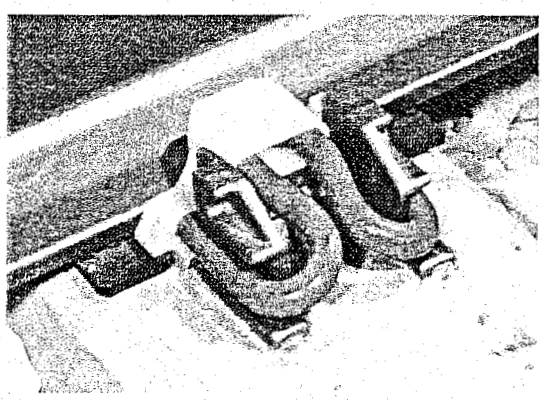


Figure 8.24: Pandrol Fastclip

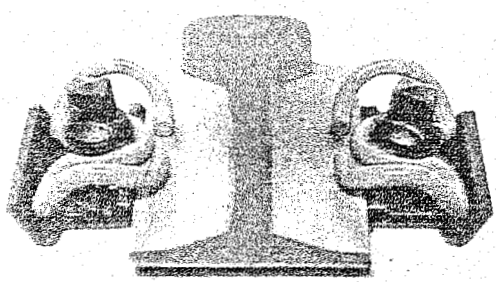


Figure 8.25: Vossloh fastening system

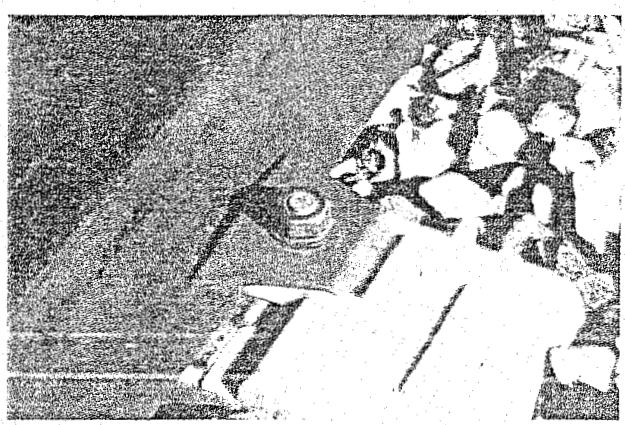


Figure 8.26: Nabla fastening system

	Soft	Normal	Hard	EVA
Type	(Fc584)	(Fc9)	(Fc846)	
Stiffness k [kN/mm]	970	1420	2990	3032
Damping c [kNs/m]	32	34	29	29

Table 8.1: Laboratory results of drop weight tests on some rail pad types

Throughout this book rail pads are mentioned with respect to their specific applications and important contribution to the dynamic behaviour of the track on structures with a continuous ballast bed and sleepers

8.9 Track on structures with a continuous ballast bed and sleepers

On short non-moveable bridges and viaducts it is worthwhile to continue the ballast bed on top of the adjoining subgrade onto the structure in order to avoid discontinuities at the transition. Although this causes an extra static load on the structure, the train load is distributed more uniformly. In addition, the usual methods for laying and maintaining ballast track can be applied without interruption. The ballasted bed also smoothes the differential displacements due to the differences in temperature of the continuous welded rails and the structure itself. To conclude, the ballast bed has good noise and vibration damping properties.

To improve the structural behaviour of ballasted track some developments are discussed:

- Ballast mats;
- Soffit pads;
- Super-elastic fastenings.

8.9.1 Ballast mats

Installing resilient mats between the bottom of the ballast and the tunnel, as sketched in Figure 8.29, has been employed as an anti-vibration method. Basically, three types of mat may be identified:

- profiled mats: usually compiled of single or multiple rubber layers;
- granular mats: made, for instance, from old tires bound with a high-grade elastomere;
- foam mats: consisting of single-layer or multi-layer polyurethane foam which flexibility may be altered by changing the ratio from/of open cell to closed cell pores.

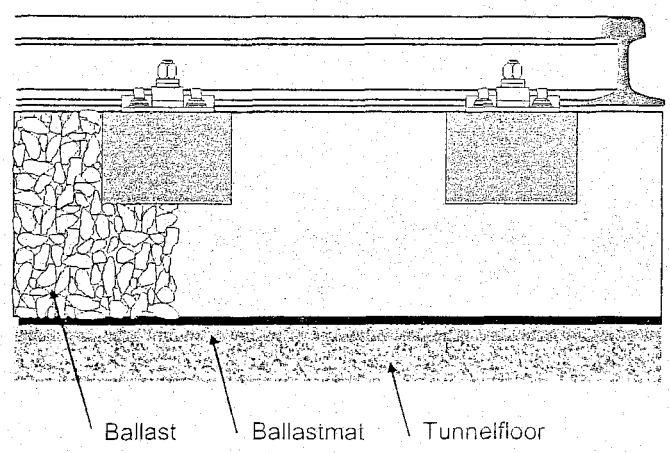


Figure 8.29: Application of ballast mat

The top layer of the mat consists of a hard protection layer which prevents ballast grains to penetrate into the layer. The mat's thickness is about 3 cm.
The rubber layers produce damping by changing the motion of the ballast and not as a result of significant energy absorption within the layer itself.
The object is to introduce resilience and hence generate a mass-spring effect carrying away energy. Maximum attenuation is obtained when the mats are as soft as possible.

However, if the mats are too soft problems may arise because of increased rail stresses and destabilisation of the ballast which requires frequent tamping. The ballast mat also acts as a protective layer to prevent damage to the structure surface by penetrating ballast particles.

Under a 25 tonne axle load, the quasi-static track deflections should be about 3 mm at 3 Hz with a limiting deflection of 4 - 5 mm under a static load. Higher values would cause intensive wear and fatigue of track components and a quick deterioration of the track geometry.

Ballast mat type	Found. modulus [N/mm³]	Track deflection [mm]
No mats	–	1.3
Mat A1	0.030	2.2
Mat A2	0.025	3.0
Mat B	0.025	3.5

Table 8.2: Static track deflection of track with ballast mat. (preload 210 kN)

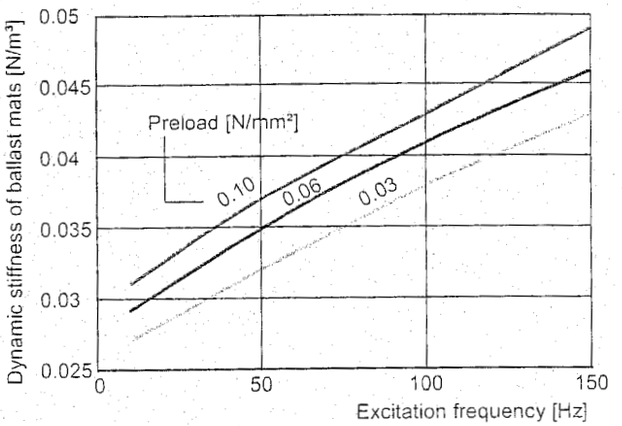


Figure 8.30: Dynamic stiffness of ballast mats

Table 8.2 lists some results of static deflection measurements of a track preloaded by a 210 kN axle load. Two kinds of ballast mat and two degrees of stiffness were involved in the test.

The dynamic stiffness of ballast mats depends on the preload, amplitude, and frequency as demonstrated by Figure 8.30, which shows the almost linear relation between dynamic stiffness and frequency.

Comparative measurements revealed that the stiffness amplification factor, which is the ratio between dynamic and static stiffness, is of the order of 3.

Soffit pad

Another way to attenuate track vibrations is to use sleeper soffit pads. The pads are placed between the sleeper and the ballast according to Figure 8.31. The composite pad comprises 22 mm of rubber-bonded cork with a hard facing to prevent damage by the ballast particles.

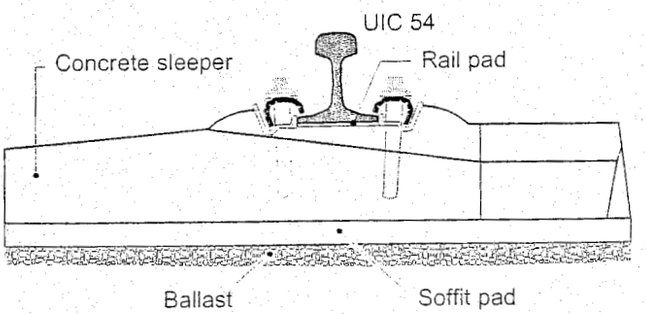


Figure 8.31: Sleeper soffit pad

Super-elastic fastening

Some special types of baseplate system, such as the "Cologne Egg" depicted in Figure 8.32, were developed to have a very soft vertical spring stiffness in order to effectively reduce vibration levels. The dimensions of the hole pattern are such that they can easily replace existing baseplates. An improvement of 7 dB in the frequency range 31.5 to 45 Hz, rising to 18 dB between 63 to 80 Hz, may be obtained compared to ballastless track. The low vertical stiffness limits the use of these systems in metro vehicles with static axle loads of 10.5 tonnes.

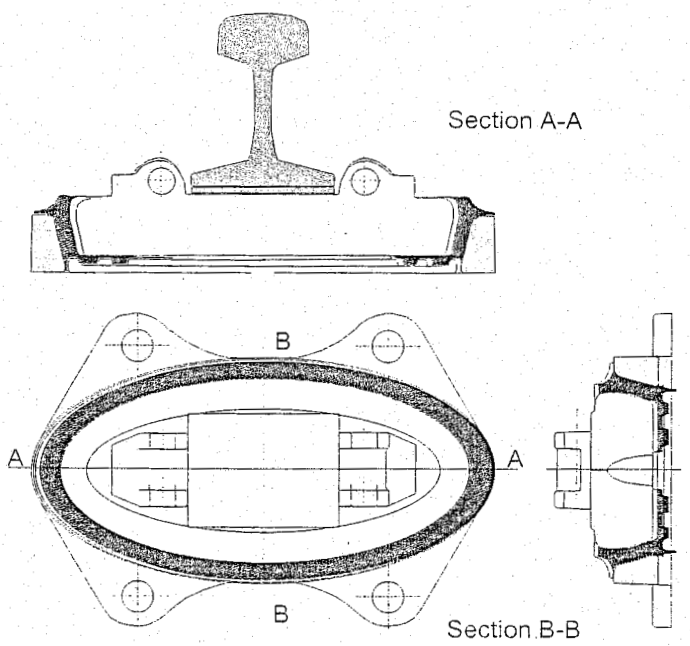


Figure 8.32: Elastic support of rails by means of "Cologne Egg"

8.10 Reinforcing layers

Softening of the subgrade can cause major problems, especially in combination with vibration. High-speed lines in Japan and Italy are therefore laid on a waterproof asphalt layer between 5 cm and 8 cm thick. In order to distribute - and hence reduce - subgrade stresses, this bituminous concrete layer can be increased to 15 cm or 20 cm (see Figure 8.33). Easy maintenance of the track geometry which is inherent in classic ballasted track is thus retained.

Asphalt layers may offer major advantages when constructing new track designed for relatively high axle loads and high gross annual tonnage. In addition, the use of reinforcing layers on conventional track designed for passenger services could lead to a significant reduction in the frequency with which the track geometry has to be maintained.

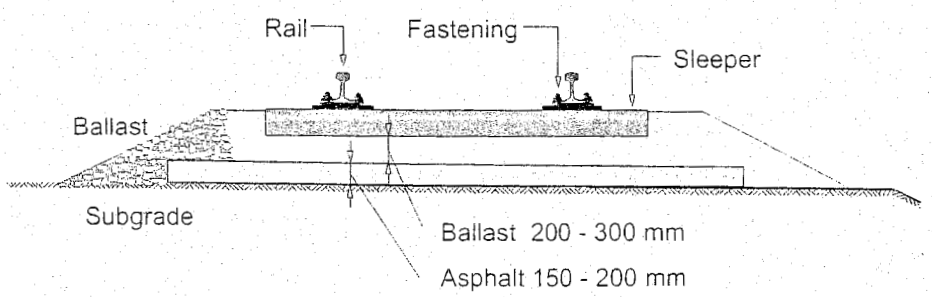


Figure 8.33: Reinforced bituminous concrete layer

8.11 Level crossings

These are constructed when road and rail cross at the same level. Only the most important types used on main line track are considered here.

Universal plates.

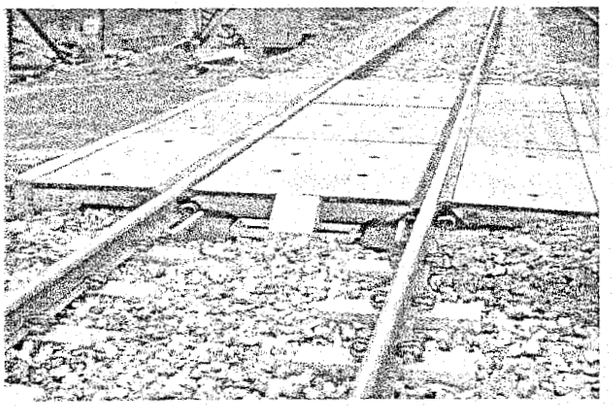


Figure 8.34: Level crossing with light-weight universal plates

In addition to train loads, level crossings have to withstand road traffic load. If the road traffic is not too heavy this load is transmitted via the standard track structure to the formation. This has the advantage that no special track structures are required and that standard maintenance methods can be applied. In such situations often preference is given to level crossings with lightweight universal plates. Variants exist for both concrete (Figure 8.34) and timber sleepers. The 15 cm thick plates are supported elastically on the sleepers and connected to the road surface via a curb. During track maintenance the plates are taken out.

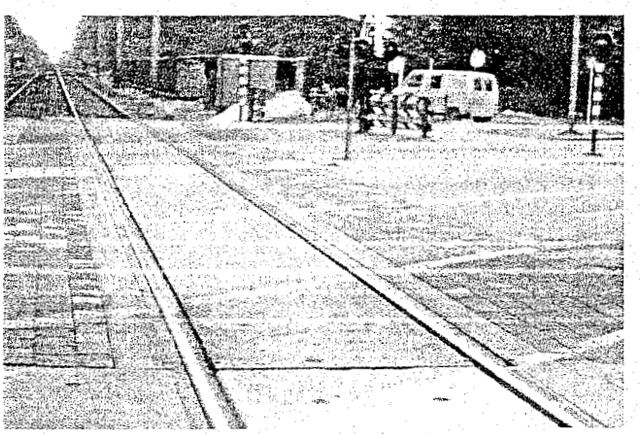
For moderately heavy road traffic the universal structure is provided with heavier plates with a thickness of 19 cm.

For heavy road traffic the universal structure is too light, thus entailing a huge increase in maintenance. In such situations it is worth installing much heavier and more expensive structures which require very little maintenance.

Harmelen crossing.

This monolithic structure, shown in Figure 8.35, was developed on Netherlands Railways to cope with moderately heavy to heavy road traffic. The slabs are prefabricated. The covering is made of a material which has the required surface roughness and is resistant to the impact loads produced by road traffic at the locations of the grooves provided for the train wheels.

During installation the slab is positioned as required and the space underneath the slab is filled with a fast-setting mortar which is injected through holes in the slab so that within a few hours the structure is ready for operation. The rails are sunk into gutters lined with cork-rubber pads and filled with a special polyurethane casting compound. For details reference is made to the embedded rail structure described in Section 9.8.



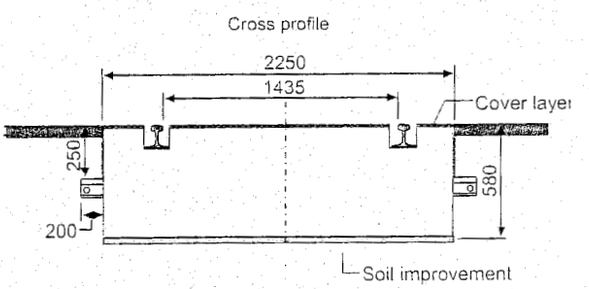


Figure 8.35: Harmelen level crossing

8.12 Tramway Track

8.12.1 Tramway track characteristics

Tramway track differs from national railway track considerably. The railways mainly operate on open track, generally consisting of continuous welded rail on sleepers embedded in stone ballast. The only places where the railways have to deal with the combination of road and rail are ports and at level crossings; in those situations a modified type of track structure is used (see 5.10). Most of a city tram operators' network consists of paved-in structures, i.e. tram structures where trams and road traffic share the same space. Grooved rails are used in this situation, to ensure that the tram wheel flange has sufficient space in the road surface. The same goes for light rail (in fact an upgraded tram system).

One can distinguish three types:

– *Exclusive tramway track*
This paved-in track or 'open' track consists of concrete sleepers in ballastbed or is carried out as lawn track. It resembles standard railway track. Road vehicles cannot use this track;

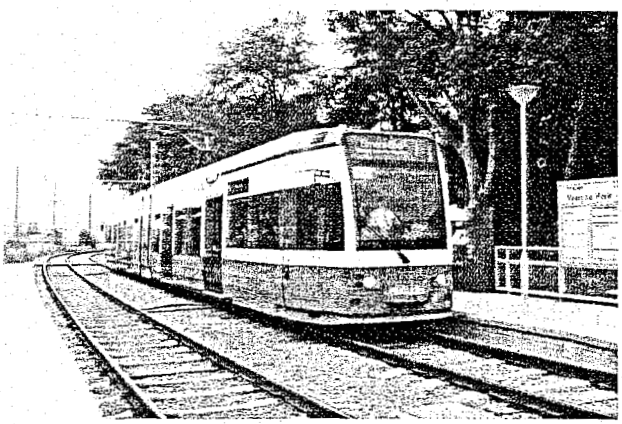


Figure 8.36: Example of exclusive tramway track: ballasted track. Tramlink, London

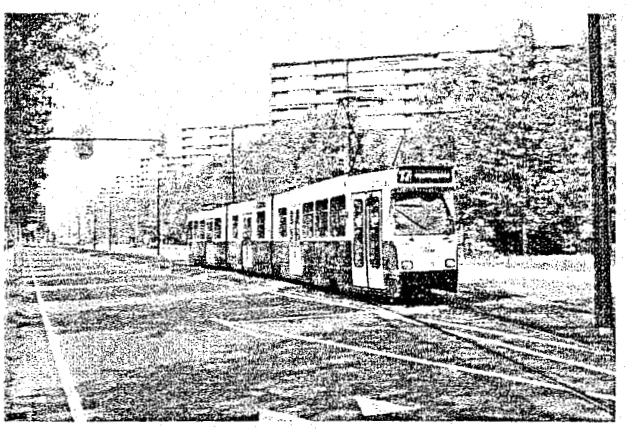


Figure 8.37: Example of exclusive tramway track: lawn track. HTM, The Hague

– *Free tramway track*

This type is built as 'closed' track, which means that only specific categories of road vehicles such as buses (public transport), public services and sometimes taxies are allowed to make use of this tramway.

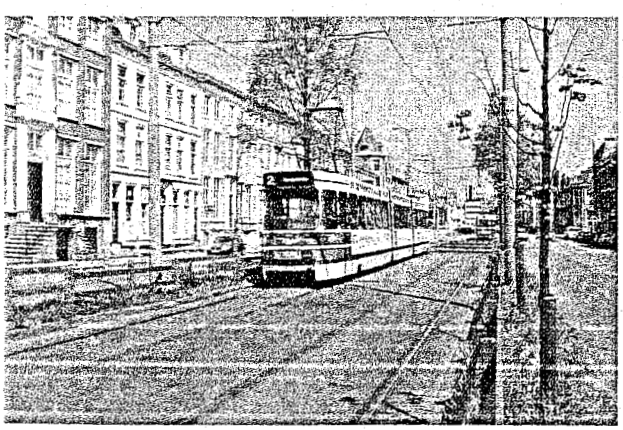


Figure 8.38: Example of free tramway track situated in the centre of the street, HTM, The Hague

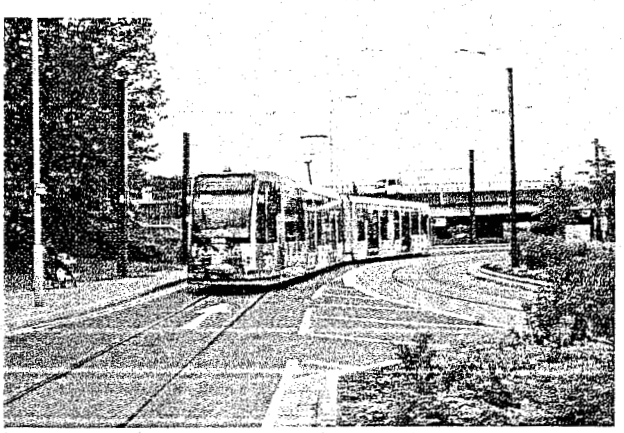


Figure 8.39: Example of normal tramway track. Tramlink, London

– *Normal tramway track*

Trams and road vehicles use the same traffic surface. This is always a closed track structure.
For tram, speed and axle loads are much lower comparing with railways. Light rail is a name for a system that is in between normal railways and trams. You could see it as an upgraded tram system. Trams usually have an average speed of 20 km/h, while light rail in urban areas aims for an average of 25 to 35 km/h or even higher.

Track geometry for tram and light rail differs from the common rail system: the curves of the track will have a much smaller radius, making it possible for the tram/light rail vehicle to be integrated in the existing city infrastructure.

For trams, normally a minimum of 25 m radius is used for tracks used in exploitation, while even smaller curves are possible near the depots. Of course the absolute minimum radius for the horizontal curve is depending of the construction of the vehicle, especially the nowadays commonly used articulated ones. For light rail, speed is an important issue and therefore curves need to be as large as possible. When the local situation won't let this large curve to be constructed, the technical conditions of the vehicle will be decisive.

For vertical curvature, also the vehicle construction is decisive for the minimum curve. But passengers are demanding comfort and this should be taken into account when designing the track. The maximum acceleration for the passenger should be determined, and depending on the speed of the vehicle, the minimum curve then can be calculated. For tram, a radius of 250 m for vertical curves is very low but in most cases technically possible, better is to aim for a radius of 1000 to 2000 m as a minimum.

Instead of using a tram system, many transport companies switch to the use of Light Rail. The advantage is a much higher average speed, for which more free track or better, more exclusive track is needed. The capacity of light rail is higher to, as is the maximum speed. When the track is for 100% exclusive and there is no possibility for the public to walk on the track, it is even possible to make the system fully automatic (for instance like the Docklands in London).

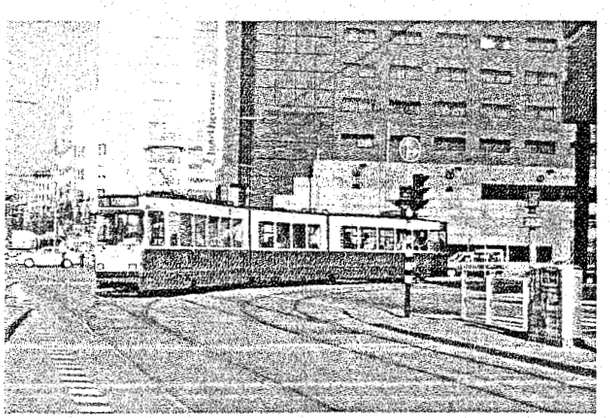


Figure 8.40: Horizontal curves need to have a radius as great as possible but should fit in the city infrastructure. HTM, The Hague

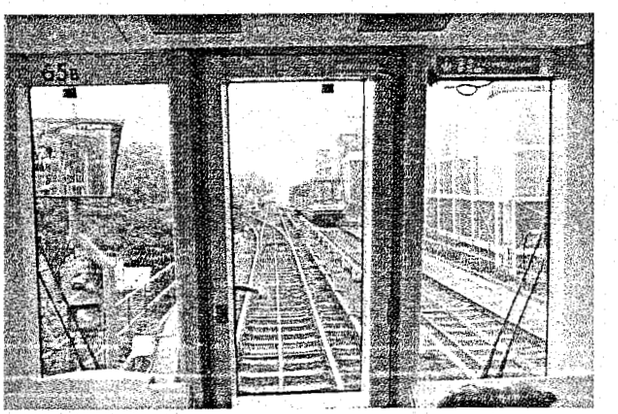


Figure 8.41: A view on the tracks from the front 'cabin' of a Docklands light rail vehicle. Docklands, London

8.12.2 Examples of paved-in tramway track

Figure 8.42 shows a frequently used construction. The supporting structure consists of a 30 cm thick continuous slab track. The grooved rails are fastened on synthetic plates with an elastic rail pad in between. Vossloh rail clips are used.

Because of the asymmetric profile of the grooved rail considerable geometric variations may occur during the rolling process of the rail. Therefore in order to attain an accurate track gauge the lateral position of the rail can be adjusted by means of wedge-shaped inserts.

The road pavement can be chosen freely and is independent of the supporting structure.

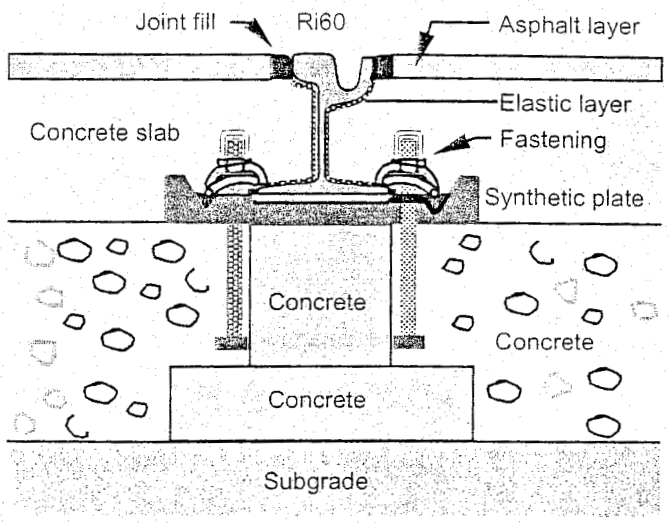


Figure 8.42: Paved-in track: Grooved rail Ri60 on synthetic plate

Before casting the upper concrete layer the rail is covered with an elastic coating. The fastenings are spaced 100 cm and are provided with a plastic cap to facilitate the unfastening at rail replacement. The gap between railhead and adjacent pavement is filled with a bituminous product, the joint fill.

Figure 8.43 shows the construction in progress. Clearly visible are the synthetic plates with the Vossloh clamps already mounted, but not yet under tension.
A technically outstanding but expensive solution is the embedded rail principle, given in Figure 8.44. Here the rail is secured accurately in a steel or concrete gutter after which an elastic mixture is poured in. This method results in a perfect separation between rail and pavement. The building procedure should be carried out very accurately and requires special treatment of the surfaces. The noise and vibration reduction is considerable.

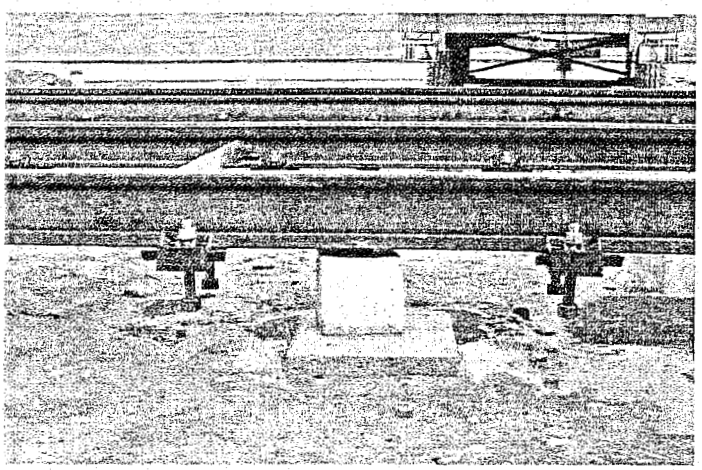


Figure 8.43: Paved-in track: Construction in progress

The structure is supposed to be maintenance free. This construction principle is also used in the Harmelen level crossing in The Netherlands. Moreover the embedded rail principle is considered as a strong candidate for high-speed railway track (Figure 9.46).

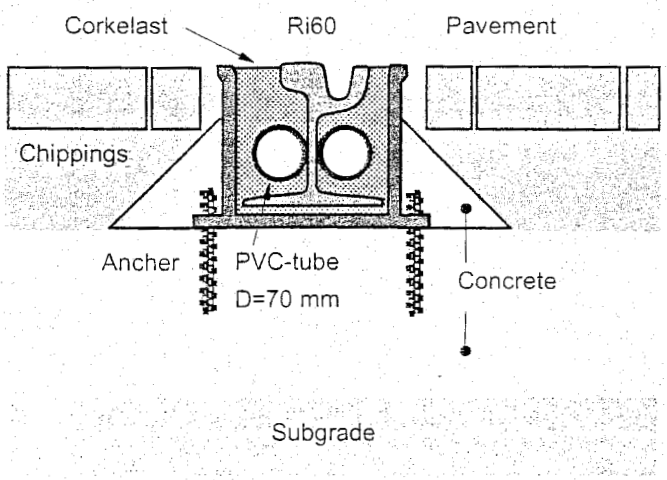


Figure 8.44: Paved-in track: embedded rail

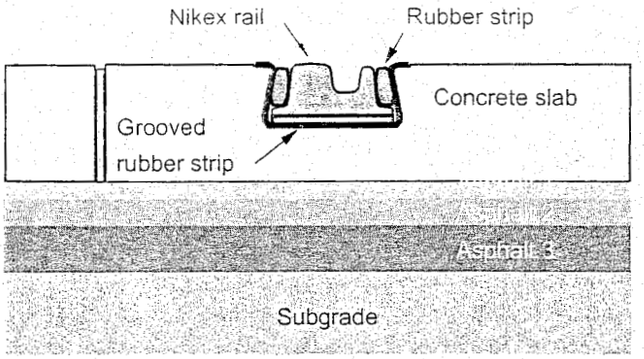


Figure 8.45: Paved-in track; Block rail in prestressed concrete slab

The last example is the so-called Nikex-structure (Figure 8.45). Here a special shaped block rail is inserted in the dovetail groove of a concrete slab. The rail is supported by a grooved rubber strip and is secured by rubber strips, which are put into place under pressure. In the Netherlands there are test tracks in the tramway network of HTM (the Hague) and RET (Rotterdam). After some initial problems they behave very well. The transition from Nikex track to standard tramway track should be carefully designed to avoid impact loads due to stiffness differences.

8.13 Crane Track

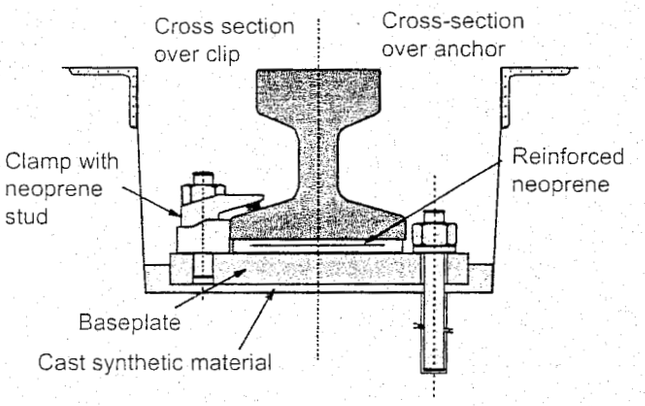


Figure 8.46: Continuously supported crane track meant for intensive use; Rail 127 kg/m concrete slab

The technical field 'crane track structures' can be regarded as a special application of railway engineering. It traditionally finds its place between the disciplines mechanical engineering and civil engineering because each crane track is strongly related to the type of crane which will use the track.

The following definition will be used: crane tracks are rail tracks on which cranes and all sorts of transshipment machines ride with wheel loads of 200-1000 kN and speeds of up to 150 m/min. Crane tracks may be fully or partly embedded in the supporting structure which can exist of wooden sleepers, heavy I-beams, or a concrete foundation.

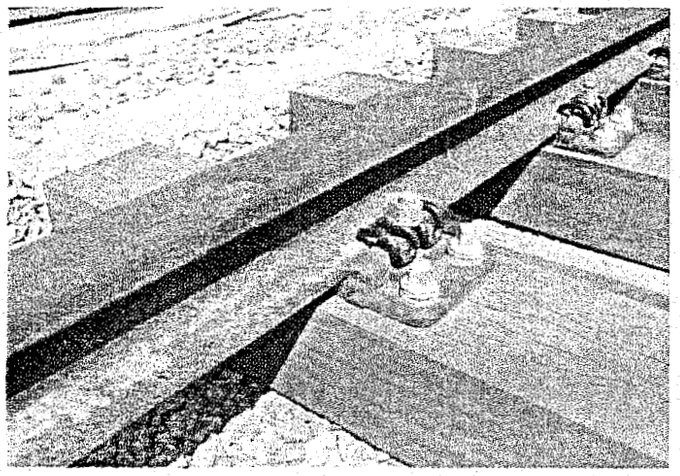


Figure 8.47: Crane rail at ECT

To withstand the high vertical load, the rail head and the wheel thread are made cylindrical and a steel quality of 700 N/mm² with a vanadium addition is used. An example of such a construction is given in Figure 8.46. Contrary to rail vehicles the wheels of a crane have flanges on both sides. The 'track width' may be as high as tens of meters. The horizontal load should also be considered as it is sometimes as high as 25% of the vertical wheel load.

The permitted tolerances for cranes in operation are of major importance and are, for instance, laid down in the Dutch standard NEN 2017. It is also interesting to note that a more flexible crane may more easily follow the crane track deformations than the stiff crane designs.

Finally, Figure 8.47 shows a picture of a crane track at the European Container Terminal in Rotterdam. The wheel load in this case is 600 KN, the rail has a mass of 125 kg/m.

9 SLAB TRACK

9.1 Introduction

Although most of the current railway tracks are still of a traditional ballasted type, recent applications tend more and more towards non-ballasted track. The major advantages of slab track are: low maintenance, high availability, low structure height, and low weight. In addition, recent life cycle studies have shown, that from the cost point of view, slab tracks might be very competitive.

Experiences in high-speed operation have revealed that ballasted tracks are more maintenance intensive. In particular, due to churning up of ballast particles at high-speed, serious damage can occur to wheels and rails, which is of course prevented in the case of slab track.

With the design of railway lines factors like life cycle cost, construction time, availability and durability play an increasingly important role. In this respect non-ballasted track concepts offer good opportunities. With the growth of traffic intensity it becomes more and more difficult to carry out maintenance and renewal work. On European networks, night time possessions often last no longer than 5 hours, or even less. Seen against this background, the current increase in the popularity of low-maintenance track designs is evident.

In the past new projects were mainly assessed on the basis of investment costs, whereas today the principle of life cycle costing is strongly emerging. As a result of these new attitude, ballasted track concepts will loose attractiveness in favour of slab track systems.

9.2 Ballasted track versus slab track

The general problem which occurs with ballasted track is the slow deterioration of the ballast material due to traffic loading. Ballast consists of packed loose granular material of which the grains wander, wear, and break up causing increasing geometrical unevenness and clogging of the ballast bed by fine particles which cause drainage problems. Therefore, regular maintenance is time after time needed to restore the track alignment.

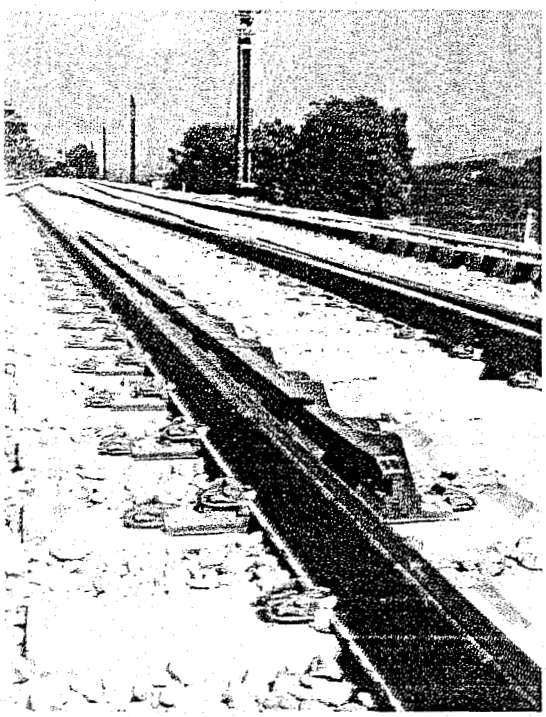


Figure 9.1: Ballasted track ...

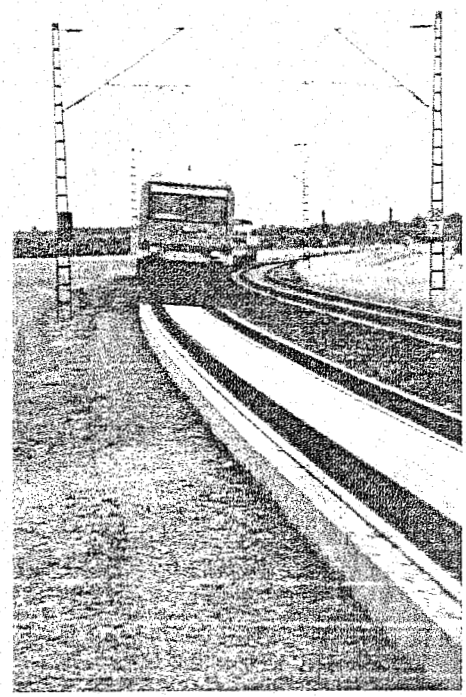


Figure 9.2: ... and slab track

In ballasted superstructures the sleepers support the rail at every 60 to 80cm distance (Figure 9.1). Each sleeper can give cause to aberrations in the track geometry. In case of a slab track (Figure 9.2), the rail supports or rail fasteners are in fixed position.

It is clear that slab track provides far higher lateral and longitudinal stability, and deviations of the track alignment are smaller and less likely to occur, given a good slab foundation. The condition of the track geometry is, therefore, very good and will remain likewise, improving passenger comfort and considerably decreasing the amount of maintenance.

9.2.1 Ballasted track

In comparison to slab track, disadvantages of ballasted track are (see also Chapter 8, [122], [161]):

- The tendency of the track to "float" in both longitudinal and lateral direction after a period of time;
- Limited non-compensated lateral acceleration in curves caused by the limited lateral resistance provided by ballast;
- Pulverisation of the ballast grains in the ballast bed resulting in particles damaging the rail and wheels;
- Reduced permeability due to contamination, the wear of the ballast, and intrusion of fine particles from the subgrade;
- A relatively heavy and high track structure requiring stronger constructions for bridges and viaducts.

The rate at which the track deteriorates is closely related to the quality of the track components and the way they are assembled (at construction), the track geometry, the homogeneity of the subgrade layers, and the supporting capacity of the subballast.

In the case of continuing ballast beds on bridges (or in tunnels), additional elasticity must be provided by:

- Application of ballast mats;
- Increased elasticity of the rail fastenings.

Nevertheless, maintenance will be needed at regular basis.

9.2.2 Slab track

In comparison to ballasted track, the advantages of slab track (or ballastless track) are in general reduction of maintenance and a higher stability of the track [42], [242]:

- The track is to a large extent maintenance free. Maintenance work like tamping, ballast cleaning, and track lining is unnecessary. Maintenance costs amount to 20 - 30% of the maintenance costs of ballasted track;
- Increased service life, and possibility of almost full replacement at the end of the service life;
- Near maximum availability and hardly any hindrance to residents due to scarce nightly maintenance works;
- No drag forces at ballast with the passing of high-speed trains;

- The unconditional use of electro-magnetic wheel brakes;
- The excess of superelevation and cant deficiency of the track with mixed use of freight- and passenger trains does not cause altering of the track position;
- Simple corrections up to 26 mm in vertical position and 5 mm in horizontal position are possible to counteract small deformations;
- Reduced structure height and weight.

Other reasons for applying slab track can be:

- The lack of suitable ballast material;
- Track accessibility to road vehicles;
- The abatement of noise and, in particular, vibration nuisance;
- Preventing the release of dust from the ballast bed into the environment.

In comparison to ballasted track, the disadvantages of slab track or ballastless track are in general:

- Higher construction costs;
- Higher airborne noise reflection;
- Large alterations in track position and superelevation can only be made possible by substantial amounts of work;
- Adaptability to larger displacements in the embankment is relatively small;
- In case of derailment, repair works will take much more time and effort;
- Transitions between ballasted track and slab track require attention.

The application of slab track may require extensive measures concerning the preparation of the foundation. The sublayers must be homogenous and capable of bearing the imposed loads without significant settlements. The resulting high construction costs have prevented widespread use of slab track for main lines so far [119], [242].

9.3 Designs of slab track superstructures

Currently, slab track is mainly being used on high-speed lines, light rail, and civil structures.

In Table 9.1 below an outline is given of the different types of existing superstructure. A more elaborate description of each kind will follow. A schematic overview is also presented in [23], [60].

There is a large variety of designs of slab track systems, each with specific features [14], [42], [197]. In Table 9.1 slab track systems are divided into six main types of design which can be applied as slabs with a high and a low flexural stiffness (seeTable 9.2).

SLAB TRACK SYSTEMS					
DISCRETE RAIL SUPPORT				CONTINUOUS RAIL SUPPORT	
With sleepers or blocks		Without sleepers			
Sleepers or blocks embedded in concrete	Sleepers on top of asphalt-concrete roadbed	Prefabricated concrete slabs	Monolithic in-situ slabs (on civil structures)	Embedded rail	Clamped and continuously supported rail
Rheda Rheda 2000 Züblin LVT	ATD	Shinkansen Bögl	Paved-in track On civil structures	Paved-in track Light rail Road crossing DeckTrack	CoconTrack ERL Vanguard KES
§ 9.4	§ 9.5	§ 9.6	§ 9.7	§ 9.8	§ 9.10

Table 9.1: An overview of the possibilities regarding the construction methods for ballastless track

SLAB TRACK SYSTEM	Flexural stiffness	
	Low	High
§ 9.4 Sleepers or blocks embedded in concrete	<——————>	
§ 9.5 Sleepers on top of asphalt-concrete roadbed	<——>	
§ 9.6 Prefabricated concrete slabs	<————>	
§ 9.7 Monolithic in-situ slabs (on civil structures)	<——————>	
§ 9.8 Embedded rail (paved-in track, DeckTrack®)	<————————>	
§ 9.10 Clamped and continuously supported rail	<——————>	

Table 9.2: A possible division between superstructures with high and low flexural stiffness

Slab track systems with a low flexural stiffness depend solely on the bearing capacity and the stiffness of the soil. The superstructure, either prefabricated or monolithic slabs, can hardly be exposed to bending forces. In case of unreliable and soft soils, a flexural stiff, reinforced slab can provide extra strength as it will be able to act as a bridge across weaker spots and local deformations in the substructure. In the case of reinforced monolithic slabs, it will be possible to apply embedded rail (Section 14.2).

Examples of several specific slab track systems will be given in the next 6 sections referred to in Table 9.2. In Section 9.9 the general and topical subject about flexural stiff slabs on top of soft soils will be discussed.

9.4 Sleepers or blocks embedded in concrete

Structures of concrete sleepers embedded in concrete are most commonly used in Germany for newly constructed high-speed lines. In general, the track consists of sleepers cast into concrete inside a concrete trough or directly on top of a concrete roadbed. Examples are the "Rheda"-system and the "Züblin"-system.

Other examples are the Low Vibration Track (LVT) or the Stedef system [60], illustrated in Figure 9.3, which is most often used in tunnels, for instance on the French TGV-tracks. This system is also commonly applied to metro systems. A rubber "boot" or an elastic encasing under the sleeper or an individual block provides a high degree of elasticity, which ensures good noise and vibration insulation (Chapter 15).

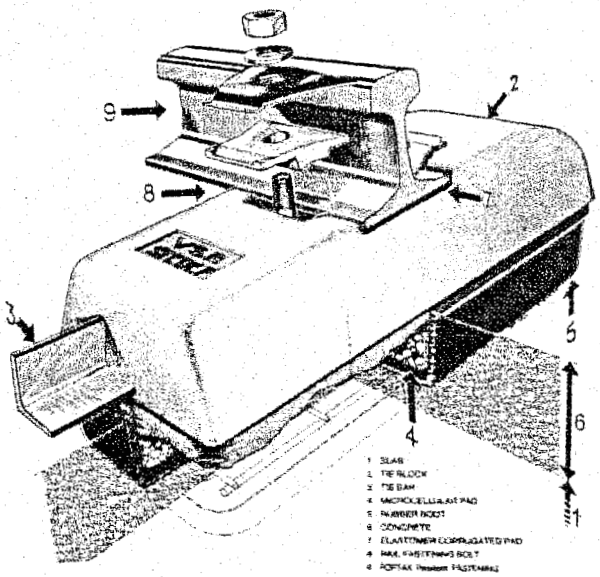


Figure 9.3: Stedef twinblock track system

9.4.1 Rheda 2000

The Rheda system has been under continuous development since its first application in 1970 [8], [42], [119], [197], [226].

The first design existed of complete monoblock sleepers embedded in-situ concrete in a reinforced concrete trough.

Figure 9.4 shows the same principle with twinblock sleepers. The sleepers were longitudinally connected with reinforcing steel to prevent them from loosening.

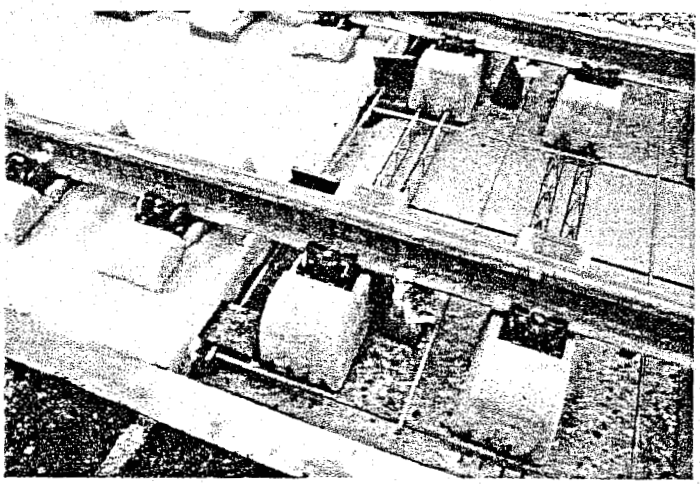


Figure 9.4: Twinblock sleepers adjusted inside a concrete trough and cast into concrete

The track framework was adjusted by means of the top-down method through spindles in lateral and longitudinal direction (Figure 9.5).

This structure might result in a stiff slab if sufficient reinforcing steel is applied inside the paved trough or at the bottom of the slab to make it able to bridge weaker spots in the supporting layers.

The Rheda 2000 Slab Track System has resulted from the ongoing development of the well-known German Rheda system. Early 1994, the modified system became known as the Rheda Berlin System.

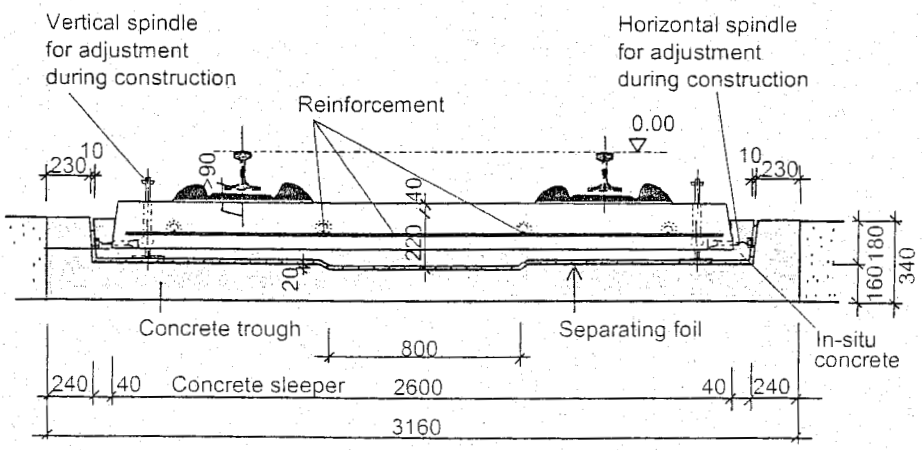


Figure 9.5: The location of the spindles for adjustment of the sleepers

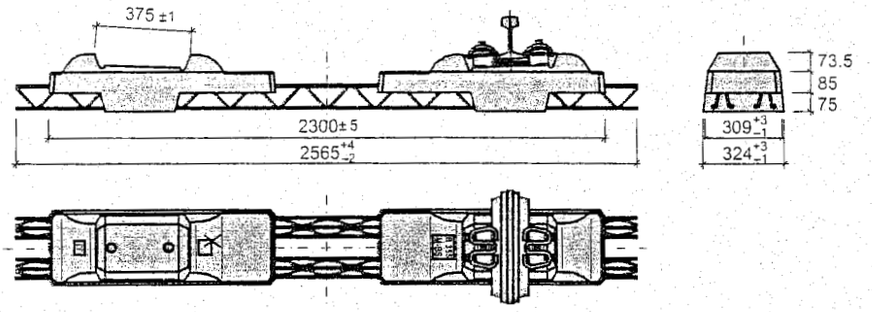


Figure 9.6: Twin-block supporting point (B 355 W60M-BS)

The Rheda 2000 system was first installed in July 2000, as part of the German high-speed section between Leipzig and Halle. Moreover, variants have been designed for commuter traffic applications (known as Rheda City).

The Rheda 2000 superstructure requires a settlement-free foundation because its reinforcing bars are applied at the centre of the concrete slab with as main function to regulate crack-width and transmit lateral forces, not with the purpose of providing a stiff slab.

Development goals and characteristics of Rheda 2000

In the Rheda 2000 system the known weak spots of previous designs have been eliminated, whereas the principle of a cast-concrete track layer, consisting of rails and a framework of sleepers (including elastic rail fastenings), was retained. The track is installed top-down using the top and inner edge of the rail as reference surface to achieve high precision track geometry. This technique furthermore neutralises the effects of unavoidable tolerances of track components. In addition, the rail fastening system allows for adjustments to compensate for long term differential settlements of the substructure.

The primary development goals for the Rheda system can be listed as follows:

- Improvement of the monolithic quality;
- Design optimisation;
- Uniform system architecture;
- Integrated techniques for slab track installation.

Improvement of the monolithic quality

In order to enhance the monolithic quality of the concrete slab, the sleepers in the Rheda 2000 system were changed from a monoblock design to a precisely dimensioned concrete filigree twin-block design as seen in Figure 9.6. The geometrically exact supporting points precisely secure the rails in their required position. In addition, they guarantee an optimal bond in the cast concrete of the track layer, due to the considerable amount of steel reinforcement, which protrudes from the concrete supporting points.

The high concrete quality of the supporting points furthermore ensures a strain-resistant bond between the rail fastening system and the concrete track layer. This solution consequently guarantees safe, long-term transfers of the high static and dynamic service loads encountered in high-speed traffic.

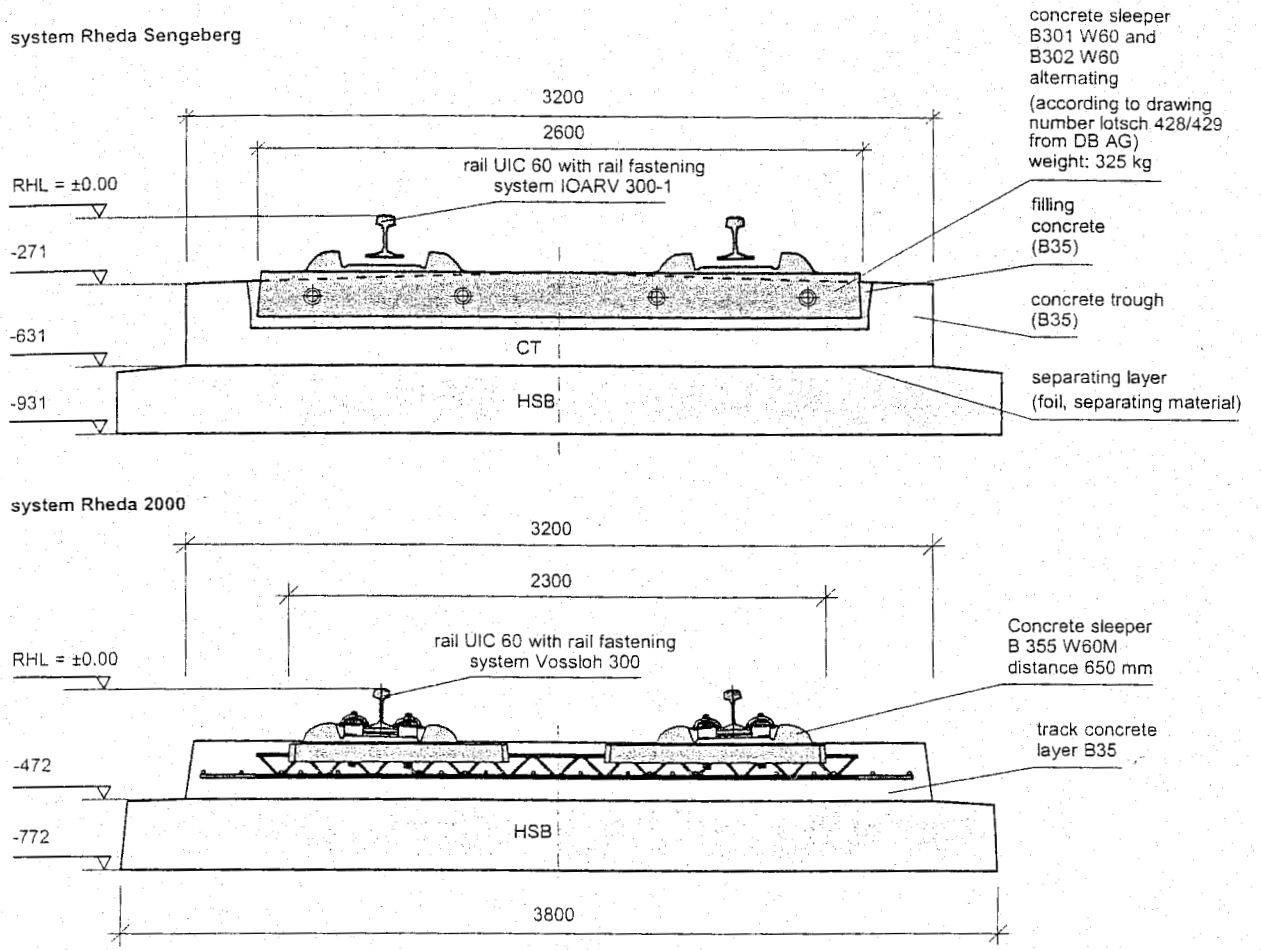


Figure 9.7: Comparison of cross sections: Rheda 2000 compared to Rheda Sengeberg

Monolithic qualities have furthermore been enhanced by elimination of the concrete trough used in earlier Rheda applications. Ideally, the longitudinal interface between trough and filler concrete should be fully closed and impervious to surface water. In actual practice, however, several cases have arisen in which a gap has in fact formed which allowed water to penetrate. If such occurrences continue, Rheda systems which have already been applied will require the necessary maintenance measures to prevent reduction of service life.

Design optimization

By eliminating the conventional concrete trough, a significant simplification of the overall system configuration was achieved. As a result, the entire cross section of the slab has become one monolithic component, whereas in the earlier trough configuration the filler concrete did not play a structural role. Due to the elimination of the trough and the use of twin-block sleepers, a considerable reduction of the structural height could be achieved as is revealed by the height figures in Figure 9.7.

Uniform system architecture

For reasons of cost-effectiveness in planning, engineering, and dimensioning, the Rheda 2000 system design had to fit a variety of track situations encountered in practice. From this objective, solutions have been developed for track systems on earthworks, on long and short bridges, in drilled tunnels, and for cut & cover tunnels. Figure 9.8, Figure 9.9, and Figure 9.10 present examples of system cross sections which in actual practice can also be effectively adapted to different conditions.

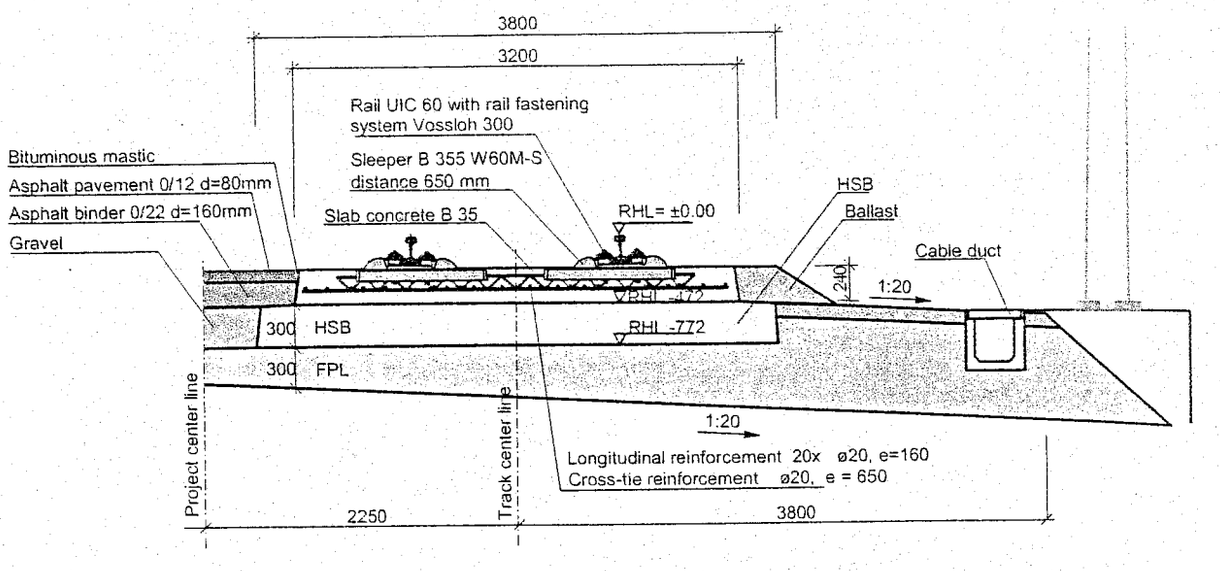


Figure 9.8: Rheda 2000 on earthworks (without cant)

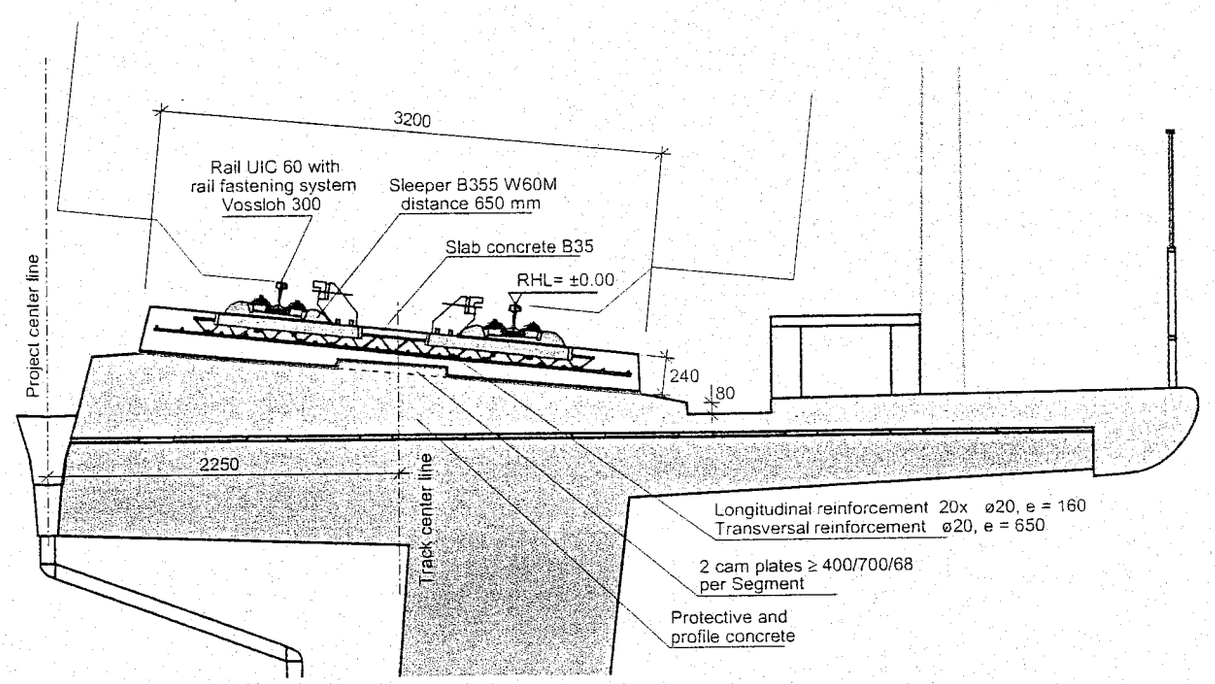


Figure 9.9: Rheda 2000 on long bridges (with cant)

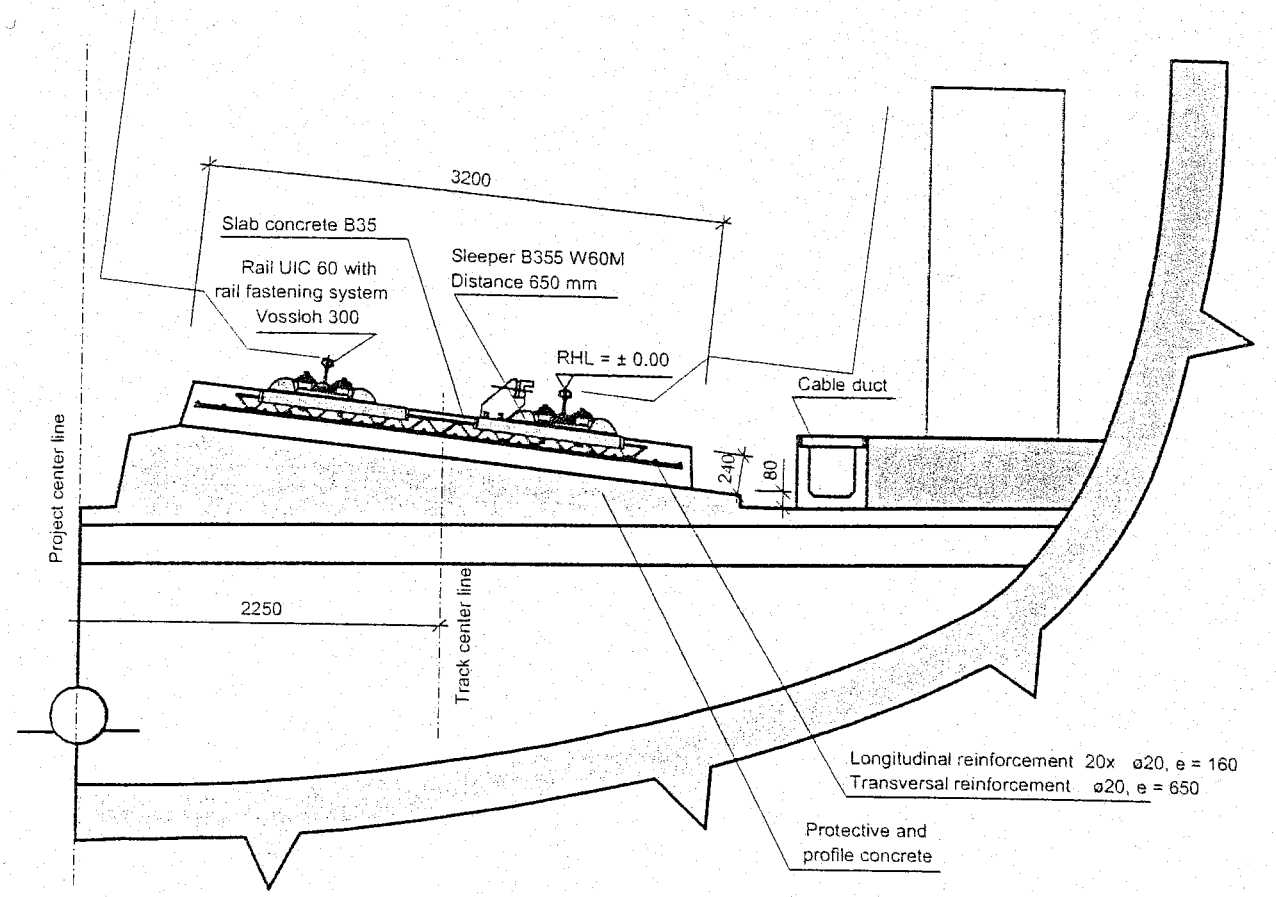


Figure 9.10: Rheda 2000 in drilled tunnels (with cant)

Rheda 2000 technology for turnouts

In order to maintain the equal system height in track standard sections with turnouts, a concrete turnout sleeper has been developed as an adaptation of the twinblock sleeper (Figure 9.11 and Figure 9.12). Due to the high lateral forces acting on turnouts, additional reinforcement in the sleepers and in the slab was necessary. All other components of the turnout system for slab track remained unmodified. These components have proved especially effective.

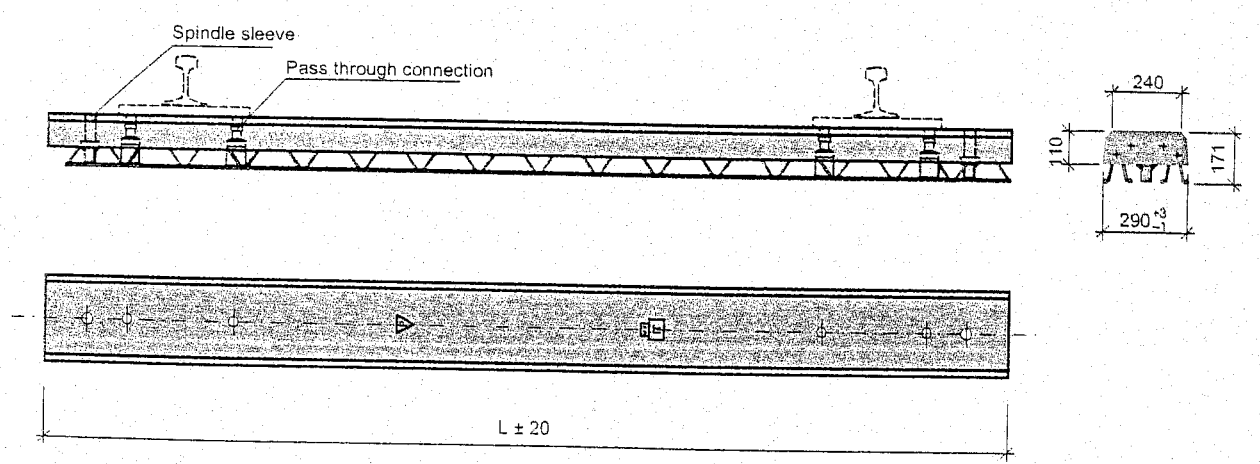


Figure 9.11: Turnout sleeper for system Rheda 2000

Transitions between ballasted track and Rheda 2000

Transition structures have been designed to provide transitions between ballasted track and Rheda 2000. Figure 9.13 and Figure 9.14 show examples from executed projects.

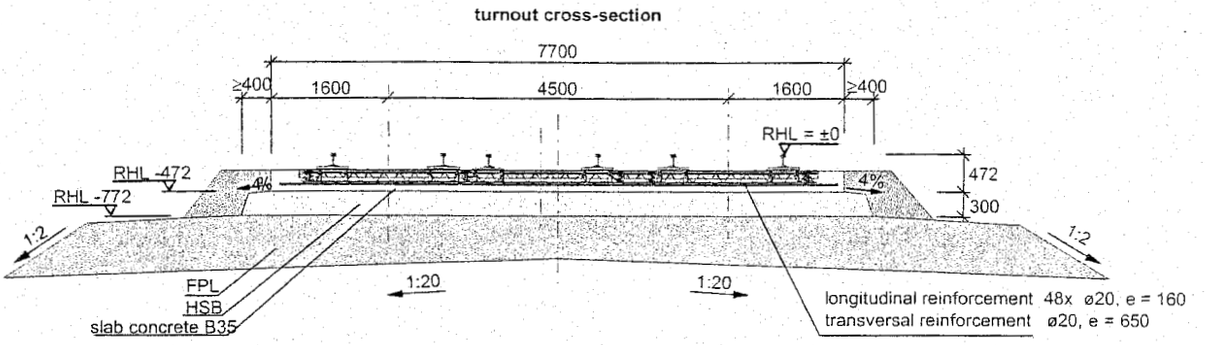


Figure 9.12: Cross section of turnouts

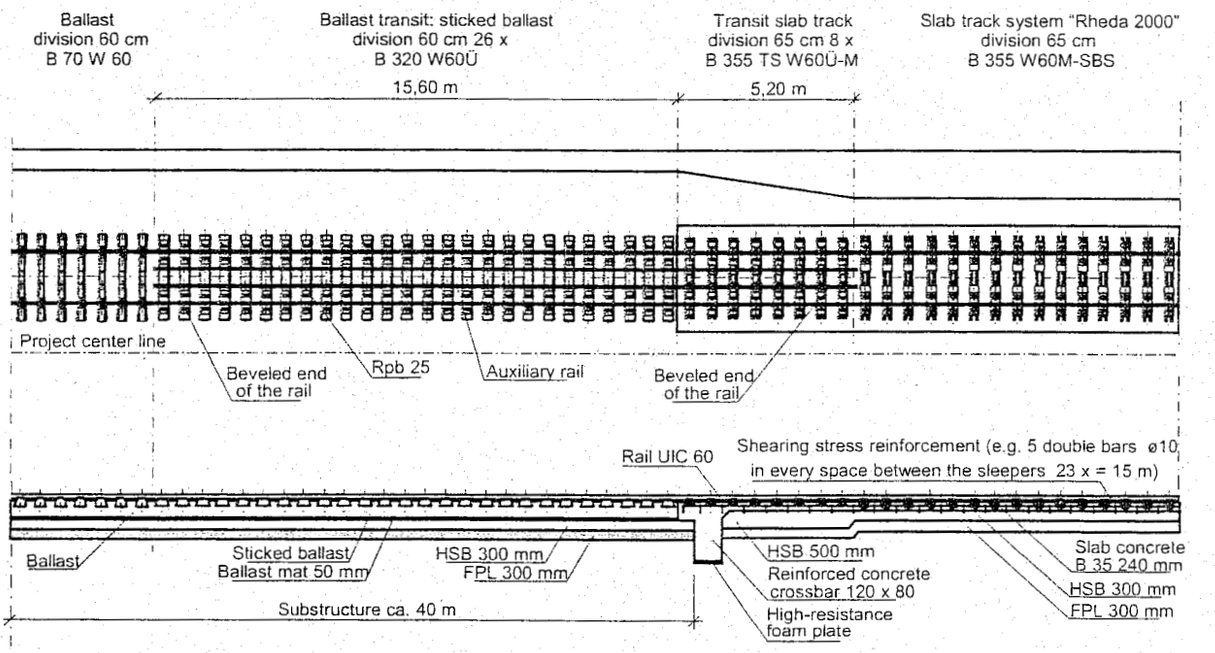


Figure 9.13: Transition between ballasted track and Rheda 2000 on an embankment

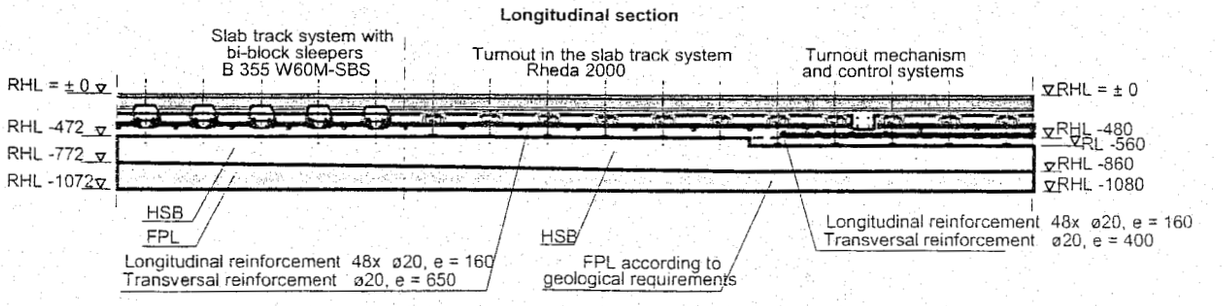


Figure 9.14: Transition between Rheda 2000 track and a turnout

Integrated techniques for slab track installation

In order to reduce the expensive construction costs a new installation concept was developed for Rheda 2000. By omitting the concrete trough, a complete step in the construction work sequence was eliminated. Application of the light twinblock sleepers significantly simplified their use at the construction site and at the same time enabled the mechanised installation of prefabricated track panels. Specially developed surveying techniques enhanced the cost effectiveness of the track installation process.

The installation of the Rheda 2000 system on earthworks begins with placement of a concrete roadbed by means of a slipform paver. In the case of engineering structures, the required protective and profile concrete is generally laid instead.

Application of the twinblock sleeper allows use of conventional track-installation processes. The foundation provided by the concrete base-sockets enables loaded construction vehicles to use the rails without difficulty before they are accurately positioned and secured in place. As a result, it is possible to lay the track in single-sleeper mode or in the form of assembled track panels.

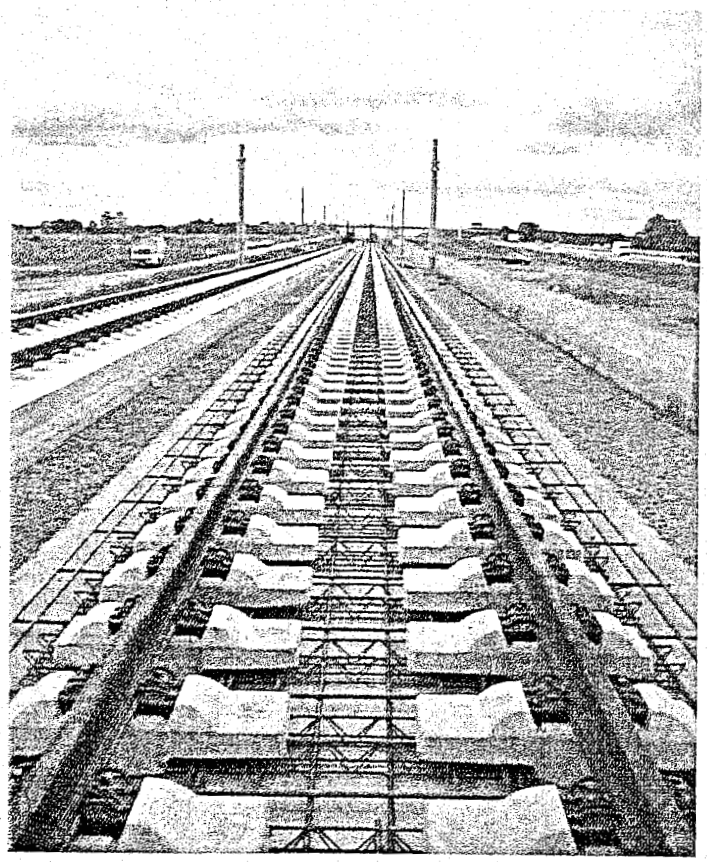


Figure 9.15: Track assembly, track on top of the concrete roadbed on the concrete roadbed (for the project Leipzig-Gröbers)

The arrangement of the slab layer reinforcement within the sleeper lattice-truss makes it possible for installation of the reinforcement to take place at the exact same time the track is laid. In this process, the construction crew places the required reinforcing rods on the concrete roadbed and inserts them section at a time through the lattice-girder compartments as shown in Figure 9.15.

Coarse and fine alignment of the track can take place with the aid of two techniques:

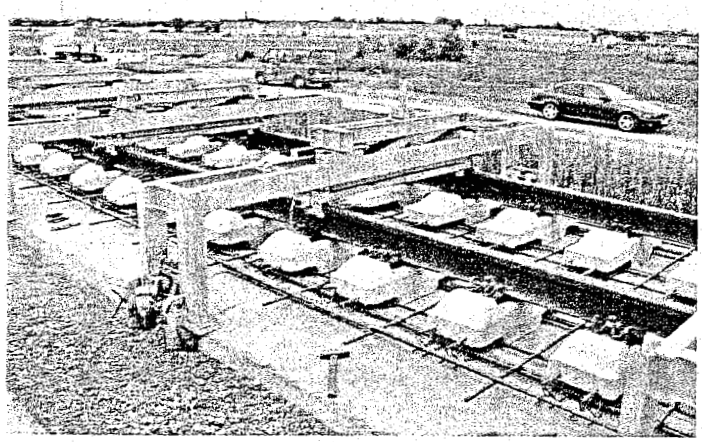


Figure 9.16: Alignment portals in the Leipzig-Gröbers project

- By means of alignment portal (see Figure 9.16): the portal units are first put into position with their feet anchored securely into the concrete roadbed after which the formwork elements are secured. The crew checks the installation for correctness. Next, the rail head clamps are lowered into place and fixed onto the rail as the track panel will be lifted approx. 9 cm and roughly aligned to ± 0.5 mm. The surveying crew gives instructions for the necessary settings to be made by the respective portal spindles for the superelevation (cant). After the final adjustments the track panel is secured and cleared for the pouring of concrete.

– By means of spindle and spreader bar alignment (see Figure 9.17). This method involves a combined technique of vertical adjustment by means of spindles installed at the ends of the sleepers and horizontal alignment by spreader bars. The spreader bars are attached to pipe bolts anchored in the concrete roadbed in the middle of the track between the sleeper units. Alignment takes place with the help of an eccentric adjustment unit. With the vertical spindles at the end of the sleepers, the required superelevation is adjusted first. Afterwards, these spindles are used to adjust the vertical position of the rails. Turning the eccentric units aligns the rails. After adjustment and final checking of the alignment, the track is cleared for the pouring of concrete.

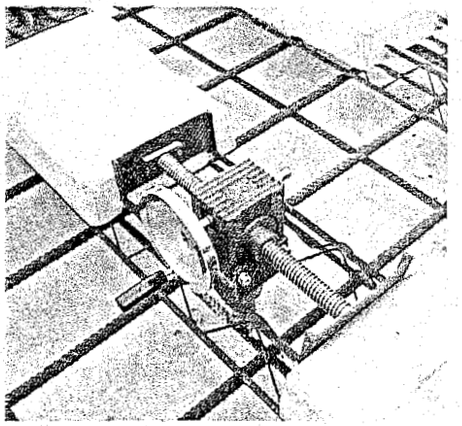


Figure 9.17: Alignment by spreader bars

Casting the track in concrete

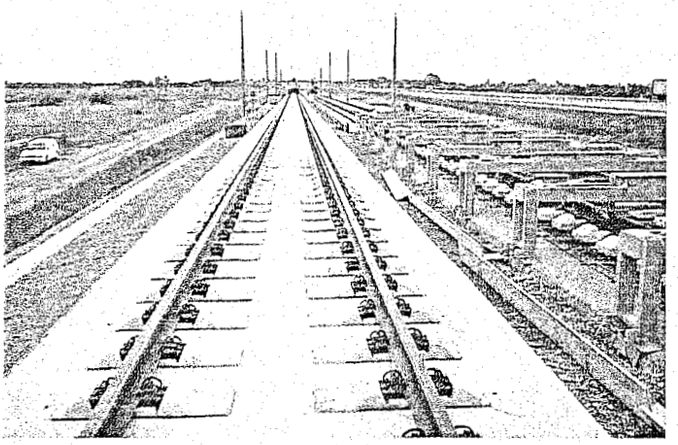


Figure 9.18: Finished track

The accurately aligned and secured track is then cast into concrete. Strict requirements for the quality of the concrete material and the execution of the concrete emplacement work are absolutely necessary. The structure is subdued to high dynamic loads which occur on high-speed train services. Concrete with a grade of B35, in the required consistency, is continuously poured in one direction in order to prevent any air entrapment under the sleepers. At the same time, it is compacted by means of vibrating equipment. Post-treatment and compacting of the fresh concrete ensures a high quality concrete structure which guarantees an optimal bond with the sleepers.

After this post-compacting the track panel is released from fixation and after another 6 hours the formwork is removed from the monolithic slab with sleepers (Figure 9.18). A cycle time harmonised with the individual operations enables a daily track production of 500 m.

9.4.2 Züblin

Development goals of the Züblin system

The Züblin system also consists of concrete twinblock or monoblock sleepers which are embedded in a monolithic concrete slab [42], [197], [293] (Figure 9.19). In 1974, developments started and directly focused on finding a highly economic constructing method. One has tried to achieve a high level of mechanisation combined with a low cost of labour and a high daily turn out of track.

This resulted in the fabrication of a continuous concrete slab without being bothered about a pre-positioned track framework. The main difference with other methods of slab track construction is the fact that the concrete sleepers are pushed into the fresh concrete by means of vibration. The concrete must be of a consistency which allows the sleepers to be pushed and vibrated into it, but does not allow the sleepers to sink (due to their mass) in.

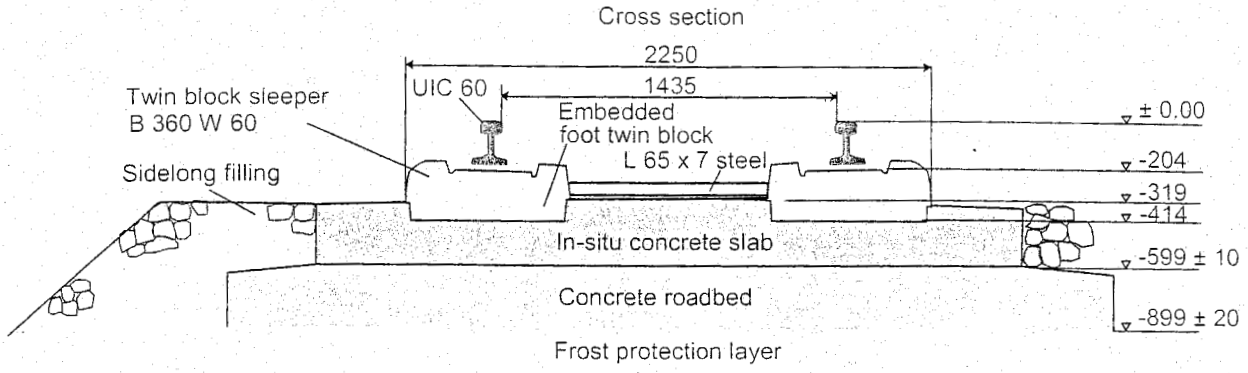


Figure 9.19: Cross-section of a Züblin superstructure

The wish for a high level of mechanisation lead to the development of a succession of machines; a slipform paver, a machine to vibrate panels of 8 sleepers into the fresh concrete, and a wagon for the supply of sleeper panels attached to frames.

Construction of Züblin slab track

After soil improvement a frost-protective layer is placed. On top of that a concrete roadbed is constructed upon which the monolithic concrete slab is made with a slipform paver. The production machines move over a rail which is integrated with the formwork elements necessary for the side of the slab (Figure 9.20). The reinforcement is placed on top of the concrete roadbed between the formwork yet before the slipform paver levels and fills the shuttering with concrete (Figure 9.21). With this system there is no need for a concrete trough.

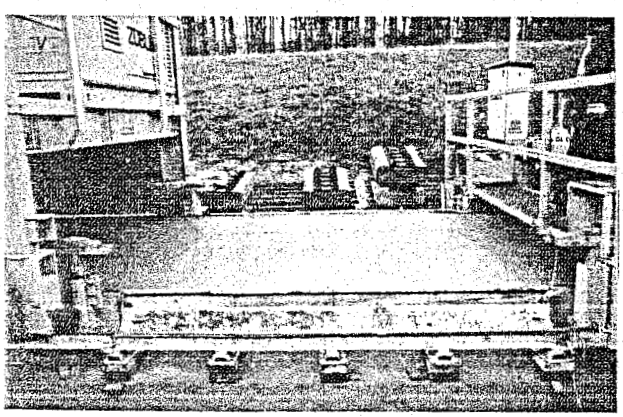


Figure 9.21: The freshly cast slab behind the slipform paver

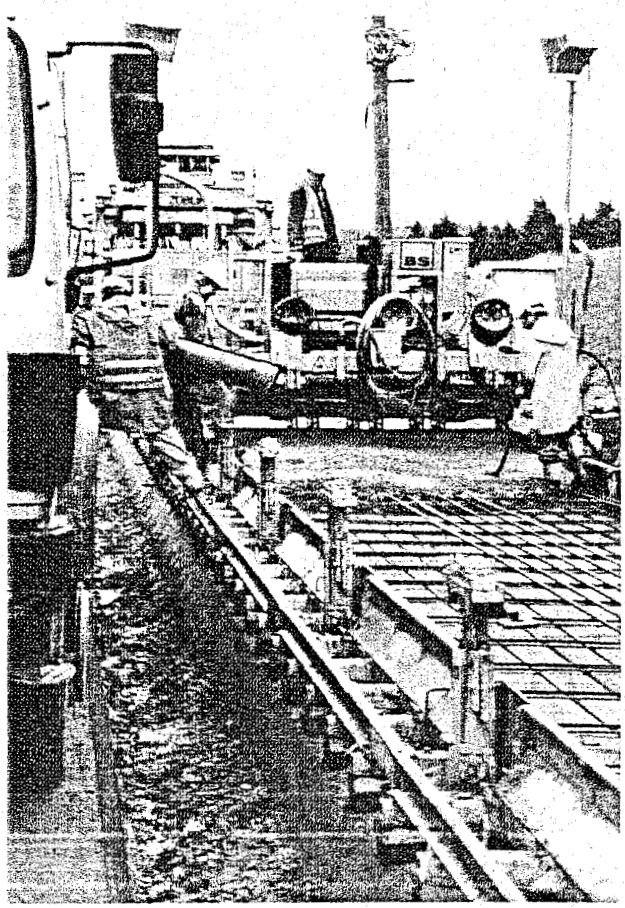


Figure 9.20: The formwork elements for the concrete slab act also as rail for the track laying machines

After the paver has laid the concrete slab, a machine follows which picks up the frames to which the sleepers are connected and pushes them into the freshly poured concrete by means of vibration to an exact, predestined position (Figure 9.22 and Figure 9.23). The preceding frame acts as a fixed point for the following frame. The exact position is determined by computerised theodolites which steer the placement of the frames.

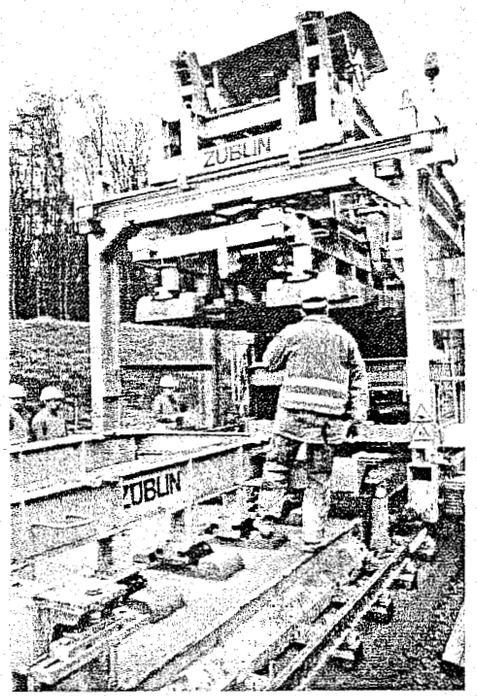


Figure 9.22: Panels of 5 sleepers at the time are pushed into the fresh concrete

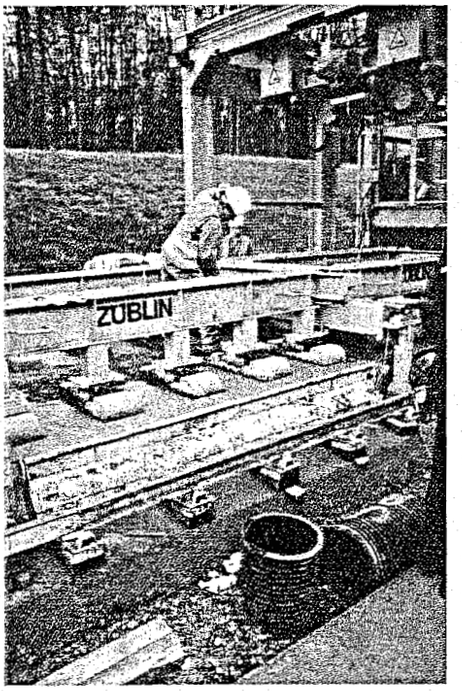


Figure 9.23: Newly placed sleepers

The surface of the slab is finished off by hand and after a given period, during which the concrete has hardened sufficiently, the frames are disconnected from the sleepers (Figure 9.24 and Figure 9.25). Afterwards the elastic rail fastenings and the rails are put into place. Within an 8-hour shift a production of 150 to 200 m can be achieved.

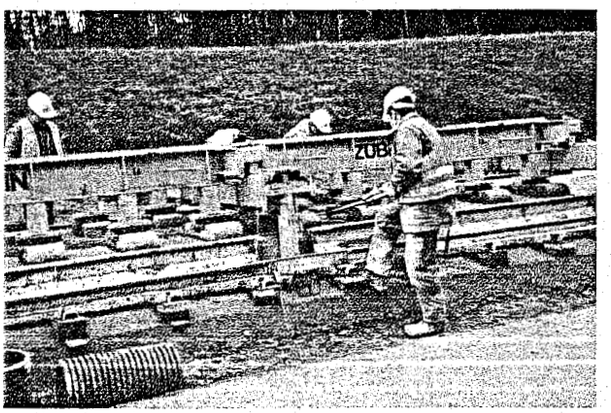


Figure 9.24: The concrete surface is levelled and smoothed by hand

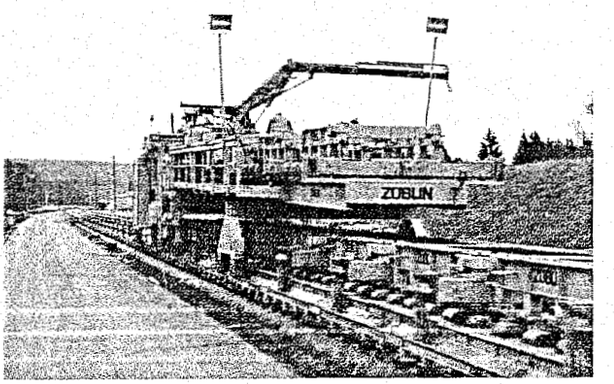


Figure 9.25: After sufficient hardening, the frames are disconnected from the sleepers, collected, and reloaded with new sleepers

Reinforcement principle

The use of the steel reinforcement in the Züblin slab is mainly to achieve a regular crack pattern and acceptable crack-widths. With the Züblin method, the reinforcing rods are not placed in the neutral line, but more to the bottom (Figure 9.26) of the slab giving it some flexural stiffness too. By increasing the height of the slab and the amount of reinforcement, the Züblin method lends itself to be constructed as a rigid slab in areas where uncertainties exist regarding the bearing capacity of the soil.

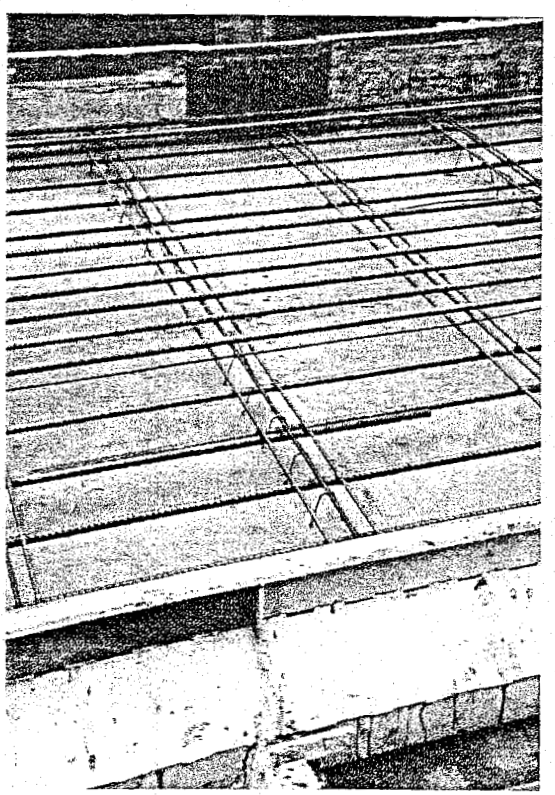


Figure 9.26: Steel reinforcement of the slab

9.5 Structures with asphalt-concrete roadbed

With this kind of ballastless structure, classified in the group of slab track structures though without a slab, sleepers are placed on top of a layer of asphalt-concrete roadbed. This roadbed is made with ordinary paving machines. It is possible to reach level accuracy of ± 2 mm and cant-deficiencies of up to 180 mm. Hence, the asphalt provides a stiff and level base for the sleepers. The space between the sleepers can be filled with ballast providing extra stability and reducing noise production. There is also the possibility of applying slabs on top of a layer of asphalt-concrete. Asphalt does not require hardening and can be subjected to loading immediately after cooling, so high construction productivity can be achieved.

Due to its specific material properties, asphalt can be suitable for slab track structures. The used asphalt mixtures do need alteration compared to the mixtures used for roads. In road construction, the important properties regarding the surface (high friction-resistance and resistance to wear and tear) are subordinate to resistance to deformation and stability. In the case of ballastless superstructures, resistance to deformation and stability are the required most [160].

Asphalt will adapt itself to stresses caused by loading and temperature changes which cause yielding. Therefore, asphalt can be applied as a continuous roadbed. Noise and vibrations are less compared to concrete due to the internal damping properties of asphalt. The material is moreover reusable. In several countries promising experiences have been had with this kind of track.

In Germany, several tests have been conducted with stretches of sleepers directly applied on top of asphalt: 1993, 14500 m at Nantenbach; 1994, 4000 m at Gotha-Eisenbach; 1996, 10260 m at Berlin-Hannover.

The so-called ATD-system consists of a concrete roadbed with a layer of asphalt on top and directly applied sleepers on top of that [42], [160], [162], [248], [293].

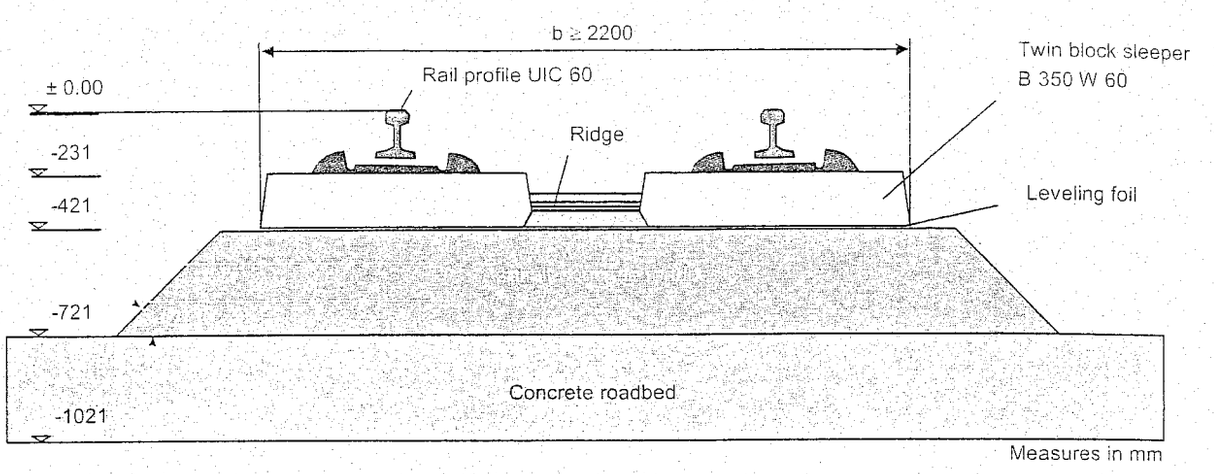


Figure 9.27: Cross-section of a superstructure with an asphalt concrete roadbed

The asphalt is applied in a similar fashion to roads with a +/- 2 mm accuracy. Special mixes have been developed to guarantee a 50 to 60 year life span under extreme weather conditions. The thickness of the layer varies on average between 25 and 30 cm.
In the middle, a ridge provides resistance to lateral forces (Figure 9.27 and Figure 9.28).
After horizontal adjustment of the rail, the remaining free space between the ridge, the roadbed and the sleeper is filled with an elastic synthetic material. Results have shown that asphalt layers can provide a stable and supportive layer suitable for ballastless superstructures.

In general, these tracks have proved to be satisfactory. A positive feature is the slight plastic adaptability that these tracks have. When higher pressures occur below certain sleepers than below others, asphalt will deform because of its visco-elastic properties until a new equilibrium has been reached leaving the pressures more levelled. Moreover, the system allows exchange of sleepers in case of damage by derailments.

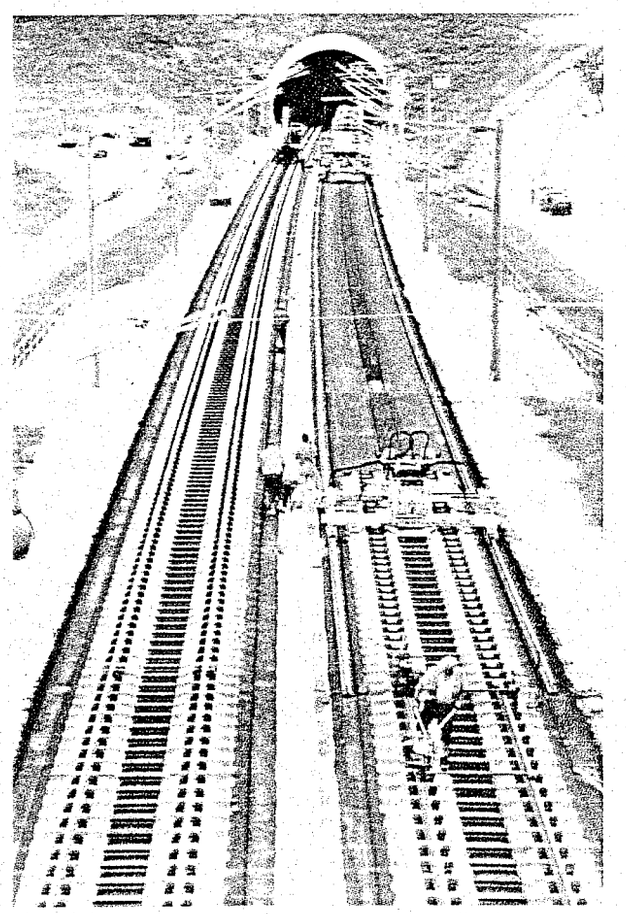


Figure 9.28: Asphalt-concrete superstructure under construction

9.6 Prefabricated slabs

Prefabrication of track components can contribute to quality. Prefabricated slabs are applied in several places around the world.

Other advantages of prefabrication in general and slab track in particular are:

- High level of mechanisation possible;
- Labour-saving construction at site;
- The rail can be directly adjusted and fixed;
- Less immune to failing workmanship;
- Repair and renovation friendly.

Prefabricated slabs with direct rail fastening have mainly been used in Japan on the Shinkansen-lines and in Italy with the IPA-system derived from the Japanese system. The use of prefabricated elements avoids having to process wet concrete during construction.

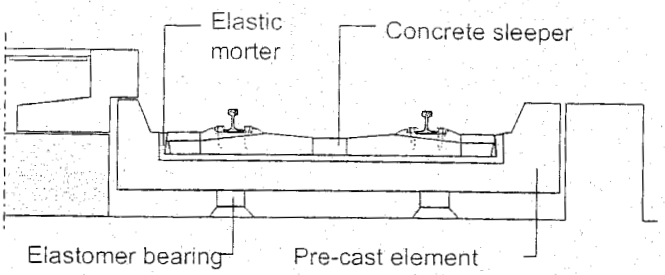


Figure 9.29: Floating slab (or 'Eisenmann track') installed on London Underground

The "Eisenmann track", displayed in Figure 9.29 [60] and - among others - applied in Munich and Frankfurt am Main S-Bahn systems, features pre-cast reinforced concrete slab-segments by elastomeric bearings supported on the tunnel floor.

Many urban rail systems have comparable prefabricated slab systems installed on particular network segments. This system is being applied in vulnerable areas with respect to noise and vibration nuisance. Although this system is very effective, it consumes considerable height and is expensive.

9.6.1 Shinkansen slab track

In Japan, bad experiences with the 516-km long Tokaido-line which was opened in 1964 and consists of traditional ballasted track led to the development of slab track in prefabricated slabs. In 1975, the Sanjo-line (562-km) was opened as an extension from the Tokaido-line to Hokkaido. In 1982, design improved with the 470 km Tohoku-line and the 275-km Joetsu-line. The amount of earth embankments in those lines is less than 5 % [161].

The Japanese Shinkansen slab track consists of a sub-layer stabilised with cement (concrete roadbed), cylindrical bollards or dowels to prevent lateral and longitudinal movement, and reinforced prestressed concrete slabs measuring 4.93 m x 2.34 m x 0.19 m and only 0.16 m thick in tunnels. The slabs weigh approximately 5 tonnes. The slabs are adjusted on top of the concrete or asphalt-concrete roadbed with spindles. Afterwards, a bituminous cement mortar is injected underneath and in between the slabs (Figure 9.30).

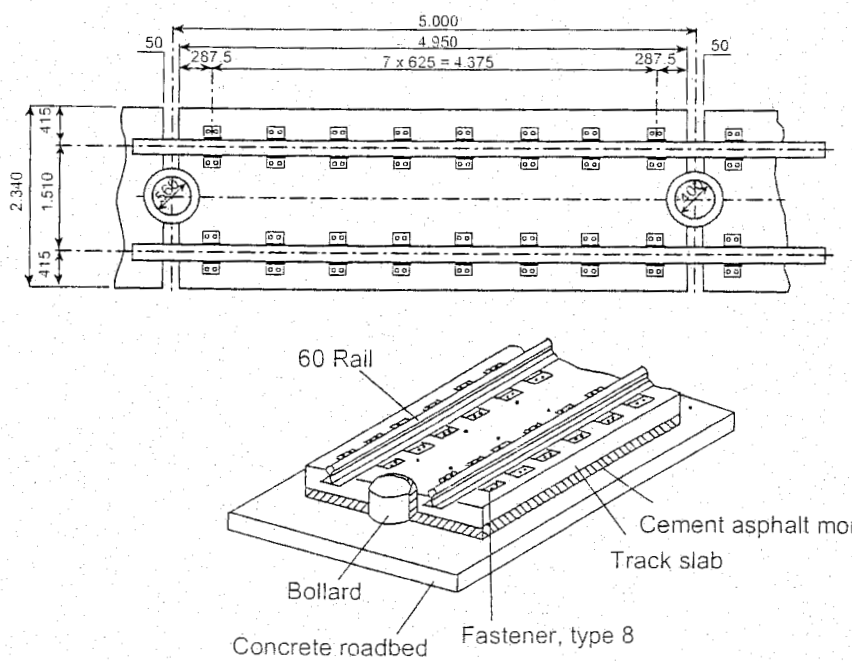


Figure 9.30: Shinkansen slab track

9.6.2 Recent design of Shinkansen slab track

The track structures shown in Table 9.3 refer to Hokuriku Shinkansen which was inaugurated in October 1997.

Track structure	Length (km)
Slab	210.452
Ballast	37.184
Total	247.636

Table 9.3: Track on Hokuriku Shinkansen

For the slab track two kinds of slab are used. The first is the type of flat slab which has been used so far but with a smaller width as shown in Figure 9.31. The width is narrowed from 2340 mm to 2220 mm.

The second kind is the frame type of track as shown in Figure 9.32.The latter is mostly used in tunnels. The first kind is supported on the slab mat in residential areas as shown in Figure 9.33. The slab mats are arranged according to Figure 9.34. The elasticity of slab mat is guaranteed with grooves.

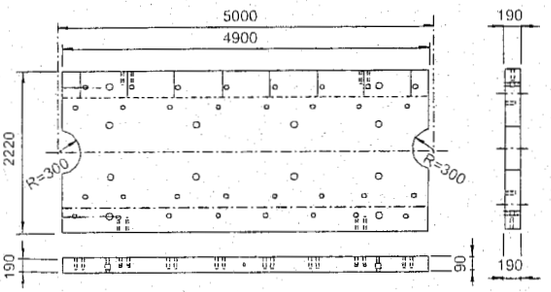


Figure 9.31: Normal track slab (A-55C) used on Hokuriku Shinkansen

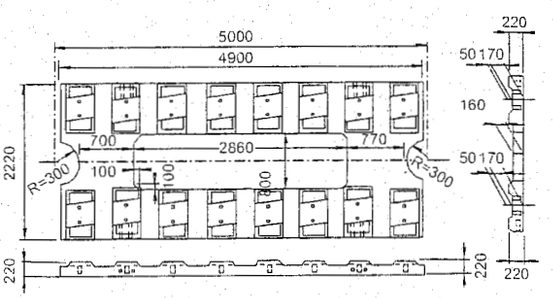


Figure 9.32: Track slab used in the tunnel on Hokuriku Shinkansen

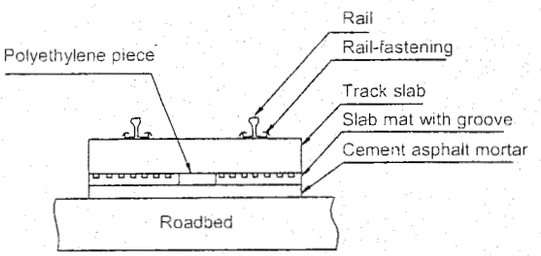


Figure 9.33: Slab mat supporting normal track slab

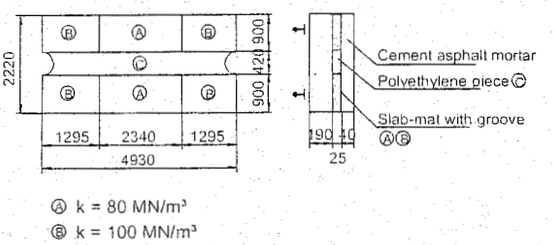


Figure 9.34: Arrangement of slab mat under track slab

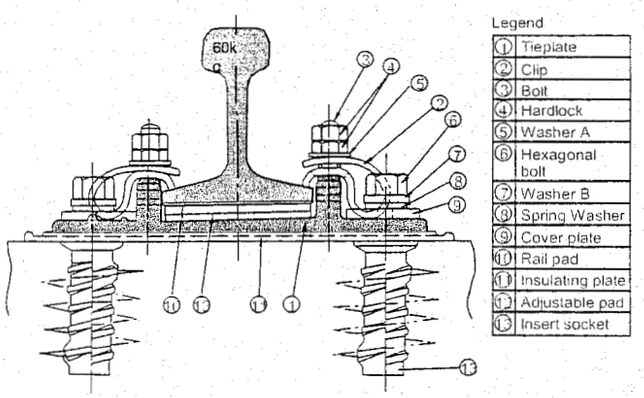


Figure 9.35: Rail fastening device of type 8 ameliorated for Hokuriku Shinkansen

The cement asphalt (CA) mortar is put in the long tube made of unwoven textile which is placed under the frame type of slab track.

The rail fastening devices are of type 8, but the socket is changed from collar type to insert type as shown in Figure 9.35. Hard locks are also added to prevent nuts from loosening. In the tunnel type 4K is used. The diameter of a bolt is 25.4 mm. The pitch is 6 mm. The fastening torque is 60 Ncm.

Some sections are put to trial using Pandrol rail fastening devices instead.

In Japan, only the static values of rail pads are included in the specification for 10-50 kN.

The normal value for all fastening types type 8 is 60 MN/m.

The machine which tests the rail fastening device for loading in two directions is shown in Figure 9.36.

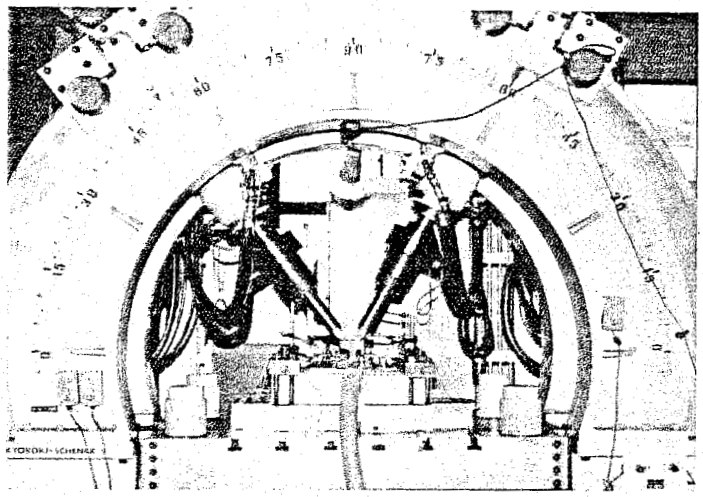


Figure 9.36: Two axial loading machine

The following items are tested by this machine:

Static performance:

- Creep resistance of the rail;
- Spring coefficient at the tip of the clip;
- Vertical spring coefficient of the rail pad;
- Lateral spring constant and the lateral strength;
- Static oblique loading test.

Dynamic performance:

- Fluctuating stress of the clip;
- Lateral displacement of the rail;
- Tilting angle of the rail;
- Dynamic test in oblique loading;
- Endurance test of rail fastening device.

Installing machinery

Laying the slab track on Hokuriku Shinkansen is illustrated in Figure 9.37.

This method called "Moving plant on running rail method" was initiated in the construction of Sanyo Shinkansen and it confirmed in the construction of Oshimizu Tunnel on Joetsu Shinkansen that it is the fastest method for laying slab track in long lengths in one direction.
The rails, which are welded in the depot and are 200 m long, are brought to the location where they will be laid on trolleys. These trolleys are driven by a motor car running on the temporarily laid rails which have a gauge of 1435 mm. The temporarily laid rails are used to follow the slabs as they are being laid, and for putting CA mortar under the slab in a neighbouring line. In this method, the temporarily laid rails are also used on the completed track too.

The rails are also laid on the road bed of the neighbouring track. Slabs are transported to the site and laid by widening the rail gauge to 3000 mm. As the rails with a 3000 mm gauge are used to transport slabs from the trolleys onto the rails laid with a gauge of 1435, it is necessary to double the rails at the location of the trolleys using additional rails for a section of 75m long.

The CA mortar is transferred from the moving plant running on rails with a gauge of 1435mm onto the neighbouring track. This moving plant fabricates CA mortar at the site by mixing liquid asphalt, cement, sand, etc. After putting CA mortar under the slabs and solidifying it, the rails are put on the slabs and fastened with the normal rail fastening devices.

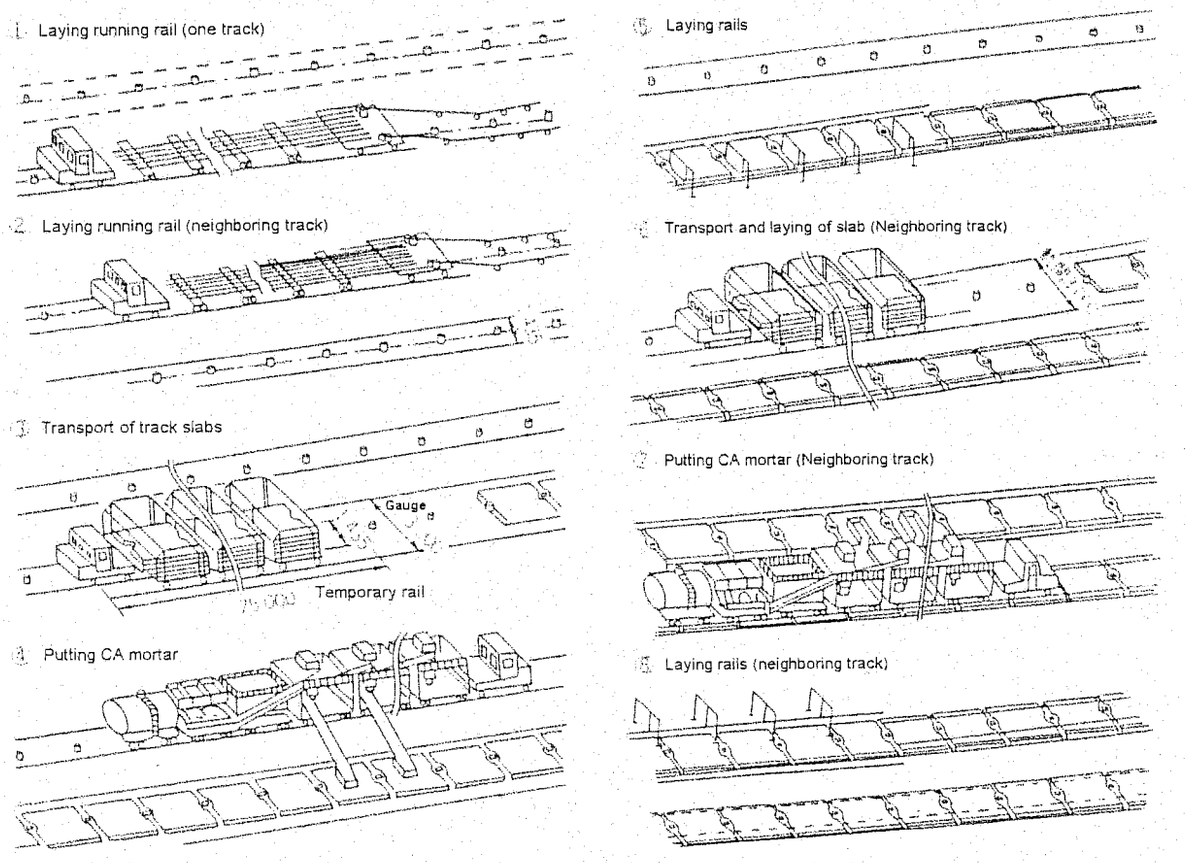


Figure 9.37: Slab track laying on Hokuriku Shinkansen

Putting CA mortar under the slabs is shown in Figure 9.38.

The installation speed on Hokuriku Shinkansen was 200-280 m/day.

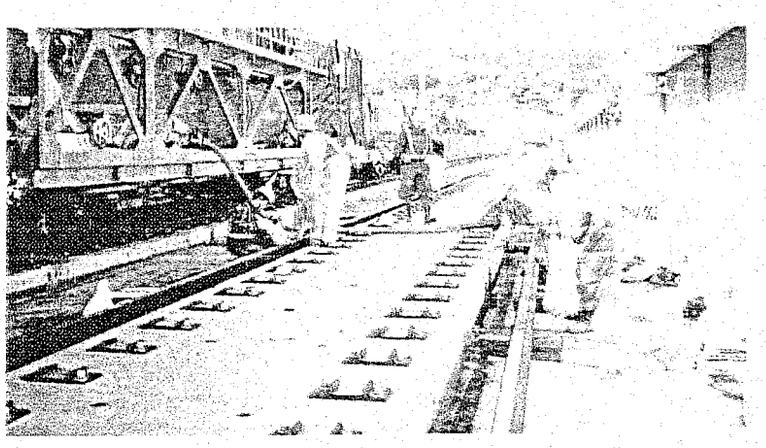


Figure 9.38: Putting CA mortar under track slab

9.6.3 Bögl slab track

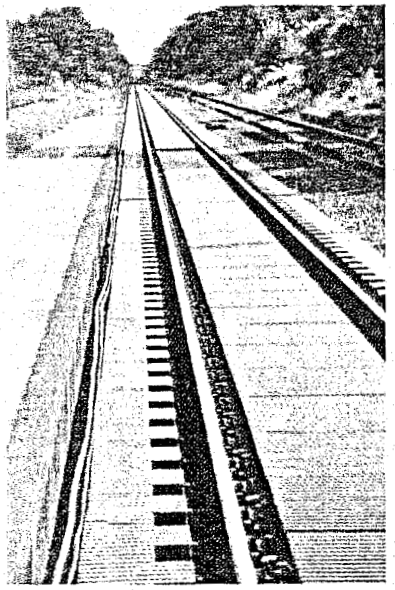


Figure 9.39: Bögl slab track with noise absorbing concrete cover

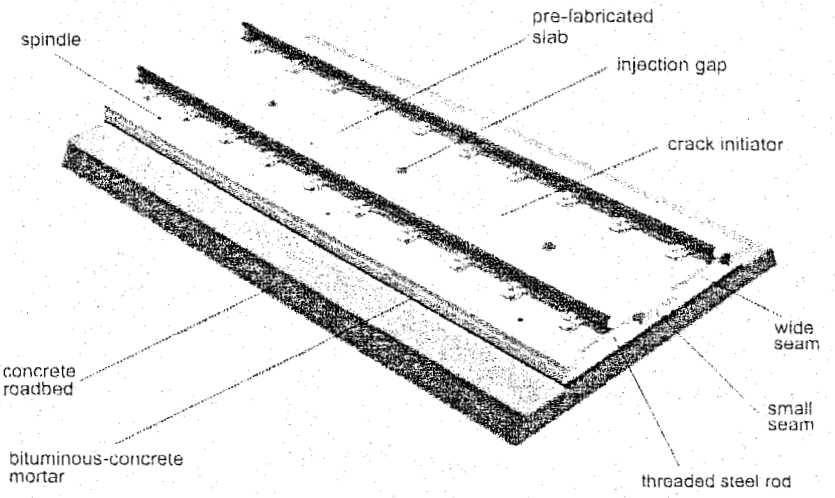


Figure 9.40: Bögl slab track system

The first application of the Bögl slab track was at Karslfeld, Germany, in 1977. The goal was to produce a long-endurance, maintenance-free track specifically meant for high-speed traffic with an enduring high track quality and the highest possible track availability. Ongoing developments have lead to a high quality slab track system (Figure 9.39).

Characteristics of Bögl slab track

The Bögl slab track system is largely similar to Shinkansen slabs except that the Bögl slabs are made of B55 steel fibre reinforced concrete and are 20 cm thick, 6.45 m long, and 2.55 or 2.80 m wide. The slabs are pre-stressed in lateral direction; in longitudinal direction traditional reinforcement is applied. Spindles integrated in the slabs provide an easy and quick adjustment of the slabs (Figure 9.40).

Construction of Bögl slab track

A crane places the slabs on top of a concrete roadbed or an asphalt-concrete top layer. The longitudinal threaded reinforcement rods, which extend from the end of the slabs, are loosely connected to each other with the help of an extended nut.
After that the rails are put into position and fastened. Afterwards the slabs are adjusted with the help of six integrated spindles.

Once the slabs have been rightly adjusted, a mortar is used to cover up the void at the edge between the slab and the roadbed before a bituminous cement mortar is injected underneath the slab.

After hardening of this mortar, the slabs are rigidly connected to each other at the edges in two steps. The heads of the slabs are toothed and by first filling the smaller void in between the slabs with concrete, it becomes possible, after hardening of the concrete, to wrench the extended nut (Figure 9.41) over both rods in order to achieve pre-stress force across the joint.

After this the wider void in between the slabs is filled with concrete, leaving a continuous and rigid slab (Figure 9.42). This slab track system has no cylindrical dowels to prevent longitudinal or lateral movements, yet depends solely on the frictional resistance between the slabs, the injected bituminous cement mortar, and the concrete roadbed.

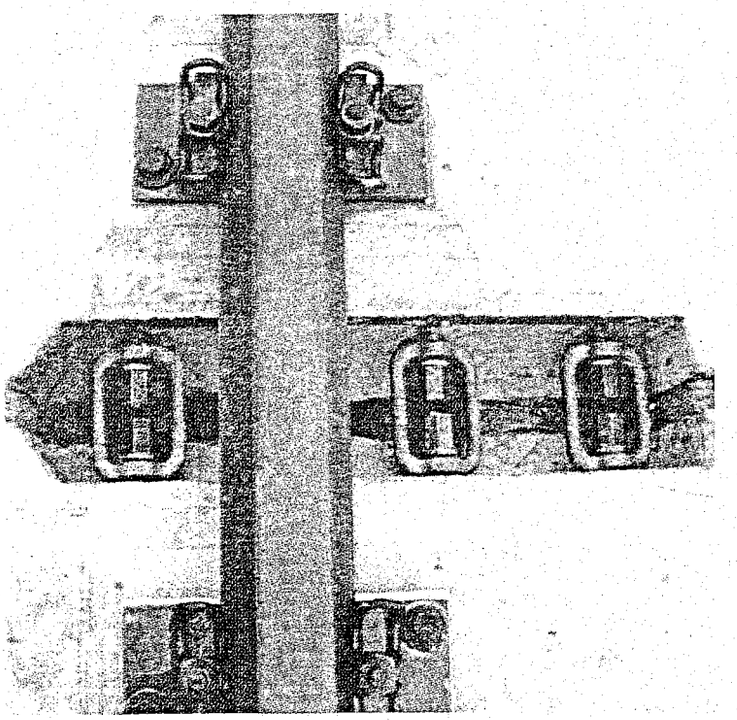


Figure 9.41: Connection of the longitudinal steel rods between two slabs

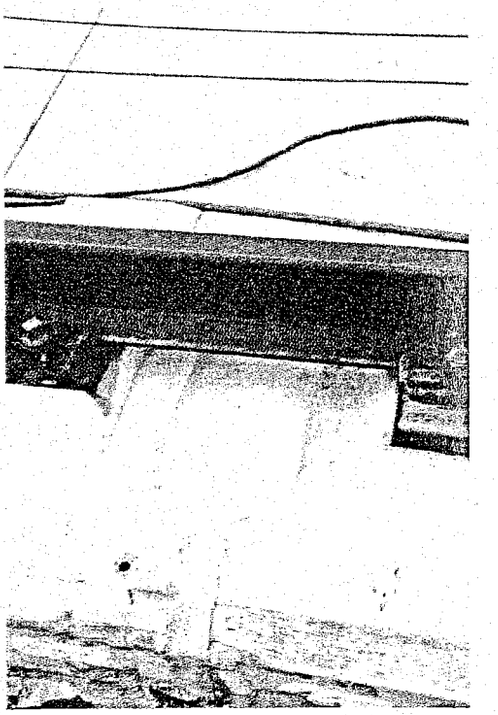


Figure 9.42: Detail of the filled joint between two slabs

9.7 Monolithic slabs and civil structures

For the application of direct rail fastenings, one could think of track on civils like bridges or of track on continuous monolithic slabs without sleepers. The direct rail fastening (Section 8.9) systems are bolted straight onto the concrete or steel decks or floors.

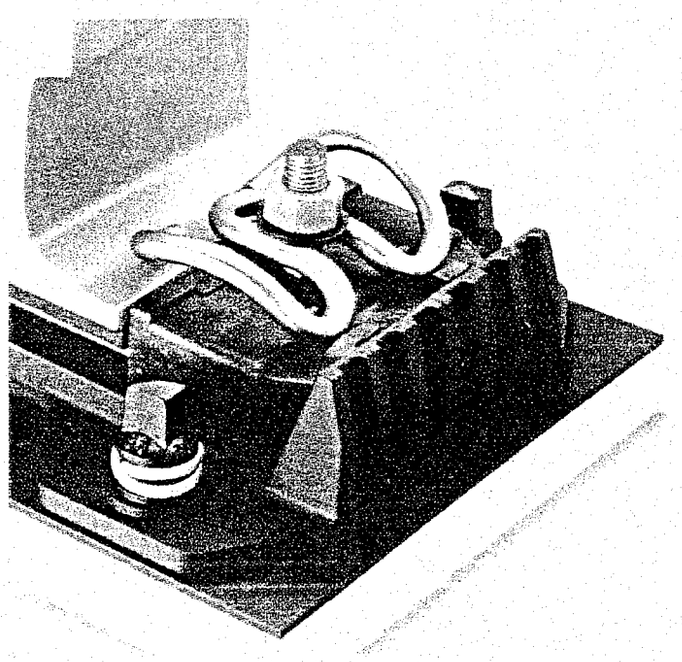


Figure 9.43: Vossloh DFF 300 rail fastening

An example of a direct fastening system is the Vossloh DFF 300 system with clip Skl 15 (Figure 9.43).

This system can be adjusted by approximately 50 mm in vertical and by 46 mm in lateral direction. The rail is permanently fastened with two clips which have a long run elastic spring deflection. The steel base plate itself is fastened on the concrete slab with four screws in plastic dowels or, alternatively, with anchor bolts in steel decks.

With the application of direct rail fastening instead of sleepers in ballasted track, bridges could be made lighter. Moreover, problems with sleepers on the bridges were solved. The direct rail fastenings are applied directly on top of the steel or concrete bridge deck.

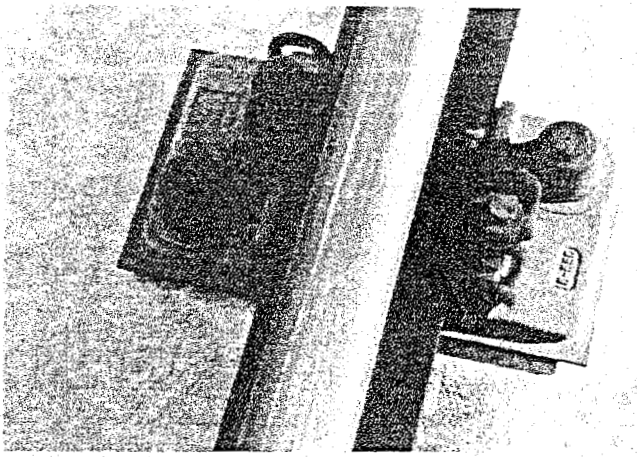


Figure 9.44: Direct rail fastening on concrete slab

A typical example of a direct rail fastening in The Netherlands consists of a base plate on top of base plate pad which is connected to the bridge. The pad in between has high elasticity as substitute for the elasticity of the ballast bed. In the case of track alignment, the base plate is horizontally adjustable with regard to the slab or the deck.

On steel bridges the anchors are bolted to the steel deck. On concrete bridges, holes are drilled in which the bolts are chemically anchored in the concrete deck (Figure 9.44) [200].

Several slab track systems with direct rail fastening have been developed over the years (Figure 9.45). In the case of these systems, a slipform paver places a monolithic slab on top of a concrete roadbed.

Similar to the application on concrete bridges, bores are afterwards drilled inside this slab in order to anchor the base plates.

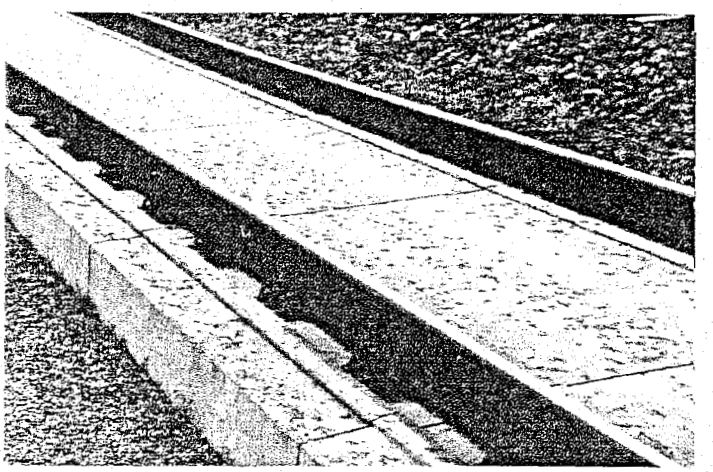


Figure 9.45: An example of a slab track structure with direct rail fastening (Heilit-Wörner)

Alternatively, track panels are aligned together with base plates and fasteners and poured into concrete.

This method of construction lends itself perfectly to create a continuous stiff and rigid slab, which will react like a continuous supported elastic beam under traffic loading.

The advantage of this is that it will become possible to apply slab track on top of weak and soft layers. The rigid slab spreads the loads across a much wider and longer surface and across local defects.

The minimum requirements for the substructure can be lowered, saving costs for ground improvements in certain situations. Such a structure can also be applied on soils with a great risk of local discontinuities in the bearing capacity of the subsoil.

9.8 Embedded Rail

9.8.1 The characteristics of embedded rail

Embedded Rail Structure (ERS) involves continuous support of the rail by means of a compound e.g. consisting of cork and polyurethane. The rails are fixated by means of this elastic compound which surrounds almost the entire rail profile except for the rail head. Characteristic of this concept is the lack of additional elements to secure track gauge. This concept can cover the full range from light rail to high-speed tracks.

The method of rail fixation is characterised by the following principles:

- Continuous support of the rail on an elastic strip;
- Guiding of the rail by elastic fixation in a groove;
- Top-down alignment of the rails;
- Fixation of the rail profile by an elastic poured compound;
- Design optimization of the groove dimensions, elastic compounds, and strips for specific elasticity.

The advantages of an elastic continuous supported rail are the absence of dynamic forces due to secondary bending between single rail supports, reduction of noise production, increase in life span of the rails, and further reduction of maintenance. The construction height can be reduced for road crossings, so that embedded rail provides a smooth and obstacle free surface for crossing traffic.

9.8.2 Construction of embedded rail track

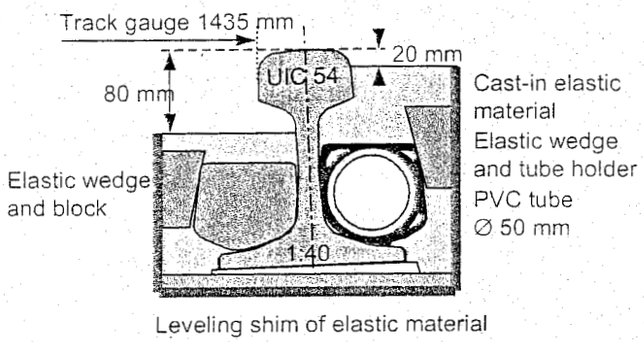


Figure 9.46: Detail of an embedded rail inside a groove

Embedded rail needs a groove to contain the poured elastic compound and the rail. Such a groove can be provided by concrete or steel. In the latter case, one could think of steel bridges or U-steel profiles cast in concrete. For paved-in track or main-line track, a concrete slab provides the support for the embedded rail. Figure 9.46 shows the cross-section of a detail of a slab containing the embedded rail.

Using a slipform paver, a continuous concrete slab is made. Because there is hardly any possibility to readjust the rail afterwards, the soil must be free of settlement.

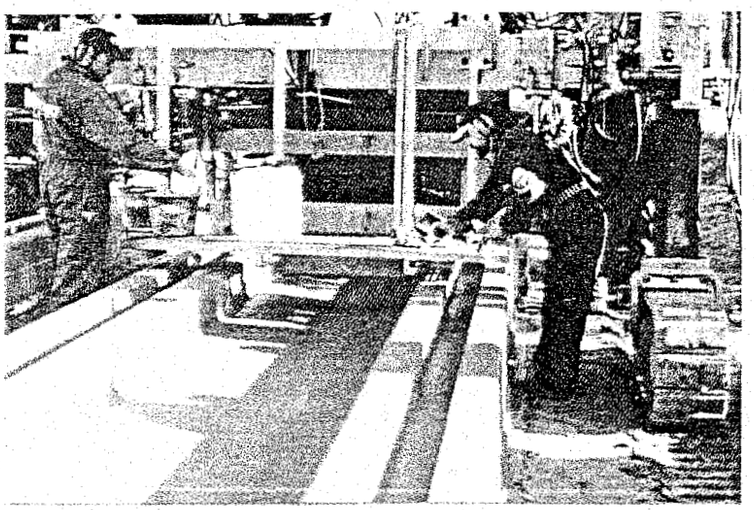


Figure 9.47: Slipform paver

The slab contains the grooves; the position of the slab determines to a very large extent the horizontal alignment, the direction, the track gauge, and the twist of the track. The production of the slab, therefore, demands high precision. Figure 9.47 shows the final stage of construction of a continuous reinforced concrete slab using a slipform paver.

At the bottom of the groove an elastic strip is laid on top of which the rail is positioned in the groove with filling pads underneath the rail foot and elastic wedges to the sides. The rail is adjusted by means of a top-down method.

After adjustment, the rail is made stress free by means of heating before the compound is poured into the groove (Figure 9.52).

9.8.3 Experiences with embedded rail

Many pilot tracks of embedded rail have been applied over the last 30 years, 246 m paved-in as well as main-line track. Nearby Deurne (The Netherlands) in 1976, a pilot was constructed in heavily used track with speeds up to 160 km/h. The track existed of a series of 6 meters of prefabricated slabs containing the gullies supported underneath with old NP46 rails [201]. The experiences were qualified as positive. In 1994, the rails were renewed, but the wear of those rails was considerably less compared to the adjacent track.

Another large pilot concerned 3 km track nearby Best (The Netherlands) which came into operation in October 1999 and is currently being monitored.

The superstructure consists of a 42 cm thick slab with longitudinal reinforcement providing the slab with a high flexural strength. The slab lies on top of a concrete road bed and stabilised subsoil.
The cross section of this track structure is shown in Figure 9.48, while the construction process of this test track is illustrated in Figure 9.49 up to and including Figure 9.54.

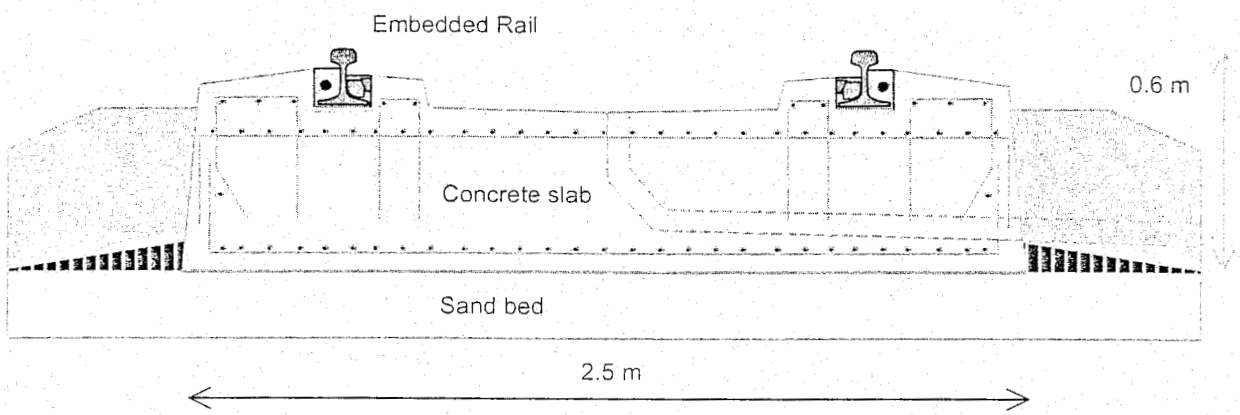


Figure 9.48: Cross section of the embedded rail superstructure near Best

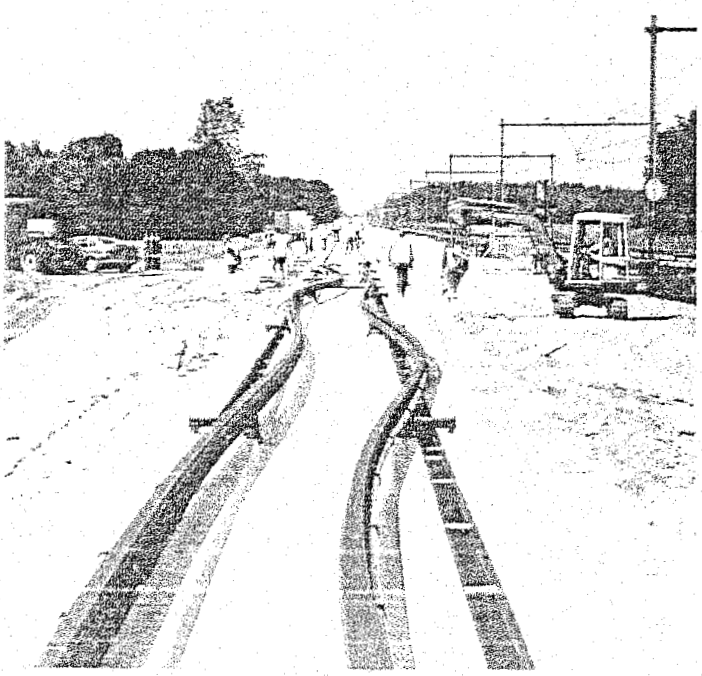


Figure 9.49: Installing of the rails

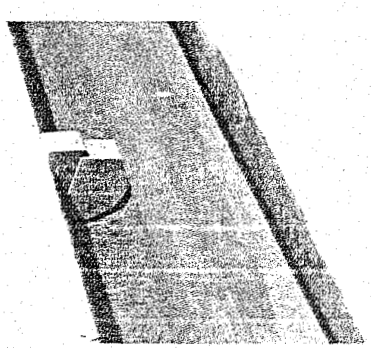


Figure 9.50: Positioning of the rails by means of wedges

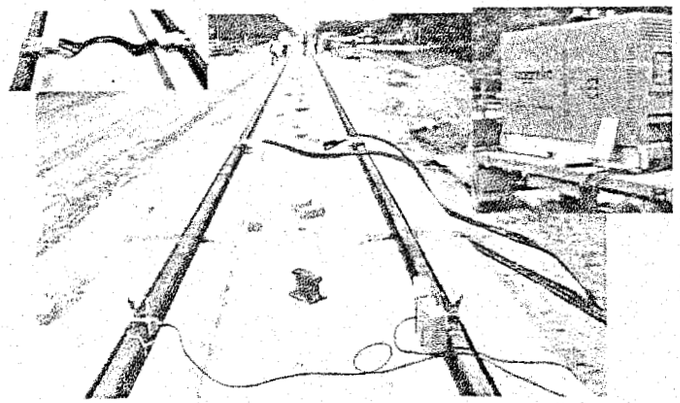


Figure 9.51: Electrical heating of the rails (17 °C)

Explanation:

As the rails are fully embedded the maximum rail temperature remains substantially lower than with a built-up structure. This, in combination with the large lateral rail resistance, was the reason to lower the neutral temperature from normally 25 °C to 17 °C.

Figure 9.52: Pouring of the compound into the groove

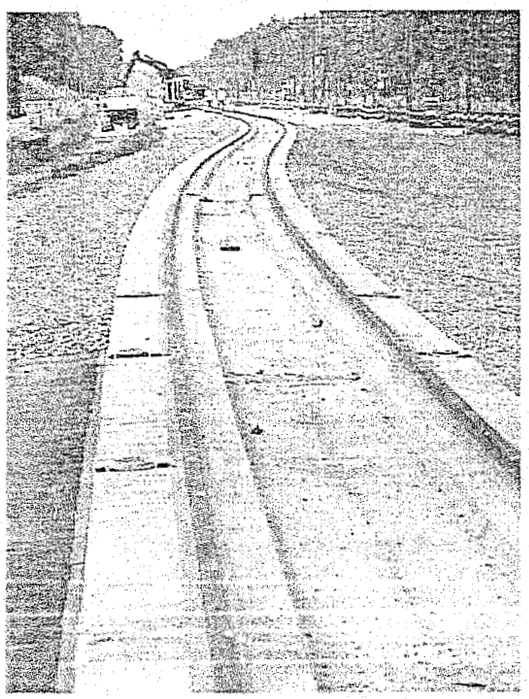


Figure 9.53: Track after completion

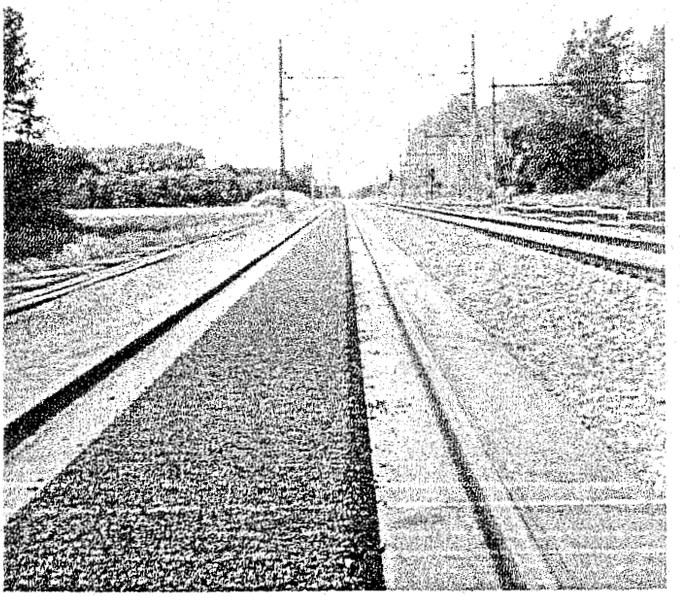


Figure 9.54: Slab covered with ZOAB asphalt for noise absorption

As part of this project an innovative rail design was also tested using the so-called low-noise Rail. As a result of an optimization process and component testing a -5 dB(A) noise reduction was achieved. Moreover, this design uses only 40 % of the elastic compound normally used with UIC54 rails to secure the rails. (refer also to Section 5.11, Section 13.4, and Section 14.2).

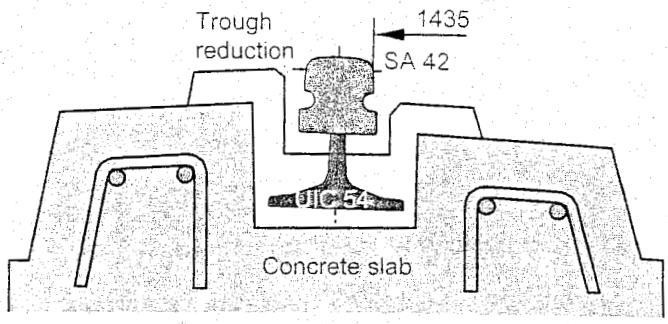


Figure 9.55: Low-noise rail SA42

In 2000, a 256 m paved-in double tramway track was constructed in The Hague (The Netherlands) on top of soft soil. The local situation posed the designers for certain problems because of plans for future earthworks adjacent to and underneath the slab, such as conveyance of pipes by pressing.

The slab was therefore reinforced to such an extent that it would be able to withstand local subsiding of the supportive layers of soil over a length of 2 meter. A slab with a high flexural strength was necessary because of uncertainties due to future building activities next to the track (Figure 9.57). Time consuming extensive soil improvements could be omitted.

Another specific application of embedded rail, in use since 1974, are the "Harmelen" level crossings (Figure 9.56). These level crossings consist of a single slab with a length of 6 or 9 meters which is lowered into a ditch partly filled with sand and adjusted in height.

Afterwards, the remaining void between the sand and the slab is injected with quickly hardening hydraulic mortar. Embedded rail thus provides an obstacle and maintenance-free level-crossing [201].

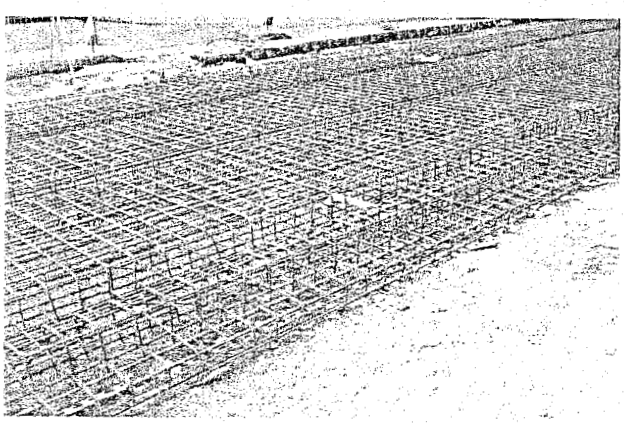


Figure 9.57: Reinforcement inside the slab for embedded rail meant for tramway application in The Hague

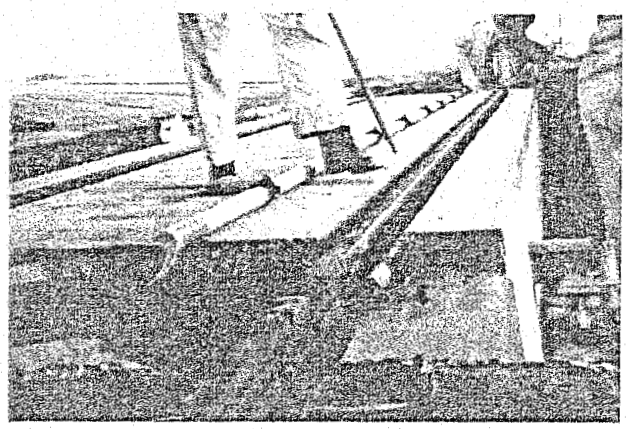


Figure 9.56: Installation of a Harmelen level crossing

9.8.4 DeckTrack

An example of an embedded rail structure with very high flexural stiffness is called "DeckTrack" [23]. DeckTrack has been specifically designed for use on soft soils and consists of a continuous in-situ or prefabricated concrete bearer laid into the ground (Figure 9.56).

On top, the rails can be directly applied as embedded rail or by means of direct fastening. The concrete deck can be considered as a hollow tube weighing more or less the same as the mass of the removed soil. No settlement will occur as a result of the structure's weight.

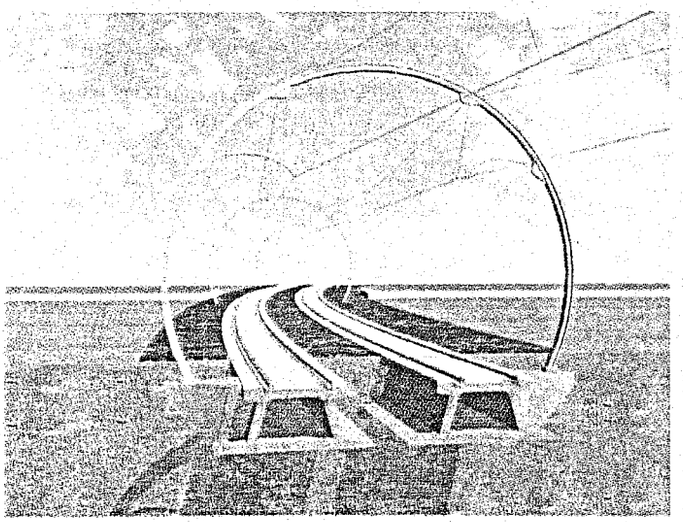


Figure 9.58: Artist impression of the Deck Track system

The high bending stiffness of the structure avoids differential settlements and reduces vibrations. High torsion stiffness provides a stable basis for the track, even on soft soils. Local subsidence of the track is precluded. Local lack of bearing strength in the subsoil will not cause problems as DeckTrack is able to act like a bridge and cover these weak spots. The high stiffness over the mass ratio causes little vibration to be transmitted into the soil and makes DeckTrack a suitable structure for high-speed traffic and for application on soft soils.

DeckTrack is perfectly suitable for the use of embedded rail. With the use of the dedicated rail profile SA 42, (Section 5.11.2), noise emission can be reduced to the level of conventional ballasted track. The structure can be cast in situ as a continuous beam, assembled from prefabricated elements linked at the joints, or as a combination of both. Both methods lend themselves to a high level of mechanisation.

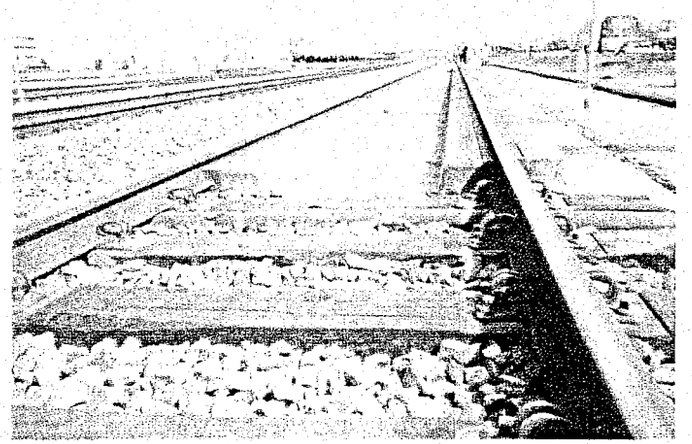


Figure 9.59: Test track in Rotterdam

Currently, a 200 m test track in Rotterdam (The Netherlands) is being monitored regarding the long-term settlement of the track (Figure 9.59). Traffic loading of the track causes uniform but limited soil settlements. While the track structure is directly connected to the concrete structure, the same positioning and deformation requirements that apply to a track will apply to the concrete structure as well. The test has shown that the track geometry remains the same quality as after construction for a lengthy period and that the track quality does not deteriorate significantly in spite of settlements of the structure.

9.9 Flexural stiff slabs on top of soft soil

Most of the currently applied slab track systems are based on the principle of a slab of relatively low flexural stiffness on top of a stiff substructure. The adaptability of slab track to differential settlements in the substructure is relatively small which is why extensive soil improvement is conducted, which is time-consuming and expensive.

In Section 9.13, an overview of the German geotechnical requirements concerning the stiffness of the substructure is given. The required stiffness on top of the frost protection layer E_{v2} = 120 N/mm^2 is equal to a foundation modulus of K = 0.15-0.20 N/mm^3. Besides soil improvement, a method to decrease the risk of differential settlements is to increase the flexural stiffness of the superstructure, i.e. by applying a concrete slab.

In general, the reinforcement in slabs is applied in the neutral line to provide a regular crack pattern due to the shrinkage stresses. Bending reinforcement increases the flexural stiffness of the slab and must be able to withstand long-term normal stresses due to shrinkage and dynamic bending stresses caused by traffic loading.

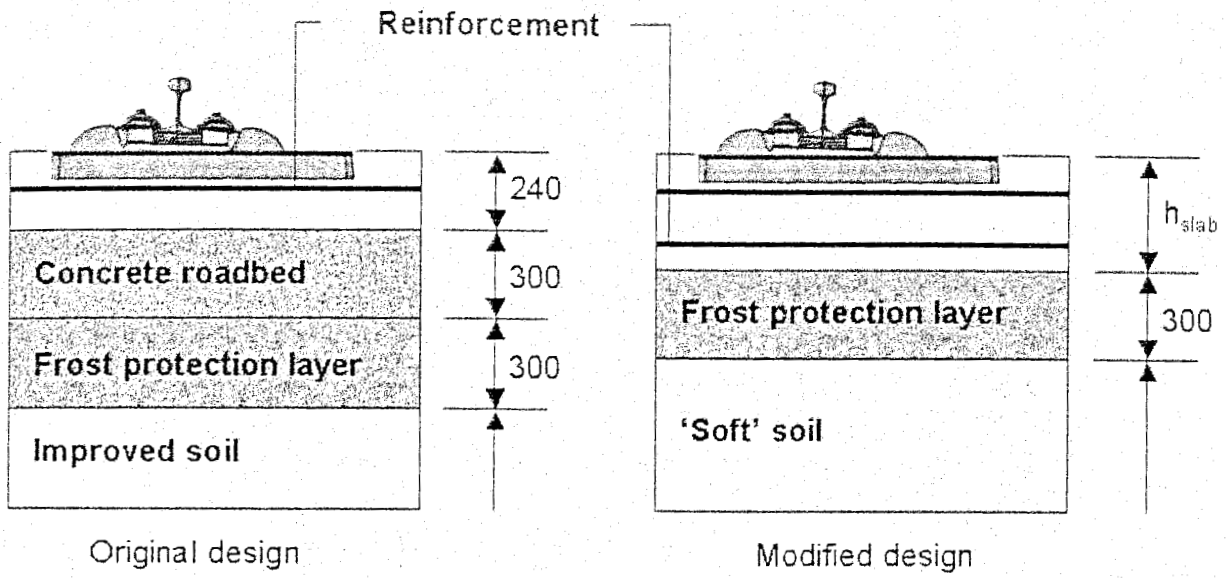


Figure 9.60: The original and the modified slab track design

Important issues concerning the application of slabs on soft soils are fatigue of reinforcement, the vertical deflection of the slab, and the level of soil stress. These factors have been considered in a research article [307]. An existing slab track system was modified and considered as if it were applied on soft soil (Figure 9.60). Calculations were made using varying slab and soil stiffness. Of paramount importance turned out to be the resistance of the slab reinforcement against fatigue.

For the reinforcement, a fatigue stress limit applies which depends on the expected number of dynamic loading cycles. The higher the number of cycles, the lower the permitted stress level is. The design stress consists of a minimum stationary stress level $\sigma_{s;d;min}$, in this case the initial stress in the reinforcement after cracking, and the fluctuating bending stress $\Delta\sigma_{s;n;d}$ caused by the dynamic loading (Figure 9.61). The sum of the stresses gives the total stress level $\sigma_{s;d;max}$, which must remain below the fatigue limit $\sigma_{s;u}(n)$.

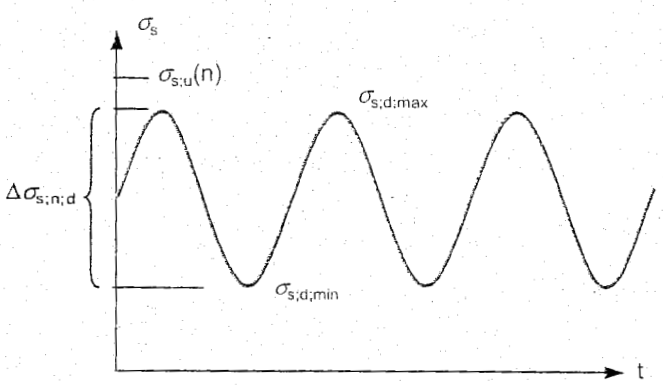


Figure 9.61: Typical stress levels in reinforcement steel under dynamic loading

For an arbitrary slab of B35 concrete, the initial steel stress after cracking ($\sigma_{s;d;min}$) and the fatigue limit for 2 million dynamic loading cycles ($\sigma_{s;u}(n)$) are given in Figure 9.62 plotted against the percentage of eccentric reinforcement. The slab is designed to withstand additional stresses caused by temperature loading and shrinkage.

The shaded part marks the difference between the initial steel stress (inside cracks) and the fatigue limit for steel; it marks the amount of stress available for bending. The figure clearly shows that 1.0 % of reinforcement is not sufficient.

Slabs with a high bending stiffness are able to bear bending moments and to spread traffic loading over a larger track length, thus decreasing the stresses on the subsoil. The slab is able to bridge weaker spots or local subsidences without alteration of the track geometry.

Besides fatigue of steel, an important constraint is the maximum permissible stress on the subsoil which is determined using the equation of "Heukelom and Klomp" (Section 5.8). The empirical rela-

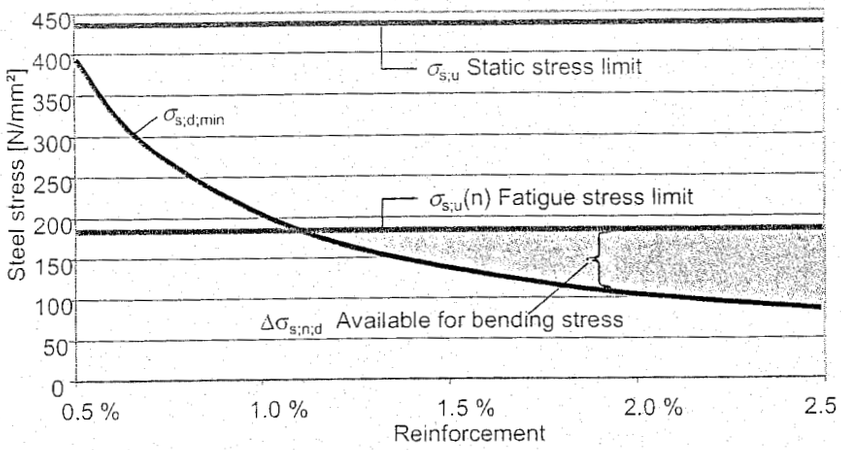


Figure 9.62: The available stress for bending the slab

tion of Heukelom and Klomp gives an indication of the maximum allowed soil stress depending on the dynamic elasticity modulus of the soil E_{dyn} and the number of cyclic load changes n_i:

$$\sigma_{soil;max} = \frac{0.006E_{dyn}}{1 + 0.7log(n_i)} \tag{9.1}$$

Using the maximum allowed soil pressure, the maximum allowed vertical deflection u_{max} of the slab is determined according to the linear relation:

$$u_{max} = \frac{\sigma_{soil;max}}{K} \tag{9.2}$$

In Figure 9.63 the maximum vertical deflections of five slab designs of varying height are shown relative to the foundation modulus K. The loading case consists of two static axle loads of 225 kN at a distance of 2.5 metres. Deflections are limited by the relation between E_{dyn} and K. Above certain deflection levels, permanent soil deformation will occur as shown in Table 9.4.

Figure 9.63 shows that any of the 5 slabs will meet the requirements regarding the maximum deflection if:

- The dynamic elasticity modulus E_{dyn} of the bearing layer is at least $E_{dyn} > 25$ N/mm^2 and
- The foundation modulus of the entire substructure is at least $K > 0.01$ N/mm^3.

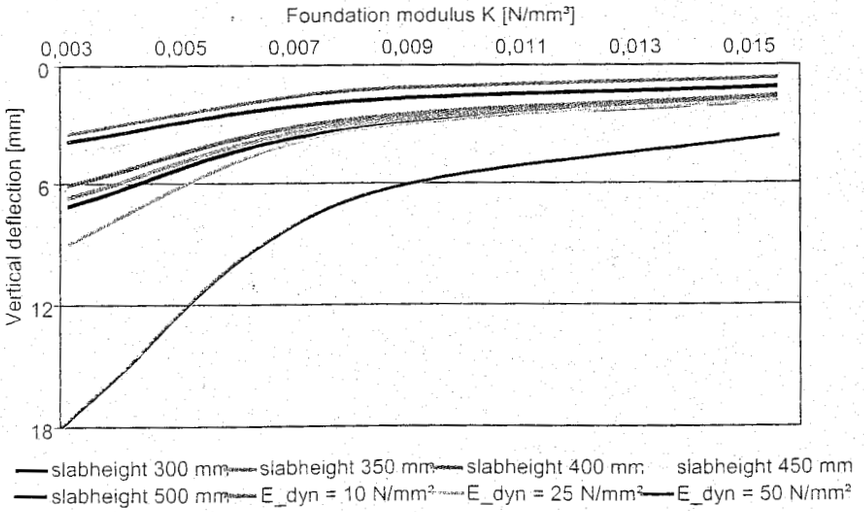


Figure 9.63: The allowed vertical deflection in relation to the foundation modulus K

Type of soil	Cone resistance q_c	Foundation modulus $K^{1)}$	Elasticity modulus E_{dyn}	CBR-value$^{2)}$	a	b
	N/mm^2	N/mm^3	N/mm^2	%		
Peat	0.1-0.3	0.01-0.02	10-35	1-2	0.5	10.0
Clay	0.2-2.5	0.02-0.04	15-60	3-8	1-2	7.5
Loam	1.0-3.0	0.03-0.06	50-100	5-10		5.0
Sand	3.0-25.0	0.04-0.10	70-200	8-18	4-7	2.5
Gravel/sand	10.0-30.0	0.08-0.13	120-300	15-40		1.5

1) In case of equally distributed loading the value of K is at least a factor 3 smaller (K/3)
2) CBR = Californian Bearing Ratio

Table 9.4: Indicative values of different soil characteristics

Research has shown [308] that it is possible to apply continuous concrete slabs with bending reinforcement on soft soil. The regular amount of 1 % of reinforcement is not sufficient; the necessary percentage will amount to up to 1.5 % spread over the top and bottom layer of reinforcement. A B35 concrete slab with a height of 350 mm can be applied to soft soil with a foundation modulus of $K \geq 0.01$ N/mm^3 and at least 1.35 % of reinforcement by slightly modifying of the traditional concrete slab track design.

9.10 Clamped and continuously supported rail structures

This section highlights some examples of non-conventional slab track structures. The advantages of continuously supported rail have already been mentioned in Section 9.8. Embedded rail provides a continuously supported rail: a groove is needed to contain the elastic compound when it is poured. Besides embedded rail there are other alternatives that can provide a continuous supported rail without a groove. Attention is paid to two alternatives for paved-in light rail track: Cocontrack and Continuously supported grooved rail.

Besides these two structures, two examples of rail fastening systems are given which contain rails that are clamped in their webs. A large benefit of these systems is their very high resilience. The systems are in general applied in slab track and light rail.

9.10.1 CoconTrack

The CDM-CoconTrack system shows a combination of a continuously supported rail with traditional elastic rail fasteners and concrete elements (Figure 9.64). The system consists of several specific elements. There is the Cocon sleeper: a concrete sleeper system with a combined transverse and longitudinal sleeper function which is H-shaped (Figure 9.65). This sleeper has an integrated continuous rail support surface and an integrated system for X, Y, Z-alignment of the track. The fasteners are placed at the crossing of the transverse and longitudinal sleepers. The distance between the fasteners is 1200 mm, and the concrete elements are designed for curves R > 20m.

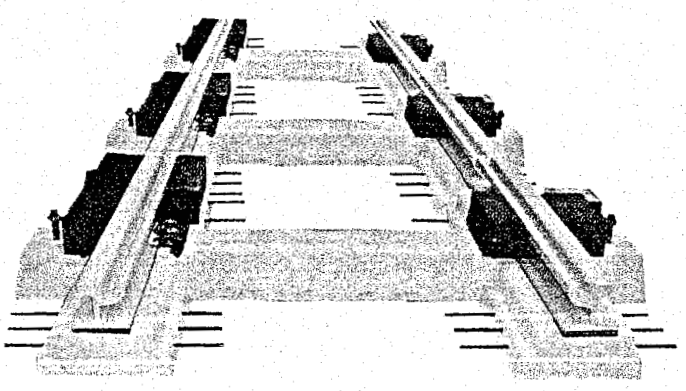


Figure 9.64: Impression of the CoconTrack system

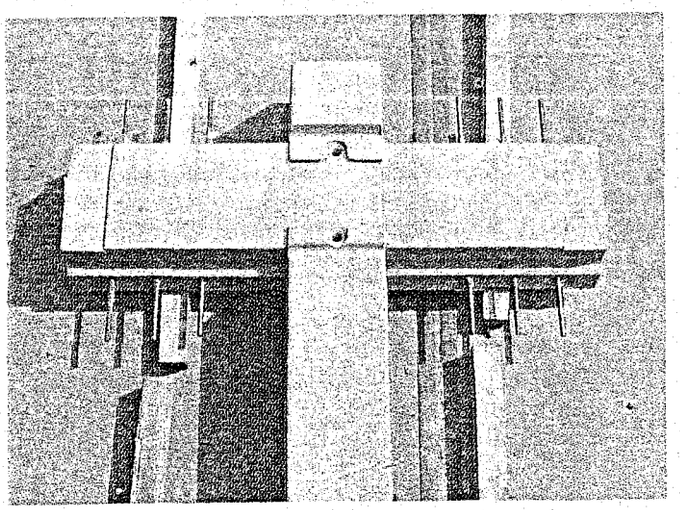


Figure 9.65: Detail of the H-shaped sleeper

On top of the longitudinal sleeper the CDM-Bistrip is applied. This strip consists of two different layers, each with a different resilience. The first layer is the adaptation strip (AS): a soft layer which will be compressed to approximately 40% of its original thickness when the rail is fastened. The compression has to be sufficient for the automatic adaptation to the rail height tolerances (2-3 mm) and possibly future height adjustments (3-4 mm). Practically, a minimum thickness of 10 mm is required. The second layer is the load bearing strip (LB): this layer is stiffer than the adaptation strip. The stiffness characteristics of this layer depend on the requirements of the maximum rail deflection; this layer is at least five times as stiff as the AS.

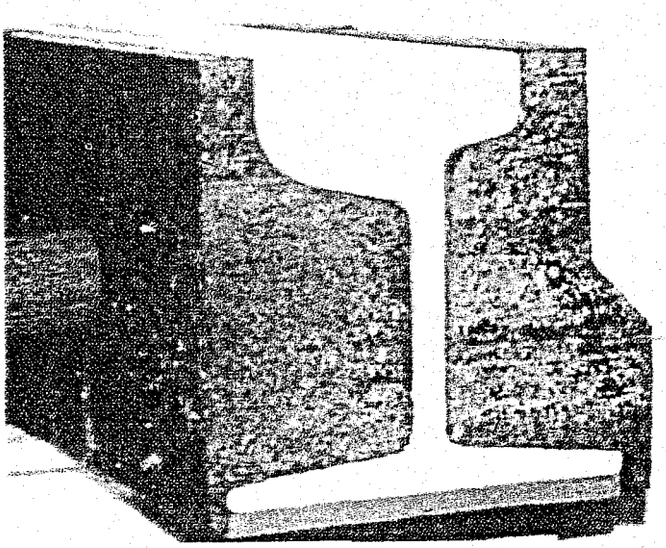


Figure 9.66: Detail of the grooved rail, the CDM-Bistrip, and the rail web chambers

In combination with the CDM-Bistrip, a sound stop of cork/rubber elastomere is applied between the fastener springs and the rails to eliminate direct contact. And rail web chambers (CDM-Flexiweb) of high quality recycled rubber, shaped according to the type of rail, are glued to the rail to avoid contact between the rail and surrounding material like asphalt or concrete (Figure 9.66). Inside the web chamber elements spaces are left open to integrate all types of rail fasteners.

This type of ballastless superstructure has been specifically designed for paved-in tramways and the main goal has been to decrease the level of airborne noise production. Depending on the stiffness of the foundation, the CoconTrack system can also contribute to high insertion losses of track vibrations.

CoconTrack is constructed similarly to ordinary paved-in track, in lengths of 18 meters. The H-shaped sleeper elements with integrated spindles are laid head to head. These elements are constructed in such a way that a continuous and flat surface is created on top of which the bi-strip is glued. Afterwards, the rail is placed on top of the bi-strip and fixed with the rail fasteners. The track panels are then lifted and placed over a thin concrete roadbed providing a dry and level support for further construction. The concrete elements are adjusted approx. 10 cm above the concrete roadbed with spindles (top-down method). Underneath, the elements reinforcement is applied: the amount depends on the stiffness of the soil. Afterwards, the rail web chambers are placed and the concrete sleeper elements with the steel rods extending sideways are partially cast into concrete. The shape of the sleepers is thus that an optimal connection with the in-situ concrete is achieved.

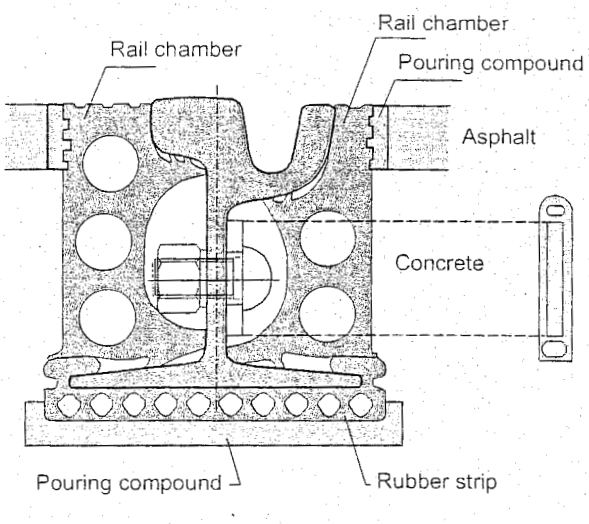


Figure 9.67: Continuously supported rail by Phoenix

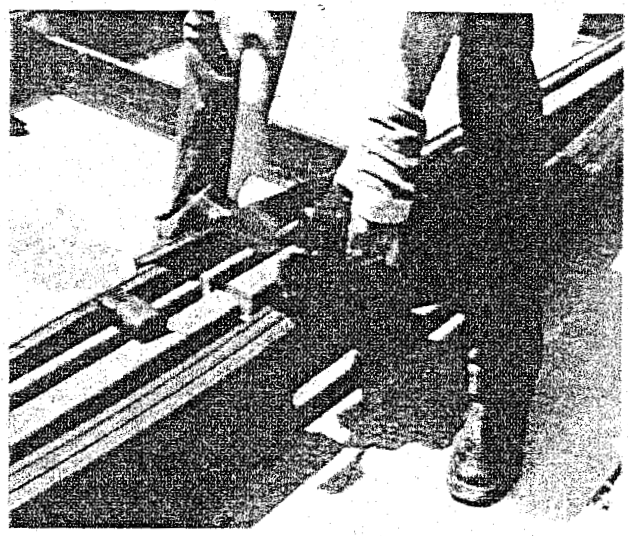


Figure 9.68: Fitting of the rubber chamber elements

9.10.2 Continuously supported grooved rail

The continuously supported grooved rail system (called ERL) developed by Phoenix consists of continuously and elastically supported rail with or without rail fasteners. Underneath the foot of the grooved rail a rubber strip with air chambers provides the elastic support. Both rails are adjusted on top of a high-grade concrete roadbed and mutually connected with steel bars to guarantee track gauge (Figure 9.67).

In the side chambers of the grooved rail, rubber chamber profiles are inserted (Figure 9.68). The purpose of those rubber chamber profiles is to reduce airborne and structure-borne noise production. Moreover, they act as a transition between the rail and the pavement on the road.

Beforehand, a concrete roadbed is constructed to act as a level and rigid bearing layer. On top of this roadbed the entire track framework is assembled with the elastic footing and the rail chambers. The framework is adjusted (top-down method) with wooden wedges.

The void between the slab and the elastic footing is filled with a filling-compound. Afterwards, the pavement is applied (Figure 9.69).

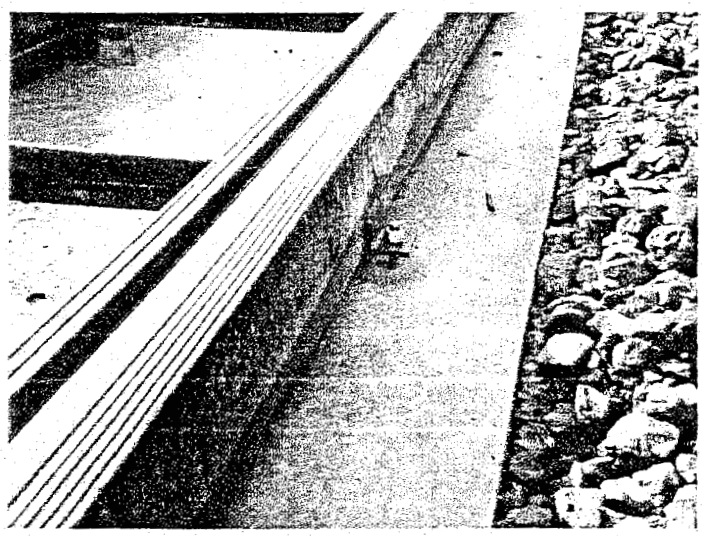


Figure 9.69: The assembled track framework is ready to be paved-in

9.10.3 Web-clamped rails

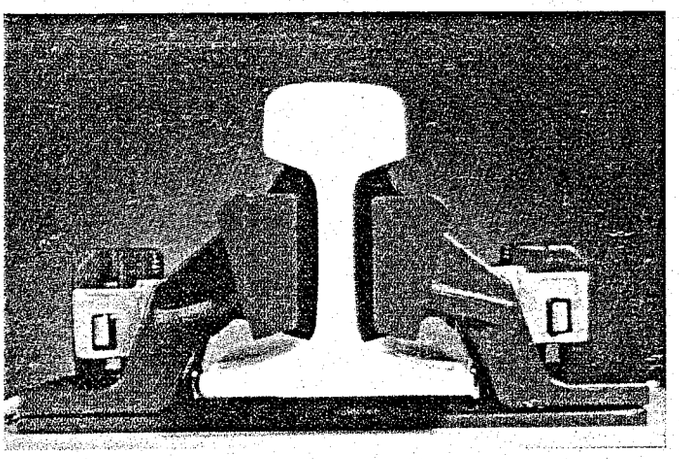


Figure 9.70: Impression of the Vanguard system

The Vanguard system developed by Pandrol and the continuous elastic support system (called KES) by Phoenix are new proprietary types of rail fastening systems in which the rail is supported by elastic wedges or strips under its head, leaving the rail base unsupported, as demonstrated in Figure 9.70.
The elastic wedges are held in place by side brackets which are fastened to the track foundation. The principal feature of this system compared to conventional rail fastenings is that it allows significantly greater vertical deflections under traffic without an unacceptable accompanying degree of rail roll.

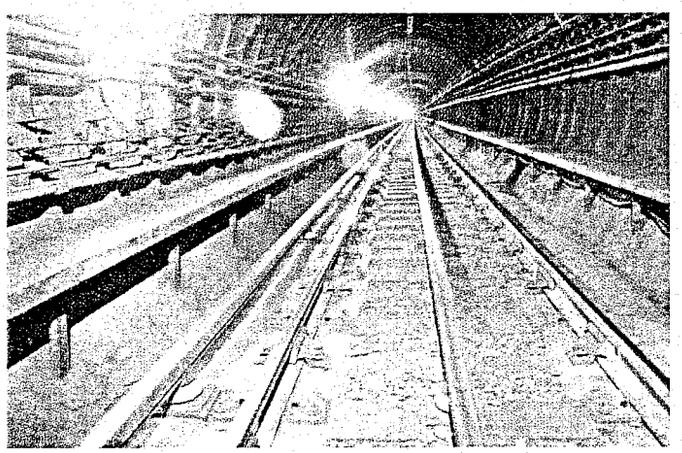


Figure 9.71: Vanguard system retro-fitted on concrete slab track

The low vertical stiffness gives improved attenuation of the dynamic forces generated at the wheel/rail interface, reducing the level of dynamic forces transmitted through the fastening and into the supporting structure.
First developed for use in deep tube tunnels, the Vanguard concept, as shown on the Lon don Underground, achieves a significant reduction in low frequency vibrations which would otherwise propagate into buildings along the track. It also has the advantage of a low overall assembly height - there is only 14 mm between the rail base and the surface to which it is attached.

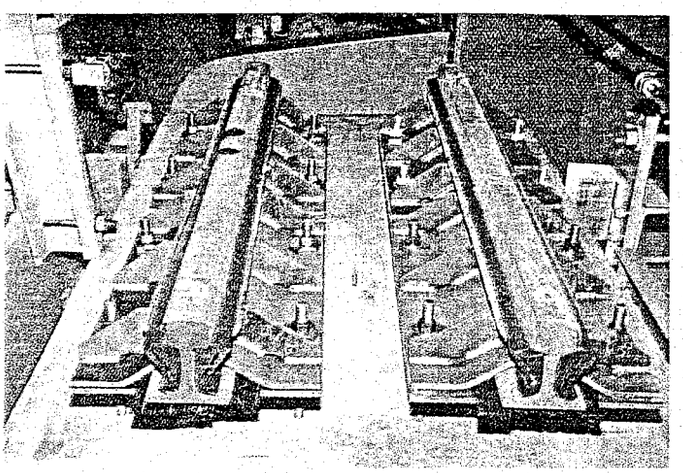


Figure 9.72: KES system during laboratory testing

A base plate version exists which can be installed on existing wooden sleepers or blocks or can be retro-fitted to concrete slab track as demonstrated in Figure 9.71. The static stiffness of the system is in the range 4-7 kN/mm per individual support.

Starting in the 1990's, applications of KES in light rail have taken place in several German cities. In contrast to the Vanguard system, the elastomere strips are present all along the rails. In principle, the systems looks similar to the ERL system described in Section 9.10.2, but steel or concrete profiles are installed by means of tension bolts in order to continuously secure and pre-stress the elastomere strips. Stability, durability, and elasticity tests in Figure 9.72 have shown satisfying results [227], so that a wider application is investigated such as installation on bridges and in main lines. Static stiffness of present light rail systems ranges between 2.9 and 4.0 kN/mm per meter rail.

Compared to the Vanguard system, KES requires more material and very accurate installation. The retro-fit option in the Vanguard system allows easier installation at points where nuisance occurs after a certain period of service. KES offers a suitable solution in combination with paved-in tracks.

9.11 EPS as subbase material in railway slab track structures

9.11.1 Introduction

The construction of relatively heavy slab track structures on soft soils requires improvement of the bearing capacity. The conventional approach consists of replacing part of the poor soil by sand as well as soil improvement. Even if pre-loading is applied, relatively large settlements are likely to occur during the initial phase of the structure's life. With the application of ultra-light materials, such as Expanded PolyStyrene (EPS), a so-called equilibrium structure can be created (see also DeckTrack, Section 9.8) which will prevent the increase of grain stresses in the subgrade and reduce the chance of differential settlements. Globally speaking, the weight of the track structure plus the light-weight material should balance the weight of the excavated material. Due to the small value of Young's modulus of EPS, a concrete slab will be required in railway structures with an EPS subbase to obtain sufficient stiffness and strength.

9.11.2 Slab track structures with an EPS subbase

In light of the positive experiences with heavy-duty light-weight pavement structures, the TU Delft has done research into the possibilities and conditions for the application of an EPS subbase in both ballasted and slab track [200]. The density of EPS is directly related to its Young's modulus and other material characteristics. The EPS types examined in this study were EPS20 and EPS35. The mechanical properties of EPS were taken from [54].

Compared to traditional subbase materials, EPS has very low values for density, for Young's modulus, for water absorption capacity, and for thermal conductivity. As EPS has a relatively low strength, a concrete slab on top of the EPS layer was inevitable. In fact, this made an integrated slab track solution very attractive. For this reason the slab track was comprised of a concrete slab with an embedded rail structure as presented in Figure 9.73.

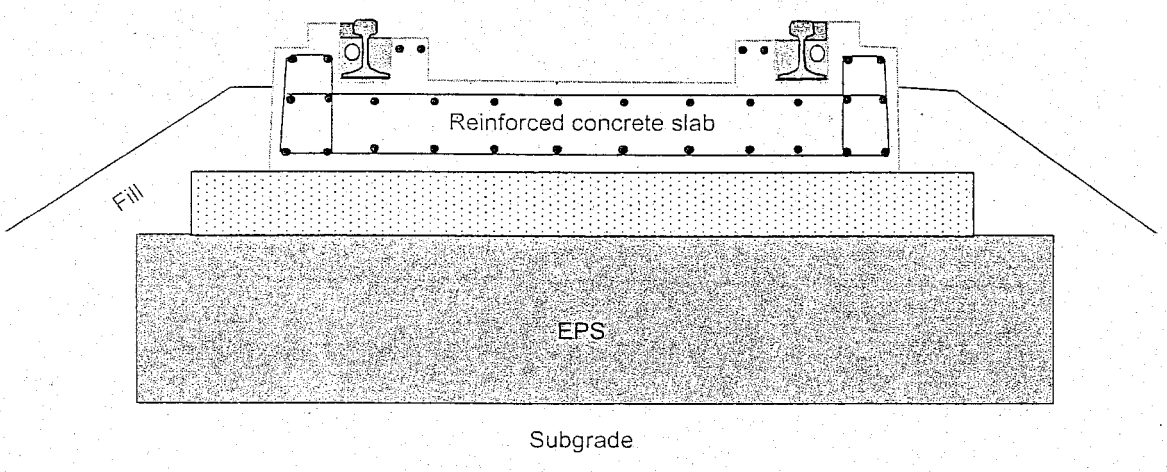


Figure 9.73: Embedded rail structure with an EPS subbase

9.11.3 Static performance

The minimum thickness of the concrete slab on top of the EPS layer was determined by a static analysis carried out with GEOTRACK. On the basis of the maximum permissible rail deflection and the maximum acceptable stresses and strains in the different layers of the structure, the minimum thickness of the concrete slab was found to be 25 cm. The stress in the EPS subbase remained within the linear-elastic range.

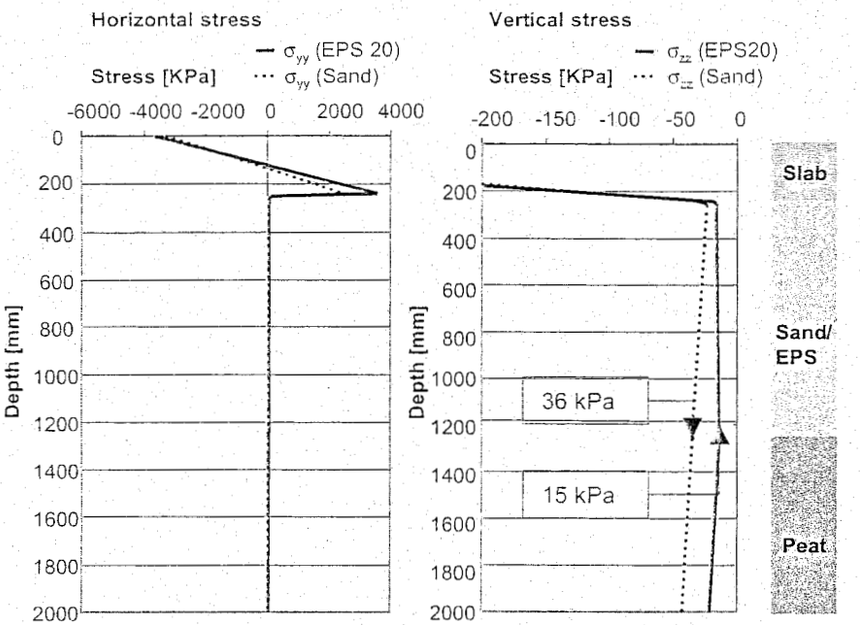


Figure 9.74: Stress distribution in an embedded rail structure under a static load of 112.5 kN

Figure 9.74 shows that in the case of an EPS subbase the vertical stress at the top of the subgrade is maintained at the original value of 15 kPa, while in the case of a sand subbase the stress is more than doubled.

This clearly reveals the potential of such lightweight solutions. Three years of experience with an EPS subbase in a tram track in Rotterdam confirmed the limited settlement behaviour.

Another important result, shown in Figure 9.74, was that the difference in stress between EPS20 and EPS35 is negligible under all practical conditions. This means that EPS20 suffices.

9.11.4 Dynamic performance

The influence of an EPS subbase on the track dynamic behaviour was investigated using the RAIL program of the TU Delft [253]. In this program real track geometry which was obtained by downloading track recording car data from the NS can be modelled. Model calculations have produced response values for car body accelerations, wheel-rail forces, track deflections, and stresses in the track structure caused by moving train sets. The damping coefficient of EPS was not determined experimentally but was estimated using literature on the damping ratio.

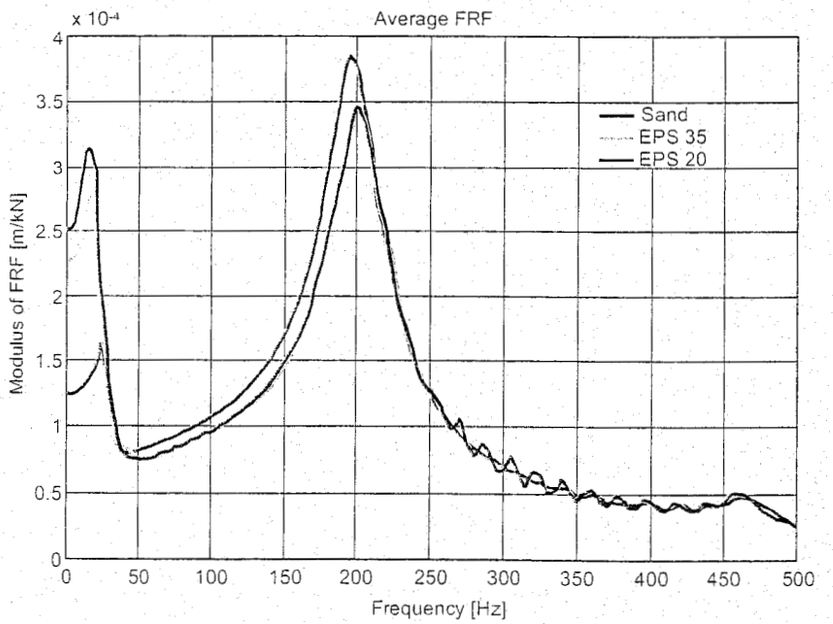


Figure 9.75: FRF of an embedded rail track with three different subbases, x = 0.25 m

The overall dynamic behaviour of the track structure, expressed in the so-called Frequency Response Functions (FRF), was determined by applying a pulse load. As an example Figure 9.75 shows the FRF functions for the Embedded Rail Structure with the concrete slab resting on a base of sand, EPS 35, and EPS 20 respectively.

Obviously the static deflection of the EPS subbase was greater than that of the sand subbase as EPS is much weaker than sand. Regarding the EPS subbase, the first natural frequency was lower than that of the sand subbase. The second natural frequency at 200 Hz was the same for all three subbases, and is mainly associated with the polyurethane filling material of the embedded rail. The two natural frequencies for the three different subbases (EPS20, EPS35, and sand) are shown in Table 9.5.

subbase	1st natural frequency [Hz]	2nd natural frequency [Hz]
EPS20	16	200
EPS35	18	200
Sand	27	200

Table 9.5: Natural frequencies of the three subbases

The results mean that frequencies above 20 Hz may be more easily filtered out with an EPS subbase than with a sand subbase. However, the first resonance peak is somewhat higher for an EPS subbase because of the higher static response. It should be noted that the damping coefficient of the EPS material was determined theoretically and may differ from an experimentally determined damping coefficient. This might have a major influence on the height of the first peak of the FRF which seems to correspond with the vibration damping characteristics of EPS referred to in [242]. In the scarce literature on the dynamics of EPS, it is pointed out that an EPS subbase might show some vibration damping in the frequency range between 20 Hz and 40 Hz.

Another series of analyses consisted of responses caused by passing vehicles. The rail geometry was set at the maintenance intervention levels of Netherlands Railways (NS) which allows a maximum vertical deviation of 3 mm for construction and 12 mm for maintenance. The analyses showed that, under these loads, stresses and strains in the EPS subbase as well as deflections of the rail remained within the admissible values. In the dynamic analysis, the difference in results for the two different EPS densities was also small. Based on the frequency response function, EPS20 was indeed slightly better which again indicates that this material is an acceptable foundation material for track structures.

9.11.5 Applications

EPS can be applied in any track structure, but significant advantages will be achieved when used on subsoil with a poor bearing capacity. In two special cases - for transitions between civil structures and plain track and for constructing a track doubling - the advantages of EPS may be even more apparent as they allow differential settlements to be avoided [42].

In [5] the different track structures were also compared with respect to life-cycle costs. In the case of very compressible subsoils, an EPS subbase was found to be amongst the cheapest solutions as maintenance costs would be reduced significantly. This subbase type would certainly be better for the environment, both during construction and during service.

9.12 Track resilience

On conventional track, the ballast bed provides approximately half the resilience needed to absorb dynamic forces; the other half is provided by the subgrade. The stiffness of the overall track structure can be of the order of 100 kN/mm per sleeper which makes the rails deflect approximately 1 mm under a 20-t axle load.A rail pad inserted between the rail and the sleeper filters out high frequency vibrations. The elastic rail pad and, if present, the baseplate pad replace the ballast bed regarding its load-distribution and the damping functions. Therefore, the importance of the elastic pads is paramount for they become the only components in the track with elastic and damping properties.

In the case of ballasted track, the response of the rail in the low-frequency domain is heavily dependent on the elasticity of the ballast and of the substructure. The response of the rail on slab track is in a large part of the frequency domain almost entirely dependent on the material properties of the railpad and the baseplate pad. A slab track in the low- to middle-frequency domain can therefore considerably vary from the traditional ballasted track [153].

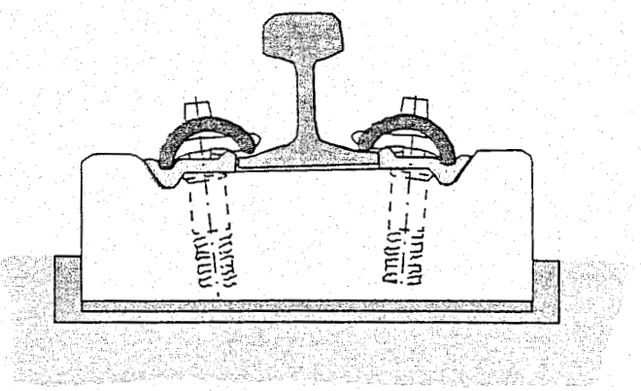


Figure 9.76: Track consisting of blocks

For ballasted track, the stiffness of the embankment is of major importance. In the case of slab track the stiffness of the embankment only matters in the low-frequency (static) domain. In the case of slab track and bridges on which the rails are directly fixed, additional resilience must be added to the system to compensate for the absence of ballast. In principle there are two ways of achieving this:

- Add extra resilience under the rail by, for instance, inserting extra thick rail pads or by using ERS;
- Insert a second resilient layer under the supporting blocks or sleepers.

In the second case, a two-mass spring system is effective, with a primary and secondary spring, analogous to a vehicle. Figure 9.76 shows an example of a block system.

9.13 System requirements

System requirements for the permanent way vary per country. For slab track there are specific requirements concerning the superstructure: the application in tunnels, embankments (the substructure), bridges, and transitions as well as specific requirements concerning signalling and electronic systems, safety, noise production, and dynamic effects. This depends as well on the applied system and materials and the type of traffic.

In case of high-speed traffic, dynamic forces will increase considerably due to aberrations in track geometry and alignment. In order to avoid progressive deterioration, to improve passenger comfort, and to decrease the level of maintenance, slab track systems are judged based on requirements and features like [14], [190], [197], [248]:

- The use of standardised construction elements;
- Rail fastening systems which are adjustable in height and sideways as well;
- Repair and construction friendly;
- Possibility to install noise absorbers;
- Easy access for rescue services;
- Labour-saving construction work;
- Maximum availability.

9.13.1 Requirements for the substructure

Concerning slab track the requirements for the substructure are in general very high. The ability to make adjustments to the track geometry after construction is finished is relatively limited, therefore, any settlement must try to be avoided.

Looking at applied substructures, for instance in Germany, a lot of effort is being made to obtain a stable embankment. The regular composition of layers consists of improved ground (through compacting or hydraulic stabilising) followed by a frost-protection layer of granular material.

On top of that a concrete roadbed or a layer of asphalt-concrete is applied, depending on the system, before the actual slab track is constructed. In advance of the preparations for the embankment, extensive geotechnical research is necessary into the ground conditions and the deformation behaviour. For the geotechnical assessment, at least every 50 m ground-probes are required ([14], [42]).

The criteria for the E_{v2} modulus (Section 5.8.4) at the top-surface of the substructure are:

- Newly constructed track: $E_{v2} \geq 60$ N/m^2;
- Existing track: $E_{v2} \geq 45$ N/m^2.

This is achieved by dynamic compacting and by mix-in-place ground-improvement, for instance with chalk and cement or by ground-replacement. The geotechnical requirements for the embankment are satisfied for a depth below the rail head level when the Proctor-densities are:

- Newly constructed track ≥ 3.0 m with $D_{pr} = 0.98 - 1.00$;
- Existing track ≥ 3.0 m with $D_{pr} = 0.95 - 1.00$.

	Requirements regarding quality		
Bearing layer	Newly constructed track	Existing track	Layer thickness
Reinforced Concrete roadbed	Concrete quality: B 35 Reinforcement percentage[1]: 0.8 – 0.9 % of the concrete cross-surface		depends on the measurements of the E_{v2} value (approx. 200 mm)
Asphalt roadbed[2]	Binder B 80 or B 65, top layer PmB 65		depends on the measurements of the E_{v2} value (approx. 300 mm)
Concrete roadbed[3]	The necessity must show in the design calculations for the superstructure		if necessary (approx. 300 mm)
Frost protection layer	$E_{v2} = 120$ MN/m^2	$E_{v2} = 100$ MN/m^2	necessary geo-technical coverage below level of rail foot: - newly constructed track: ≥ 3.00 m - existing track: ≥ 2.50 m
Embankment	$E_{v2} = 60$ MN/m^2	$E_{v2} = 45$ MN/m^2	

1) With bridges according to static design.
2) Construction in multiple layers.
3) A bearing double frost protection layer with the top layer of ballast material can replace the lower concrete roadbed.

Table 9.6: Overview of the major requirements for the bearing layers of the substructure

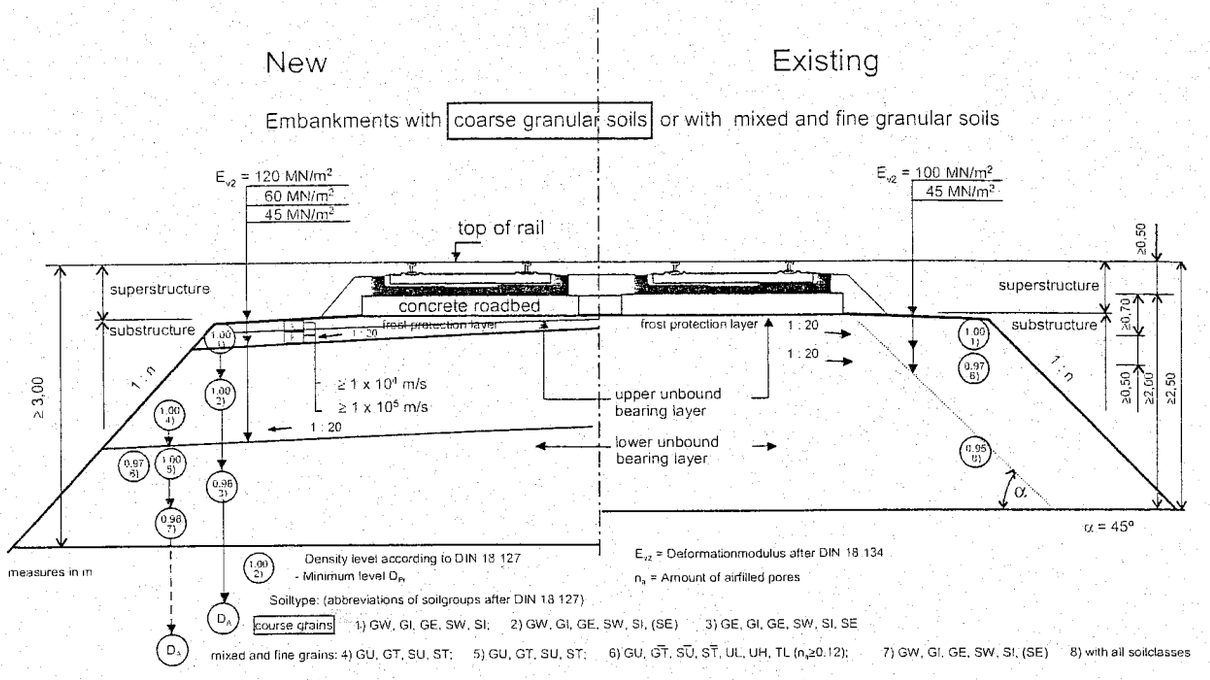


Figure 9.77: Requirements for the unbound bearing layers

The German requirements concerning substructures of existing tracks and newly constructed tracks are collected in Table 9.6 [42] (Figure 9.77).

In Japan as well as in Italy, extensive efforts with improved ground and a reinforced concrete roadbed contribute to the high construction costs of slab track systems.

Especially in areas where the soil has low bearing capacity, such as in The Netherlands and in parts of Japan, the subgrade often consists of weak soils with critical speeds far below the intended operational speed (Section 6.10). In such cases measures to increase the vertical stiffness are unavoidable. Possible solutions include soil improvement, deep mixing, grouting, and piling.

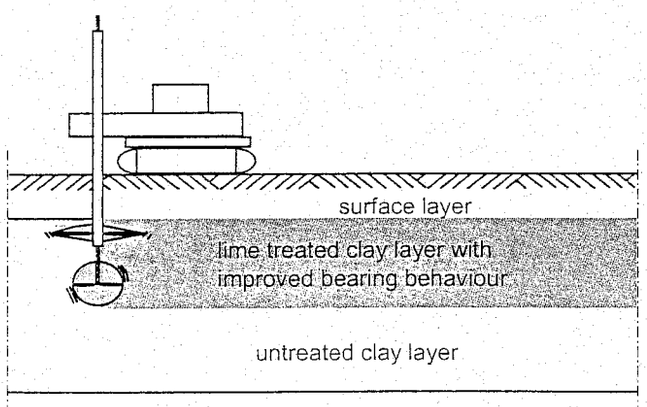


Figure 9.78 illustrates the principal of lime treatment for large-scale stabilising of the subgrade. Stabilising using grouting or deep mixing can be applied to the soil very precisely. Grouting involves injecting liquid mortar (based on cement or limestone) which then hardens, while deep mixing consists of mixing the material already present with a liquid or dry mortar based on cement, limestone, fly-ash, etc.

Figure 9.78: Principle of soil improvement by lime treatment

In Japan, considerable experience has been acquired applying this technique, especially where houses and business premises have to be made earthquake-proof. In the case of ground improvements, it is possible to achieve a considerable increase in supporting capacity and a reduction in the subsidence of various thick layers incapable of supporting loads.

9.13.2 Requirements for slab track in tunnels

Slab track is, in general, very suitable for application in tunnels. In certain tunnels in geological unstable formations (where there is the chance of subsidence, earth slides, the appearance of wells) slab track could, however, be unsuitable. When slab track is applied in tunnels, specific requirements have to be met: drainage should be guaranteed and in case of calamities and safety the slab must be accessible for rescue-vehicles [42], [190].

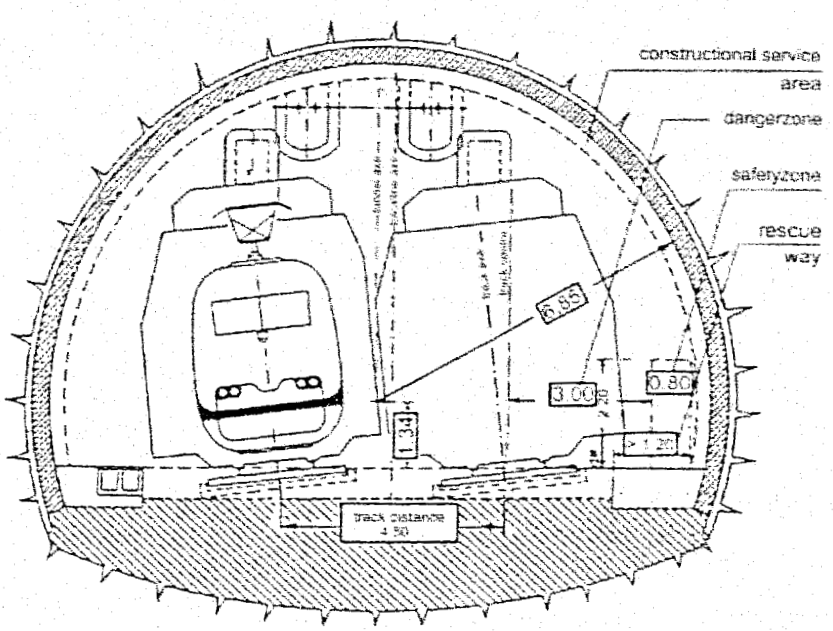


Figure 9.79: A tunnel cross-section with the required dimensions of free space

In tunnels the total construction-height is restricted often by omitting or decreasing the thickness of the concrete roadbed beneath the trough or asphalt-concrete layer. The available free space in the tunnel cross-section may influence the decision which type of slab track to apply (Figure 9.79).

9.13.3 Requirements for slab track on bridges

Continuation of slab track across bridges could pose problems if certain typical mechanical behaviour is not considered. A bridge provides a solid foundation for slab track, but acts in fact as a discontinuity. Due to temperature changes longitudinal movements of the bridge-structure will occur. Traffic loading causes the spans of the bridge to bend and the edges to twist over the supports. The superstructure must be able to cope with these movements [42], [190].

By applying slab track systems on short bridges, several solutions are possible:

- Fasteners with reduced clamping force: the movements of the bridge are compensated in the rail fastenings with reduced clamping force if the sleepers on top of the reinforced concrete roadbed are rigidly connected to the bridge deck or direct rail fastening systems are used.

- Embedded in bridge decks: in case of a continuous rail-support rigidly connected to the bridge, maximum active extendable bridge-spans up to 15 m are permitted. Larger spans are possible by applying extension devices and joints.

- Sliding slabs: the bridge structure can freely move underneath the slab track which "glides" on top. This option is limited to freely extendable bridge-spans up to 25 m.

- Track frame on roadbed: the track lies freely movable on top of a concrete or asphalt-concrete roadbed. This solution exists due to possible motions and twisting of the sleepers on top of bridge-structures and spans up to 10 m with framespans limited to 25 m.

9.13.4 Requirements for transitions

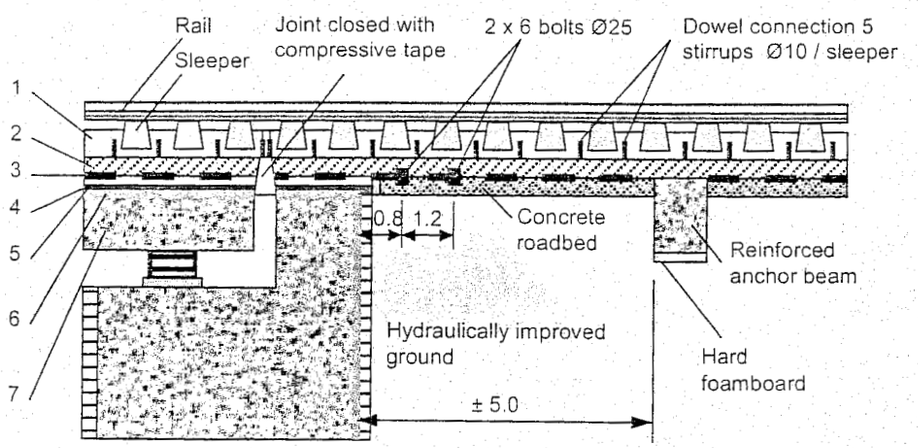
Transitions in the substructure and superstructure occur, but also between different types of each sort. In the substructure transitions between embankments, bridges, and tunnels occur. In the superstructure there are transitions between slab track and ballasted track. Transitions demand special measures to smooth out the difference in the elastic properties of both track structures.

With transitions from one type of track to another, forces caused by traffic loading and temperature must be transmitted smoothly. Problems can occur with differences in construction height: it could even be necessary to use a joint in between and apply both ends with a special anchoring. In the case of transitions between a track structure with sleepers applied on top of an asphalt-concrete roadbed and a slab track system with direct rail fastening or embedded sleepers, the sleepers must be anchored to the substructure/roadbed over a distance of at least 10 meters.

Regarding transitions from slab track to ballasted track, two structures must be connected which are principally different regarding stiffness, flexibility, and settlement behaviour. This adaptation can be achieved in several ways:

- Stiffness: a gradual decrease of the elasticity of the rail fastening;
- Flexibility: the application of two extra rails over a length of 20 metres;
- Settlement: stabilising of ballast by chemical binders;
- Settlement and flexibility: extension of the concrete roadbed instead of the subballast over a length of at least 10 metres;
- Settlement: the application of a reinforced anchor at the end of the slab track in combination with a horizontal slab as buffer at the beginning of the ballasted track.

The transition between bridges and slab track involves specific measures too. There are differences between steel and concrete bridges, but the chosen solution also depends on the length of the bridge. In Figure 9.80 an example is given of a transition with a Rheda-system crossing a bridge. Figure 9.81 shows a transition between two prefabricated slab structures in which the discrete supports (sleeper blocks) are cast into the slab.



1 Reinforced in-situ concrete
2 Continuous reinforced concrete slab
3 Foil + 2 layers of bituminous cardboard
4 Hard foamboard, Styrodur 5000, h = 5cm
5 Concrete cover of 6 cm
6 Sealing
7 Bridge structure

Figure 9.80: A transition with an elastic-plastic intermediate layer (Rheda system)

9.14 General experiences with slab track systems

As one of the concluding sections of this chapter, some experiences concerning the maintenance of high-speed lines are given. Only in Japan, slab track has been applied on a large scale for high-speed lines. In Germany, the amount of slab track is still relatively small. On the other hand, more than thirty slab track systems have been designed and tested there throughout the years. Many of them with satisfactory results yet only a few of them with potential for further, extensive use. Those systems have meanwhile been improved and are increasingly being applied on new high-speed tracks.

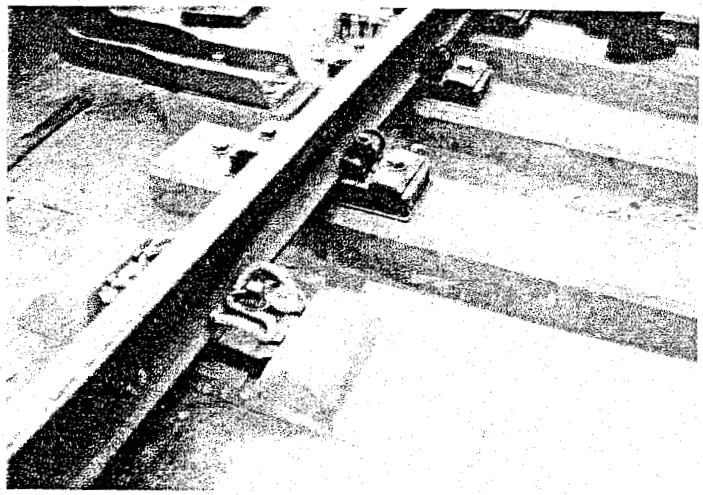


Figure 9.81: A transition between two prefabricated slab structures

In Germany, a new high-speed track is currently being constructed as slab track. Some of the requirements have been mentioned before. Long term experience with slab tracks has shown that these requirements result in a very good track quality directly after construction as well as over many years, as there is hardly any maintenance required.

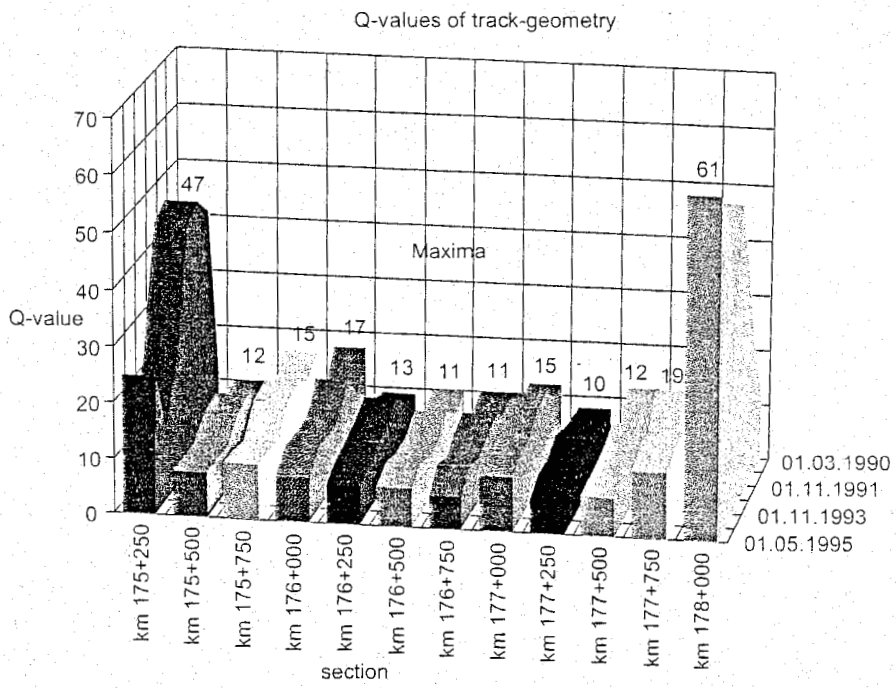


Figure 9.82: Measured Q-values of a length of slab track adjacent to ballasted track over a 5-year period

In Germany, the track geometry is qualified with a value Q. Q stands for a quality-value which is derived from the five parameters measured regarding the track geometry over a 250 m length of track. Concerning ballasted track, a Q-value of 10 can be considered as extremely good track geometry. A Q-value of 100 will indicate that periodic maintenance is needed.

Figure 9.82, for example, shows the difference in Q-levels concerning the quality of the track-geometry. Measurements are shown of twelve sections over a period of time. The ten graphs in the middle are of slab track structure and the outer two are of ballasted track [40], [41], [42].

From the experiences with ballasted track on high-speed lines, it is obvious that slab track systems require hardly any maintenance and provide high availability.

9.15 Maintenance statistics of slab track

The constructed sections of slab track in Germany have, in general, shown what was expected of a track in need of little maintenance. The lasting good track-quality throughout the periods observed does not only guarantee a minimum of required maintenance, the improved driving-comfort for the passengers as well as the highly available track have proven to be benefits of slab track.

If the track stability is provided by means of a rigid slab, the amount of maintenance will decrease dramatically or become almost superfluous. Although the overall experiences concerning maintenance are very promising, due to the limited application in relation to traditional ballasted track, too little is still known to provide empirical rules concerning predictions for progress of deterioration in the case of slab track systems.

In the network of the DB AG, approximately 360 km of track and about 80 switches are constructed in slab track systems [40], [123]. Experience with maintenance of slab track goes back to 1972 when the Rheda pilot-project was accomplished. Concerning the Rheda-system, in general the maintenance work was restricted to changing of the synthetic rail pad. A further maintenance measure was preventive rail grinding.

By means of preventive rail grinding irregularities in the rail and at the rail surface are removed. This is an important condition for smooth running and a decreased noise production. Preventive grinding will also stop the beginning of ripples, further growth of headchecks, and damage to the rail surface by imprints. Ripples develop much slower.

Compared to ballasted track, failures like dents near welds or near (insulated) rail joints will evolve considerably slower as the dynamic forces are lower and hardly cause progressive deterioration at the rail supports. The cost of maintenance of a Rheda-track is approximately only 10 % of the cost involved with ballasted track.

Bad maintenance experiences with the 516 km long Tokaido-line in traditional ballasted track on embankments, which was opened in 1964, lead to the development of slab track [119]. The main part of the Shinkansen-track which was constructed after 1972 has been applied on top of bridges, viaducts, and tunnels. The amount of earth embankments in those lines is less than 5 %.

Maintenance costs in the Japanese experiences amounted to 20-30 % of the costs involved in the case of ballasted track, and yet had a better track-geometry and alignment [4], [161]. The improved track design of the Tohoku-line which was opened in 1982 and had an increased slab-height, application of prestressing, improved frost-proof cement-bituminous under-pouring, and a better drainage in tunnels, made it nearly maintenance-free. The main requirement in using this design is a settlement-free foundation.

10 THE RAIL

10.1 Introduction

As the rail is the most important part of the track structure a separate chapter is devoted to it. In Chapter 8 several basic functions have been discussed. In this chapter some fundamental aspects of the quality of rails are discussed, such as the rail manufacturing process, acceptance procedures, mechanical properties, flash butt and Thermit welding, control of weld geometry, required standards, rail failure types and rail defect statistics.

10.2 Modern rail manufacturing

Modern rail manufacturing technology is considered in the new standard EN 13674 of the European Community. Different to existing specifications, it is a performance based standard. Some of the manufacturing techniques are defined in order to ensure that the rail shows good service properties. The steel may be produced either by the basic oxygen process (BOF) or in an electric arc furnace, although the latter is currently not used in Europe. Ingot casting is no longer allowed. Secondary metallurgy is more or less standard practice. Vacuum degassing is mandatory in order to avoid rail breakage caused by flakes and non-metallic inclusions. The manufacturer has to apply a quality management system to ensure consistent product quality and to pass a qualifying procedure to become approved for delivery.

The rail manufacturing process consists of the following main parts as indicated in Figure 10.1:

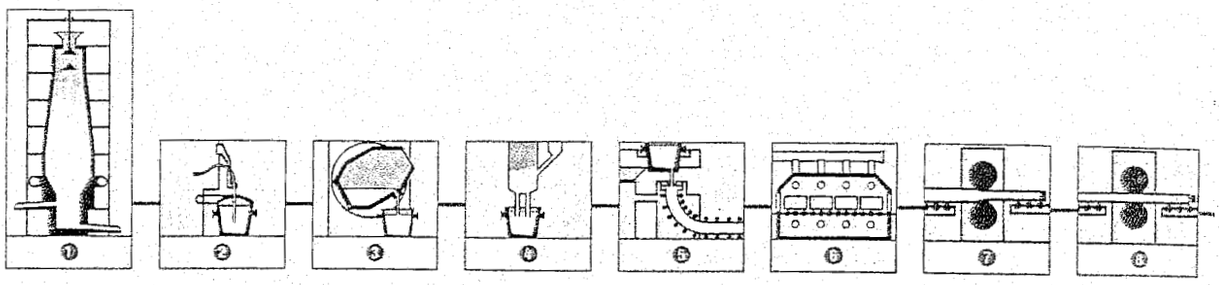
- Blast furnace;
- Steel-making;
- Continuous casting;
- Rolling;
- Straightening;
- Measurements (ultrasonic, geometry, manual inspection);
- Final acceptance.

In the next part of this section some of these processes will be discussed in greater detail.

10.2.1 Blast furnace

Steel is in fact iron which has been refined with carefully measured amounts of other elements added to it. Iron is found as iron oxide in rocks, known as iron ore. This only occurs in sufficiently large quantities and with reasonable accessibility in a few scattered areas of the world, for the most part in Scandinavia, the Americas, Australia, North Africa, and Russia.

The ore is graded and crushed and some of the finer ore is taken to the sinter plant where it is mixed with coke and limestone and heated to form an iron-rich clinker known as sinter. This sinter is fed into the top of the blast furnace together with more iron ore, coke and limestone in controlled proportions, and the whole is fired. Great heat is generated and fanned to white hot intensity by blasts of superheated air.



The rail manufacturing process at Thyssen.

1. *Blast furnace*
2. *Hot metal desulphurization facility*
3. *Basic-oxygen steelmaking shop*
4. *Degassing plant*
5. *Continuous casting plant*
6. *Walking-beam reheating furnace*
7. *Roughing mill*
8. *Finishing stand 1*
9. *Finishing stand 2*

Figure 10.1: The rail manufacturing process at Thyssen

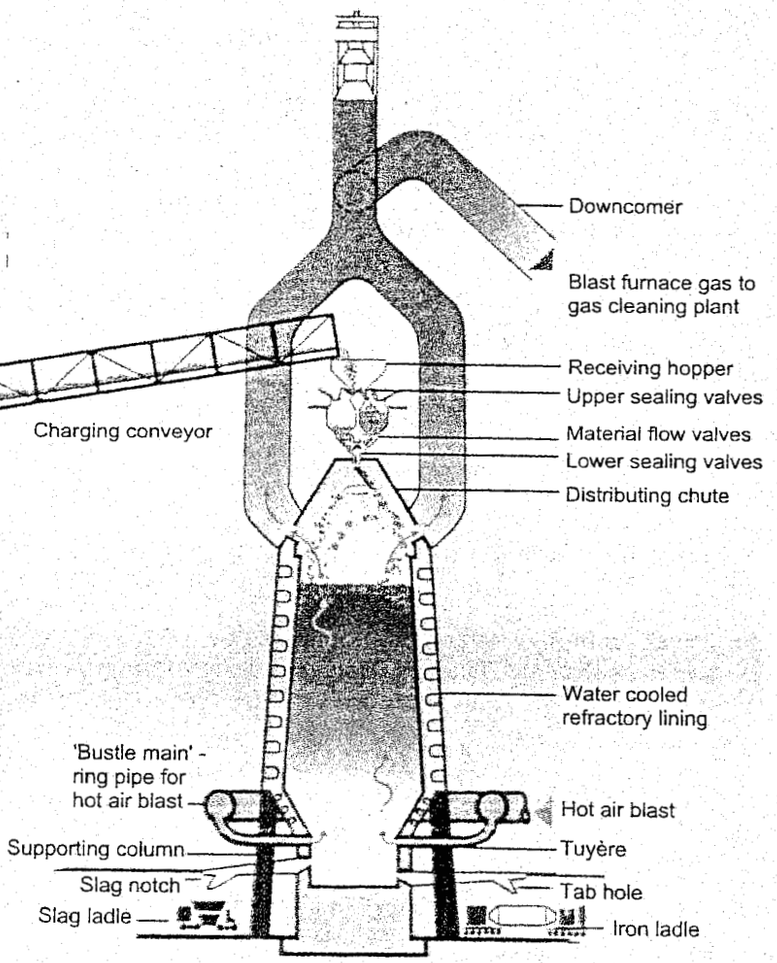
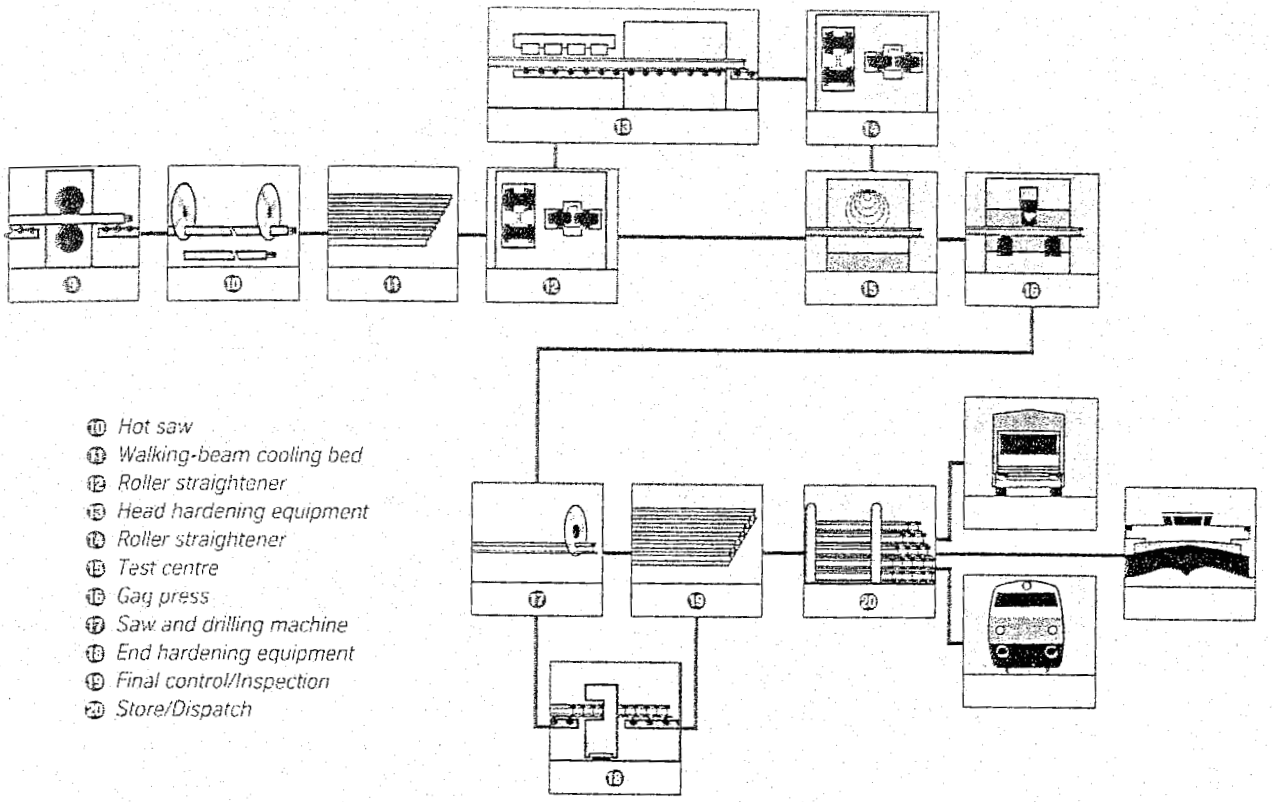


Figure 10.2: Blast furnace schematic



By means of this process the iron ore is reduced. The iron in the sinter and ore melts to form a pool of molten metal at the foot of the furnace which has an air temperature of approximately 1500 °C. The limestone combines with impurities in the iron ore, forming a liquid slag which, being lighter than the metal, floats on top of it. Figure 10.2 is a schematic representation of a blast furnace section (Corus).

The blast furnace process is continuous. As the slag builds up on the surface of the molten iron it is periodically tapped off. Similarly, as a sufficient quantity of molten iron accumulates beneath the slag it too is tapped off ready for the steel furnaces.

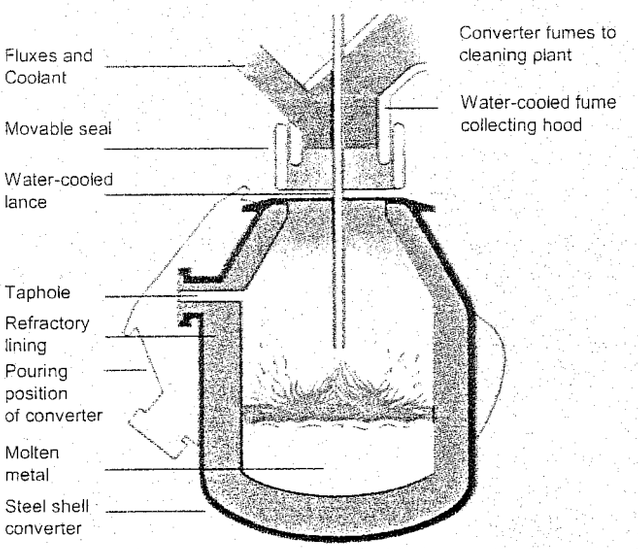
Meanwhile, the raw materials, sinter, ore, coke and limestone, are continually poured into the top of the furnace and heated air is blasted in at the bottom. This process goes on continuously, often for four years or more, before the heat-resistant brick lining of the furnace begins to deteriorate.

10.2.2 Steel-making

The Basic Oxygen Furnace process, known as BOF, has become the major method of making steel. Modern furnaces will take a charge of 150-350 tonnes or more at a time and convert it into steel in 40 minutes (Figure 10.3).

Hot metal is the principal material used in the Basic Oxygen process. First, the hot metal is desulfurized.The furnace is tilted and charged first with scrap and then with the molten iron, and is then returned to the upright position. The molten iron accounts for about 70 per cent of the charge, the scrap 30 per cent.

A water-cooled oxygen lance is lowered into the furnace and high-purity dry oxygen is blown onto the metal at very great speed. The oxygen combines with carbon and other unwanted elements, thus eliminating these impurities from the molten charge. The carbon is blown down to less than 0.1%. During the "blow" lime is added as a flux to help carry off the oxidised impurities as a floating layer of slag.



The positioning of the lances, the determination of the volume of oxygen to be injected, the additions to be made and the corrective steps required are computer-controlled and fully automated.

When all the steel has been tapped out into a ladle the converter is turned upside down and the residual slag is tipped into a waiting slag ladle for removal to a slag pool.

In the ladle the molten steel is carburized and alloyed. By means of the so-called secondary metallurgy the chemical composition is refined, the temperature adjusted and the cleanliness improved through decantation of inclusions.

Figure 10.4 shows a photograph of the filling process.

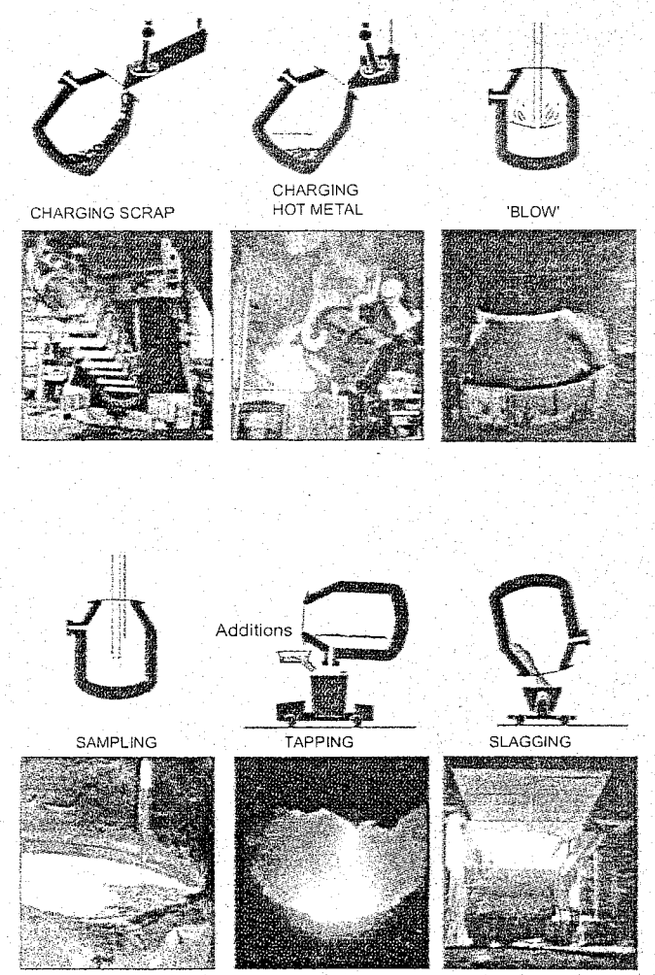


Figure 10.3: Basic Oxygen Furnace (BOF) (Corus)

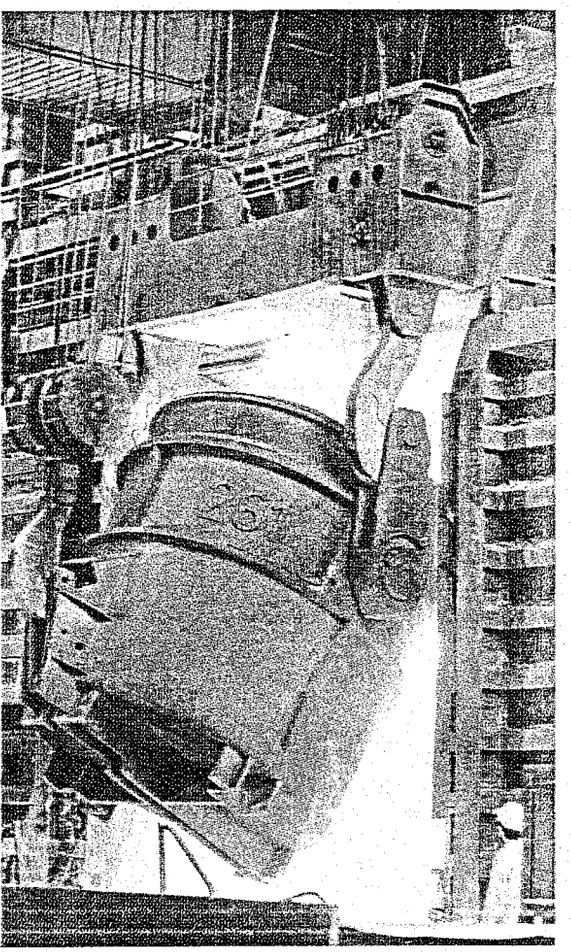
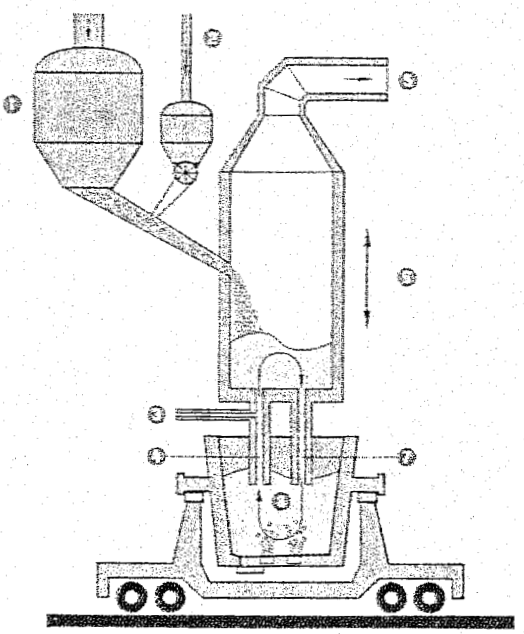


Figure 10.4: Filling proces of converter

10.2.3 Vacuum degassing and argon flushing

In the modern steel-making process several other steps are implemented to improve steel quality. Argon flushing facilities are used to homogenise temperature and chemical composition. Vacuum degassing units, such as depicted schematically in Figure 10.5, reduce the hydrogen content to less than 2 ppm and improve the oxidic cleanliness of the steel. With a hydrogen content of less than 2 ppm in the liquid steel no particular cooling measures are required to prevent the development of flakes. Under train loads these flakes can initiate fatigue cracks, an example of which is represented in Figure 10.69. (UIC code 211)

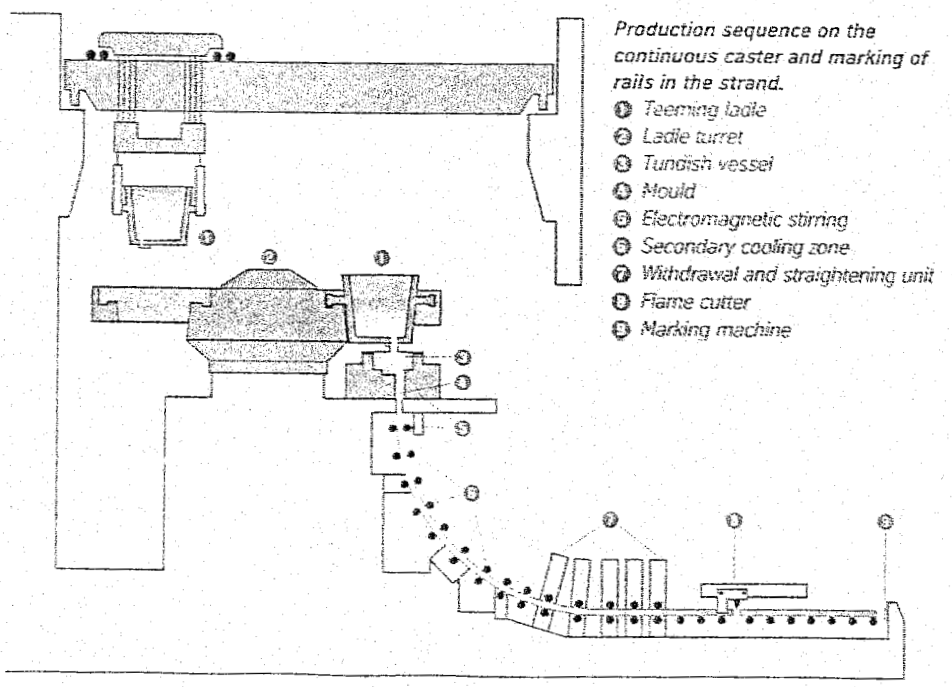


Schematic representation of an RH vacuum facility.
1. *Alloying installation*
2. *Cellular alloying facility*
3. *Scavenging gas*
4. *Suction pipe*
5. *Vacuum producer*
6. *Vacuum vessel in three parts*
7. *Outlet pipe*
8. *Ladle with porous plugs*

Figure 10.5: Schematic representation of a vacuum degassing unit (Thyssen)

10.2.4 Continuous casting

The principle of continuous casting, which is at present used by most of the steel works, is annotated in Figure 10.6. The liquid steel is supplied in a 150 - 350 tonne ladle which is placed in a turret. This turret can contain 2 ladles to practice sequencing of ladles. When the molten steel is poured from the ladle into the tundish the next ladle can be prepared.
In this way teeming may proceed continuously.



Production sequence on the continuous caster and marking of rails in the strand.
1. *Teeming ladle*
2. *Ladle turret*
3. *Tundish vessel*
4. *Mould*
5. *Electromagnetic stirring*
6. *Secondary cooling zone*
7. *Withdrawal and straightening unit*
8. *Flame cutter*
9. *Marking machine*

Figure 10.6: Principle of a Continuous Casting machine (Thyssen)

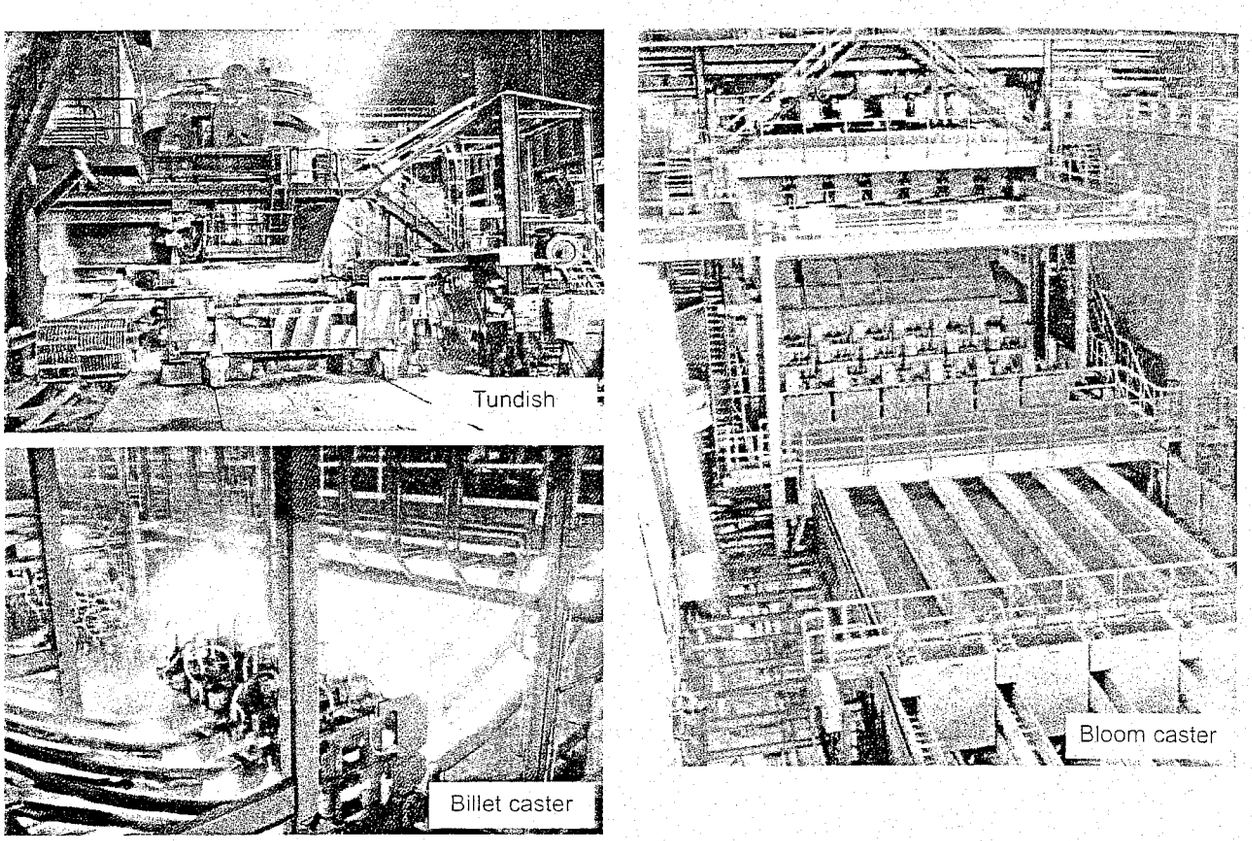


Figure 10.7: Some pictures of the continuous casting process

Figure 10.7 shows some pictures of a continuous casting facility. The liquid steel is cast from the ladle into the tundish using submerged pouring techniques. Metering nozzles are used to deliver precisely the right amount of steel to the 6 to 8 moulds.

Figure 10.8: Hot strands passing through cooling chambers in a circular arc

All steel is protected from atmospheric oxidation by refractory tubes between ladle and tundish, and also between tundish and mould. The double-walled moulds are water-cooled. They shape the strands and may have different cross-sections.

Thyssen [18], for example, uses mould sections of 265 x 385 mm. The mould corners are chamfered to prevent corner cracking.

The amount of super-heat contained in the liquid steel has a profound effect upon the internal metallurgical quality of the cast bloom. The liquid steel temperature in the tundish is therefore maintained within the range liquidus plus 15 °C.

During casting the moulds oscillate with a frequency of 60 - 200 cycles per minute, depending on casting speed and oscillation stroke, to prevent the steel from adhering to the copper mould. The casting speed amounts to about 0.8 m/min. To improve the solidification structure the strands are equipped with electromagnetic stirring coils.

On exiting the spray chambers for secondary cooling, shown in Figure 10.8, the hot strands, curved with a radius of 10 - 13 m, enter the straightener drive units. After flame-cutting to the required length and hood or pile-cooling, the blooms are shipped to the reheating units, mostly walking-beam furnaces. At the beginning of casting a hydrogen increase may occur. If critical hydrogen values are reached, these blooms are controlled-cooled to get a hydrogen effusion.

Compared with ingot casting continuous casting has the advantages of producing a cleaner steel with smaller and more uniformly distributed inclusions.

10.2.5 Rolling mill

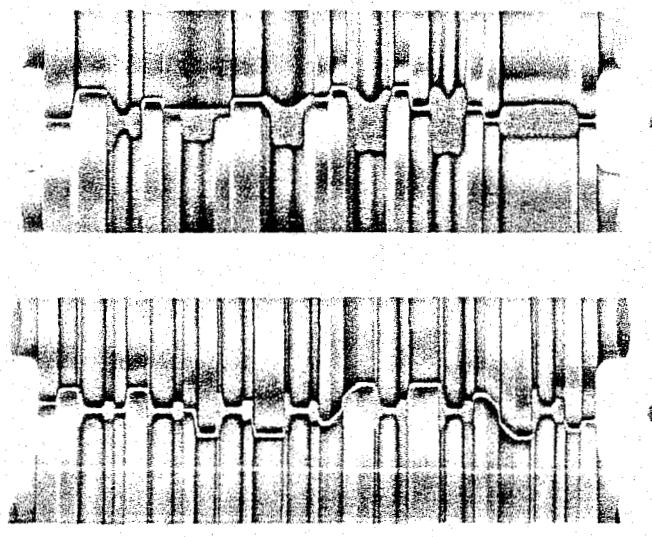
The blooms are inspected and any defects removed to ensure that finished products have a high-grade surface quality. Prior to being rolled into steel sections the blooms are reheated to 1250 °C in specially designed furnaces. After the blooms have left the walking-beam furnace they are sprayed with water at a pressure of about 200 bar to remove the furnace scale. This prevents the surfaces of the rollers and rolled steel from being damaged by the hard scale.

The blooms are rolled on a roughing stand to dimensions suitable for the initial pass on the section mill. The bloom coming from the roughing stand is rolled into the finished rail in a total of 8-11 passes.

Figure 10.9 demonstrates the principle of rolling a rail profile according to the Bartscherer roll pass design. The two knife passes (3 and 4) cause the dendrites, initially oriented perpendicular to the surface of the bloom, to be relocated roughly parallel to the bottom surface of the rail foot. Material from the interior of the bloom is also brought to the surface. The horizontal reorientation of the dendrites substantially improves the behaviour of the rail foot under dynamic bending stresses.

Some rail manufacturers apply the universal rolling process [45] in which the metal is shaped alternately in a universal stand and an associated edging stand. In this way tight tolerances and high surface quality are achieved. The improved forging action positively influences the fatigue strength.

Descaling equipment removes the mill scale, thereby permitting a consistently smooth product surface. During the final pass the rolling mark is rolled into the rail. This mark provides information on the manufacturer, section, year of manufacture, and rail steel grade. Additionally, each rail is marked with the heat and bloom numbers while still hot so as to identify the heat and bloom from which it was made.



Calibration of flat-bottom rails.
1 *Roughing mill*
2 *Finishing mill*

Figure 10.9: Bartscherer roll pass design used by Thyssen

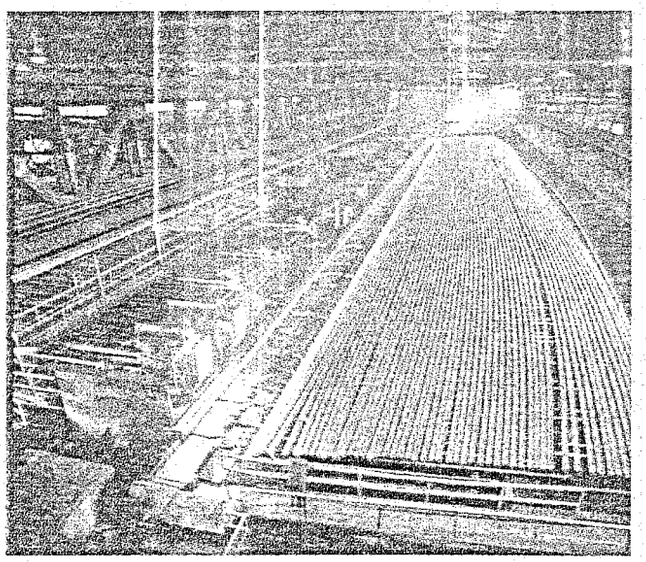


Figure 10.10: Cooling bed Voest-Alpine

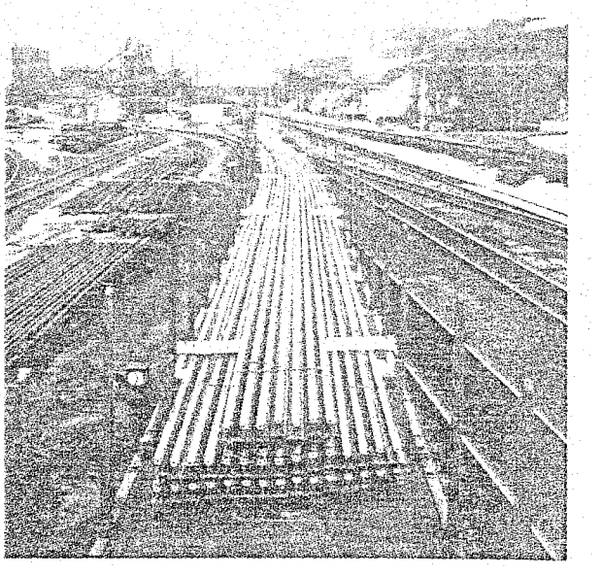


Figure 10.11: Long rail delivery

The rail is rolled up to 125 m length (Figure 10.10), hot-cropped by saws, and is either cut into standard lengths in the finishing shop, or delivered as ultra-long rail directly to the job-site (Figure 10.11).

The hot rails are placed on a walking-beam cooling bed to cool them throughout their entire length. It takes about 3-4 hours before the temperature has dropped from 800 °C to less than 60 °C. Walking-beam type of handling prevents damage occurring to the surfaces. Cooling is followed by roller straightening.

10.2.6 Finishing shop

On account of the asymmetric profile, rail head and rail foot cool at different rates as a result of which the rails curve in a vertical direction. To correct this, the rails are straightened in the vertical and also in the lateral direction by a roller straightener composed of 7 rollers as depicted in Figure 10.12. The circumference of the rollers amounts to about 3 m and this value reappears in the rail as wavelengths of geometrical deviations called rolling defects.

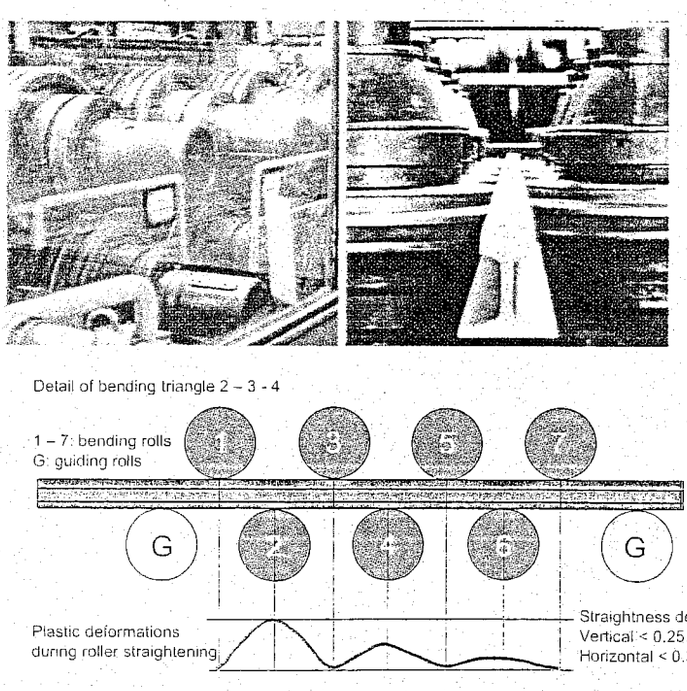


Figure 10.12: Rolling straightener

The roller-straightening process, as applied to date, provides excellent straightness but introduces detrimental residual stresses of the order of 100 to 300 N/mm², depending on the yield stress of the rail steel.

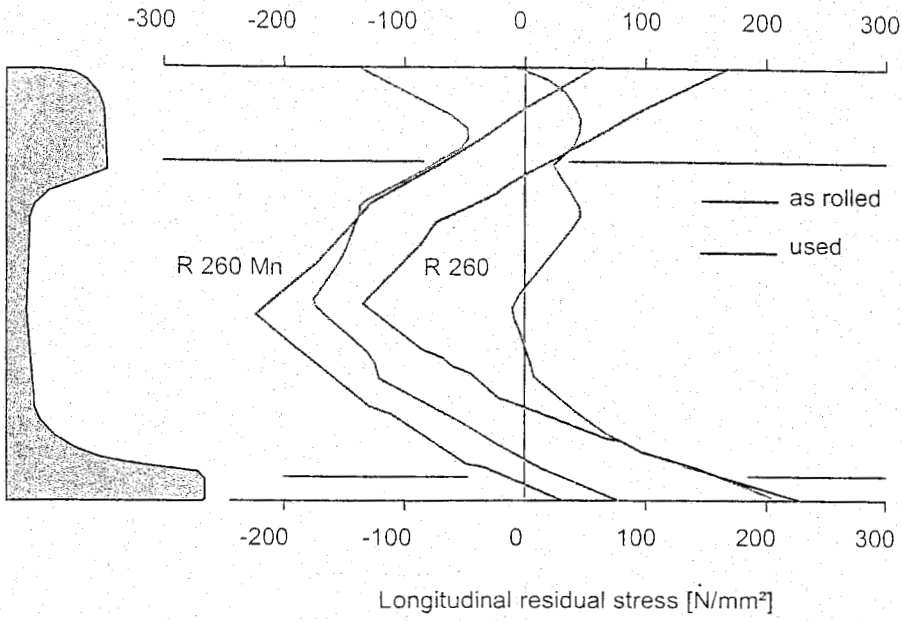


Figure 10.13: Residual stresses due to roller straightening

These stresses should be added to the load stresses. Figure 10.13 summarizes the residual stresses measured by ORE D 156 [217] for new grade R 260 (900 A) and R 260 Mn (900 B) rails and used rails. The figure shows that the residual stresses in the rail head change from tension into compressive stresses due to the rolling-out effect caused by the wheel-rail forces. The residual stresses are limited in the European Standard to 250 N/mm² in the centre of the rail foot.

Subsequent to roller straightening each rail passes through the inspection centre. The rails are checked in continuous operation against pre-specified criteria for internal defects, surface defects, dimensional accuracy and running surface straightness.

Any internal defects are detected by the ultrasonic testing method according to Figure 10.14a and Figure 10.14b. Eddy current testing equipment is used to check for surface defects at the rail head and the foot (see Figure 10.15).

The straightness of the rail is measured automatically by laser equipment in the 0.5 - 3 m waveband. The rail ends are then checked for their straightness either in the automatic straightness plant or manually by means of a straightedge. If necessary, they are additionally straightened with a gag press.

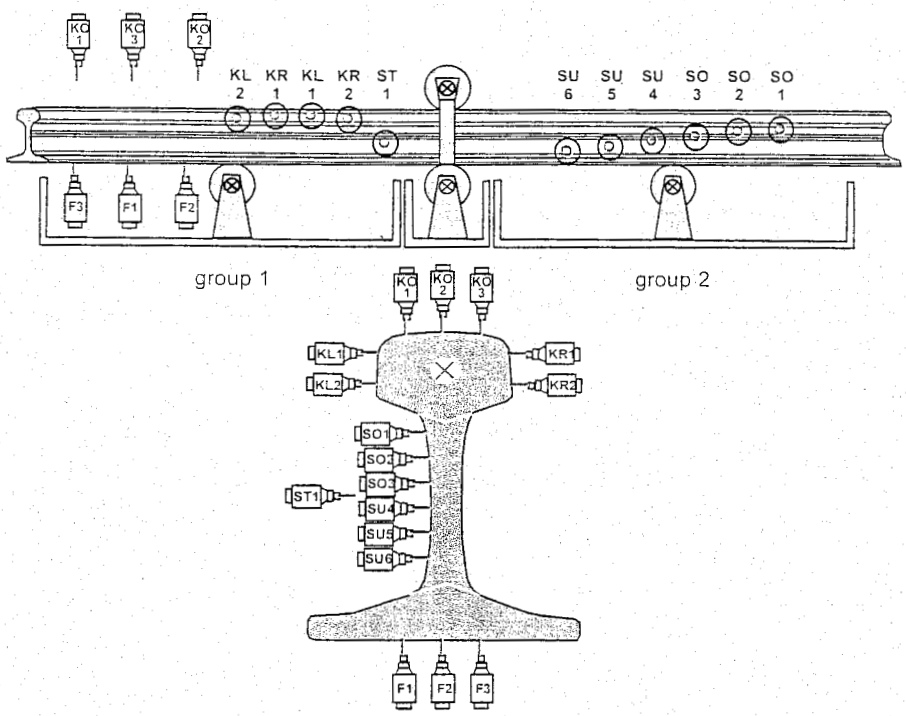


Figure 10.14: (a) Ultrasonic testing process

The manufacturing process ends with cutting the rails exactly to the ordered lengths by means of carbide saws. If required, the rails are also provided with fishplate holes. The rails are once again checked and sorted in a finishing shop. It is here that the final inspection and acceptance are carried out by mill agents and/or an inspection agency.

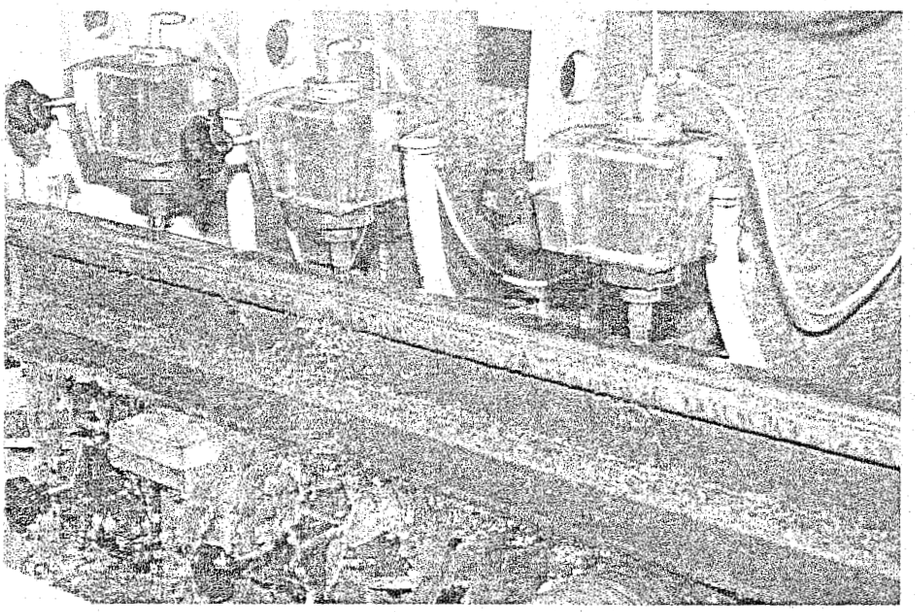


Figure 10.14: (b) Ultrasonic testing device

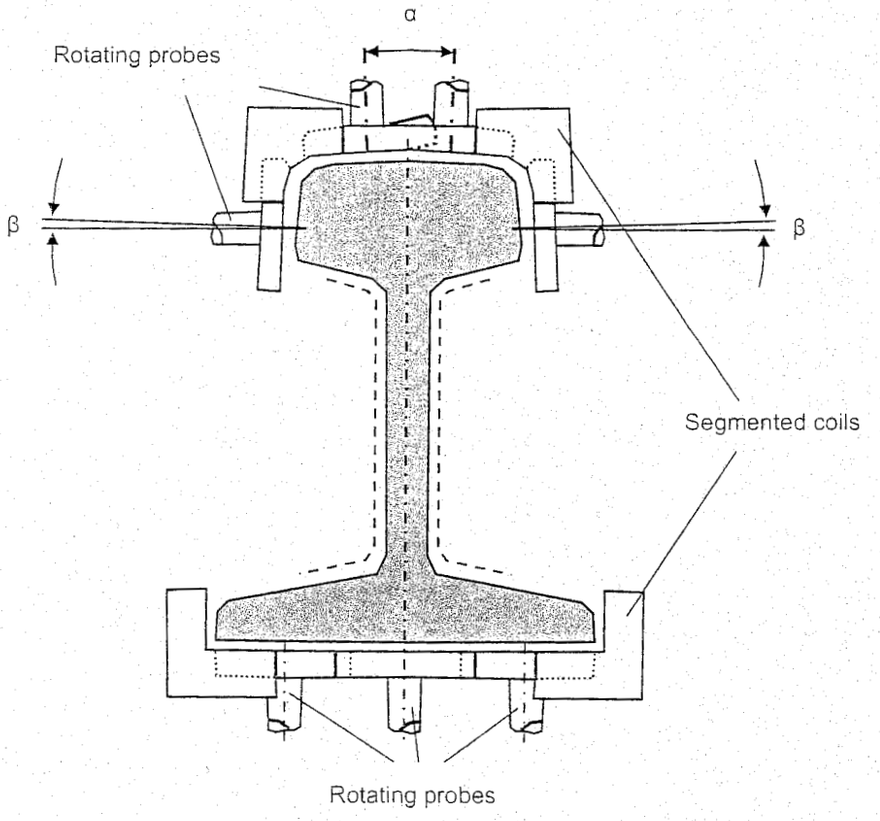
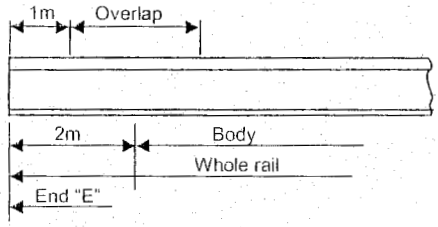
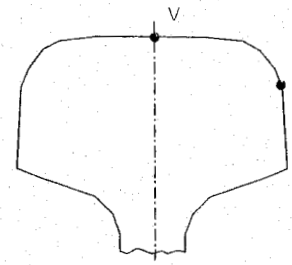


Figure 10.15: Eddy current testing

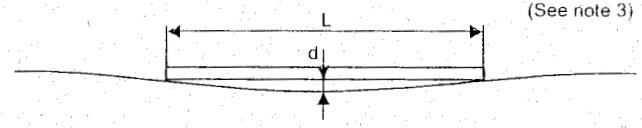
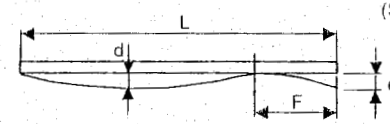
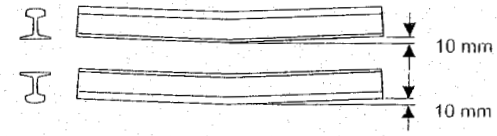




V } location of
} flatness
H } measurements

The position of H is nominally 5-10 mm below the gauge corner on the side of the head

Table 10.1: Straightness, surface flatness and twist tolerances according to EN 13674 - 1

		Class B		Class A (1)		
		d	L	d	L	
BODY (see note 2)	Vertical flatness	≤ 0.4 mm	3 m (see note 4)	≤ 0.3 mm	3 m (see note 4)	(See note 3) 
		and		and		
		≤ 0.3 mm	1 m (see note 4)	≤ 0.2 mm	1 m	
	Horizontal flatness	≤ 0.6 mm	1.5 m (see note 4)	≤ 0.45 mm	1.5 m (see note 4)	
ENDS (see note 2)	End "E"	1.5 m		2 m		(See note 3)  if e > 0 F ≥ 0.6 m
	Vertical straightness	≤ 0.5 mm	1.5 m	≤ 0.4 mm	2 m	
				and		
				≤ 0.3 mm	1 m (see note 5)	
		and		and		
		e ≤ 0.2 mm		e ≤ 0.2 mm		
	Horizontal straightness	≤ 0.7 mm	1.5 m	≤ 0.6 mm	2 m	(See note3)
				and		
				≤ 0.4 mm	1 m (see note 5)	
OVERLAP (see note 2)	Length of overlap	1.5 m		2 m		(See note3)
	Vertical flatness	≤ 0.4 mm	1.5 m (see note 4)	≤ 0.3 mm	2 m (see note 1)	
	Horizontal flatness	≤ 0.6 mm	1.5 m (see note 4)	≤ 0.6 mm	2 m (see note 1)	
WHOLE RAIL	Upsweep and downsweep	10 mm (see Note 6)		10 mm (see Note 6)		
	Side sweep	curve radius R > 1500 m		curve radius R > 1500 m		
	Twist	(see notes 7 & 8)		(see notes 7 & 8)		

10.2.7 Inspection and acceptance

The annual purchase of rails involves huge amounts of money. For the European administrations a global figure of 1.5-5 tonnes per km of main track can be indicated as the amount of rails to be renewed each year. For the NS the annual rail demand amounts to about 23.000 tonnes (1987), corresponding to a value of EUR 15 million. It is not surprising that the railway administrations are very keen on the inspection and final acceptance procedure, all the more since the 5-year guarantee period is restricted to mere refunding of the material costs, which are only a fraction of the total costs involved in replacing a rail.

Inspection and acceptance can be done either by a special independent department at the manufacturer's, or as a concerted effort on the part of manufacturer and customer, in which case the customer takes care of the final acceptance. The latter used to be common practice in Europe. However, the acceptance philosophy has changed more and more towards the supplier having full responsibility for good quality. If the manufacturer applies a certified quality management system, the inspection is replaced by periodic audits by the customer. With this approach the railways are provided with all necessary information according to their specifications and only take care of checking procedures.

Additional requirements

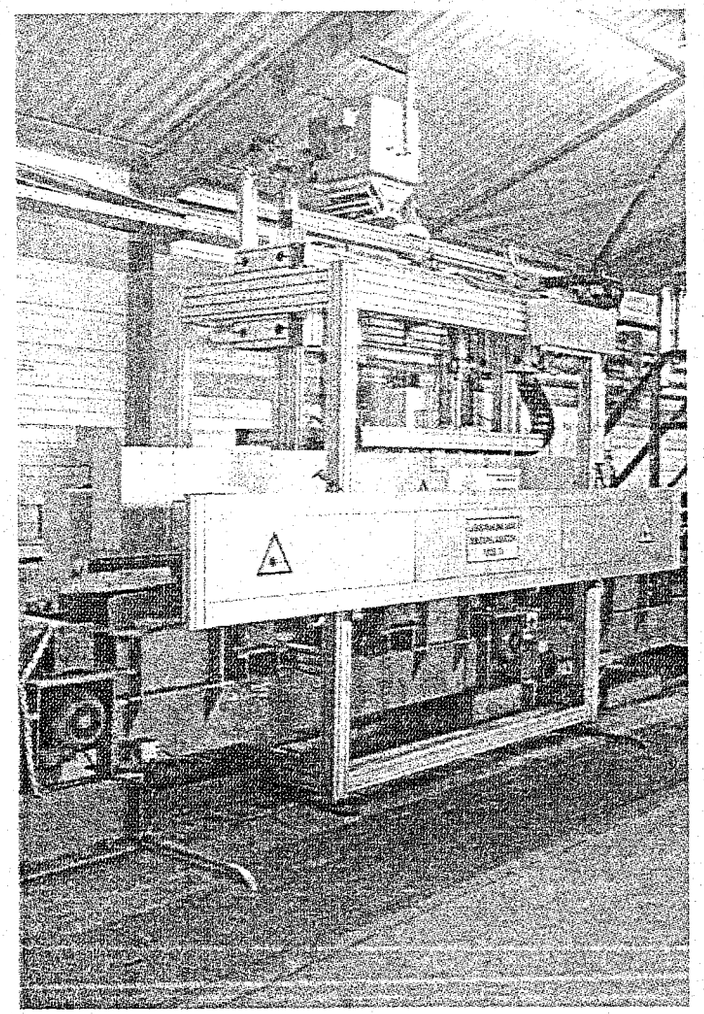


Figure 10.16: Rail surface flatness measuring device

The railways may set additional specifications for rail supply. The NS, for instance, requires that the vertical geometry is measured for the entire rail in the 0.5 - 3 m waveband, whereby 5% of the peak-peak values, i.e. twice the amplitude, may exceed a level of 0.3 mm, with a maximum of 0.4 mm. This was considered in the European standard EN 13674-1 as class A straightness.

Figure 10.16 shows a photograph of the Voest-Alpine installation for continuous measurement of both vertical and horizontal rail surface flatness.

With every delivery the supplier must report the percentage of rails rejected during mill inspection as well as the reasons why.

For continuous cast rails of grade R 260 (900) and higher, the test piece for the tensile test has to be taken from the waste piece at the end of the rail rolled from the first block of a cast sequence. Figure 10.17 shows the tensile test arrangement.

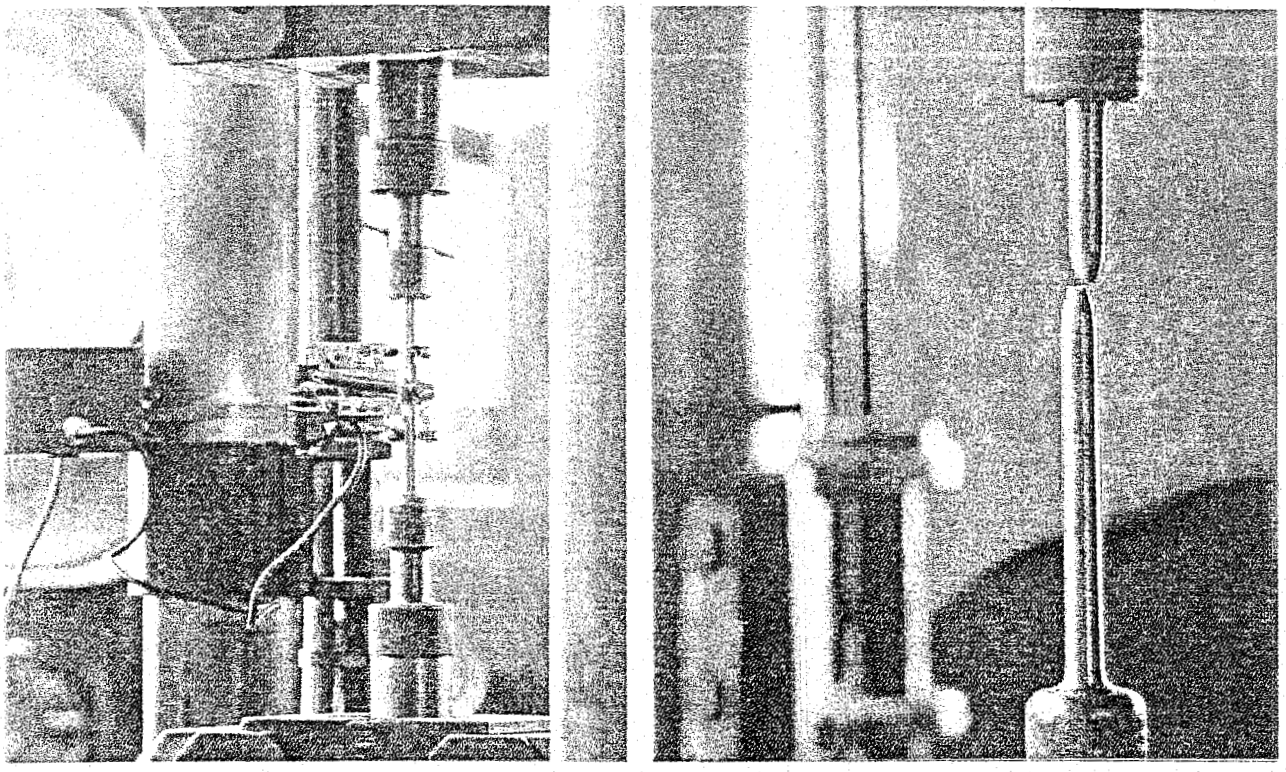


Figure 10.17: Determination of tensile strength of a standardized specimen

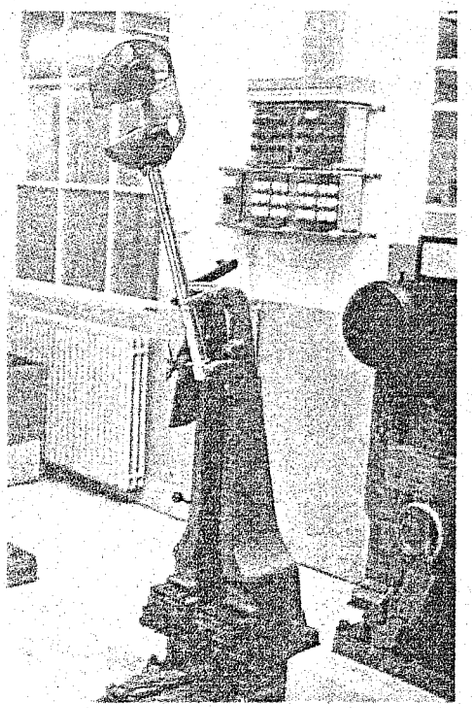
The waste piece for the drop weight test must be taken from the top end of the rail rolled from the last block of the last sequence.

For the notched bar impact test a block in the middle of a cast sequence should be taken. The test pieces must be manufactured in accordance with DVMF and should be taken from the rail foot as shown in Figure 10.18. The mean impact value for 6 pieces should be determined at -20 °C and at 20 °C.

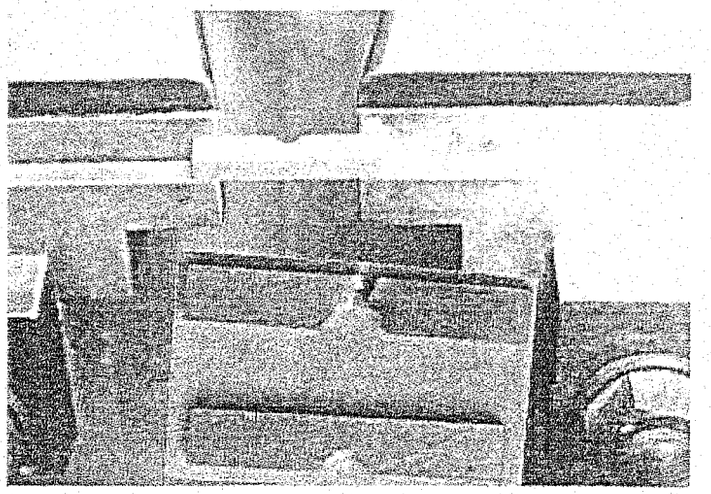
Although no minimum requirements are laid down, it is expected that the mean values for the grade R 260 Mn will be at least 34 Joule and 40 Joule respectively in accordance with the bandwidth presented in Figure 10.18. For the grade R 260 the minimum requirements can be set at 25 Joule at -20 °C and 35 Joule at 20 °C. The tolerances specified in Table 10.1 are applicable.

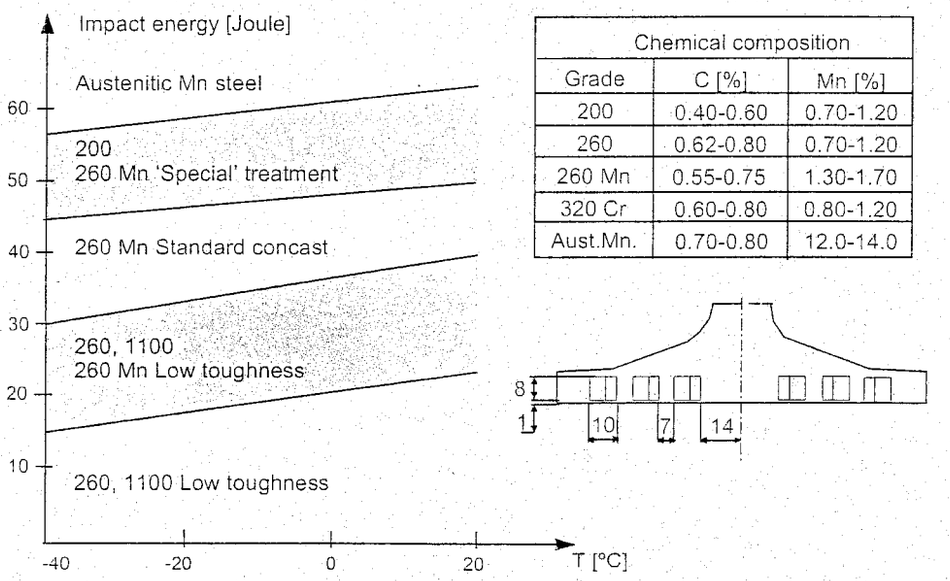
The rail dimensions are checked with special templates for upper and lower tolerances. An example of this examination is shown in Figure 10.19.

For checking the railhead shape and dimensions often the MINIPROF system is used, which records and analyses the profile in real time (Figure 10.20). For further details of this system please refer to Chapter 16.



Impact-test equipment





Chemical composition		
Grade	C [%]	Mn [%]
200	0.40-0.60	0.70-1.20
260	0.62-0.80	0.70-1.20
260 Mn	0.55-0.75	1.30-1.70
320 Cr	0.60-0.80	0.80-1.20
Aust.Mn.	0.70-0.80	12.0-14.0

Figure 10.18: NS guidelines concerning DVMF impact values

10.2.8 Rail profiles

Worldwide many different rail profiles are in use. Table 10.2 gives a survey of the different sections rolled by Thyssen Stahl in Germany. The table also contains dimensional data and cross-sectional properties. Although the drafting group of CEN tried to reduce the number of profiles, still 21 different profiles with sections heavier than 46 kg/m are used in Europe.

10.2.9 Indication of profile types according to CEN

According to the new CEN standards, rail profiles are indicated by the mass per meter followed by the letter E and a serial number starting with 1. The old profiles UIC 54 and UIC 60 are now indicated as 54 E1 and 60 E1, respectively.

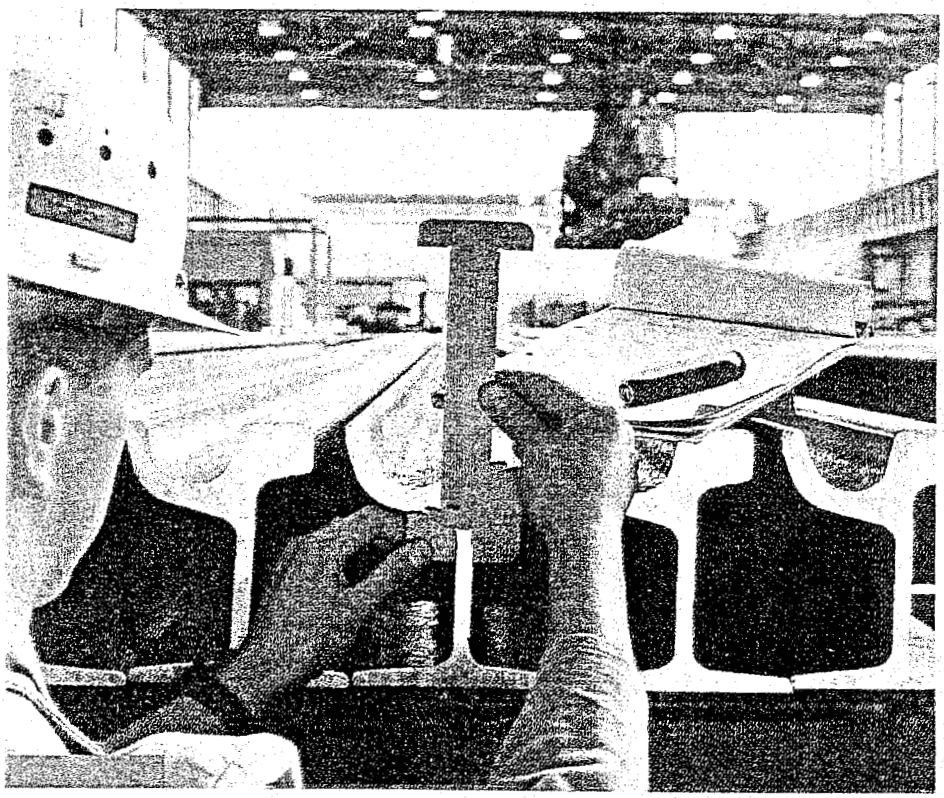


Figure 10.19: Checking of rail dimensions with special templates

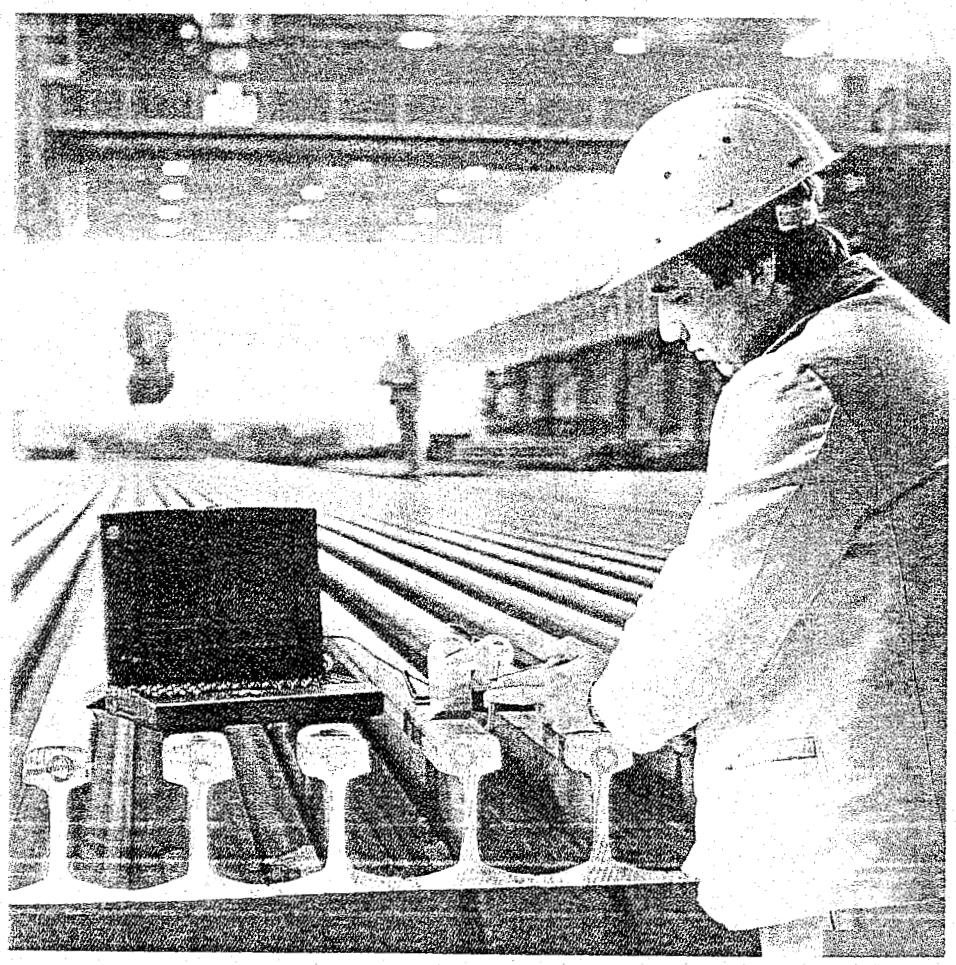
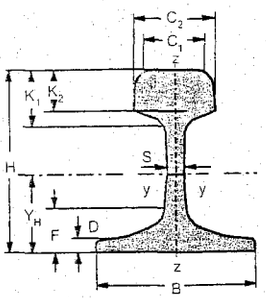


Figure 10.20: Rail head profile measurement by means of MINIPROF



	H Rail height		B Base width		C_1 Head width		C_2 Head width		S web		K_1 Total head height		K_2 Head height	
	mm	inch	mm	inch	mm	inch	mm	inch	mm	inch	mm	inch	mm	inch
German sections														
SMR 29	115.00	4.528	90.00	3.543	47.00	1.850	55.00	2.165	12.00	0.472	40.00	1.575	31.00	1.221
S 30	108.00	4.252	108.00	4.525	60.30	2.374	60.30	2.374	12.30	0.484	31.00	1.228	24.00	0.945
S 33	134.00	5.276	105.00	4.134	58.00	2.383	58.00	2.283	11.00	0.433	39.00	1.535	31.75	1.250
S 41, 10 R	138.00	5.433	125.00	4.921	67.00	2.638	67.00	2.638	12.00	0.472	43.00	1.692	31.83	1.253
S 41, 14 R	138.00	5.433	125.00	4.921	67.00	2.638	67.00	2.638	12.00	0.472	43.00	1.692	31.83	1.253
49 E1 (S 49)	149.00	5.866	125.00	4.921	67.00	2.638	70.00	2.756	14.00	0.551	51.50	2.027	39.80	1.567
54 E3 (S 54)	154.00	6.063	125.00	4.921	67.00	2.638	70.00	2.756	16.00	0.630	55.00	2.165	43.30	1.705
UIC sections														
50 E4 (UIC 50)	152.00	5.984	125.00	4.921	70.00	2.756	72.20	2.843	15.00	0.591	49.40	1.945	36.30	1.429
54 E2 (UIC 54E) (SBB IV)	161.00	6.339	125.00	4.921	67.00	2.638	69.50	2.736	16.00	0.630	51.40	2.024	38.70	1.524
54 E1 (UIC 54) (SBB III)	159.00	6.260	140.00	5.512	70.00	2.756	72.20	2.843	16.00	0.630	49.40	1.945	36.30	1.429
60 E1 (UIC 60)	172.00	6.771	150.00	5.906	72.00	2.835	74.30	2.925	16.50	0.650	51.00	2.008	37.50	1.476
American sections														
ASCE 60	107.95	4.250	107.95	4.250	60.33	2.375	60.33	2.375	12.30	0.484	30.96	1.218	23.99	0.944
ASCE 75	122.20	4.811	122.20	4.811	62.70	2.468	72.20	2.862	13.50	0.531	36.10	1.422	28.90	1.138
ASCE 80	127.00	5.000	127.00	5.000	63.50	2.500	63.50	2.500	13.90	0.547	38.10	1.500	30.80	1.213
ASCE 90	142.90	5.625	130.20	5.125	61.50	2.421	65.10	2.563	14.30	0.562	37.30	1.469	29.20	1.150
ASCE 100	152.40	6.000	136.50	5.375	65.30	2.570	68.30	2.689	14.30	0.562	42.10	1.657	33.50	1.319
ASCE 115	168.27	6.625	139.70	5.500	69.06	2.719	69.06	2.740	15.87	0.625	42.86	1.687	34.23	1.687
ASCE 132	180.98	7.125	152.40	6.000	74.45	2.931	76.20	3.000	16.67	0.656	44.45	1.750	34.93	1.375
ASCE 133	179.39	7.062	152.40	6.000	71.10	2.799	76.20	3.000	17.46	0.687	49.21	1.937	36.51	1.437
ASCE 136	185.74	7.312	152.40	6.000	72.62	2.859	74.61	2.937	17.46	0.687	49.21	1.937	39.89	1.570
CB 122	172.21	6.780	152.40	6.000	71.14	2.801	74.68	2.940	16.51	0.650	49.02	1.930	36.00	1.417
British sections														
BS 60 R	114.30	4.500	109.54	4.312	57.20	2.252	57.20	2.252	11.11	0.437	35.70	1.406	26.20	1.031
BS 70 A	123.80	4.874	111.10	4.375	60.30	2.374	60.30	2.374	12.30	0.484	39.70	1.562	28.20	1.110
BS 80 A	133.40	5.251	117.50	4.625	63.50	2.500	65.40	2.570	13.10	0.516	42.50	1.672	30.60	1.205
BS 90 R	142.90	5.626	136.50	5.374	66.70	2.626	66.70	2.626	13.90	0.547	43.70	1.720	32.50	1.280
BS 90 A	142.90	5.626	127.00	5.000	66.70	2.626	68.75	2.707	13.90	0.547	46.00	1.811	33.50	1.319
BS 113 A	158.75	6.250	139.70	5.500	69.85	2.750	72.06	2.837	20.00	0.787	49.21	1.937	36.12	1.422
Various sections														
U 33	145.00	5.709	134.00	5.276	62.00	2.441	64.30	2.531	15.00	0.591	47.00	1.850	36.60	1.429
India 52	156.00	6.142	136.00	5.354	67.00	2.638	67.00	2.638	15.50	0.610	51.00	2.008	36.82	1.528
Netherlands SA 42	80.00	3.150	80.00	3.150	70.00	2.756	72.2	2.843	40.00	1.575	46.90	1.846	36.30	1.429
Netherlands 46 E3 (NP 46)	142.00	5.590	120.00	4.724	74.00	2.913	76.00	2.992	14.00	0.551	42.50	1.673	34.00	1.338
Denmark Form V	141.00	5.551	126.00	4.960	69.30	2.728	71.30	2.807	13.80	0.543	43.00	1.692	34.10	1.343
Denmark Form VII	172.00	6.771	150.00	5.906	72.00	2.835	74.30	2.925	16.50	0.650	51.00	2.008	37.50	1.476
Switzerland SBB I	145.00	5.708	125.00	4.921	65.00	2.559	65.00	2.559	14.00	0.551	45.00	1.772	36.88	1.452
Turkey 145/46.303	145.00	5.709	134.00	5.276	62.00	2.441	64.30	2.531	15.00	0.591	47.00	1.850	36.30	1.429
South Africa SAR 48	150.00	5.905	127.00	5.000	68.00	2.677	68.00	2.677	14.00	0.551	43.00	1.692	35.00	1.377
China 43	140.00	5.512	114.00	4.488	70.00	2.756	70.00	2.756	14.50	0.571	42.00	1.654	29.50	1.161
China 50	152.00	5.984	132.00	5.197	70.00	2.756	70.00	2.756	15.50	0.610	42.00	1.654	33.30	1.311

Table 10.2: Rail sections rolled by Thyssen

F Base height		D Base thickness		A Cross Section		G Weight		Y_h Neutral axis		J_x Moment of inertia		W_x Section modulus head		Stability index B/H	Efficiency index W_x/G		Rigidity index J_x/G	
mm	inch	mm	inch	mm^2	$inch^2$	kg/m	lb/yd	mm	inch	cm^4	$inch^4$	cm^3	$inch^3$		$cm^3/kg/m$	$inch^3/lbs/yd$	$cm^4/kg/m$	$inch^4/lbs/yd$
19.00	0.748	10.00	0.394	3794	5.880	29.78	60.04	55.00	2.165	624	14.992	103.98	6.35	0.783	3.492	0.106	20.954	0.250
19.50	0.768	7.00	0.276	3825	5.928	30.03	60.54	52.12	2.052	606	14.559	109.00	6.65	1.000	3.630	0.110	20.180	0.240
19.00	0.748	9.50	0.374	4264	6.609	33.47	67.47	67.33	2.651	1040	24.986	155.00	9.45	0.784	4.631	0.140	31.073	0.370
24.00	0.944	9.50	0.374	5271	8.170	41.38	83.42	68.23	2.686	1368	32.866	196.00	11.96	0.905	4.736	0.14.3	33.059	0.394
24.00	0.944	9.50	0.374	5216	8.084	40.95	82.56	68.23	2.686	1368	32.866	196.00	11.96	0.905	4.736	0.143	33.059	0.394
27.50	1.083	10.50	0.143	6297	9.761	49.43	99.65	73.30	2.894	1819	43.701	240.00	14.64	0.838	4.855	0.147	36.799	0.439
29.00	1.142	12.00	0.472	6948	10.770	54.54	109.95	75.00	2.953	2073	49.804	262.00	15.99	0.811	4.083	0.154	38.000	0.453
28.00	1.102	10.00	0.394	6392	9.908	50.18	101.16	76.00	2.992	1940	46.608	253.60	15.48	0.822	5.053	0.153	38.660	0.461
30.20	1.189	12.00	0.472	6855	10.625	53.81	108.48	69.47	2.735	2308	55.450	276.37	16.86	0.778	5.136	0.155	42.892	0.511
30.20	1.190	11.00	0.433	6934	10.748	54.43	109.73	74.97	2.952	2127	51.101	279.19	17.04	0.880	5.129	0.155	39.077	0.466
31.50	1.240	11.50	0.453	7687	11.915	60.34	121.64	80.90	3.185	3055	73.396	335.50	20.47	0.872	5.560	0.168	50.630	0.603
19.45	0.765	6.99	0.275	3825	5.928	30.03	60.54	52.07	2.050	606	14.559	108.50	6.62	1.000	3.613	0.109	20.179	0.240
21.40	0.842	7.30	0.287	4727	7.326	37.11	74.81	58.40	2.299	952	22.872	149.00	9.09	1.000	4.015	0.122	25.653	0.306
22.20	0.874	7.60	0.299	5070	7.857	39.80	80.23	61.20	2.409	1098	29.379	165.00	10.07	1.000	4.145	0.126	27.587	0.329
25.40	1.000	9.10	0.358	5686	8.813	44.64	89.99	64.50	2.539	1610	38.680	205.49	12.54	0.911	4.603	0.139	36.066	0.430
27.00	1.062	9.90	0.390	6414	9.942	50.35	101.50	75.40	2.970	2040	49.011	247.00	15.07	0.895	4.905	0.148	40.516	0.483
28.57	1125	11.10	0.437	7236	11.215	56.80	114.50	75.69	2.980	2730	65.588	295.00	18.00	0.830	5.193	0.157	48.063	0.573
30.16	1.187	11.11	0.437	7633	11.830	65.53	132.10	81.28	3.200	3671	88.196	368.70	22.50	0.842	6.153	0.186	61.265	0.730
30.16	1.187	11.60	0.457	8429	13.064	66.17	133.40	81.28	3.200	3576	85.913	364.70	22.25	0.849	5.511	0.167	54.042	0.644
30.16	1.187	11.11	0.437	8606	13.338	67.56	136.19	85.01	3.347	3949	94.875	391.50	23.89	0.820	5.794	0.175	58.450	0.697
31.50	1.240	11.43	0.450	7743	12.002	60.78	122.53	80.77	3.160	3080	73.997	337.00	20.56	0.884	5.544	0.168	50.674	0.604
16.70	0657	7.60	0.299	3792	5.878	29.77	60.01	55.70	2.193	677	16.265	115.40	7.04	0.958	3.876	0.117	22.471	0.271
23.00	0.906	7.90	0.311	4438	6.878	34.84	70.24	61.30	2.413	912	21.911	146.00	8.91	0.897	4.190	0.127	26.176	0.312
25.00	0.984	8.70	0.343	5071	7.860	39.80	80.23	65.60	2.583	1209	29.046	178.00	10.86	0.880	4.472	0.135	30.376	0.362
20.60	0.811	9.30	0.366	5684	8.810	44.62	89.95	68.00	2.677	1600	38.440	214.00	13.06	0.955	4.796	0.145	35.858	0.428
26.20	1.031	9.10	0.358	5735	8.888	45.02	90.76	70.00	2.756	1558	37.431	214.00	13.06	0.888	4.573	0.144	34.606	0.412
30.16	1.187	11.11	0.437	7183	11.134	56.39	113.68	84.32	3.319	2349	56.435	278.61	17.00	0.880	4.940	0.150	41.656	0.496
27.00	1.062	10.50	0.413	5898	9.142	46.30	93.34	67.20	2.646	1588	38.152	204.00	12.44	0.924	4.406	0.133	34.298	0.409
29.00	1.142	9.00	0.354	6610	10.245	51.89	104.60							0.872				
30.6	1.181	25.0	0.984	5352	8.296	42.01	84.70	38.2	1.504	295.8	7.107	57.10	3.48	1.000	1.359	0.041	7.041	0.084
25.00	0.984	10.00	0.394	5930	9.191	46.55	93.84	70.30	2.767	1605	38.560	224.00	13.67	0.845	4.612	0.146	34.479	0.411
24.00	0.944	8.30	0.327	5791	8.976	45.46	91.64	68.00	2.677	1520	36.518	209.00	12.75	0.894	4.597	0.141	33.436	0.398
31.50	1.240	11.50	0.453	7687	11.915	60.34	121.64	80.90	3.185	3055	73.396	335.50	20.47	0.872	5.560	0.168	50.630	0.603
25.00	0.984	9.40	0.370	5880	9.114	46.16	93.05	69.47	2.735	1631	39.185	216.30	13.19	0.862	4.686	0.142	35.334	0.421
27.00	1.062	10.50	0.413	5898	9.142	46.30	93.34	67.20	2.646	1588	38.152	204.00	12.44	0.824	4.405	0.133	34.298	0.409
25.00	0.984	11.00	0.433	6114	9.476	48.00	96.76	72.20	2.843	1822	43.773	234.00	14.27	0.846	4.875	0.147	37.958	0.452
27.00	1.062	11.00	0.433	5688	8.816	44.65	90.01	68.80	2.709	1489	35.773	208.30	12.71	0.814	4.665	0.141	33.348	0.397
27.00	1.062	10.50	0.413	6562	10.170	51.51	103.80	70.90	2.791	2037	48.939	251.30	15.37	0.868	4.879	0.148	39.546	0.471

10.3 Rail properties

10.3.1 Metallurgical fundamentals

In pearlitic steels, the structure of which is depicted in Figure 10.21, the mechanical properties are largely governed by the distance between the cementite (Fe_3C) lamellae, their thickness and by the grain size. The influence of the interlamellar spacing on yield point, tensile strength and reduction of area is, according to [112], clearly demonstrated by the structures shown in Figure 10.22. These examples display microstructures of the same steel subjected to different cooling rates, which lead to different rates of diffusion of carbon at the austenite pearlite growth front. The same effect can be achieved by controlling the diffusion rate via the alloying contents.

Yield point and tensile strength increase as the distance between the lamellae decreases. The quantitative interrelationships are shown in Figure 10.23.

As tensile strength increases the yield point increases even more, as shown in Figure 10.24, due to a greater influence of the distance between cementite lamellae on the yield point. This means an improvement in the elastic ratio.

Toughness is primarily governed by the thickness of the cementite lamellae and the grain size. Figure 10.25 shows the influence on the reduction in area which is improved as the thickness of the cementite lamellae and the grain size decrease. A similar interrelationship is obtained for the results from impact tests.

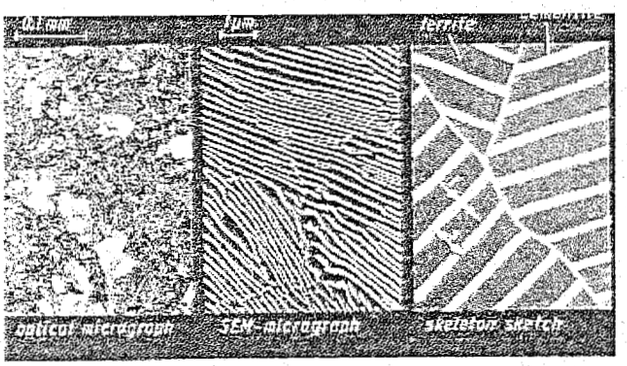


Figure 10.21: Microstructure of pearlite

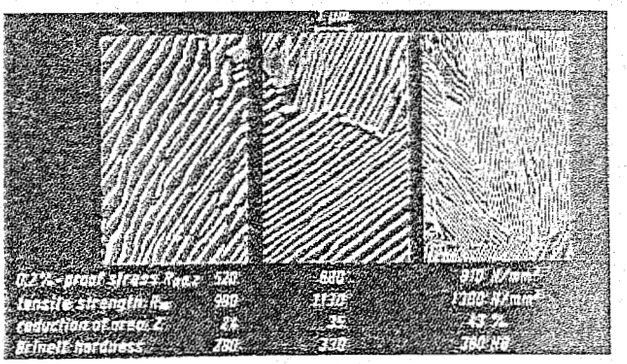


Figure 10.22: Scanning electron micrographs of pearlite

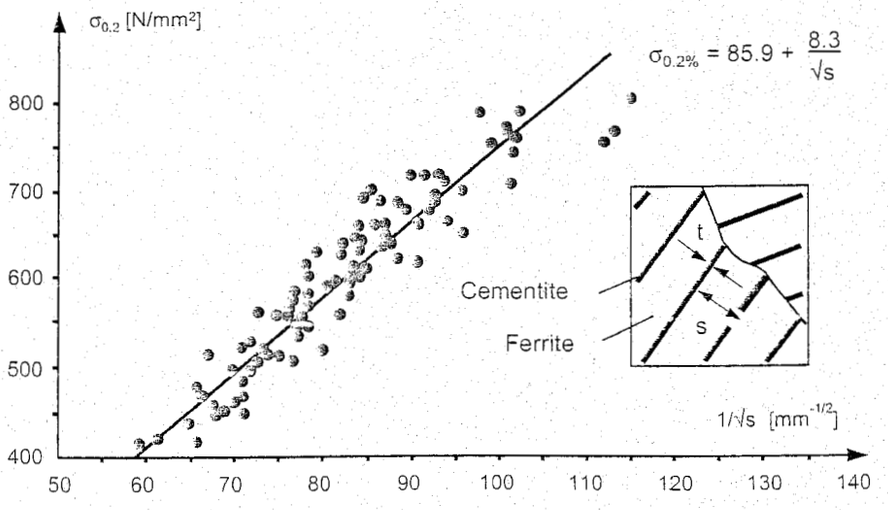


Figure 10.23: Relation between 0.2% proof stress and interlamellar spacing

10.3.2 Heat treatment

In addition to alloyed premium rails, tensile strength can be improved by heat treatment. Heat treatment is mainly applied to rails used in heavy duty tracks which have sharp curves and to turnouts. In order to prolong the life-span of fishplated rail joints, end-hardening of the rail head is also carried out in separate plants. Compared to alloying, heat treatment has the big advantage that toughness is also improved. The method can be applied to either the whole section or to just the rail head. In the latter case, a tensile strength of 900 N/mm² is retained in the rail web and base, while the rail head has a tensile strength between 1200 - 1350 N/mm² (350-400 BHN) which provides a very large wear resistance.

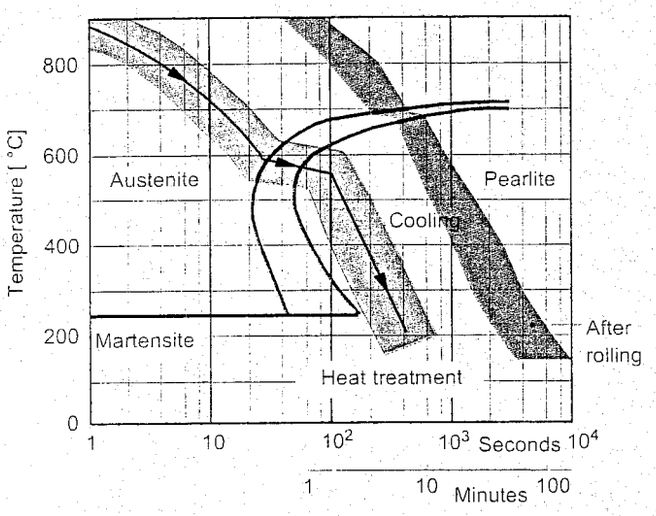


Figure 10.26: TTT-diagram for grade R 260 (900 A) with cooling curve for heat treatment

The rail can be heat treated either after cooling and straightening outside the rail production line (off-line) or when it comes directly from the rolling mill (in-line). The TTT-diagram with the cooling curve is presented in Figure 10.26. The closer the cooling curve approaches the right side of the nose of the TTT-curve the finer the pearlite structure is and the better the properties are. In any case, the material structure shall be fine-pearlitic, and no martensite or bainite shall occur. Additionally, the microstructure and hardness shall change gradually from the head to the web.

Off-Line Hardening

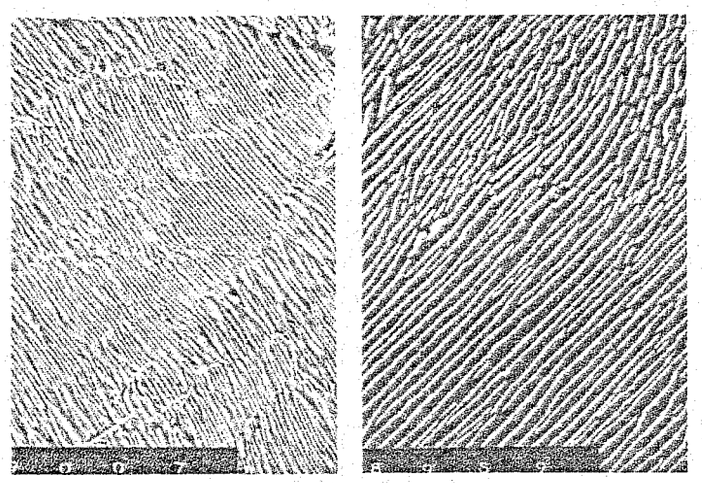


Figure 10.27: Microstructure resulting from heat treatment

The cold finished rail is taken to the hardening plant and the head or the total cross section is heated by induction in 2 - 6 minutes to the austenitising temperature of 850 - 950 °C. This is followed by accelerated cooling by compressed air, water spray or water mist down to 650 - 500 °C, this temperature is maintained until a fully pearlitic structure is attained according to Figure 10.27 [250].

Additionally to the refinement in the pearlite interlamellar spacing, the pearlite block size becomes smaller. This process is continuous, but, because the rail was reheated and quenched, it must be roller straightened again in order to achieve the straightness tolerances. The microstructure resulting from the Thyssen head-hardening facility is illustrated in Figure 10.28. The hardness distribution is presented in Figure 10.29.The cold rail moves continuously through the hardening plant. The time consumed for reheating limits the production capacity to about 10.000 tons per year.

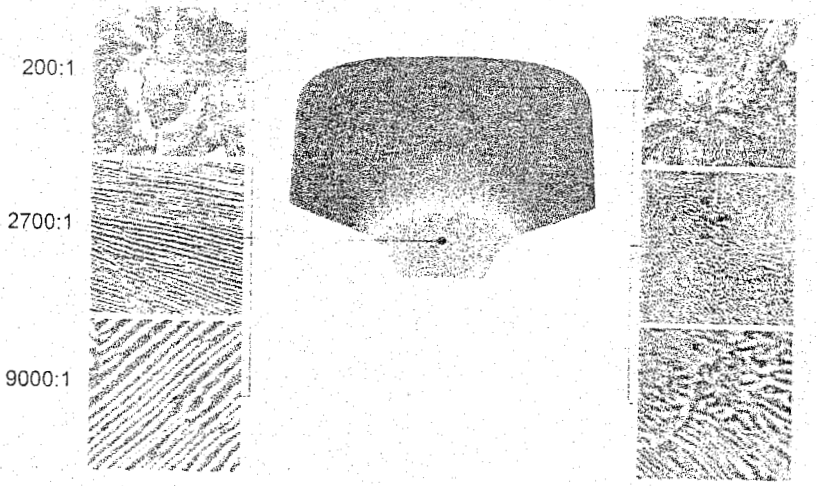


Figure 10.28: Microstructure resulting from head-hardening (Thyssen)

In-Line Hardening

The rail is taken directly from the rolling mill and brought to the hardening plant at a temperature beyond 800 °C. The quenching may proceed continuously similar to the off-line plants, or the rail is quenched in its entire length at once.

Such a process was developed by Voest-Alpine, which cooled the rail head in a hardening bath (Figure 10.30). The rail is turned upside down and dipped into the quenching medium leading to a hardness increase in the entire rail head (Figure 10.31). Rails up to a finished length of 120 m can be heat treated in this plant which has a capacity of 260,000 tonnes annually. When the entire rail has cooled down to 60 °C, it is roller straightened and finished like other rails.

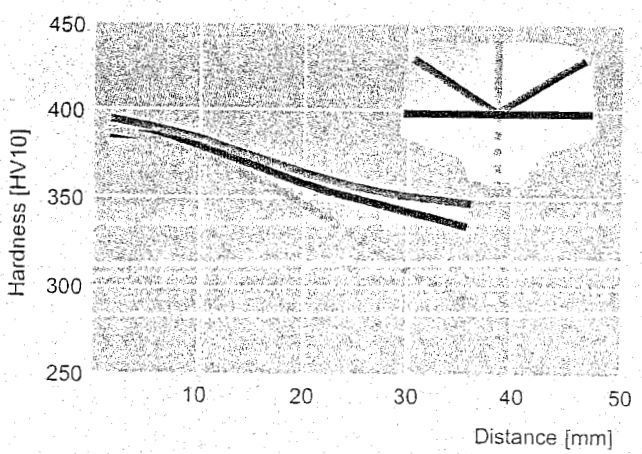


Figure 10.29: Hardness distribution (Thyssen)

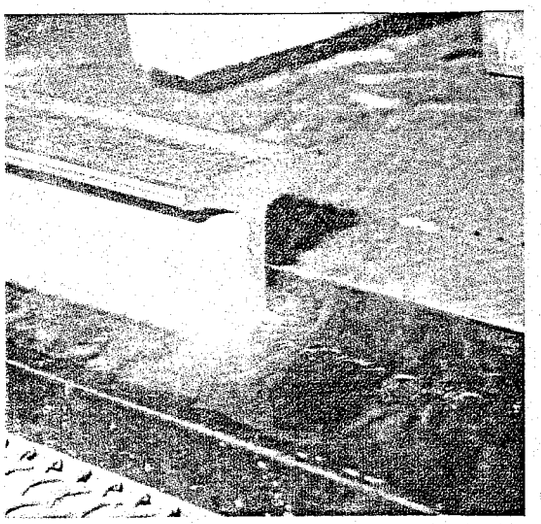


Figure 10.30: Cooling of the rail head

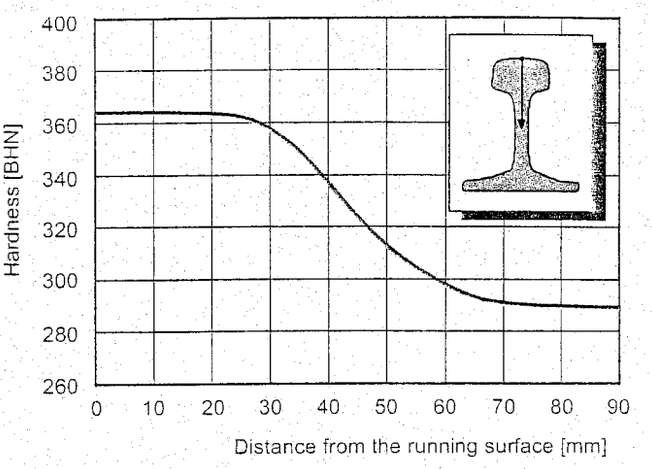


Figure 10.31: Hardness distribution after head-hardening

10.3.3 Rail grades

In Europe normally rails of the UIC standard grades according to UIC leaflet 860-0 were employed. In the year 2001 a new European standard for rails was introduced, referred to as EN 13674-1, which would replace the UIC leaflet. In the new standard the rail grades are no longer indicated by the tensile strength but by the Brinell hardness BHN. For the relationship between tensile strength and Brinell Hardness see (10.4) and (10.5).

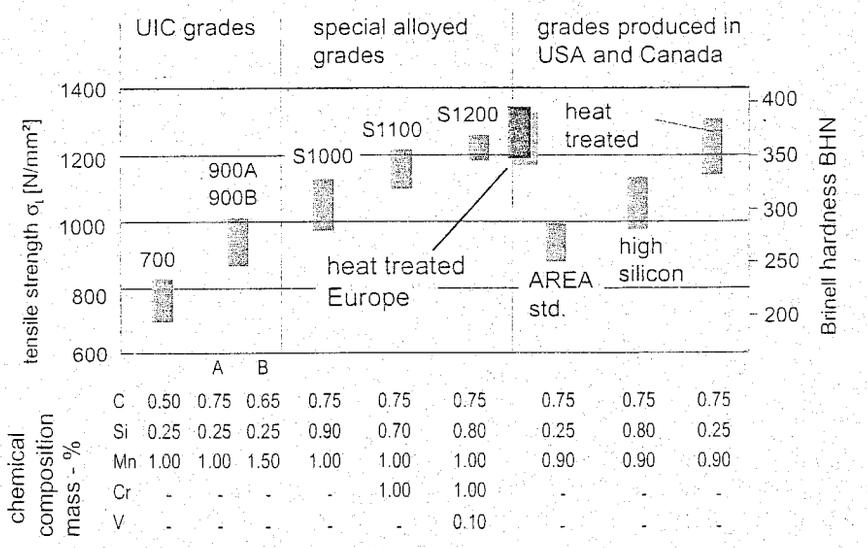


Figure 10.32: Rail grades used in Europe and America

Figure 10.32 presents a compilation of the modern rail grades presently used in Europe and America. The specifications according to UIC, BS and AREA standards are listed in Table 10.3. The standard grades could be designated as R 200 (700), R 260 (900 A), R 260 Mn (900 B), and R 320 Cr (1100), represented by the Brinell hardness. For completeness the minimum tensile strength is indicated between brackets, varying from 700-1100 N/mm². Their characteristic elements are carbon and manganese as well as chromium for the high-strength grade R 320 Cr. The basic mechanical properties are related to the alloying elements according to formulae obtained from regression analyses. Examples are described e.g. in [111]:

- tensile stress σ_u [N/mm^2]:

$$\sigma_u = 227 + 803\%C + 87\%Si + 115\%Mn + 133\%Cr + 891\%P + 614\%V \pm 19 \quad (10.1)$$

- yield stress σ_y [N/mm^2]:

$$\sigma_y = 101 + 469\%C + 36\%Si + 85\%Mn + 116\%Cr + 0\%P + 634\%V \pm 21 \quad (10.2)$$

- elongation after fracture [%]:

$$\delta_5 = 30.8 - 22.6\%C - 1.7\%Si + 0\%Mn - 2.3\%Cr + 0\%P + 4.4\%V \pm 0.9 \quad (10.3)$$

Rail steel grades	Chemical composition %							Mechanical properties		
	C	Si	Mn	P_{max}	S_{max}	Cr	V	R_m N/mm²	A_5 %	BHN
UIC 700	0.40–0.60	0.05–0.35	0.80–1.25	0.05	0.05			680–830	14	
UIC 260 (900 A)	0.60–0.80	0.10–0.50	0.80–1.30	0.04	0.04			880–1030	10	
UIC 260 Mn (900 B)	0.55–0.75	0.10–0.50	1.30–1.70	0.04	0.04			880–1030	10	
UIC 1100	0.60–0.82	0.30–0.90	0.80–1.30	0.03	0.030	0.80–1.30		> 1080	9	
R 200	0.40–0.60	0.15–0.58	0.70–1.20	0.035	0.035			> 680	14	200–240
R 220	0.50–0.60	0.20–0.60	1.00–1.25	0.025	0.025			> 770	12	220–260
R 260	0.62–0.80	0.15–0.58	0.70–1.20	0.025	0.025			> 880	10	260–300
R 260 Mn	0.55–0.75	0.15–0.60	1.30–1.70	0.025	0.025			> 880	10	260–300
R 320 Cr	0.60–0.80	0.50–1.10	0.80–1.20	0.020	0.025	0.80–1.20	0.18	> 1080	9	320–360
R 350 HT	0.72–0.80	0.15–0.58	0.70–1.20	0.020	0.025			> 1175	9	350–390
R 350 LHT	0.72–0.80	0.15–0.58	0.70–1.20	0.020	0.025	< 0.30		> 1175	9	350–390

Table 10.3: Tensile strength and chemical composition of rails

Compared with grade R 200 (700) the R 260 (900 A) grade obtains its increased strength mainly from carbon, and R 260 Mn (900 B) mainly from manganese. This results in better welding properties for R 260 (900 A), but in greater toughness for R 260 Mn (900 B). The R 260 (900 A) grade, which is more or less the standard grade in Europe, is very similar to the AREA standard grade.

For heavy duty tracks special as-rolled steel grades were developed with minimum tensile strength values between 1100 and 1200 N/mm^2. This is achieved by increasing the silicon content, adding chromium and sometimes vanadium as specified in Figure 10.32. As mentioned before heat-treated rails have a larger tensile strength and thus a great wear resistance, while at the same time they are relatively tough. Therefore, they replace the alloyed steels in heavy haul applications and in heavily loaded curved tracks like in Switzerland and Austria.

The UIC grades are found in the new European Standard EN 13674 with minor changes in the chemical composition (Table 10.3). Especially phosphor and sulphur are limited to lower values because of their influence on the wear properties and inclusion formation. The rail grades are named by the hardness because this is directly related to the wear properties. EN 13674 also considers two heat treated grades with different chromium-content. Grade R 350 HT (HT for heat treated rail steel) has basically the same composition as grade R 260, the former UIC grade 900 A. R 350 LHT has a chromium-content up to 0.30% (LHT low alloy heat treated). It can be flash-butt welded without the need for a post-welding heat treatment.

Since the pearlitic structure is limited to about 400 BHN hardness, corresponding to 1350 N/mm², harder rail steels need to have another type of microstructure. Tempered Martensite with a hardness exceeding 400 BHN was tested, but showed a comparatively low wear resistance. The bainitic microstructure offers a potential to combine high strength and hardness with good toughness properties. Bainitic rail steels were developed as a material with a higher resistance against rolling contact fatigue defects rather than to have better wear properties than pearlitic rails. They perform best e.g.

as balls in roller bearings, where both wear as well as rolling contact fatigue is the concern.

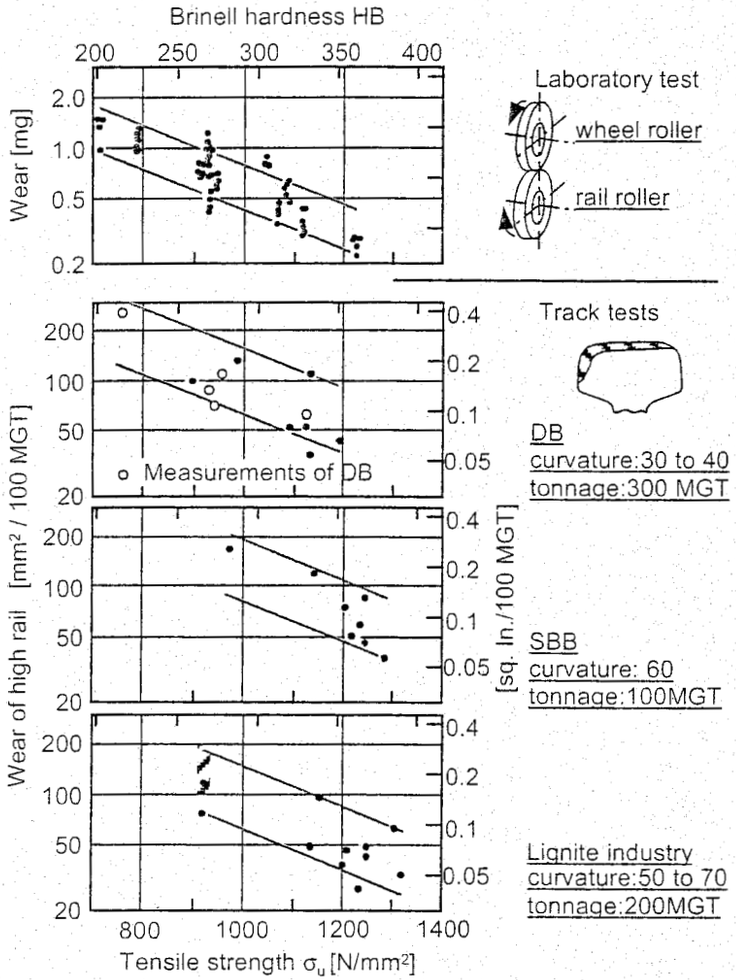


Figure 10.33: Wear of pearlitic rail steels

Generally, the chemical composition of bainitic steels shows a wide variation in different alloying elements, e.g. C = 0.05 - 0.80 % and Cr = 0.01 - 3.0 %. Depending on the chemical composition and the heat treatment, different microstructures occur. Steels with a low carbon content (C = 0.10 %) exhibit a carbide-free microstructure of granular bainite. If manganese is added, also an acicular structure with highly-dispersed fine carbide precipitations is formed. With a higher carbon content (C = 0.25 %), an acicular structure of upper or lower bainite with carbide precipitations appears.

Currently, bainitic rails are not industrially produced because their manufacturing technology is more complex compared to head-hardened rail [20], [49], [111], [121], [130], [180], and [278]. The rail must be either quenched and kept on an elevated temperature (isothermal transformation) or annealed to avoid untempered martensite which is a brittle component of the microstructure. The rail must also be annealed for internal stress relief.

Bainitic rails have been tested in track since 1980 and have shown different behaviour depending on their properties. The rails performed better regarding both wear and fatigue defects e.g. at the Ofotbahn in Norway, but not as good as R 260 (900A) or head hardened rail steels e.g. at tests in the USA. The metallurgy of bainitic rail is still a topic of development with the goal to achieve a high resistance to rolling contact fatigue together with a wear resistance at least comparable to head hardened rail steels.

Austenitic steels are used for switch frogs because they work harden under the impact load of the wheel. They are not used for regular rail steel because the resistance to wear (without work hardening during operation) is lower than that of pearlitic steels.

10.3.4 Wear resistance

Resistance to wear is governed by the content of carbon and a few other alloying elements as well as the interlamellar spacing. Under normal conditions this means that wear resistance increases with hardness which is proportional to the tensile strength. Outside Europe, and since 2001 also inside Europe, hardness is used to specify the rail grade or tensile strength. The most important hardness scales are those of Vickers (HV) and Brinell (BHN). The relationship with the tensile stress σ_u equals:

$$HV = 0.3\sigma_u \tag{10.4}$$

$$HV \approx 15 + BHN \tag{10.5}$$

The top of Figure 10.33 shows the results of laboratory wear tests, and the bottom of the figure shows the results of many years of trials under different service conditions [113]. The first graph presents the results of roller tests and reveals that abrasive wear decreases by almost 50% when the tensile strength increases by 200 N/mm^2.

This trend has been confirmed by service tests. The loss of rail head area in mm^2 of high rails is given with respect to a loading of 100 million gross tonnes on curved tracks on DB and SBB, with axle loads of 200 to 220 kN, and on tracks belonging to the German lignite surface-mining industry, with axle loads of 340 kN. In all cases the wheel flanges were lubricated. As expected, the results under actual service conditions scatter more widely than those determined in laboratory tests because of local track conditions. The graphs show, however, that, even in small-radius curves and under high axle loads rail wear can be kept low by using high-tensile rails so that loadings in excess of 500 million gross tonnes can be reached. The use of high-tensile steels has the additional advantage of increasing the resistance of the rail head to plastic deformation, for instance, such as in the case of shelling.

The influence of local conditions is clearly demonstrated by lateral rail head wear measurements carried out by the NS. Figure 10.34 monitors lateral wear versus tonnage borne for a number of curves carrying the same traffic. The plots clearly reveal a large variation in wear rate under almost identical conditions. However, the wear rate in each curve is very consistent which enables wear predictions to be made by extrapolating the data.

The NS experience is that the lateral wear of grade R 260 Mn (900 B) is three times less than the lateral wear of the R 200 (700) grade. Grade R 320 Cr (1100) wears two to three times slower than grade R 260 Mn (900 B). In the case of lubrication R 260 Mn (900 B) is equivalent to grade R 320 Cr (1100). Lubrication of grade 1100 only produces minor improvements.

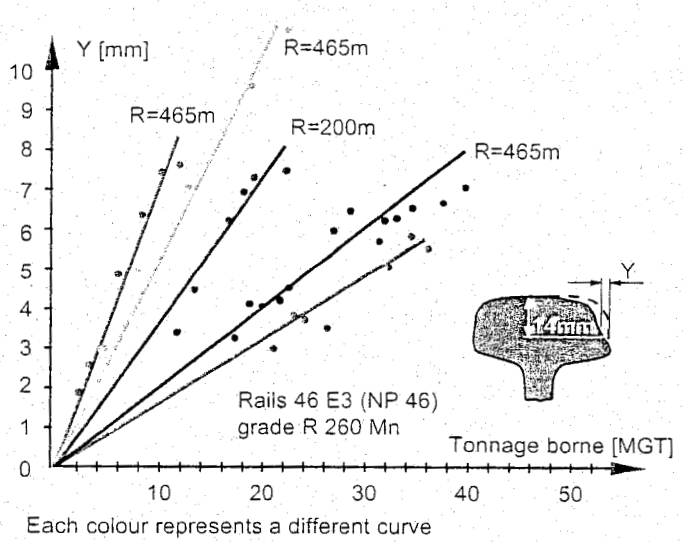


Figure 10.34: Test results for lateral rail wear on NS

The head hardened steels exhibit a wear resistance of at least three times larger than grade 260 (900 A) and are superior to grade R 320 Cr (1100) by a factor of about two.

Vertical rail head wear is approximately proportional to the accumulated tonnage. If wheel flanges are lubricated, the vertical wear for rails, with a tensile strength of 900 N/mm^2, amounts to about 1 mm per 100 MGT. Without lubrication this value is of the order of 1 to 2 mm per 100 MGT.

10.3.5 Fatigue strength

Resistance to fatigue is often expressed as fatigue strength. A rail or weld should be able to resist at least 2.10^6 stress repetitions between 0 and the fatigue stress without developing any cracks.

Fatigue strength is of outstanding importance for resistance to excessive loading. If this value is too low cracks may develop and phenomena such as shelling may occur. Figure 10.35 presents results of bending tests, which show that the fatigue strength of the as-rolled rail increases linearly with the tensile strength. During service, however, it drops again due to corrosion.

Cracks always start from notches which may originate from exterior damage or from internal imperfections. For this reason a high level of cleanness, i.e. low content of non-metallic inclusions and a low hydrogen content, are important. EN 13674 specifies limits for acceptable surface defects in order to increase the reliability of rail.

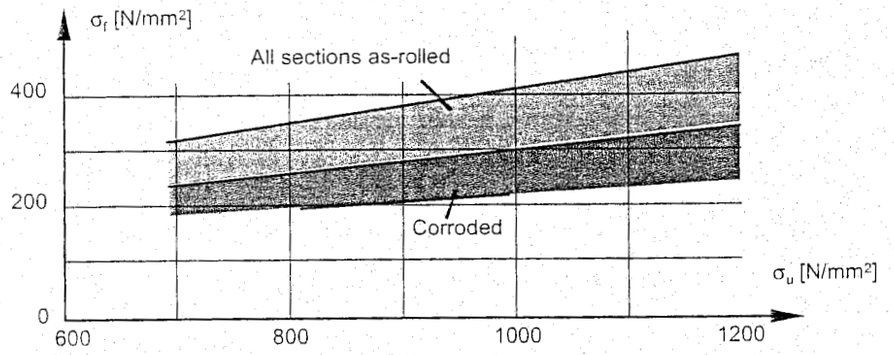


Figure 10.35: Bending fatigue strength of rails

The following indicative values for fatigue strength can be given:

As-rolled rail	: 300 ± 20 N/mm²	
Used rail	: 220 ± 20 N/mm²	(refer to Table 5.3)
Flash butt weld	: 200 ± 20 N/mm²	
Thermit weld	: 180 ± 20 N/mm²	

These values are obtained from four-point bending tests with a rail supporting span of 1.25 m and an applied minimum stress of 20 N/mm². Compared to literature values, these figures are relatively low. This might be due to shear deformation which plays a role at a height/length ratio greater than 0.1.

10.3.6 Fracture mechanics

Safety against fracture of rails is of decisive importance for railway traffic. Fracture mechanics can be used to assess this property. The basic theory of fracture mechanics is treated in [26], [5], and [6]. In fact there are two major aspects, namely fracture toughness, normally expressed as the critical stress intensity factor K_{IC}, and the rate of crack propagation, normally referred to as da/dN.

The stress intensity factor is defined as:

$$K = C_S \sigma \sqrt{a} \qquad (10.6)$$

in which:

σ = stress at half the crack depth;
a = crack depth including the initial notch depth;
C_s = shape factor (order 1.4 - 1.9).
For determination of the shape factor C_s please refer to [216].

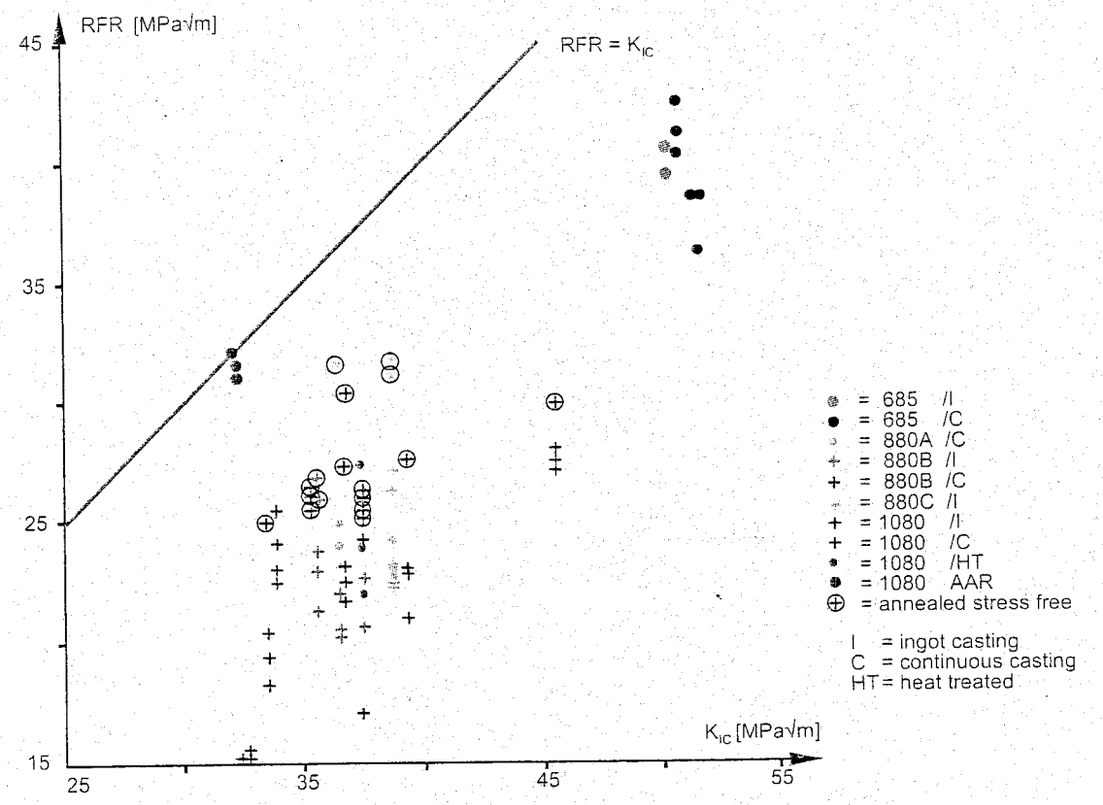


Figure 10.38: RFR-values versus K_{IC}

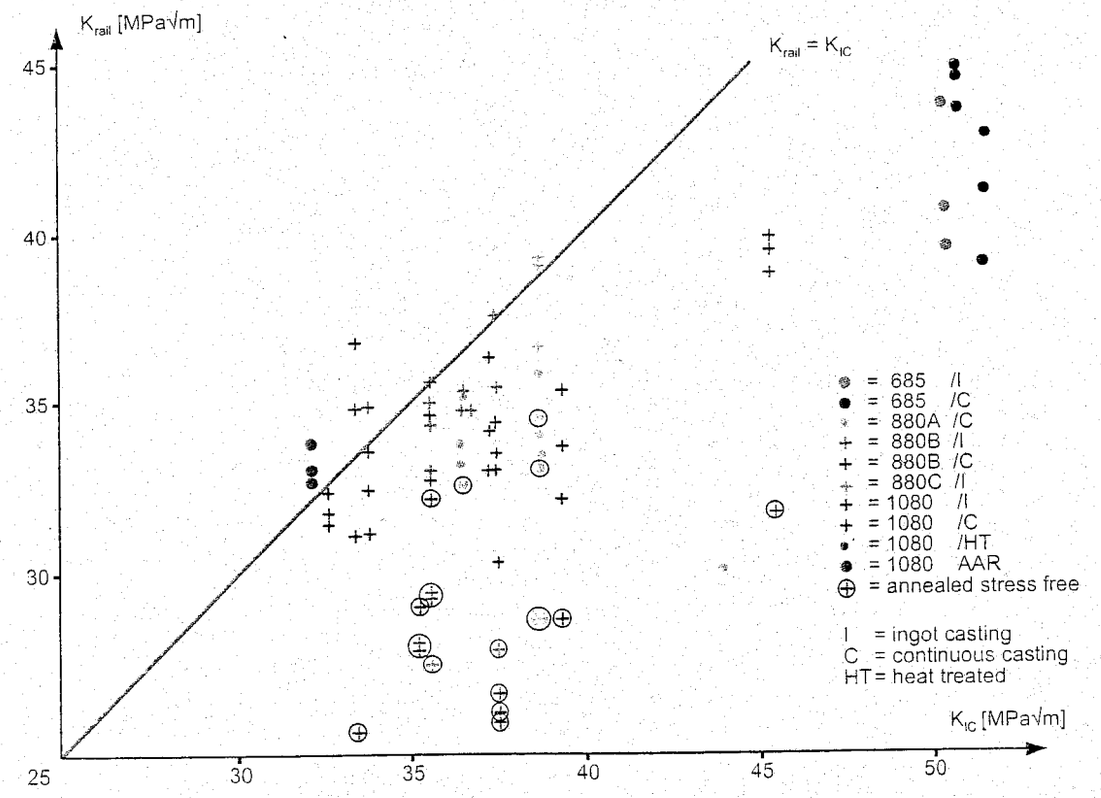


Figure 10.39: K_{rail}-values versus K_{IC}

When considering the stress intensity factor (SIF) for the full rail section the residual stresses should be taken into account. This stress was measured separately in the tests. The following values were determined:

K_{IC} = SIF for specimen from rail head;
RFR = Rail Fracture Resistance, referring to the applied stress without considering the residual stresses;
K_{rail} = SIF rail for the actual stress, i.e. applied stress plus residual stress.

The NS tests were carried out on grades 700, 900 and 1100 N/mm². The 300 HB AREA profile was stress-relieved by stretch straightening [44].

Figure 10.38 represents RFR-values versus K_{IC} for different rail grades. The K_{rail}-values versus K_{IC} are plotted in Figure 10.39.

Figure 10.40 summarises the NS results [21]. The residual stresses appear to have a dominating influence on rail fracture resistance. From this point of view it would be advantageous to improve the existing roller-straightening process in order to confine the residual stresses. For tensile strengths of 880 N/mm² and higher the K_{IC}-values hardly decrease at all.

Rail grade	K_{1c} (MPa $m^{-1/2}$)	K_{rail} (MPa $m^{-1/2}$)	RFR (MPa $m^{-1/2}$)
UIC R 200 (700)	50 - 52	39 - 47	36 - 44
UIC 260 (900 A)	36 - 39	33 - 36	24 - 32
UIC 260 Mn (900 B)	34 - 46	32 - 40	21 - 30
UIC 1100	33 - 39	28 - 37	15 - 28
R 300 HB stretch straightened	32	33	32

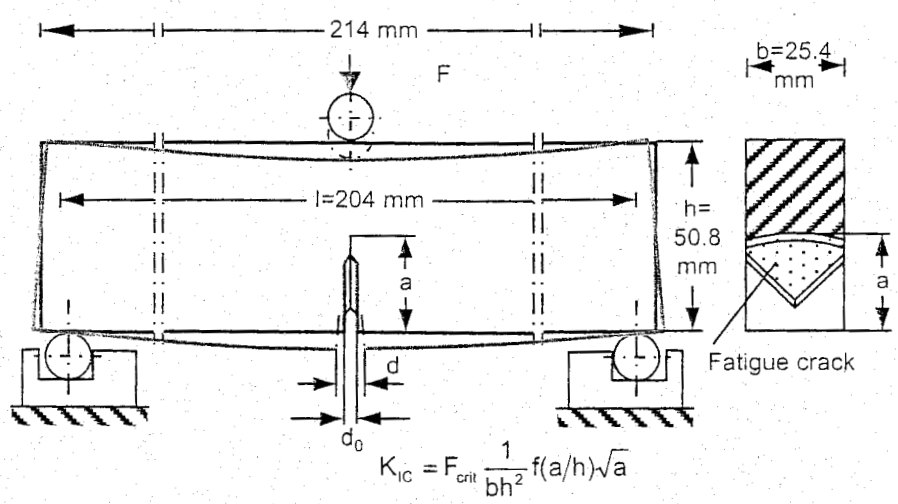


Figure 10.40: Summary of NS rail fracture test results

The da/dN values have also been determined on rail head specimens like those shown in Figure 10.41. These results are summarized in Figure 10.42. The crack propagation rate for 1080 N/mm² grades is about twice as high as for the 880 N/mm² grade. Similar values have been found in other investigations, for instance as reported in [11], [216] and [217].

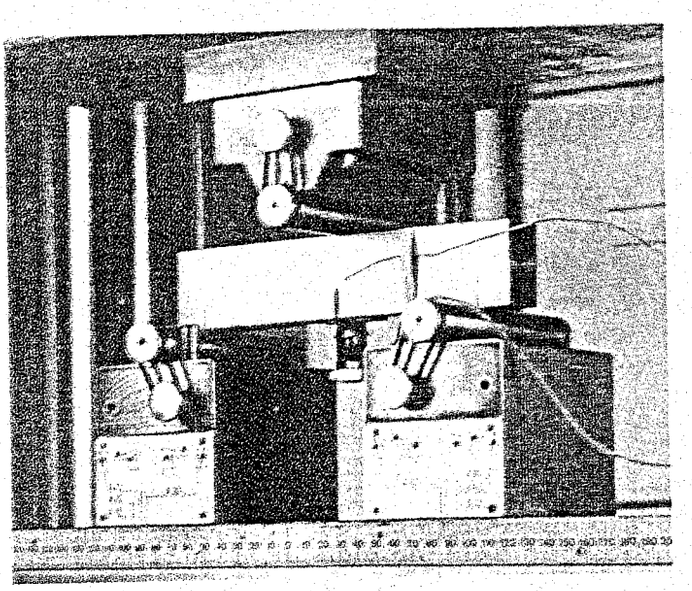
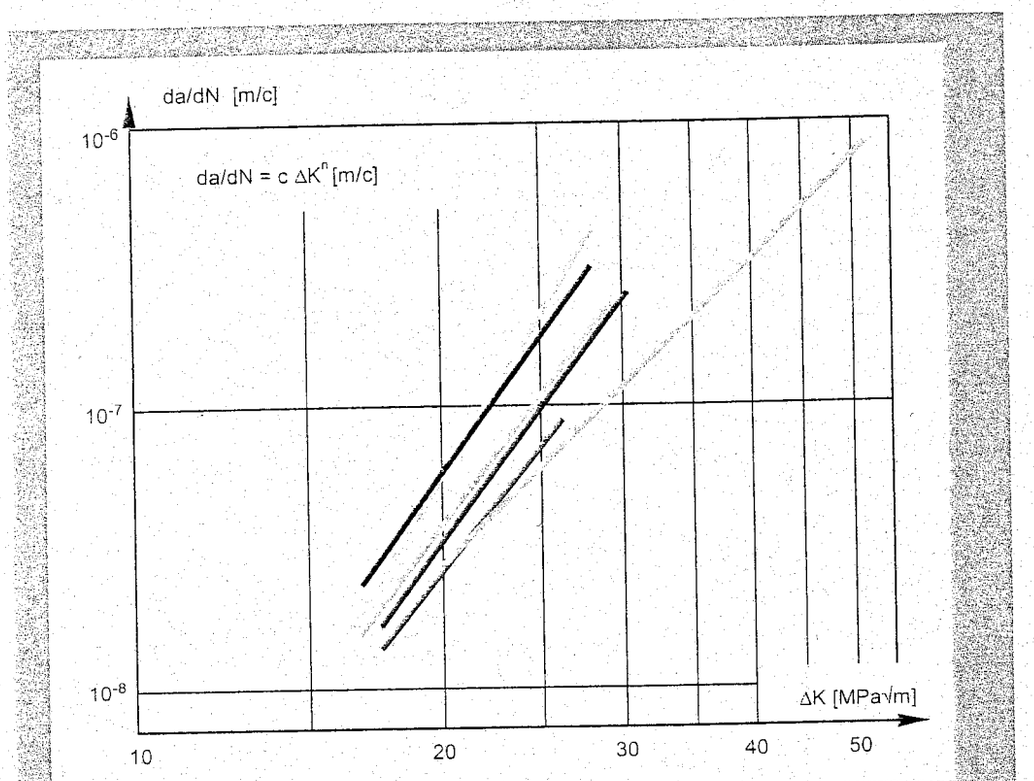


Figure 10.41: Fracture mechanics test arrangement



	Rail grade	Factor c	Exponent n
	1080 N/mm²	$1.9\ 10^{-15}$	5.7
	300 HB	$3.1\ 10^{-14}$	4.8
	880 N/mm² - A	$5.0\ 10^{-14}$	4.5
	880 N/mm² - B	$2.3\ 10^{-14}$	4.7
	880 N/mm² - C	$4.6\ 10^{-14}$	4.4
	685 - 835 N/mm²	$7.2\ 10^{-13}$	3.5

Figure 10.42: Crack propagation rate versus stress intensity factor

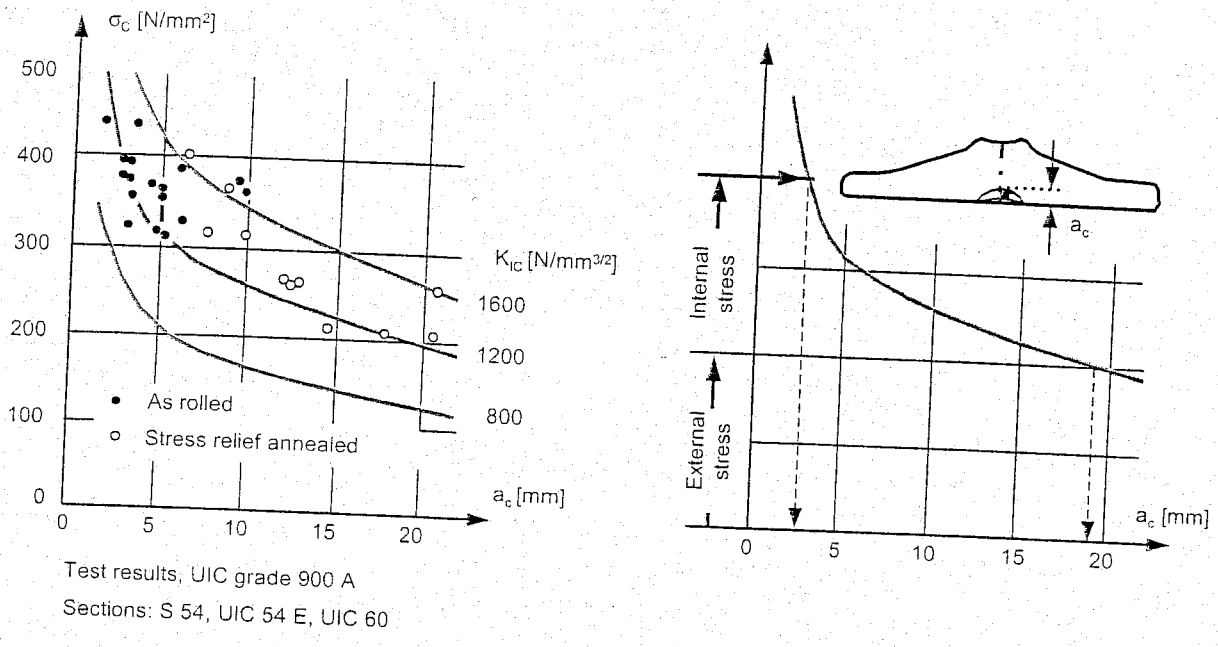


Figure 10.43: Critical stress values versus critical crack depth for UIC R 260 (900 A)

According to tests carried out by Krupp [113] on rail grade R 260 (900 A), K_{IC}-values vary between 1000 - 1600 N/mm$^{3/2}$, or 30-50 MPam$^{-½}$. Critical stress values determined for full rail sections are presented in Figure 10.43 as a function of the critical crack depth. In this example too the need for reducing the residual stresses is demonstrated. An external stress of 200 N/mm² and a residual stress of a further 200 N/ mm² will result in a critical crack depth of just 3 mm. If, however, the residual stress is reduced, as in this example by stress-relief annealing, an external stress of 200 N/mm² leads to a critical crack depth of about 20 mm. In this case the rail is more resistant to fracture because small incipient cracks, which can be more frequent, will still not lead to failure. Larger incipient cracks are easier to detect by ultrasonic inspection.

In the new standard EN 13674-1 minimum K_{IC} values according to Table 10.4 are specified, for which tests should be undertaken at a regular time interval by the manufacturer. For the grade R 260 the draft requires 26 MPam$^{-½}$ as minimum and 29 MPam$^{-½}$ as a mean value. In practice normally higher values are achieved with a mean value around 35 MPam$^{-½}$ for the grade R 260 as well as for the heat treated rails 350 HT and 350 LHT.

Steel grade	Minimum single value K_{1c} (MPa m$^{-1/2}$)	Minimum Mean K_{1c} (MPa m$^{-1/2}$)
R 200 and R 220	30	35
R 260 and R 260 Mn	26	29
R 320 Cr	24	26
R 350 HT	30	32
R 350 LHT	26	29

Table 10.4: Minimum single and minimum mean values of K_{1c}.

10.4 Rail welding

10.4.1 Introduction

Factors like increasing train frequency, higher axle loads, and higher train speeds have resulted in an increase of track loads and the associated stresses. Fishplated rail joints exhibit considerable impact forces when the wheel passes the gap of the joint. Continuous welded track avoids such weak points, leading to a higher reliability and safety of the track. Although the performance of the welds improved with further developments of the welding process, defects in welds are still a large cost factor for the railways. The concept of just-in-time delivery of unwelded rails of up to 120 m length to the building site reduces the number of welds in the track, and thus the failure events.

10.4.2 Flash butt welding

The flash butt welding, or electrical resistance welding principle, characterised by current, travel, and pressure is shown in Figure 10.44. After evenly flashing of both rail end sections preheating commences. Here the energy input should be high and the flashing time short. The heat input should be uniform throughout the cross-section. At the end of the preheating phase the temperature right at the joint should drop abruptly from a high level in towards the rail.

The subsequent flashing should, if possible be done progressively at an increasing travel rate and with growing current. In this way metal vapour is produced which protects the abutting surface against oxidation, thus preventing the formation of non-metallic inclusions.

To obtain a high quality weld the flashing operation should smooth the contact surfaces without forming craters. Surface and neighbouring rail area need to be at melting temperature to obtain sufficient travel during upsetting to prevent lenses, micro-porosity and non-metallic inclusions in the welding seam.

Most of the premium rails require post-welding treatments. The welding machine must be programmed according to the transformation behaviour of the rail steel.

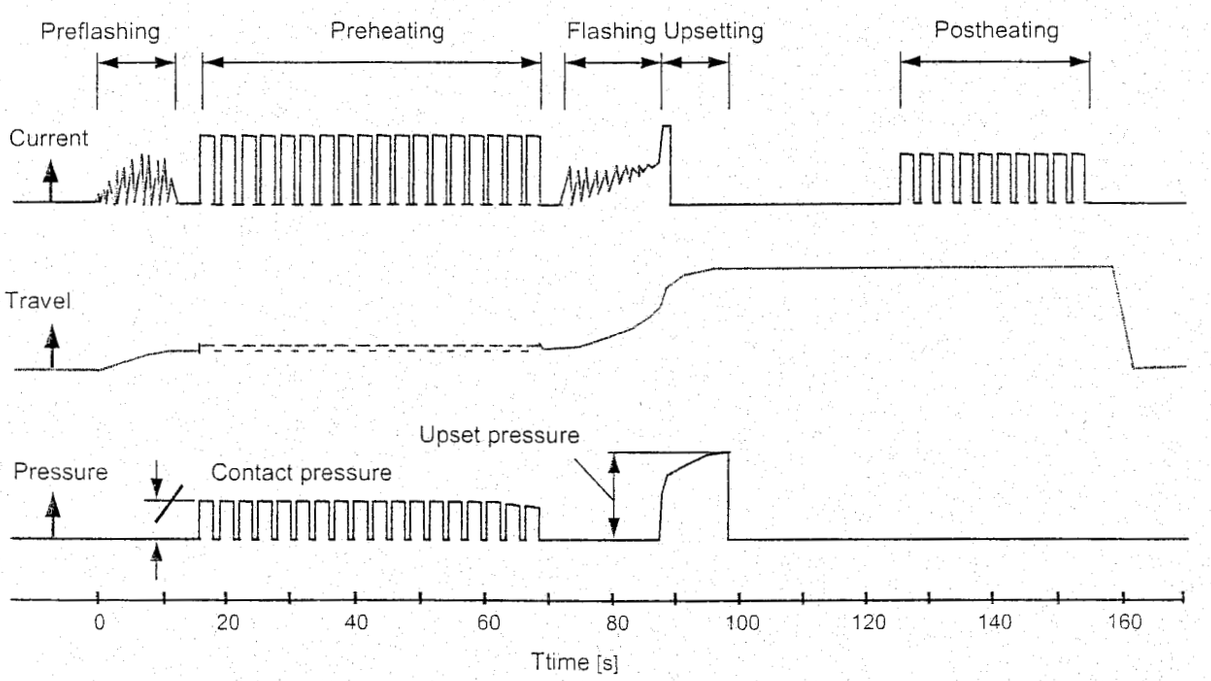


Figure 10.44: Flash butt welding principle

Until the early seventies most of the flash butt welding machines were working with alternating current (AC). In the new generation welding machines DC units are also applied. The welding process that uses DC runs very smoothly and regularly. The formation of deep craters during flashing is avoided and the weld spatter remains small in size. DC only causes minimal inductive losses in the secondary circuit and, therefore, the rectifier units can be situated away from the weld area, well-protected against flashing spatters. Moreover, this remote arrangement allows full space to install aligning equipment close to the weld to control the geometrical straightness.

Figure 10.45 shows a picture of the Schlatter rail welding plant used by KHRC for the Korean High Speed Line from Seoul to Pusan. This machine is equipped with automatic rail end alignment. A pre-set vertical overlift can also be given. Immediately upon completion of the welding process, and with the rail still clamped in welding position, the burr is sheared off all around the rail section.

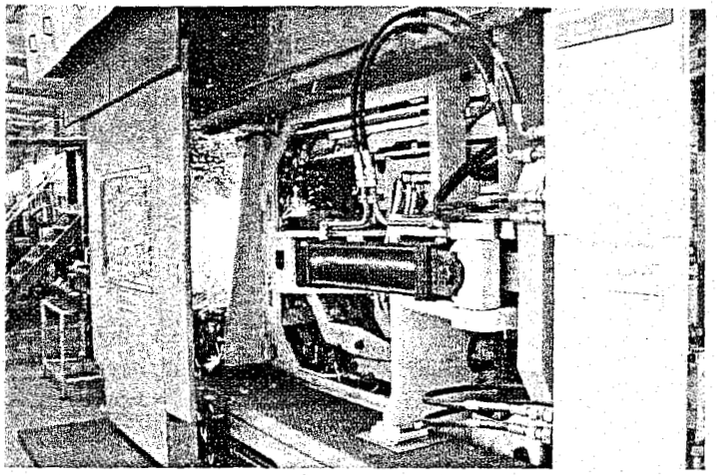


Figure 10.45: Schlatter' rail welding plant used at KHRC

Due to the symmetrical arrangement of the rectifier units with respect to the rail, a uniform heat-affected zone is built up during preheating and flashing so that the hot material is forged equally over the whole rail section.

A longitudinal cut, taken after welding of head-hardened rail 132 lbs/yd, is presented in Figure 10.46.

Mobile flash butt welding

Mobile flash butt welding was originally developed in Ukraine by the Paton institute and introduced on western railways by Plasser & Theurer in 1973. Today it is a well proven technology on high speed tracks as well as on US heavy haul railroads. Figure 10.47 shows a series of photograph of the Plasser & Theurer APT 500. Its predecessor model K 355 APT was used during the building of the new NS Flevo-line. The metallurgical weld quality, as represented by the fatigue strength, produced by the mobile machine is equivalent to that produced by the stationary machine. Although the machine has an integrated system to align the rail ends (Figure 10.47.2), site conditions demand special care of the operators in this respect.

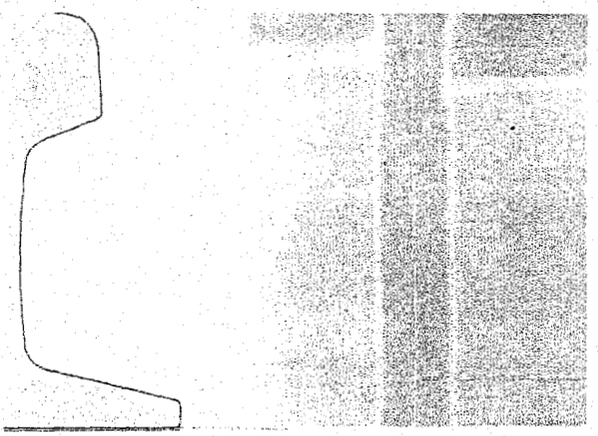
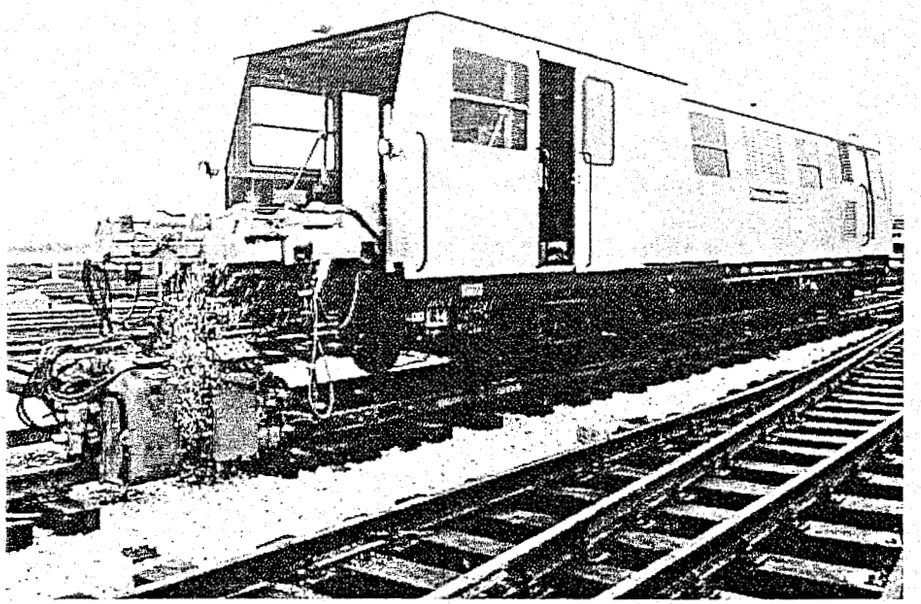


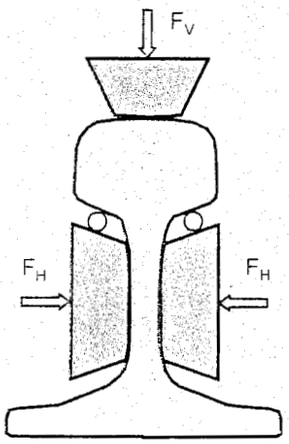
Figure 10.46: Longitudinal cut after flash butt welding of a head-hardened rail 132 lbs/yd

On the NS flash butt welds should have a fatigue strength of at least 180 N/mm². Total deflection on a 1 m base due to static bending should be at least 30 mm for the 700 grade and 20 mm for the R 260 (900) grade and the 1100 grade, without any crack development.

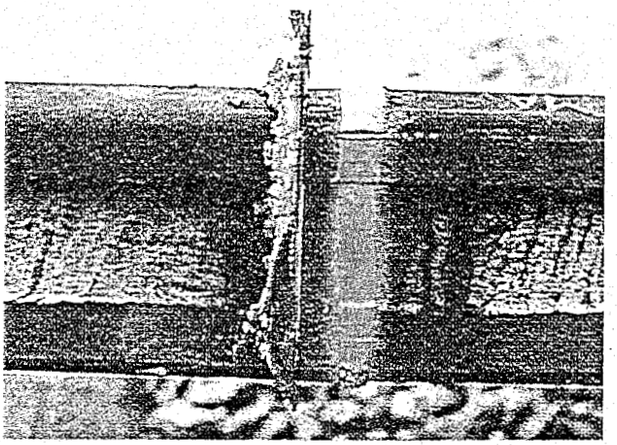
In 1989, the Super Stretch welder (Figure 10.47.4 and Figure 10.47.5) was introduced. This machine has an on board hydraulic rail puller with 175 tons towing capacity and enables continuous rail welding below neutral temperature, closure welds, plug welds and others. The most important feature of this machine is that the hydraulic rail pulling cylinders are automatically synchronised with the upset cylinders of the welding head, which is important to ensure weld quality.



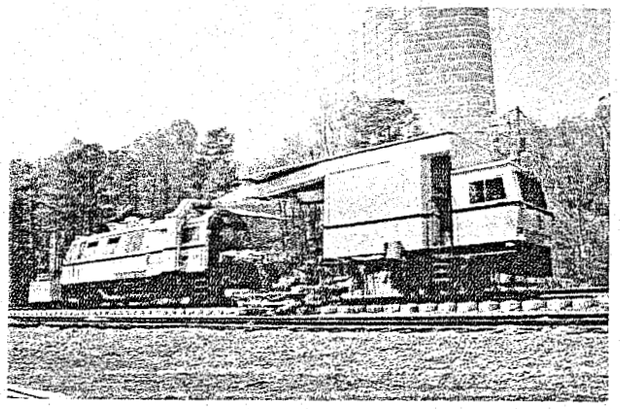
1. Mobile flash butt welding machine APT 500



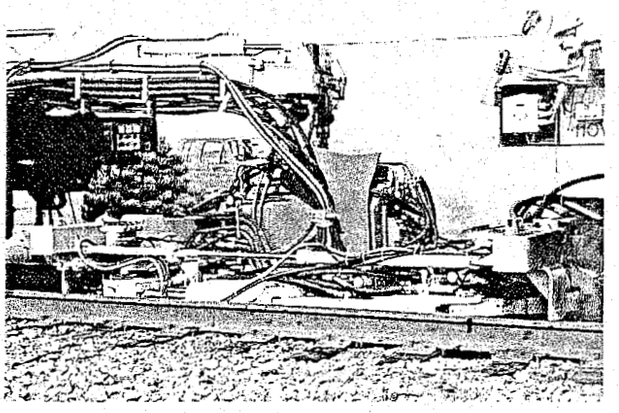
2. Rail end alignment system



3. Trimmed flash but weld



4. APT 500 S mobile flash butt welder with integrated hydraulic puller



5. APT 500 S hydraulic rail puller and flash butt welder head in operation

Figure 10.47: Mobile flash butt welding machine APT 500

10.4.3 Post-processing of flash butt welds in the NS welding depot

In order to obtain a better geometry and higher fatigue strength values a number of additional steps following flash butt welding could be introduced according to the diagram of Figure 10.48 as originally applied in the NS welding Depot.
Directly after welding and stripping the hot weld is given a stress-free overlift of 2 mm on a 1.2 m base. This is done by a special press located 36 m in front of the welding machine.
During overlifting the weld is precooled with air in order to restore enough strength to avoid plastic deformation during transport to the next location.

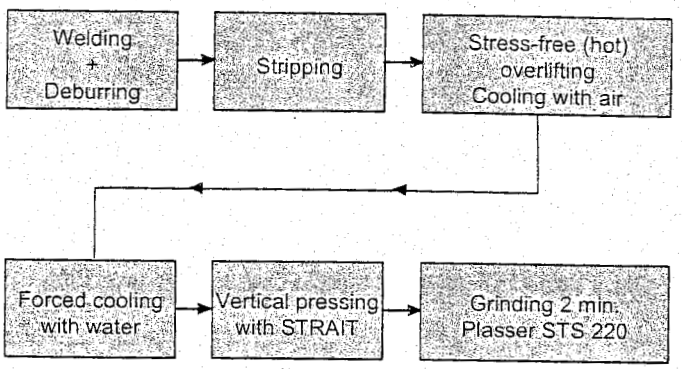


Figure 10.48: NS post-processing flash butt welds

There the weld is cooled with water which does not commence, however, until 4 minutes after welding to guarantee safe transformation for R 260 Mn (900 B) rails.

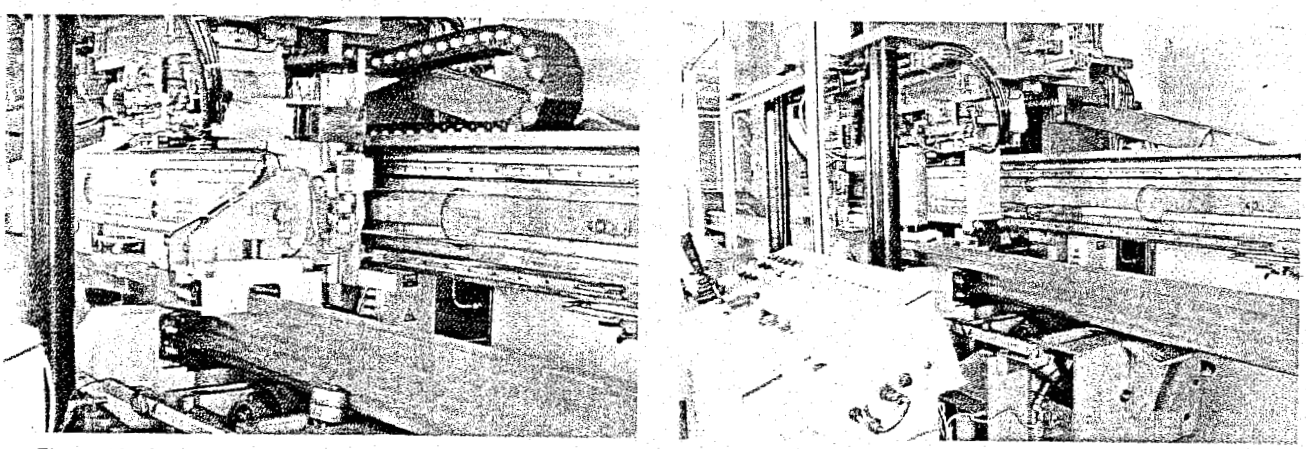


Figure 10.49: Press and operating console used at KHRC

At the next station a STRAIT-system reduces the 2 mm vertical overlift automatically to an overlift within the 0.1 - 0.3 mm interval. As a result of favourable residual compression stresses, the fatigue strength of the weld increases by about 8%. Figure 10.49 shows the press and the operating console. From this console the grinding process, which forms the finishing step, is also controlled.

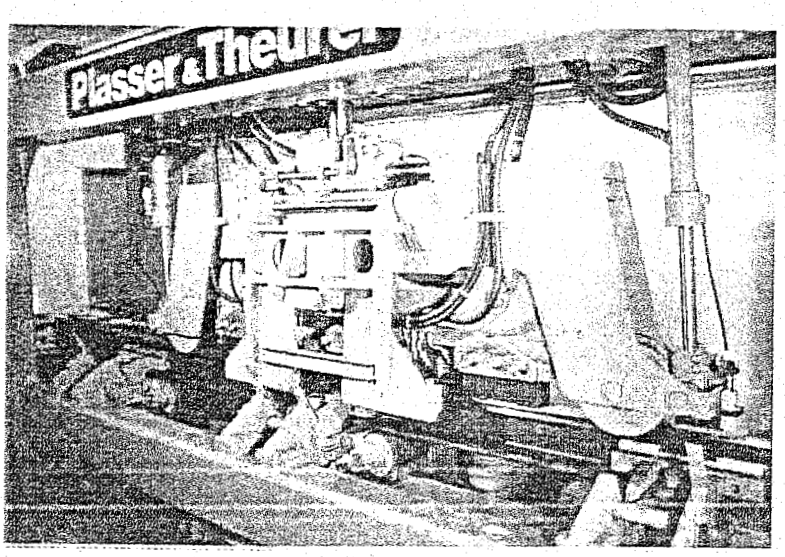


Figure 10.50: Plasser & Theurer STS 220 stationary grinder

The STS 220 grinder, presented in Figure 10.50, is located 36 m ahead of the STRAIT machine. The STS 220 automatically grinds the weld vertically and horizontally for a period of 1.5 to 4 minutes, depending on the weld geometry measured by STRAIT.

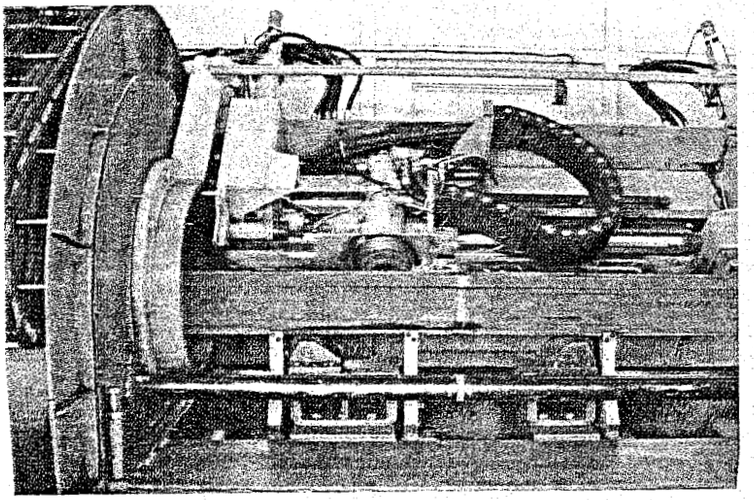


Figure 10.51: Stationary grinder used by KHRC

An alternative system is used at KHRC in Korea where the grinding is performed by moveable rotating stones. For details please refer to Figure 10.51.

Figure 10.52 shows some grinding results obtained with the STS 220. The grinding performance depends heavily on the step size (misalignment). In the welding process this value should preferably be kept below 0.2 mm.

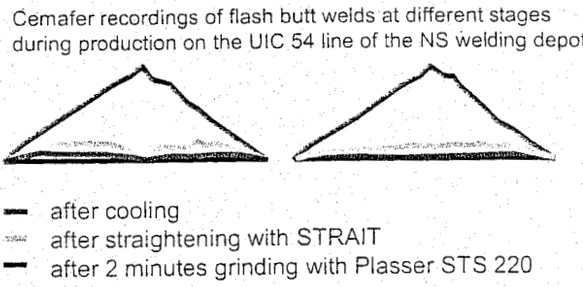


Figure 10.52: Some grinding results obtained with the STS 220

10.4.4 Thermit welding

Thermit (aluminothermic) welding [185] is counted among the fusion welding processes. The parts to be welded together are enclosed in a sand mould with a defined gap between the faces and preheated to temperatures which are suitable for both the materials and the process. The parts are then welded together by allowing the molten steel, produced by the reducing effect of the aluminium on the heavy metal oxide, to flow into the mould. The process is governed by the following chemical reaction:

$$Fe_2O_3 + 2Al \rightarrow 2Fe + Al_2O_3 + 850kJ \qquad (10.8)$$

$Fe_2O_3 + 2\,Al \rightarrow Al_2O_3 + 2\,Fe + 850$ kj

1000 g THERMIT → 476 g slag + 524 g iron + 3970 kj

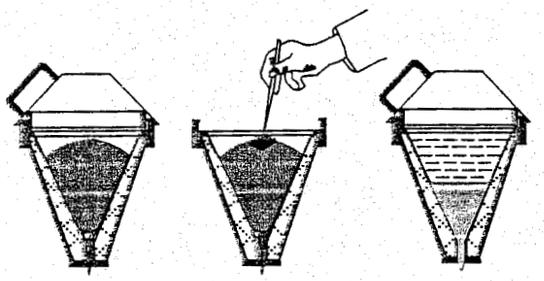


Figure 10.53: Thermit reaction

The aluminothermic mixture is ignited in a highly refractory crucible using a special igniter. After the exothermic reaction, lasting a few seconds, approximately equal quantities of molten steel and fluid Al_2O_3 are separated at a temperature of about 2400 °C. The principle is illustrated in Figure 10.53.

The Thermit iron obtained from such a reaction is, however, too soft for practical use, and therefore, steel-forming alloying additions are added to the Thermit mixture to harden and improve wear resistance, while cold metal (mild steel) reduce the reaction temperature and improve the steel yield.

SmW-F flat-collar welding method

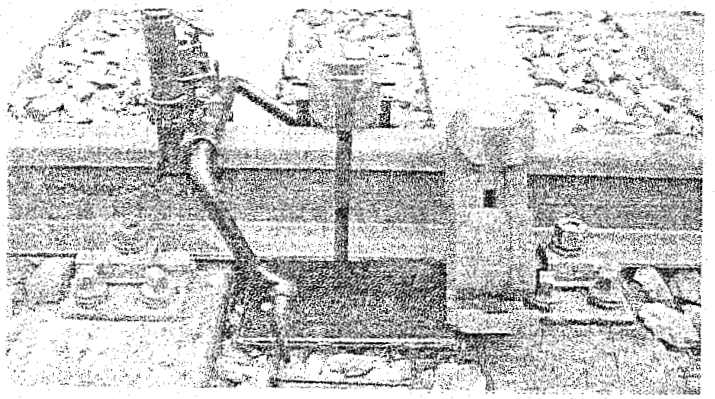


Figure 10.54: Mould for SmW-F process

Irrespective of the sections the rails are laid with a gap of 20 to 22 mm and are aligned. The preformed mould, made for example from CO_2-hardened silica sand and enclosed by mould shoes, is fixed by a specially designed device at the weld gap, as shown in Figure 10.54.

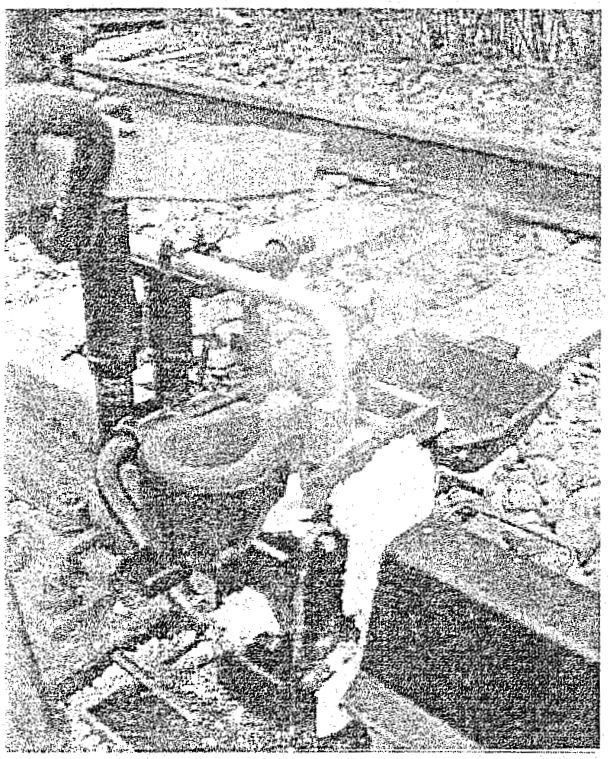


Figure 10.55: Preheating with oxygen/propane

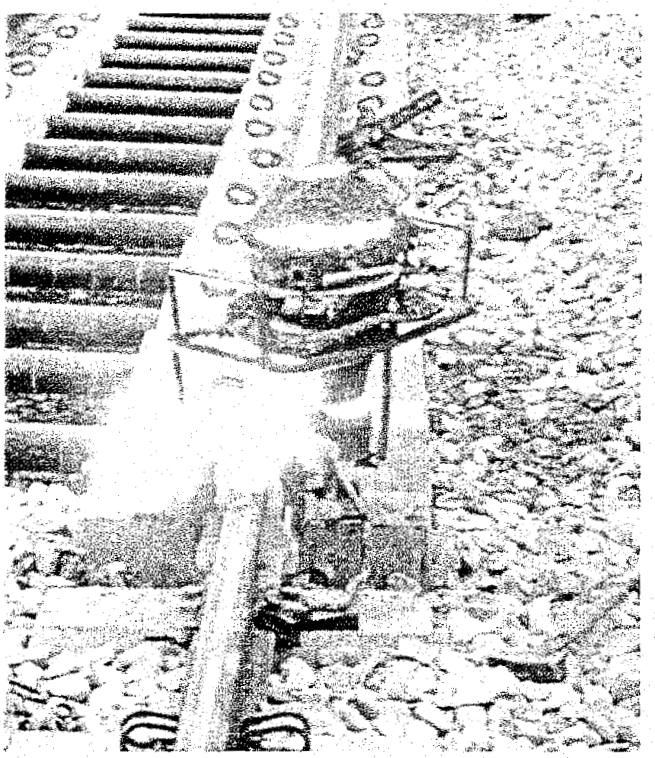


Figure 10.55a: Thermite process in action

The mould is sealed against the rail using naturally or artificially bound silica sand. The rail ends are uniformly preheated within the mould cavity to a temperature of about 1000 °C by the flat flame of a specially designed burner operating on propane and oxygen according to Figure 10.55. In the case of the 60 E1 (UIC 60) section, preheating lasts about 7-8 minutes.

When the Thermit reaction is completed, the preheating torch is removed from the mould automatically and the molten steel is tapped from the crucible into the pouring pot attached to the side (Figure 10.55a). From there the metal flows into the mould. Figure 10.56 presents the pouring diagram. When the steel has solidified, the mould is removed only in the region of the rail head (in the case of the 60 E1 (UIC 60) section after 4-5 minutes), the rest of the mould and the foot risers being left in place to reduce the rate of cooling of the joint. Figure 10.57 shows a machined weld produced by the flat-collar method.

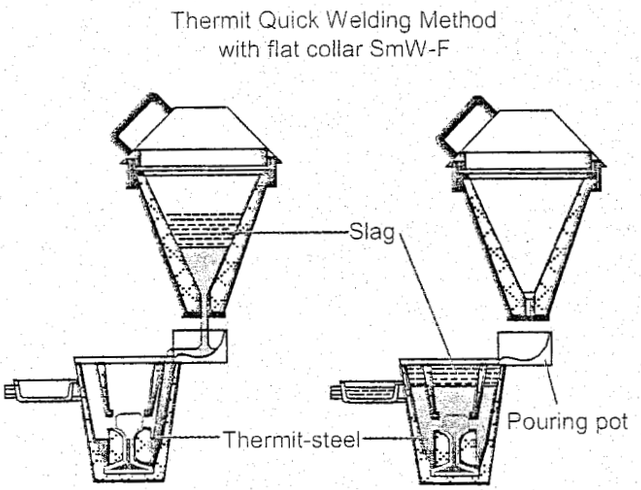


Figure 10.56: SmW-F pouring diagram

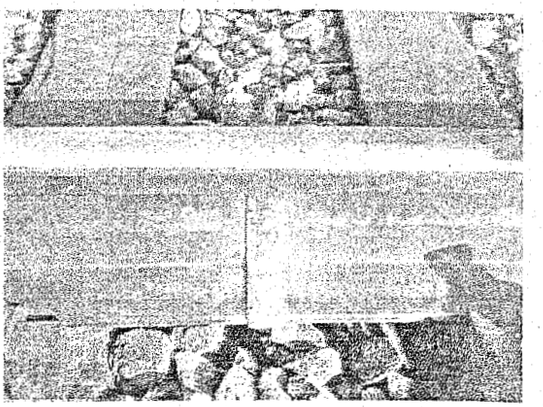


Figure 10.57: Machined Smw-F weld

SkV short-preheat welding method

The short-preheat welding method permits rail welding to be carried out during extremely short time-intervals of 12-15 minutes between trains on very busy lines.

A number of characteristic features distinguish the SkV from the SmW-F process. Preheating, for instance, only lasts between 1 and 2 minutes. It serves to dry the mould and heat the rail ends to a temperature of about 600 °C in a narrow zone of about 3-5 mm.

The larger heat quantity needed for the fusion of the rail ends is supplied by an increased amount of welding mixture. The liquid metal flows from the crucible directly into the mould, through a central inlet, without passing through a pouring pot first.

For this purpose the centre of the preheated opening in the upper part of the mould is closed with a plug after the burner is removed. The mould evenly distributes the molten steel as it flows down into the 24-26 mm gap between both rail ends. The lower heat input at preheating is balanced by using a larger SkV-portion

Figure 10.58 shows the SkV pouring diagram. For the repair of rail failures within a gap of 75 mm a special process, referred to as SkV-75, is used. This weld is an attractive alternative for repairing vertical cracks which would otherwise have required the application of a short rail section with 2 welds.

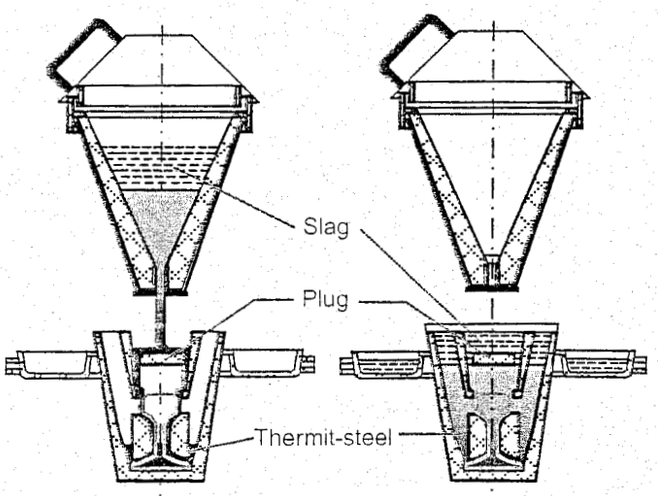


Figure 10.58: SkV-F pouring diagram

Improvements of the Thermit welding method

Besides improvements on the process consistency, one-shot light weight crucibles were introduced in order to reduce the weight of the equipment from more than 200 kg to less than 50 kg. They are made from steel or paper cans coated with highly refractory material and filled with the aluminothermic mixture. After casting, the can is deposed in accordance with legal regulations.
In 2000, Elektro-Thermit presented a method with partial alloying of the rail head. The material properties follow the idea of a head hardened rail with a high hardness in the rail head and a considerable toughness in the rail web and foot. These properties are achieved by using a soft but tough aluminothermic mixture and an alloying bar at the top of the mold. During casting, the bar melts and alloying elements increasing the hardness in the top part of the cross section are resolved in the thermite steel.

Finishing of Thermit welds

After solidification, the excess weld metal required to compensate for shrinkage can be removed almost completely in the region of the rail head while still red hot. This is normally done with a hydraulic shearing device as depicted in Figure 10.59.

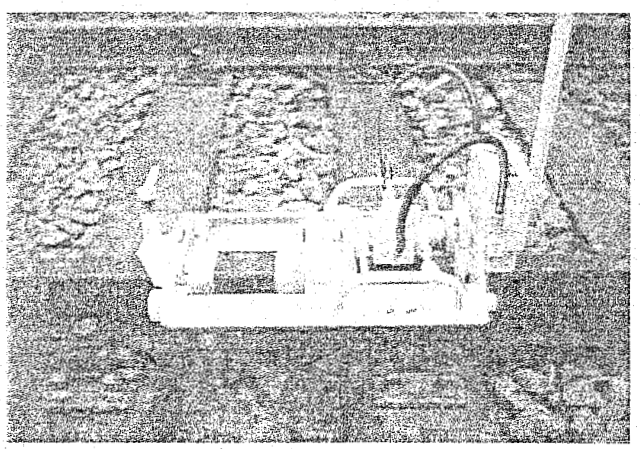


Figure 10.59: Hydraulic shearing device

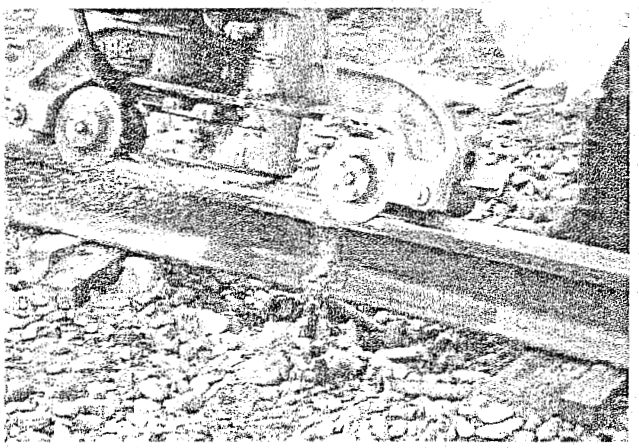


Figure 10.60: Machine for grinding Thermit welds

Subsequently, the riser on the running surface is ground by a special grinding machine, illustrated in Figure 10.60, until only 0.5 mm remains. After the weld has completely cooled, the foot risers can be broken off cleanly and with ease and the weld cleaned of the remaining mould sand. The average time needed between trains to perform a Thermit weld amounts to 15 minutes for the SkV process and 20 to 25 minutes for the SmW-F process, depending on the profile. After the weld has cooled down to ambient temperature, running surface and gauge corner are smoothed by finish-grinding to conform to the rail section.

Mechanical properties of Thermit welds

According to DB the ultimate static bending stress should be at least 750 N/mm². The NS requires that in a bending test the total deflection on a 1 m base is at least 30 mm for the 700 grade and 20 mm for the R 260 (900) grade and the R 320 Cr (1100) grade, without the development of any cracks. On DB the fatigue strength, being twice the stress amplitude according to the definition given in [185], should be at least 226 N/mm². This is for pure bending, with a minimum applied stress of -6 N/mm².

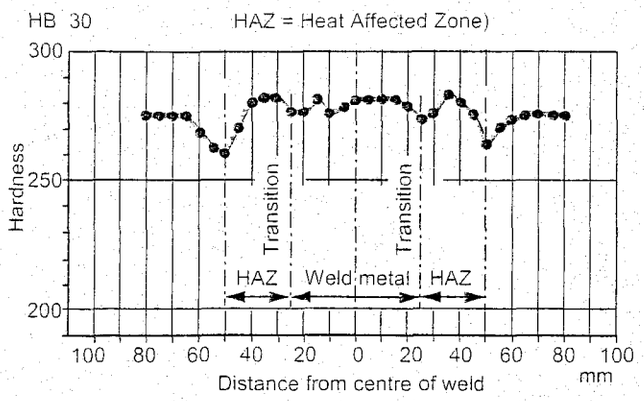


Figure 10.61: Hardness distribution on running surface of SmW-F weld

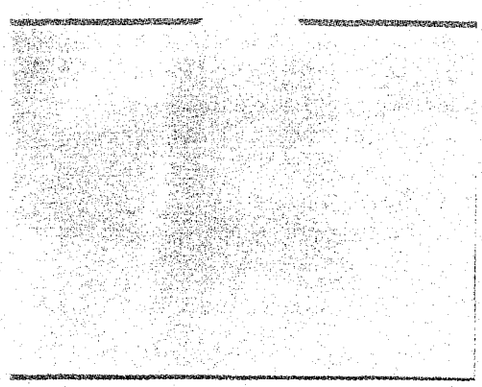
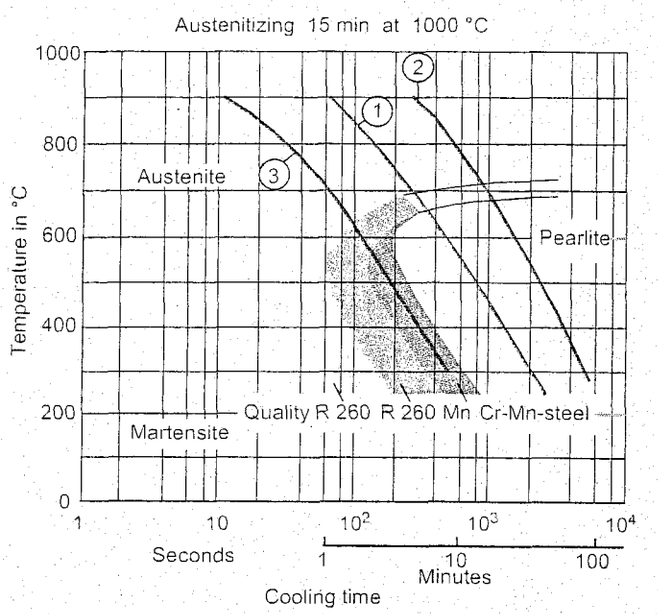


Figure 10.62: Machined Smw-F weld

The hardness of the aluminothermically produced steel must match that of the rail to avoid battering. Figure 10.61 shows a hardness profile on the running surface of an S 49 section rail with a minimum tensile strength 880 N/mm².

The Brinell hardness of the rails is thus 275 in the non-heat-affected zone, rising to 285 in the weld metal, and drops in the narrowly formed hardness trough in the fine-grained zone. This is confirmed by the longitudinal macro-section through the rail centre presented in Figure 10.62. The picture shows that at about 55 mm from the weld centre the soft-annealing temperature of 650 °C to 750 °C causes the grain structure of the steel to become finer. The weld gap was 22 mm and the width of the weld metal about 30 mm at its narrowest point.



1. Cooling curve for rail steel after SkV Thermit welding followed by accelerated cooling with compressed air.
2. Cooling curve for rail steel after SkV Thermit welding followed by retarded cooling with heat insulation.
3. Cooling curve for rail steel after flash butt welding.

Figure 10.63: TTT-diagrams of 3 rail steels

10.4.5Cooling rates

The pearlitic rail structure must also remain intact in the heat-affected zone of welds. When welding, part of this zone is heated to over 720 °C, causing pearlite to be transformed into austenite. By subsequent cooling, the austenite should retransform into pearlite. Whether or not this actually happens depends on the steel's chemical composition and the rate of cooling. The time-temperature transformation diagram (TTT-diagram) furnishes more information on this relationship [185].

By way of example, the TTT-diagrams for UIC grades R 260 (900 A) and R 260 Mn (900 B) and for special grade Cr-Mn steel are superimposed in Figure 10.63. Also shown in this diagram are two limiting curves for accelerated and retarded cooling of the rail steel following SkV Thermit welding.

When rails are welded by the Thermit method, the austenite completely transforms into pearlite. In the least favourable case, the steel passes through the decisive temperature range from 800 °C to 500 °C in 11 minutes.

According to this diagram, the minimum cooling time from 800 °C to 500 °C required to just achieve transformation to pearlite is 50 s for grade R 260 (900 A) and 130 s for grade R 260 Mn (900 B). The special grade, alloyed with chromium and manganese, requires 210 s for complete transformation into the pearlite stage. After Thermit welding the actual cooling rate is so slow that, even in the least favourable case, no critical situation will ever occur.

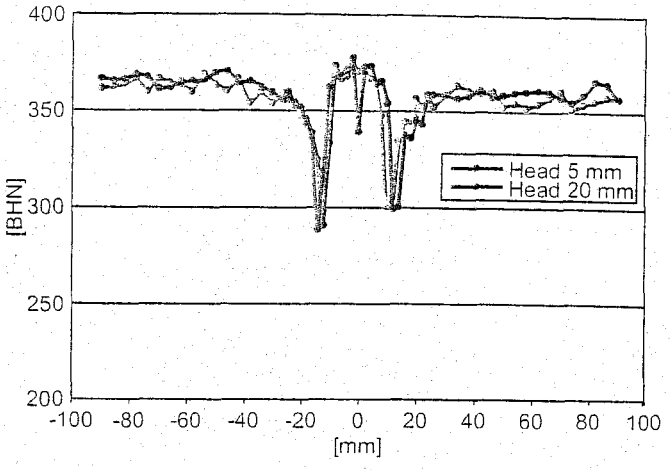


Figure 10.64: Hardness distribution over a R 350 HT flash butt weld

When using the flash butt welding method with its lower operating temperatures, special measures need to be taken such as post-heating. This is confirmed by the flash butt weld cooling curve sketched in Figure 10.63. If after welding the rails should cool down in less than the minimum cooling time, martensite may form locally. This structure is hard and brittle and may lead to rail failure. Alloyed rail steels (Cr-Mn-steel in Figure 10.63) need to be post-heated in order to avoid martensite formation in the heat affected zone.

The cooling procedure of the heat treated steel grades must be well controlled to adjust the hardness in the joint according to the hardness of the rail. For the CEN grade R 350 HT the rail head must be quenched, e.g. by using a blow-off-device. This device gives a similar heat treatment as used in the manufacturing process of the rail. The hardness and tensile properties of the weld are comparable to that of the rail, having a high hardness in the head and a high toughness in the rail foot (Figure 10.64). Grade R 350 LHT is alloyed with chromium and, in order to avoid martensite formation, no such quenching procedure is applied.

Table 10.5 provides another comparison of transformation behaviour of various rail steel grades, which are listed in order of the minimum cooling time $t_{8/5}$ required from 800 °C to 500 °C to achieve complete transformation into the pearlite stage. This list shows that manganese exerts the decisive influence on the transformation behaviour of rail steels, carbon having no influence.

Rail steel	$t_{8/5}$ [s]	Temperature (°C) at which fully pearlite	% by mass			
			C	Si	Mn	Cr
UIC grade R 200	50	560	0.5	0.2	1.0	-
UIC grade R 260 (900 A)	50	555	0.7	0.2	1.0	-
CEN R 350 HT	55	565	0.8	0.4	1.1	-
head hardened						
CEN R 350 LHT	75	580	0.8	0.4	1.1	0.2
head hardened						
UIC grade R 260 Mn (900 B)	130	562	0.6	0.2	1.5	-
Special grade	210	580	0.7	0.7	1.1	0.9
Cr - Mn - steel						

Table 10.5: Summary of minimum cooling times $t_{8/5}$

Yield stress and Young's modulus for a flash butt weld in grade R 260 Mn (900 B) have been determined by the NS as a function of temperature. Figure 10.65 shows both relationships. At 500°C the yield stress has almost dropped to zero. A value of 200 N/mm² is found at a temperature of about 400 °C. At about 200 °C half the final yield stress is attained.

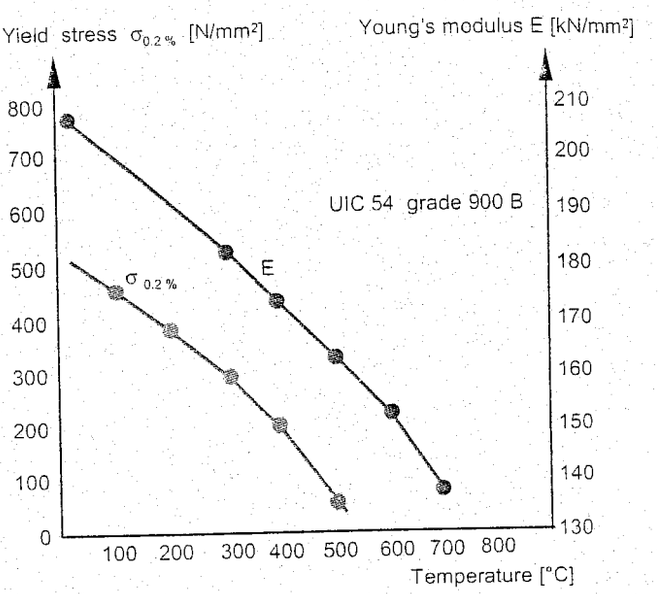


Figure 10.65: Yield stress and Young's modulus versus temperature

10.4.6 Improvement of weld geometry

In the publication on vertical rail geometry [76], problems concerning weld geometry on NS rail have been analysed and an integrated approach to the improvement of weld geometry proposed. The most important step forward consisted of the STRAIT/GWM development for in-track correction after welding.

Positioning of the rails prior to welding is normally done with the use of a 1 m straightedge. An overlift of 1 - 1.5 mm on a 1 m base is required. In order to check the geometrical quality of the finished weld use can be made of an electronic straightedge, like the RAILPROF. (For details please refer to Chapter 16). In principle, such an electronic straightedge could also be used for the initial positioning of the rails.

10.4.7 Weld geometry standards

The following acceptance specifications for the geometric quality of new flash butt and Thermit welds are recommended:

Vertical direction

- Welds which have been straightened and ground:
  - Versine on a 1 m base between 0 and 0.2 mm (hump);
  - Gradient over 200 mm less than 1:1000;
- As welded:
  - Versine on a 1 m base between 0 and 0.3 mm (hump);
  - Gradient over 200 mm less than 1:1000;

Horizontal direction

- Versine on a 1 m base between -0.5 and 0.5 mm;
- Gradient over 200 mm less than 1:500.

10.5 Rail failures

Rail failures are usually the final result of a period of fatigue crack growth. The cracks grow from small defects or stress concentrations. For this reason it is often possible to find the crack before failure occurs. It is common practice to use ultrasonic inspection. This subject will be dealt with in Chapter 16. In service, rails are known to fail in several ways. These failures are classified in UIC Leaflet 712 R [279]. The UIC codes are composed according to the simplified scheme in Table 10.6 . The following sections describe some of the most common and hence most costly types of failures.

1st digit	2nd digit	3rd digit
1 rail ends	0 full section 1 innerside rail head 2 web 5 foot	1 transverse 2 horizontal 3 longitudinal crack 5 star crack fishbolt holes
2 away from rail ends	2 rail head surface	0 corrosion 2 shelling 5 wheel burn
4 welds	1 flash butt 2 Thermit 7 resurfacing 8 other	1 transverse 2 horizontal

Table 10.6: Coding used in UIC classification of rail defects

10.5.1 Defects in rail ends

code 1321: horizontal cracking at web-head fillet radius

This crack originates in the rail end and tends to separate the head from the web. At the outset, the crack progresses parallel to the web-head fillet (Figure 10.66.a) and may, as it develops, curve either upwards (Figure 10.66.b) or downwards, possibly passing through a fish-bolt hole (Figure 10.66.c), or simultaneously upwards and downwards (Figure 10.66.d).

Not to be confused with code 135. In all cases this crack causes the head to break away or fragmentation of the rail.

The defect should be eliminated by inserting a rail section.

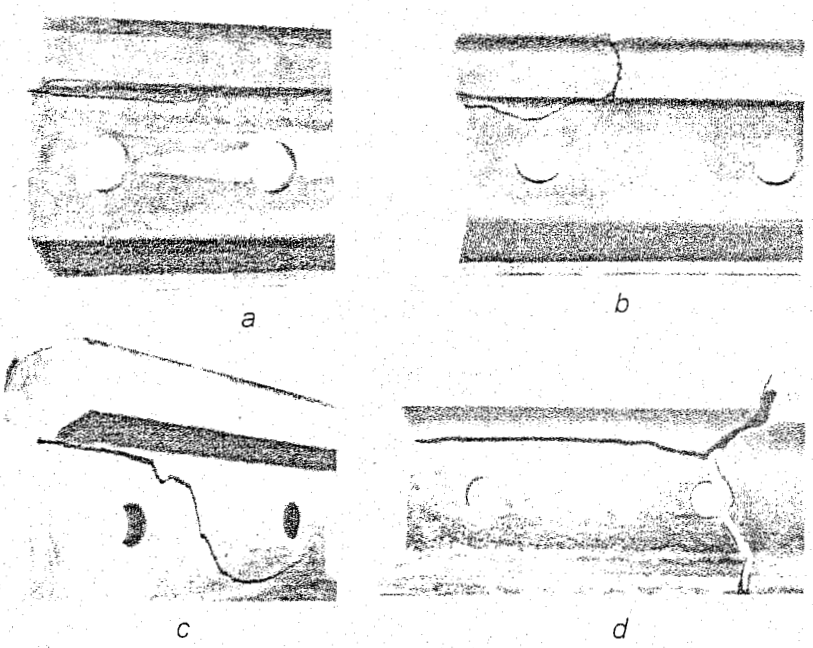


Figure 10.66: Code 1321: Horizontal cracking at web-head fillet radius

Other pictures of rail surface fatigue defects

The deterioration of hard-steel rails through surface fatigue follows a slightly different pattern. The gauge corner may initially show small, extremely superficial filiform cracks which are formed due to a combination of high loads and large slip at the gauge corner. These are transverse, oblique cracks, called head checks, which are rather close to one another (Figure 10.74, Head checks)

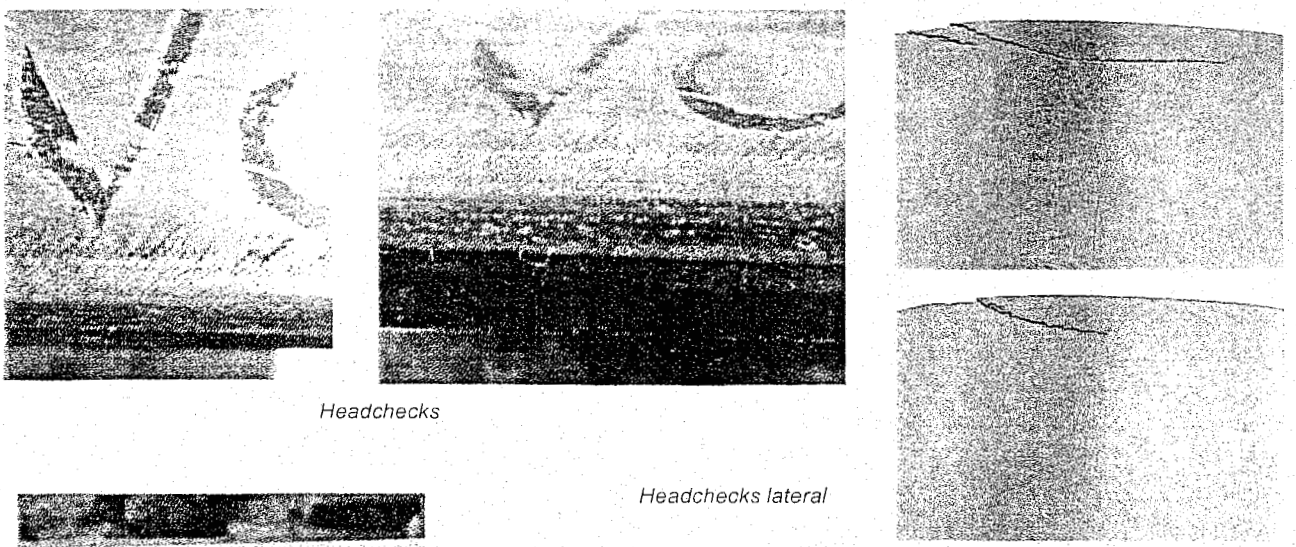


Figure 10.74: Some other rail head defect types

These cracks grow and either form transversal fractures or merge below the surface. In the latter case, small metal particles go off (Figure 10.74, Spalling), referred to as spalling. In a more advanced phase an almost general disintegration of the gauge corner is to be observed. (Figure 10.74, HC lateral). These defects are not distinguished from shelling in the UIC Leaflet 712 R.

Generally speaking, shelling, head checking and spalling occur on the outside rails in curves and are promoted by lubrication applied to reduce lateral wear. Check that this defect is not accompanied by progressive transverse cracking 211. At an early stage grinding is an option, otherwise the rails should be removed. To prevent recurrence, use harder rails or preventive grinding.
References for head checks are [53] and [104].

Code 2251: isolated wheel burn

Defect due to traffic loads: the slipping of a driving axle can cause the formation of an elliptical self-quenched layer, usually in both rails of the track (Figure 10.75). This layer may disappear or develop:

- either horizontally in the head, where it soon degenerates into local shelling which does not develop in depth but progressively causes depression of the running surface under repeated traffic loads;

– or transversely in the head, resulting in an internal crack which progresses towards the web and is accompanied by a depression of the running surface without any other local alteration.

This internal crack soon reaches the surface and invariably leads to a break, the pattern of which is very characteristic. At the early stages wheel burns can be repaired by resurfacing, otherwise a short section of rail should be welded in.

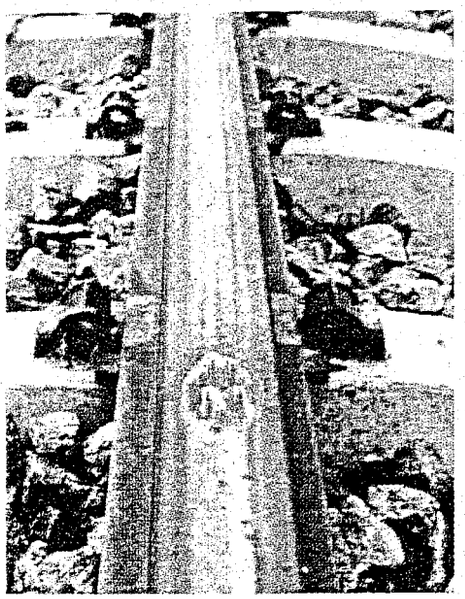


Figure 10.75: Code 2251: Isolated wheel burn

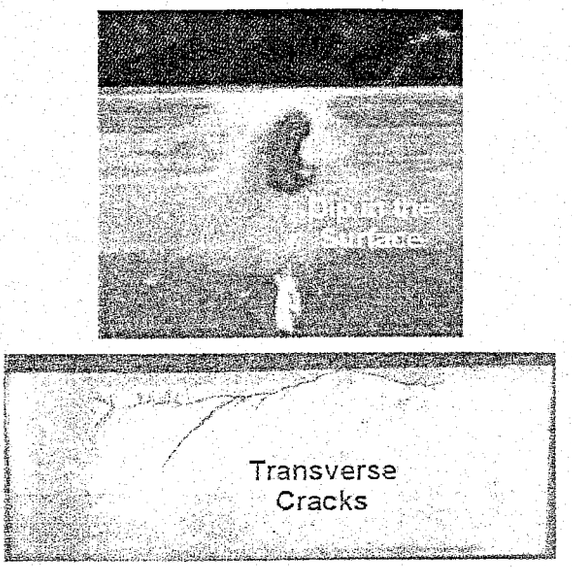


Figure 10.76: Squat defect

This defect shall not be mixed up with the so-called squat defect which was observed first in high-speed lines. Surface cracks are formed in the same way like head checks due to high tangential forces at the surface. The defect exhibits a shallow dark impression in the surface. A crack develops into the rail, preferentially in the direction of traffic (Figure 10.76).

The rail should be replaced. References for squats are [53] and [104].

Code 2252: repeated wheel burns

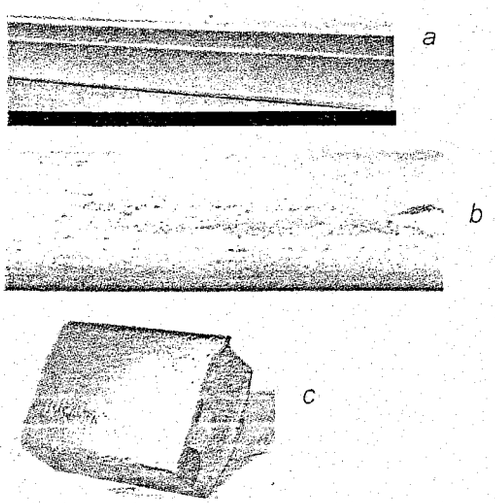


Figure 10.77: Code 2252: Repeated wheel burns

Defect caused by traffic loads: on sections where wheel slipping or braking occurs repeatedly or in places where slipping occurs while the tractive unit is on the move. The running surface may assume a characteristic wavy appearance (Figure 10.77.a).

This phenomenon may be characterised by the appearance, on the running surface, of a series of minute cracks which go deep inside the rail (Figure 10.77.b). These increase the brittleness of the rail appreciably in cold weather, and may cause a break (Figure 10.77.c).

The defect is generally encountered in rails located near stop signals. Rails should be ground at an early stage, otherwise renewed.

Code 233: longitudinal vertical crack (piping)

Manufacturing defect: this defect, commonly known as piping, is characterised by a longitudinal vertical crack in the rail web (Figure 10.78.a). If compounded with another defect it may cause a break to occur.

Swelling on both faces of the web can be observed in exceptional cases (Figure 10.78.b) combined with a slight depression of the running surface at right angles to the defect.

Rails should be renewed.

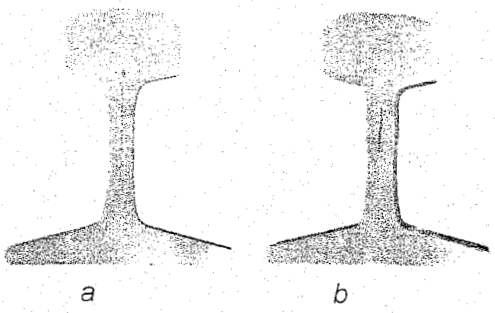


Figure 10.78: Code 233: Longitudinal vertical crack in the web (piping)

Code 253: longitudinal vertical cracking

Manufacturing defect: small longitudinal lines, which occur during rolling and which are located in the middle portion of the underside of the foot, develop into cracks due to traffic loads. These cracks ultimately cause sudden fracture of the rail, generally during cold weather (Figure 10.79.a and b).

When the crack is not located in the middle portion - which is the case particularly on rails where the foot shows only one slope in its top part - the break follows a semi-circular pattern and this tends to cause a fragment of the foot flange to break away. This crescent-shaped break is shown in Figure 10.79.c.

The break is generally very oblique, particularly in the web (Figure 10.79.b), and the longitudinal crack at the origin of it can be easily detected after breakage has occurred. This defect can develop in several parts of the same rail and results in the formation of an extensive gap due to multiple breaks.

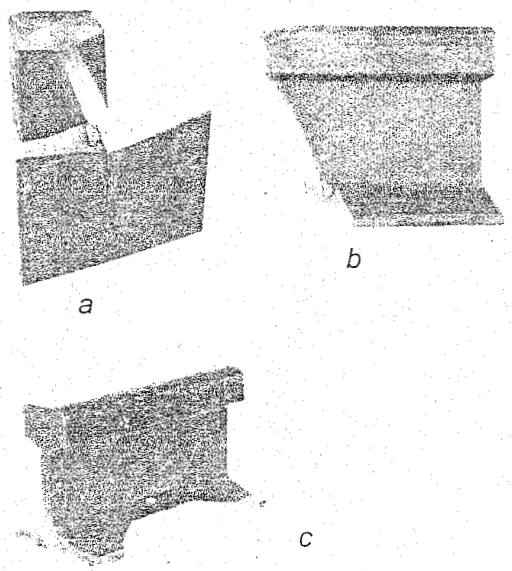


Figure 10.79: Code 252: Longitudinal vertical crack in the foot

This defect may take the form of an epidemic in rails from the same cast. Broken rails should be renewed at the earliest opportunity.

10.5.3 Weld and resurfacing defects

Code 411: transverse cracking of profile

Defect in flash butt weld: cracking develops in the weld cross-section either from an internal defect of the head in the weld (Figure 10.80.a) or from a defect located in foot of the rail (Figure 10.80.b). It ultimately leads to complete breakage of the section.

The pattern of the break is characterised either by a smooth and bright spot in the head or by a dark spot on the foot.

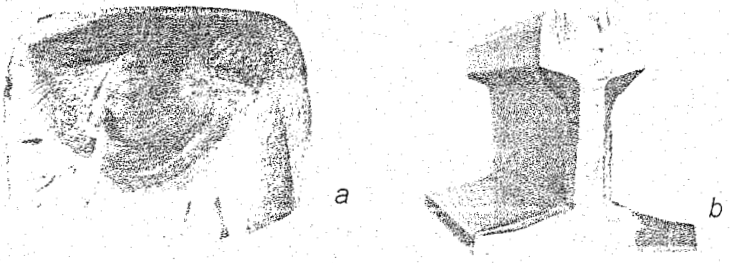


Figure 10.80: Code 411: Transversecracking of profile flash butt weld

Weld failures are repaired either by welding in a short length of rail, e.g. 3 m, or replacing the original weld with a normal gap or a wide gap Thermit weld.

Code 412: horizontal cracking of web

Defect in flash butt weld: this defect develops across the weld and generally assumes a curved shape in the weld. As it spreads, it may curve downwards (Figure 10.81.a) or upwards (Figure 10.81.b), or simultaneously in both directions (Figure 10.81.c). It ultimately leads to fracture of the rail near the weld.

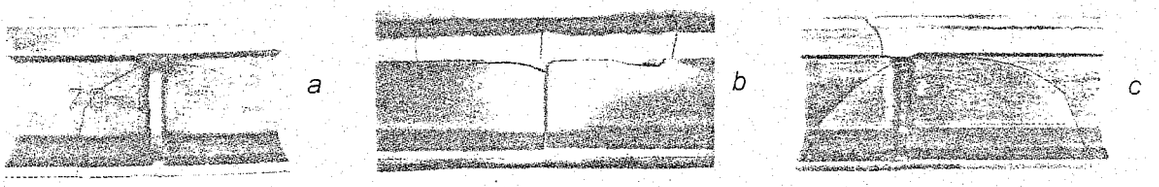


Figure 10.81: Code 412: Horizontal cracking of web flash butt weld

Repair should take place by welding in a short length of rail.

Code 421: transverse cracking of profile

Defects in Thermit weld: this defect develops in a plane near a normal cross-section of the profile. It ultimately leads to a breakage in the profile. Some of the characteristics of this defect are given below:

- Crack originating in the flange under the foot and developing in a transverse plane in adjacent rails (code 421.1, Figure 10.82.a);
- Vertical crack located the weld (code 421.2, Figure 10.82.b, c and d);
- Crack located in a vertical plane near the weld (code 421.3, Figure 10.82.e);

Repair should be done as specified under code 411.

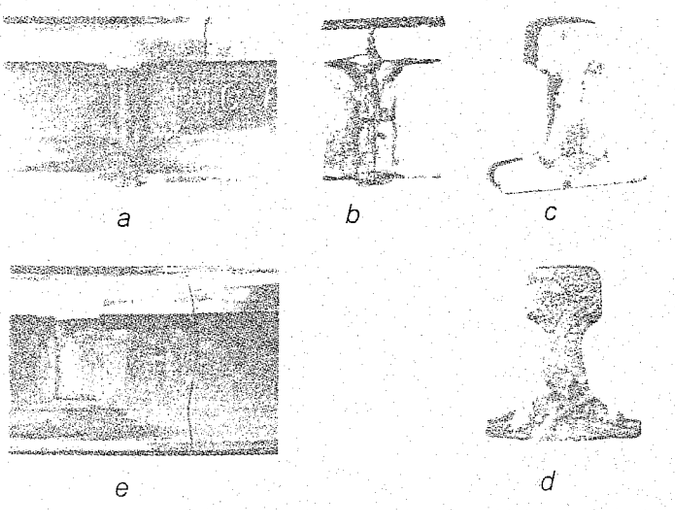


Figure 10.82: Code 421: Transverse cracking of profile Thermit weld

Code 422: horizontal cracking of web

Defect in Thermit weld: this defect is generally found in welded rails where the rail ends have not been cut off before welding. The crack, which generally links up the fishbolt holes across the weld (code 422.1, Figure 10.83.a), can spread to the foot or web of adjacent rails and lead to breakage.

This defect may also be found, although much less frequently, in fresh welding of non-drilled rails (code 422.2, Figure 10.83.b). In such cases, the defect can develop from a small crack in the vertical flange and also lead to breakage.

Repair should be carried out by welding in a short length of rail.

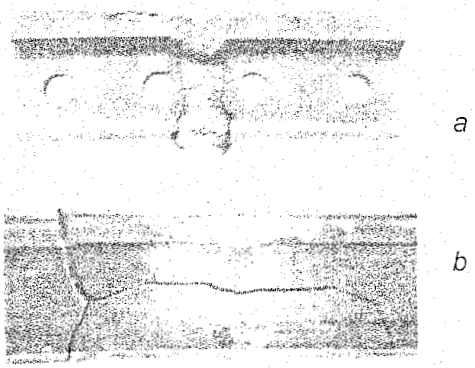


Figure 10.83: Code 422: Horizontal cracking of web Thermit weld

Code 471: transverse cracking of rail head

Resurfacing defect: this transverse fatigue cracking across the recharged part originates in a nucleus generally located at the base of the weld deposit (Figure 10.84). This nucleus may be an inclusion, a local discontinuity due to weld instability, or a flake due to improper stoving. The defect may also start on an end-bead crater or on shrinking cracks. As it develops, the cracking ultimately leads to fracture of the rail. The break is characterised by a smooth, shiny patch originating in the deposit metal.

Not to be confused with 211 or 2251.

Repair should be carried out by welding in a short length of rail.

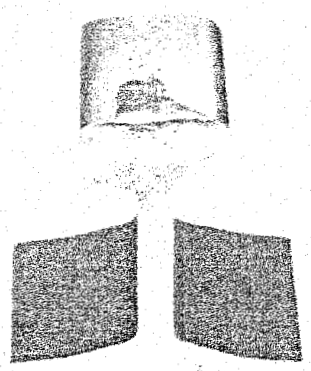


Figure 10.84: Code 471: Transverse cracking of rail head due to resurfacing

Code 481: transverse cracking under electrical connection

Progressive transverse cracking originating on the outer face of the rail head (code 481.1, Figure 10.85.a) on one of the web faces (code 481.2, Figure 10.85.b), or on one of the flanges of the foot (code 481.3, Figure 10.85.c), at right angles to an electrical connection for return current.

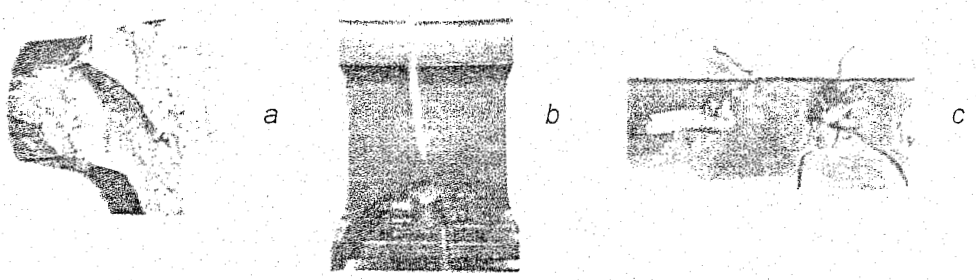


Figure 10.85: Code 481: Transverse cracking under electrical connection

The progression of this crack may ultimately lead to fracture of the rail.

Repair should take place by welding in a short length of rail.

10.5.4 Rail defect statistics

Most of the railway administrations collect information on rail failures for further analysis and interpretation. Roughly speaking, there are two ways of analysing rail defect data. One can consider the number of defects per year in different categories in order to compare different years and look at trends to see whether these comply with the applied rail maintenance strategies. The other type of analysis consists in looking at the development of rail failure types as a function of the tonnage borne. Both aspects will be discussed here. It is implicitly assumed that the data refer to long sections of track, for instance a whole network. The way in which short sections of track could be analysed will be discussed in Chapter 18.

Predicting rail failures

The expected number of failures on a route can only be predicted if enough information is known. The history of earlier failures is particularly important. For a general length of track the following information is needed:

- classification of failure;
- tonnage to failure;
- time to failure;
- method of repair;
- welding process;
- post-weld treatment;
- infra data.

For the purpose of rail failure prediction a homogeneous unit of track will have the same general details, as listed above, along its length. If this rule were applied strictly the track units would be very small. There would not be enough failures in a unit to permit analysis. Thus some relaxation is necessary to produce units of at least 5 to 10 km. Units of track with the same general definition from different parts of the network can be grouped together to form longer units.

From NS rail defect statistics it has been found that the defect rate may vary strongly per location. This is illustrated by Figure 10.86, which follows the defect rate of the NS network over 6 successive years. The scatter around the mean is quite high, with the maximum value being 4 to 6 times the minimum value.

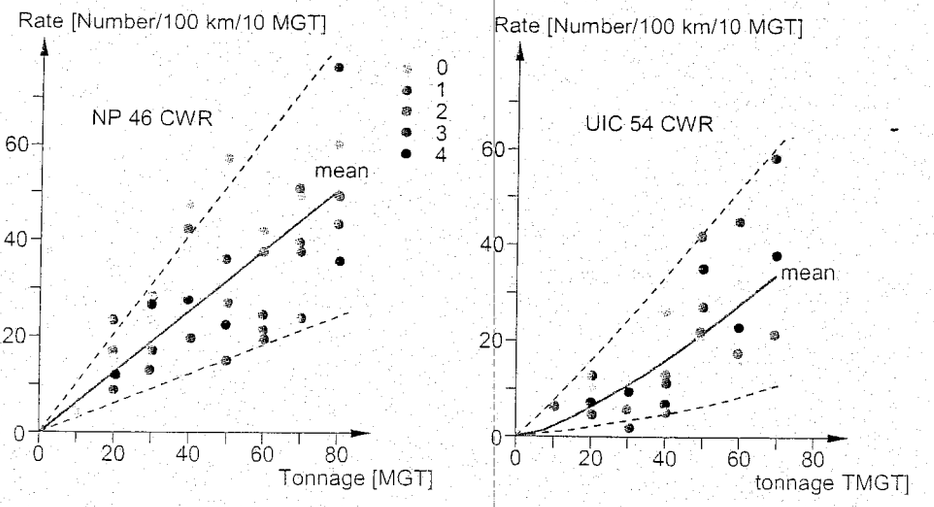


Figure 10.86: Scatter of rail defect rates on NS

Examples of the NS network

An example of the number of rail defects per year in plain tracks, as a function of time, is presented in Figure 10.87. These values refer to NS main lines, consisting of 4500 km of track. In 1988, the total number was reduced to 300 which corresponds to 6.7 failures per 100 km of track. Over 80 % of these failures were found by ultrasonic inspection.

The number of defects shows a decreasing trend of about 40 defects per year. As can be seen from Figure 10.88, the descending trend is mainly caused by glued insulated joints, Thermit welds, and the rail itself. These curves in fact reflect a number of strategies, such as application of stronger glued insulated joints, weld straightening and grinding, and improved quality control with properly defined standards for welding and rail supply.

On the other hand, it should be mentioned that the accumulated tonnages on the NS network are very low. This is shown by the distribution functions in Figure 10.89 for rail types 46 E3 (NP 46) and 54 E1 (UIC 54) on the NS CWR tracks. In fact, these defects refer to process failures rather than to fatigue failures, for fatigue occurs at much higher tonnages, at least for 54 E1 (UIC 54).

Another way of combining the defects is by means of the UIC codes discussed earlier. Figure 10.90 shows the development, as a function of time, of the most dangerous defects on NS rail. Among these, defect 132, horizontal cracking at the web-head fillet radius and defect 135, star-cracking of fishbolt holes, appear most frequently, primarily in glued insulated joints. The total number has decreased substantially and other types have almost disappeared.

The second way of looking at rail failures is to consider them as a function of the accumulated traffic load, which aims at giving more insight into the service life of rail components.

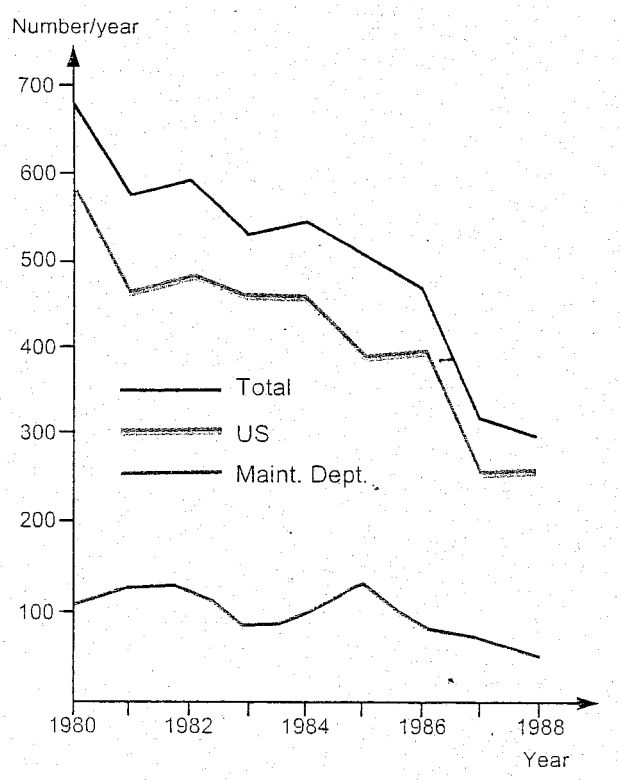


Figure 10.87: Number of rail defects per year in plain track on NS

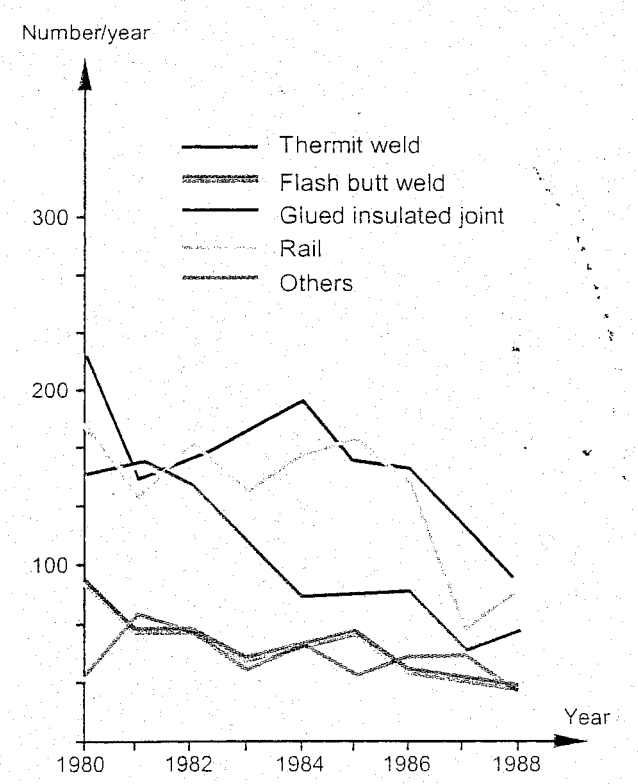


Figure 10.88: Defects split up into construction types

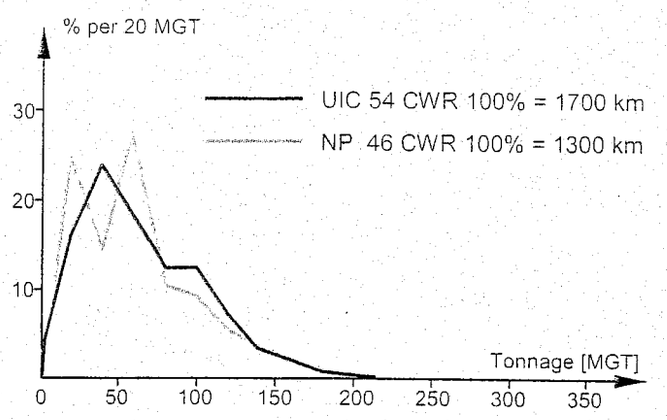


Figure 10.89: Distribution per 1.1.88 of tonnage borne for CWR tracks on NS

Figure 10.91 depicts the defects per 100 km of track per 10 MGT, as a function of the tonnage borne for rail profiles 46 E3 (NP 46) and 54 E1 (UIC 54). These data, collected over a period of 6 years, are split up into contributions from flash butt welds, Thermit welds, glued insulated joints, and the rail itself, all restricted to CWR.

The 46 E3 (NP 46) profile reveals a distinct progression in the development of rail defects, which indicates the start of fatigue. With the UIC profile, on the other hand, hardly any progression in rail defects is noticeable as yet. Regarding the behaviour of welds, the difference between the two rail profiles is much less pronounced. Particularly for Thermit welds defect development seems to depend on the welding performance rather than on the rail profile. In this case the geometrical quality is decisive. For both profiles, over 75 % of the total number of defects is caused by welds and joints. Figure 10.91a shows some examples of damage due to poor weld geometry.

To investigate the differences in defect development in tracks on wooden sleepers and in tracks on concrete sleepers, the sub- division shown in Figure 10.92 has been made. Glued insulated joints have been disregarded. When examining the total amounts of defects, concrete compares favourably with wood for both rail profiles. The wood/concrete defect ratio at a tonnage borne of 100 MGT amounts to 1.4 for NP 46 and 1.7 for UIC 54.

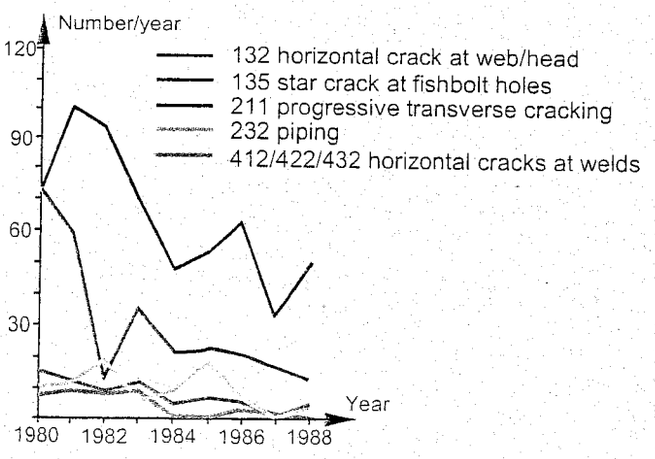
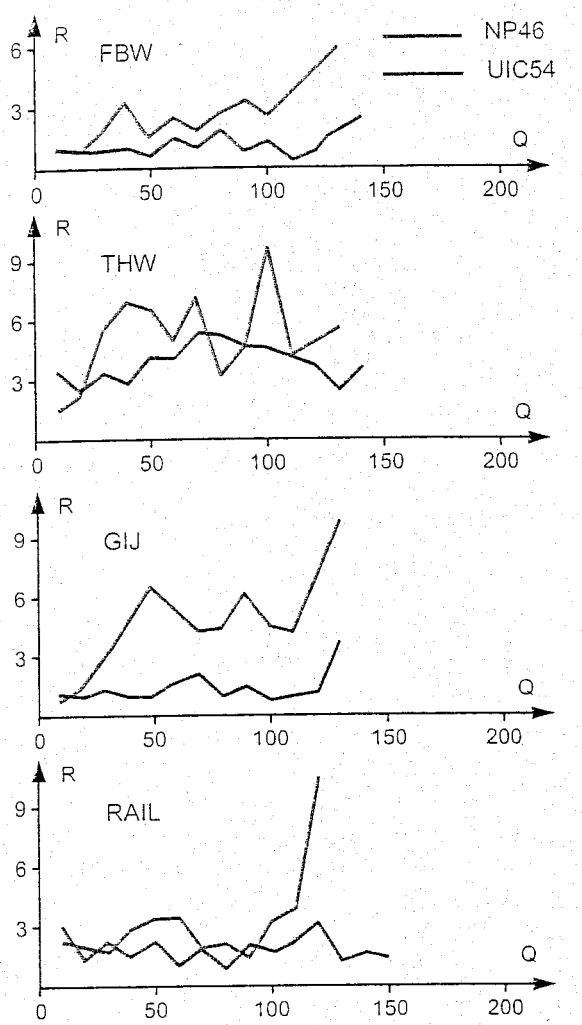
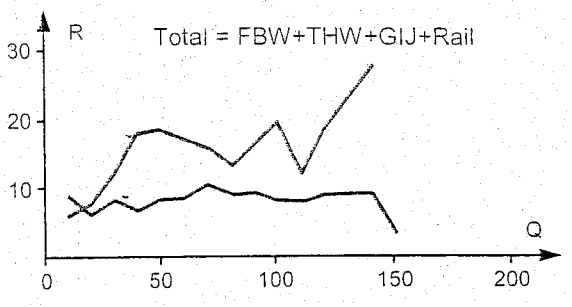


Figure 10.90: Dangerous rail defects on NS split up according to UIC code





	DEFECT RATE N = αQβ					
	NP 46			UIC 54		
	α	β	r	α	β	r
RAIL	0.05	1.00	0.87	0.02	1.00	0.81
THW	0.07	1.00	0.88	0.04	1.00	0.88
FBW	0.04	1.00	0.95	0.01	1.00	0.88
GIJ	0.06	1.00	0.95	0.02	1.00	0.87
TOTAL	0.19	1.00	0.94	0.08	1.00	0.86

FBW = Flash Butt Weld
THW = Thermit Weld
GIJ = Glued Insulated Joint
R = Defect Rate: Number/100 km/10 MGT
Q = Accumulated tonnage in million gross tonnes [MGT]

Figure 10.91: Number of defects per 100 km per 10 MGT versus tonnage borne for CWR tracks on NS

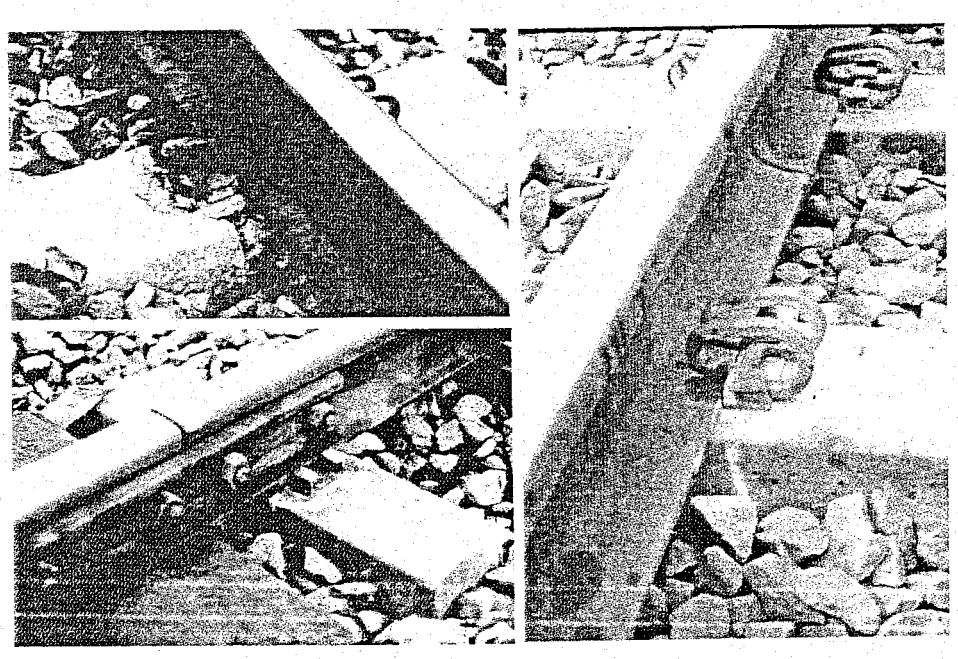


Figure 10.91a: Damage to track components due to poor weld and joint geometry

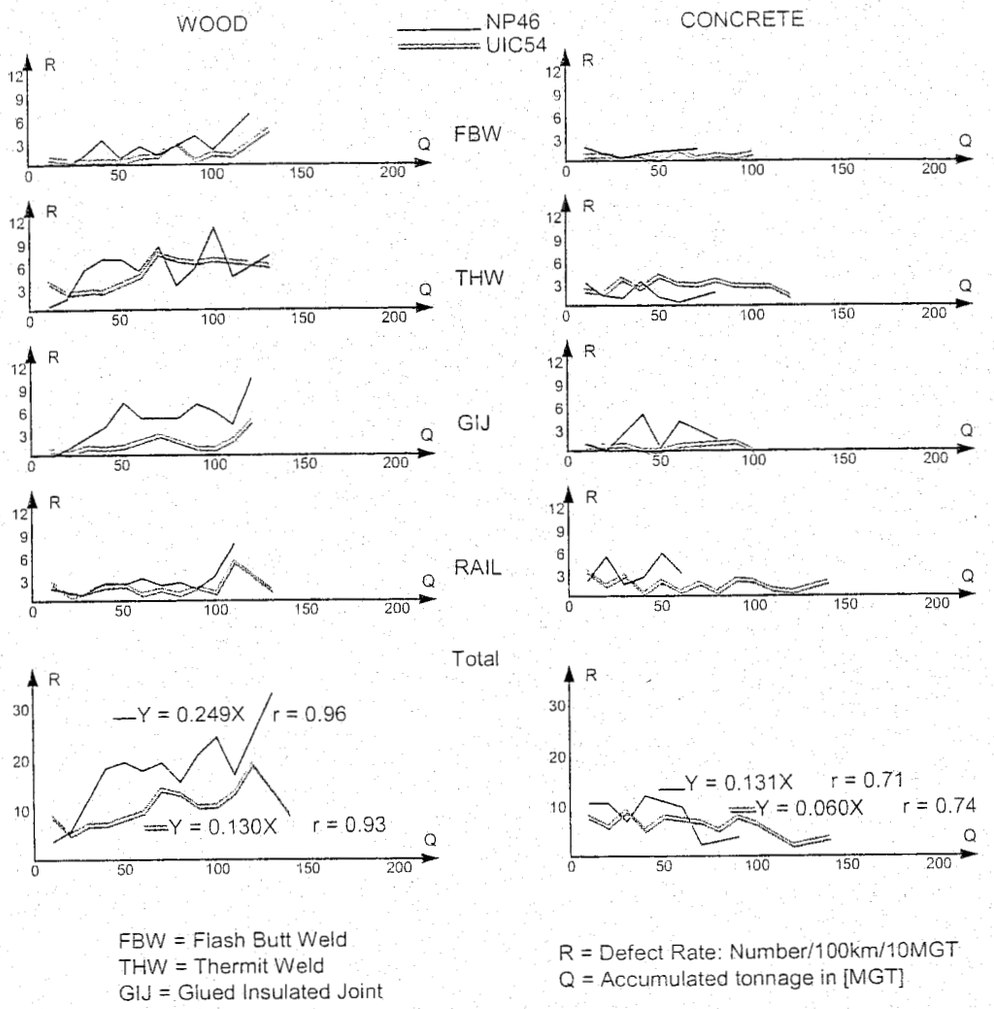


Figure 10.92: Number of defects per 100 km per 10 MGT versus tonnage borne for CWR tracks on NS

Experiences from UIC

In order to learn more about rail failures, several sites with newly laid tracks were monitored during the ORE investigation described in [219]. These sites were on PKP, MAV and CFR. In general, failures of aluminothermic welds, UIC code 421, were found to be the most common, accounting for 65 - 97 % of all failures (Figure 10.93).

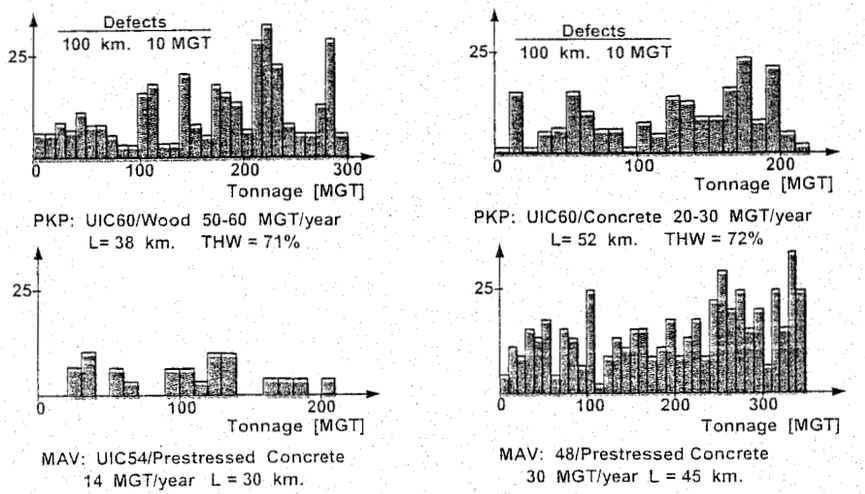


Figure 10.93: Rail defect rates from ORE D161 test sectionst

Also the cumulative probability of failure as a function of tonnage borne was investigated. In most cases this function was found to fit a Weibull distribution [9] described by:

$$F(Q) = 1 - (-Q^{\delta}/\theta) \tag{10.9}$$

$$\equiv \frac{Q^{\delta}}{\theta} \tag{10.10}$$

In this formula F(Q) is the cumulative probability of failure at a tonnage of Q million gross tonnes (MGT) and δ and θ are functions of the Weibull distribution. F(Q) is sometimes known as the unreliability or mortality. The approximation follows from a series expansion for small values of F. On a logarithmic scale the Weibull function becomes linear.

For example, a 50 km long test section on the PKP Katowice - Tczes line was found to have $\delta = 1.16$ and $\theta = 17380$. The Weibull distribution is presented in Figure 10.94.

Assuming there are 80 rails per km of track, the tonnage for any cumulative number of failures per km, k, can be found. A value of k = 6 failures/km is often chosen on PKP as an indication for renewal. In this case Q = 500 MGT. One should bear in mind, however, that over 80 % of the defects relate to Thermit welds. Weld geometry correction would therefore, in this case, be very advisable.

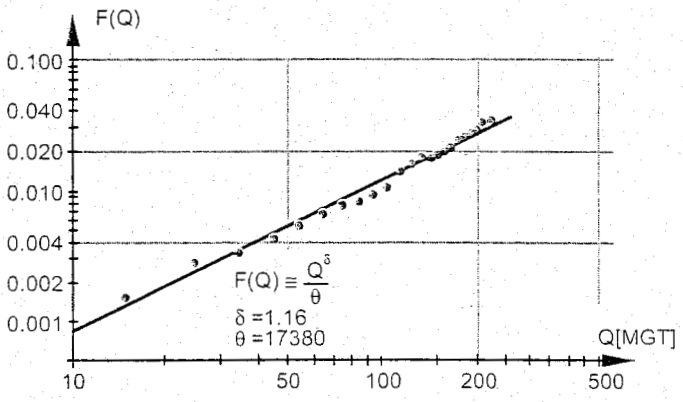


Figure 10.94: Cumulative defect occurrence on PKP test section and Weibull fit

To obtain good estimates of the Weibull parameters long sections of track should be considered. In the D 161 tests in fact only Thermit welds occurred in large enough numbers to be analysed. Subsequent analyses by BR, grouping rails together by age, have been summarised in Figure 10.95 and show that tache ovale, shelling, and flash butt weld failures also follow a Weibull distribution.

The cumulative number of failures is important when each failure is repaired by cutting out a section of rail and welding in a new piece. Every time this is done the geometry is worsened and two new potential positions for failure are created.

If care is taken of the repair then these effects can be minimised. Some types of failure, welds in particular, can be repaired with a wide gap Thermit weld process. In this case it is more important to establish the expected rate of failure, as is common practice on NS rail.

The local failure rate is calculated using:

$$\lambda(Q) = \frac{dF(Q)}{dQ}[1 - f(Q)]^{-1} = \frac{\delta}{\theta}Q^{\delta - 1} \tag{10.11}$$

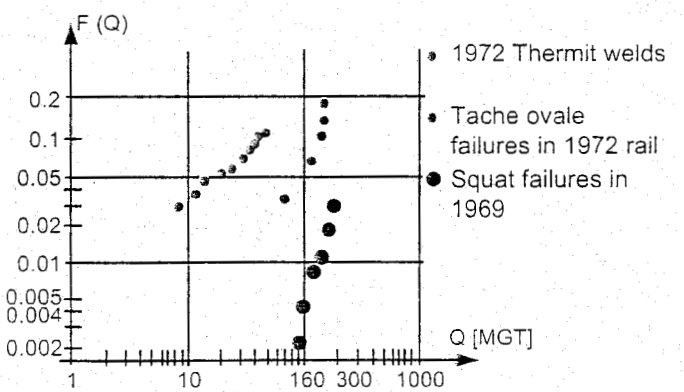


Figure 10.95: Mortality curves for some types of failure from BR sections

For the PKP test section mentioned previously the proportion of failures expected per million gross tonnes is:

$$\lambda = 6.6710^{-5}Q^{0.16} \tag{10.12}$$

Some administrations, such as NS, use the number of failures expected per 100 km in a 10 MGT band (N) as a parameter for deciding on rail renewal. In this example:

$$N = 5.34Q^{0.16} \tag{10.13}$$

Thus at 200 MGT N = 12.5 and at 500 MGT N = 14.4. It can be seen that only a small increase is expected in the rate of failure after more than twice the tonnage. This statement has been found to be true for many of the failure statistics studied by D 161. It implies that there is no point in renewing the rail while the local failure rate is low and constant. At very high tonnages the failure rate will rise as fatigue damage accumulates. This would be the time to consider renewal of the rail.

When predicting the cumulative number of failures and the failure rate from a Weibull function fitted to measured data, it is unwise to predict too far into the future. As a rule the period of monitoring should be extended by no more than 20 %.

11 SWITCHES AND CROSSINGS

11.1 The standard turnout

Turnouts are used to divide a track into two, sometimes three tracks. The purpose of crossings is to allow two tracks to intersect at the same level. If a complete train is to pass from one track to another while moving and without being subdivided, turnouts are essential in the absence of turntables or traversers.

It must be possible to run through switches and crossings in both directions. A normal or single turnout, as shown in Figure 11.1, allows movement of traffic in a straight direction on the through track or in a divergent direction. A picture of the right-hand turnout is given in Figure 11.2.

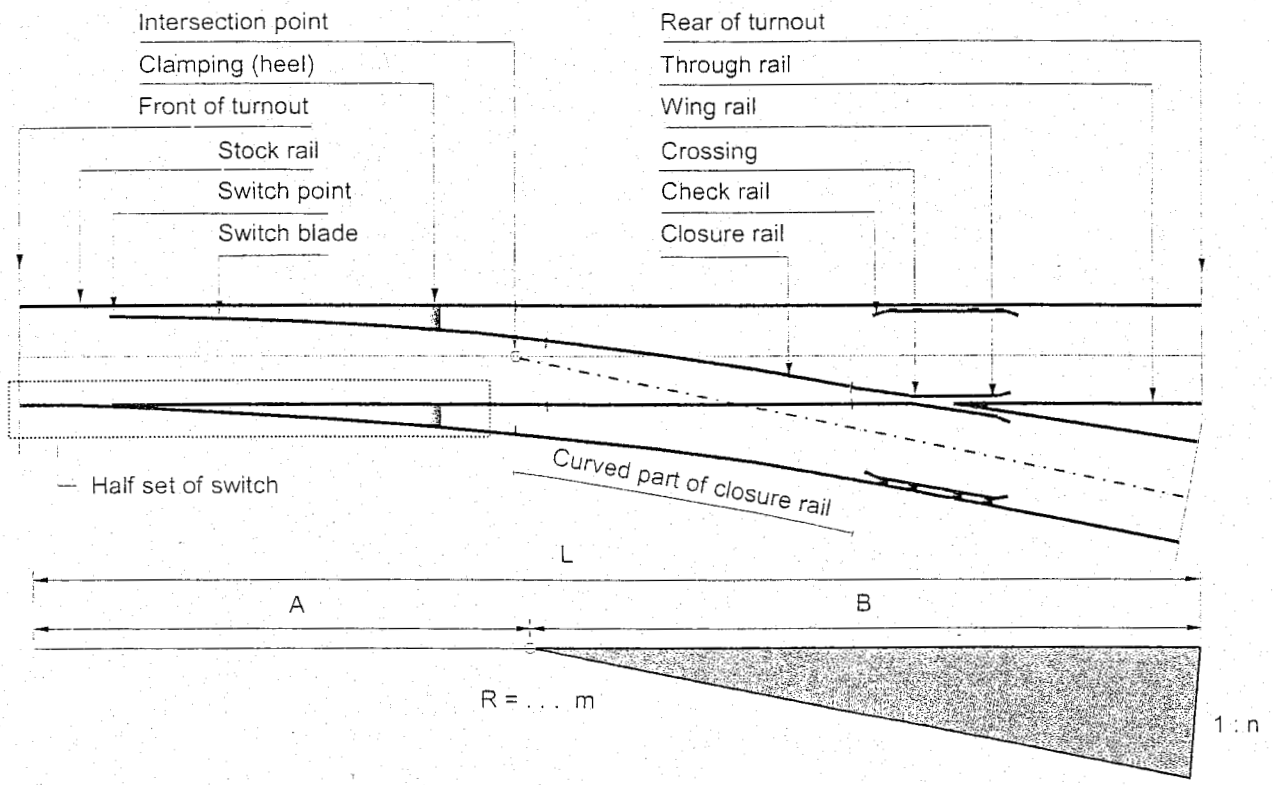


Figure 11.1: Standard right-hand turnout

The turnout consists of three major parts:

- Set of switches (switch blades);
- Common crossing;
- Closure rail.

These parts will be discussed separately below.

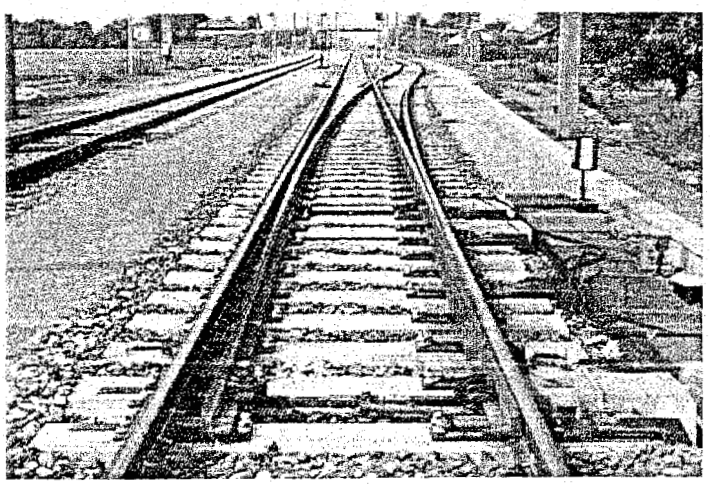


Figure 11.2: Picture of right-hand turnout

11.1.1 Set of switches

Switches consist of two switch blades and two stock rails. The switch blades can be moved and determine which of the above-mentioned tracks will carry traffic. In Figure 11.1 this is the through track.

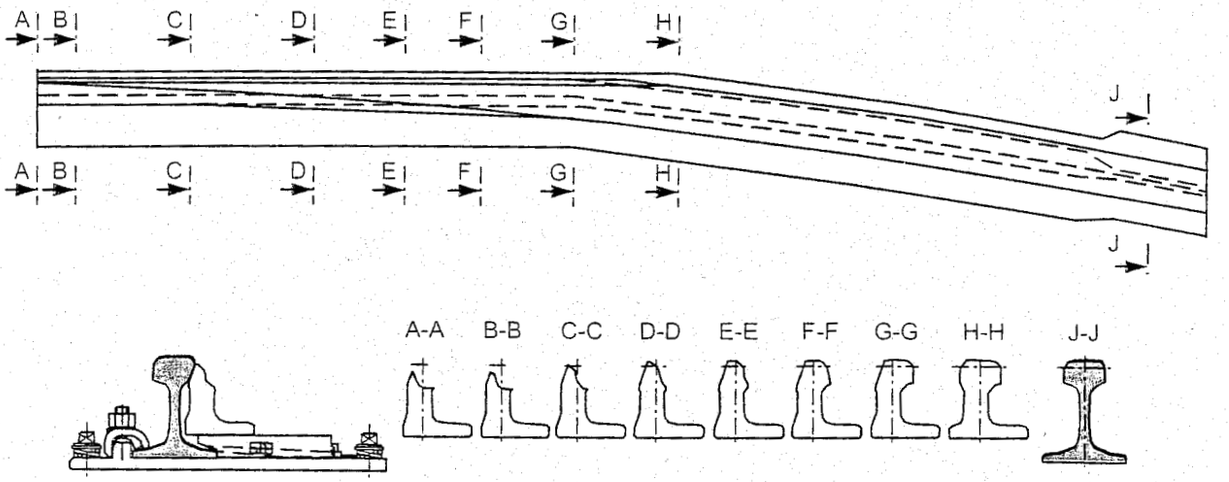


Figure 11.3: Cross-sectional drawing of switch blade and stock rail

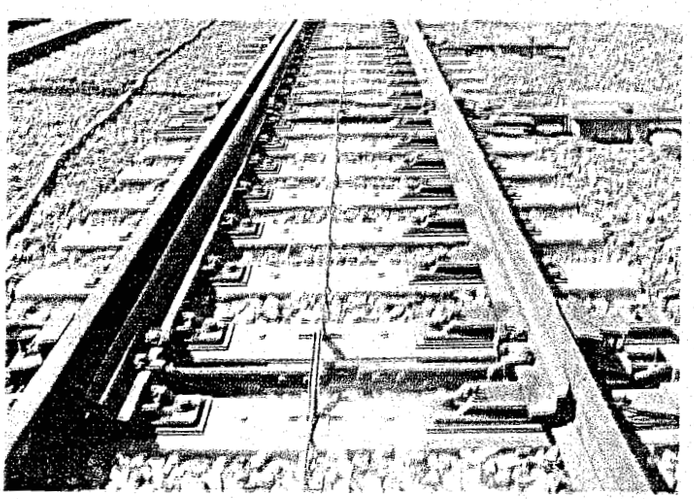


Figure 11.4: Switch blade and stock rail

The cross-section of the switch blade in modern designs is an asymmetric section that is lower than the standard rail profile. This has the advantage that there very little machining of the base of the switch is necessary.

Because of the asymmetric base, the moment of inertia is higher compared to a switch blade made of standard rail. The lower height allows the use of an elastic fastening system for the stock rail on both sides which is a must in modern turnouts. Figure 11.3 and Figure 11.4 show an example.

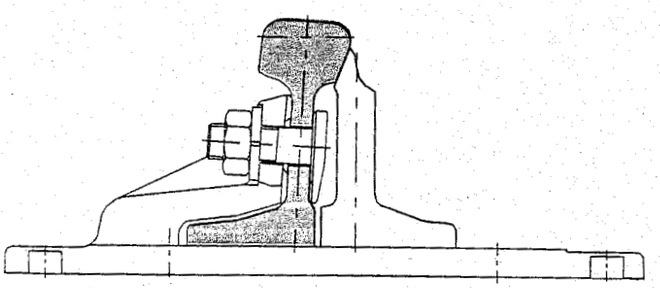


Figure 11.5: Cross-sectional drawing of T-rail switch blade

Some railways still use switch blades made of standard rails. Figure 11.5 shows the cross-section for such a turnout.

The turnout can be operated by different types of point machines, e.g. electrically, hydraulically or pneumatically. The locking system can be either in the switch machine (internal locking) or in the track (external locking). In switches for medium and especially for high speed several locking locations are necessary.

- A familiar construction is the built-up common crossing formed from two machined rail elements. If standard steel quality is used the common crossings are hardened by heat treatment in order to resist the high impact loads. Use is also made of head-hardened rails or rails made of austenitic manganese steel. This crossing is used for low end applications.

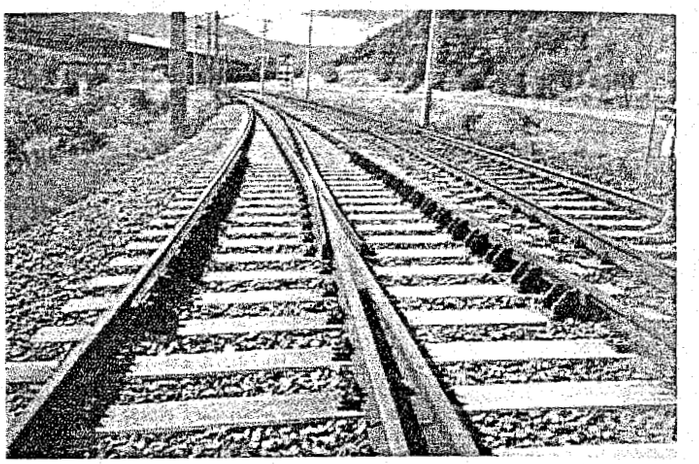


Figure 11.8: Cast manganese frog

- For higher loads crossings are used which have a crossing nose machined out of a block of heat treated steel and bolted on wing rails.

- An alternative to this built-up crossing is a crossing with a crossing nose made of cast Austenitic Manganese Steel (AMS) and bolted on wing rails. An Austrian development allows the welding of AMS to rail steel by use of an intermediate piece of special steel. This welding technique eliminates the bolted and fish-plated connection of olden days.

- If the traffic load is heavy preference is given to the monobloc AMS crossing with welded on legs on all four ends. This crossing is very wear resistant and can be welded in CWR track. Before installation these crossings can be pre-hardened by explosive hardening which reduces the initial wear and the need for maintenance considerably. Figure 11.8 shows a cast monobloc AMS crossing with welded on legs.

For higher axle loads and speeds of more than 200 km/h crossings with movable parts are used. In these crossings the gap between the wing rail and the point is closed and the impact load is reduced.

Three major types exist:

- Swing Nose Crossing SNC with a nose made out of a machined and heat treated block. Smaller crossings of this type have an expansion joint in the heel to compensate for the difference in length after switching from one position to the other.

- Movable Point Frog MPF with a nose made out of rails as main and counter point. Compensation for the length is achieved between the two points; no separate expansion joint is necessary. Figure 11.9 shows a crossing 1:38 installed on FS high-speed line.

- Crossing with movable wing rails. This type is used for small turnouts and when the length of the turnout is restricted.

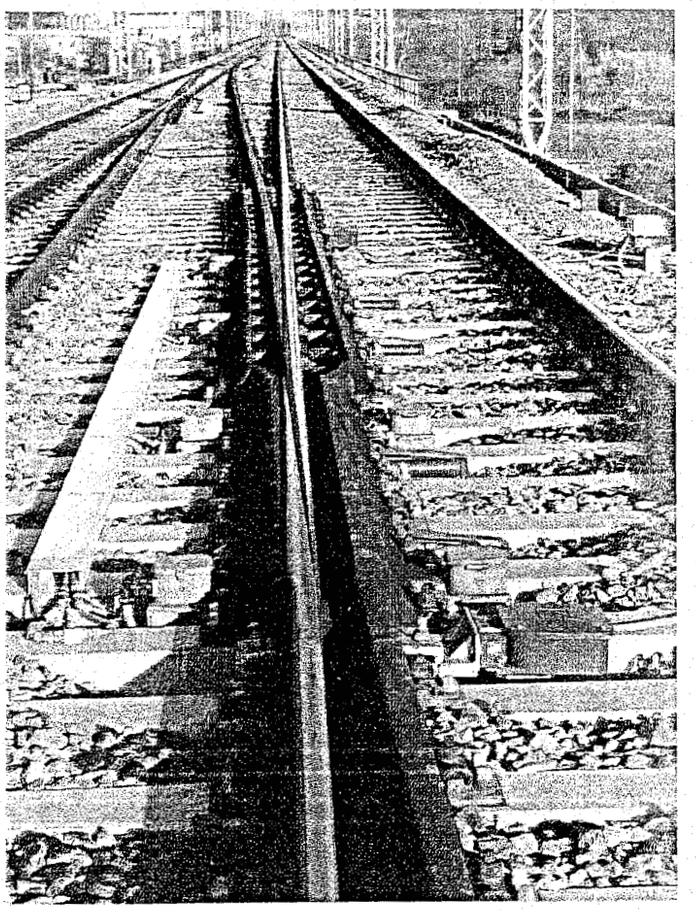


Figure 11.9: Movable Point Frog 1:38, FS

Check rails prevent the wheel from running into the wrong groove in the unguided part of the common crossing or from striking the crossing. Modern turnout design uses the UIC 33 profile for check rails that are attached to special plates and are not bolted to the running rail. This allows easy adjustment when the check rails wear.

11.1.3 Closure rail

The closure rail lies between the set of switches and the common crossing.

11.1.4 Rails and sleepers in turnouts

Turnouts are in principle constructed of normal rails. Special profiles are used for the switch blades and check rails. The steel grade of the rails usually is the same as on plain track. Special grades like Head Special Hardened HSH rails improve the wear characteristic of the turnout.

All rail joints within the turnout are welded joints or, if necessary, glued insulated joints. No bolted and fish-plated joints should be used in a modern design.

Normally long switch sleepers made of hardwood, also known as bearers, are used, although use is increasingly being made of concrete sleepers. Especially in main line and high-speed track much use is made of monobloc concrete sleepers or slab track.

11.2 Geometry of the turnout

Front of turnout: this lies in the centre of the rail joint on the side of the non-divided track;

Rear of turnout: this lies at corresponding points on the side of the divided track;

Intersection point (Mathematical point): this is defined as the point of intersection of the centre line of the straight track and the tangent to the centre line at the rear of the diverging track.

Crossing angle of the turnout: the above-mentioned centre line and tangent form an angle which is referred to as the crossing angle of the turnout. The same angle can also be seen in the common crossing. For the crossing angle 1:n whole numbers are chosen. Common values are 1:7, 1:9, 1:12, 1:14, 1:15, and 1:20.

Layout geometry of the turnout: The layout geometry of a turnout defines the speed that is possible by means of the diverging route of the turnout and the forces created between wheel and rail. A modern design uses a tangential beginning of the turnout curve. Today, mainly constant radius curves, clothoids, and a combination of these types of curves are used. With special simulation software it is possible to optimise the layout geometry for minimum forces and maximum comfort. (See also paragraph High Speed Turnout).

Transition geometry in the switch: In Germany, a special transition geometry from stock rail to switch point has been developed. This geometry called FAKOP or KGO helps the axles to steer into the diverging route and reduces the forces that are created between wheels and rails. It also increases the thickness of the switch blade and, therefore, increases the useful life. (See also paragraph High Speed Turnout).

Rail inclination in turnouts: The rails in the turnout can be placed vertically or inclined. In order to achieve a better curving behaviour a higher conicity is preferable and, therefore, vertical rails are recommended.
The transition to the normal rail inclination takes place away from the turnout by slightly twisting the rails.

11.3 High-speed turnouts

11.3.1 General

High-speed traffic is a new challenge both for the railway operators and the railway industry. Time and quality factors are steadily developing into the decisive elements in the acceptance and use of transport systems. The areas of social considerations, speed, and comfort are where the potential for the future success of the railways lies.

Components and materials like swing nose crossings, asymmetrical switch points, elastic inner fastening systems of stock rail, head special hardened rails, high wear-resistant manganese steel, vibration dampened fastening plates and concrete sleepers, integrated operating systems, lubrication, and maintenance free components should nowadays be part of any modern high-speed turnout.

11.3.2 Traditional turnout design method

Traditional turnout design methods assume that vehicle response is determined by kinematics rather than dynamics [123], [184]. For a given vehicle speed and assuming the vehicle to be a point mass, the lateral acceleration which is not compensated through the turnout can be calculated. Turnouts or crossovers - particularly for high-speed applications - are generally designed based on three parameters derived from the kinematic acceleration.

These are:

- Maximum lateral acceleration which is not compensated (m/s^2);
- Maximum rate of acceleration change (m/s^3);
- Maximum entry and exit jerk (m/s^3).

11.4 Vehicle dynamic

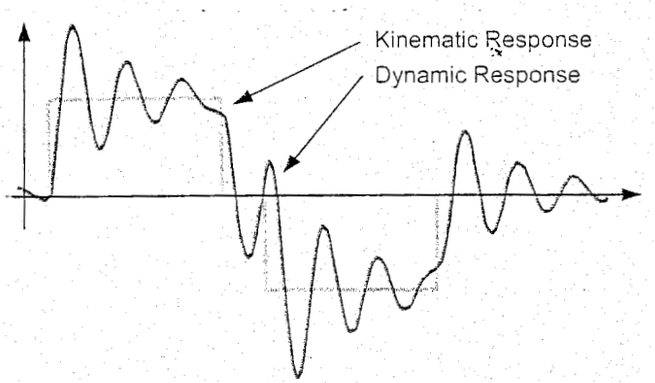


Figure 11.10: Kinematic versus Dynamic Response

A failing of the traditional design method is that vehicle response is clearly dynamic. As shown in Figure 11.10, simple kinematic considerations will underestimate actual response. Field measurements confirm that passenger compartment lateral accelerations in turnouts can go up to twice the kinematic values [291].

Recognising the shortcomings of the existing design approach, new methods for the design and evaluation of turnout geometry are now being used. Turnout design has to be considered as a vehicle/track interaction problem. The method is to use a vehicle dynamics simulation to predict forces and accelerations throughout the turnout. The optimum geometry usually is a double clothoid or vertex clothoid geometry.

One method to further reduce the forces that are created when a car enters a turnout and the wheel makes the transition from stock rail to switch blade is called "Kinematic Gauge Optimization = KGO". KGO makes better use of the rolling radius difference of the two wheels on an axle when it enters a turnout by bending the stock rail out. This increases the rolling radius difference and makes the axle steer in the desired direction. Calculated results have predicted an advantage compared to conventional turnout design.

Above-mentioned considerations and methods should form the prerequisites when developing new, advanced geometry for high-speed turnouts. Only then can we develop turnouts that meet the requirements for highest safety, best efficiency, and excellent riding comfort at the same time.

11.4.1 Examples of modern high-speed turnouts

Figure 11.11 shows a photograph of the FS Italia high-speed turnout with swing nose crossing for 160 km/h on the diverging track, geometry UIC60-7300/3000/15000-1:38-vertex clothoid, turnout geometry designed and optimised by using vehicle dynamics simulation. It was installed at Gallese, Italy on the Direttissima Rome / Firenze in 1999.

Figure 11.12 shows a photograph of a high-speed turnout with swing nose crossing for 200 km/h on the diverging track used on DB, geometry UIC60-16000/6100/infinit-1:48, switch including KGO design which is already a standard turnout at DB. Installed at Bitterfeld, Germany on the high speed line Berlin / Leipzig in 1998.

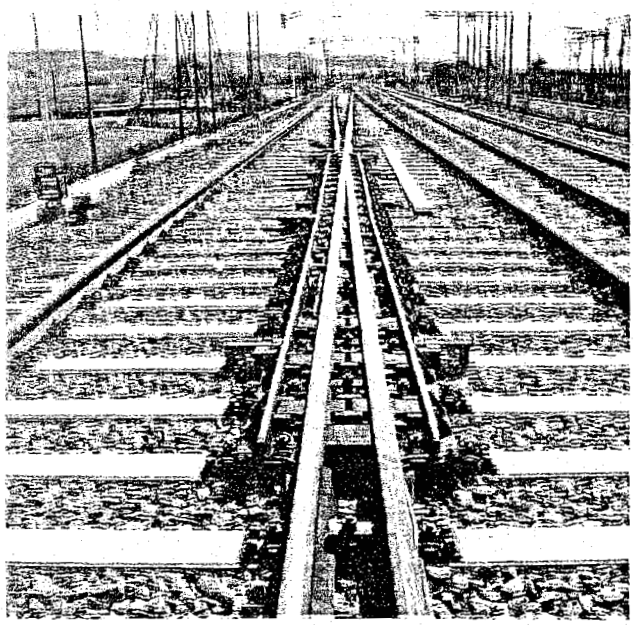


Figure 11.11: FS turnout with movable point crossing for 160 km/h

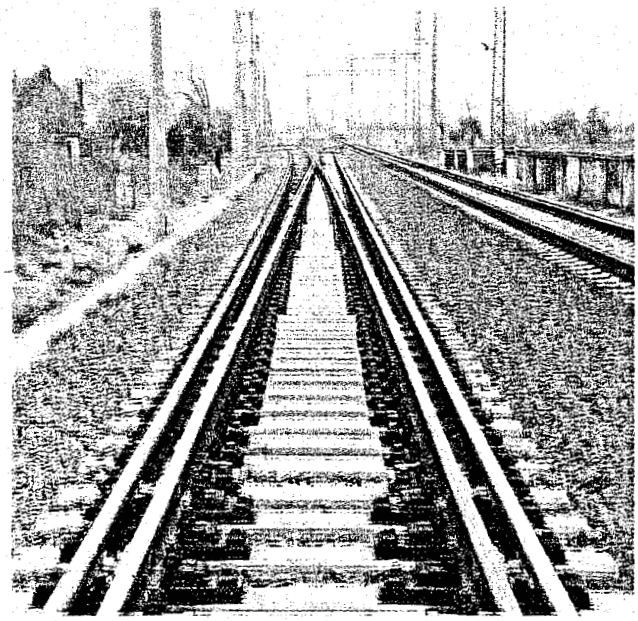


Figure 11.12: DB turnout with swing nose crossing for 200 km/h

Finally, Figure 11.13 shows a new track lay-out for High Speed crossovers, consisting of a circle combination of spiral segment-circle-spiral segment.

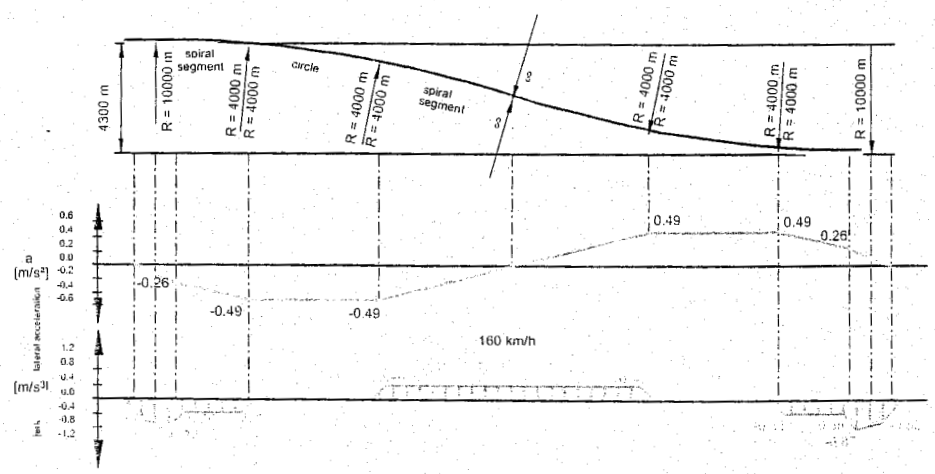


Figure 11.13: Typical diagram for a S - 160 km/h (100 mph) cross-over

11.5 Notations used for switches and crossings

Turnouts and crossings are indicated in different ways on drawings. The design drawing gives all the necessary details for the design and construction. The double line drawing with the two rails of a track drawn separately, an example of which is given in Figure 11.1, illustrates this principle.

There are also sketches in which only centre lines and tangents are shown. The mathematical point is indicated by a small circle. The most important details of the turnout or crossing can also be indicated.

11.6 Types of turnouts and crossings

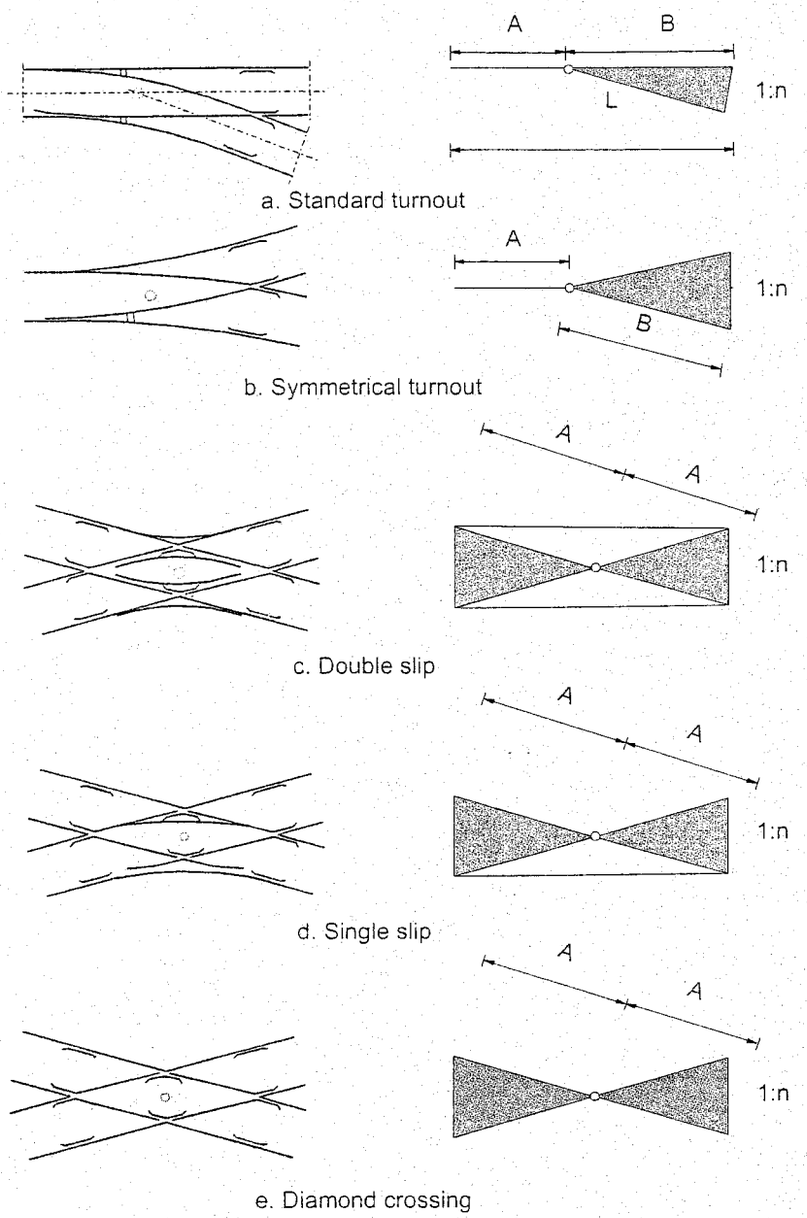


Figure 11.14: Types of switches

Using Figure 11.14 the most important types of turnouts and crossings will be discussed in more detail. In standard or single turnouts, a distinction is made between right-hand and left-hand turnouts depending on the direction the diverging track takes, seen against the direction of the switch point. The turnout shown in Figure 11.14a is a right-hand turnout.

In a curve the turnout can diverge in or against the direction of the curve (not drawn) with the straight track lying in a curve and the turnout track curved opposite the through track or in the same direction.
A symmetrical turnout (Figure 11.14b) is a special case of a turnout with two routes diverging symmetrically from the common route.

A diamond crossing with double slips consists of two diagonally intersecting tracks which are connected by two curves (Figure 11.14c). This means there are four routing possibilities.

A diamond crossing with a single slip is a crossing which only has one connection between intersecting routes (Figure 11.14d) and because of this has 3 routing possibilities. Reference is made to the difference in notation for diamond crossings with slips and crossings (Figure 11.14e).

A diamond crossing is considered normal if the angle of intersection is equal to that of a single turnout, and the intersecting tracks are straight as well (Figure 11.15f).

A crossing comprises two common crossings (acute angle of intersection) and two obtuse crossings (obtuse angle of intersection).

A curved diamond is a junction between straight and curved track or between two curved tracks (Figure 11.15g). The crossing angle is different at both ends.

Traffic running over right-angle crossings (Figure 11.15h) produces hefty impacts on the joints.

In the case of an oblique crossing (Figure 11.15i) this is much less the case. In crossings with a very small angle of intersection large unguided parts will occur. For values less than 1:12, the obtuse crossing is provided with moveable switch blades and the common crossing with a swing nose as outlined in Figure 11.15j.

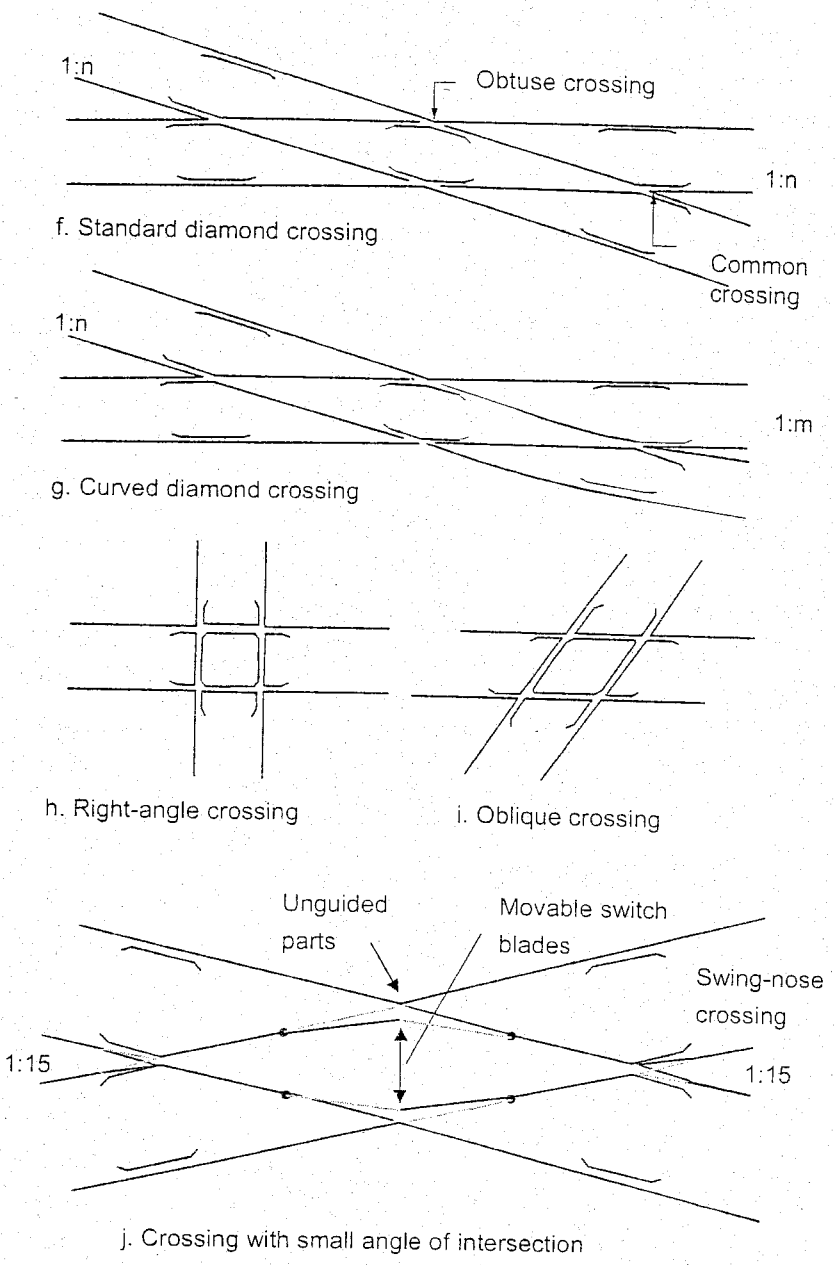


Figure 11.15: Types of crossings

11.7 Cross-overs

A sketch of the connection between two parallel tracks consisting of two single turnouts, together with the required total length, is given in Figure 11.16. Standard dimensions are used for the distance w between rails.

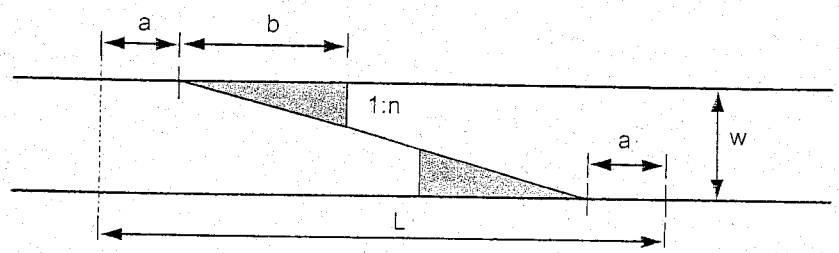


Figure 11.16: Switch connection between two parallel tracks

Connections between more than two parallel tracks can be made by joining up single connections according to figure Figure 11.17a, or by using diamond crossings with slips as shown in Figure 11.17b. System (a) is used as much as possible on main lines.

a:

b:

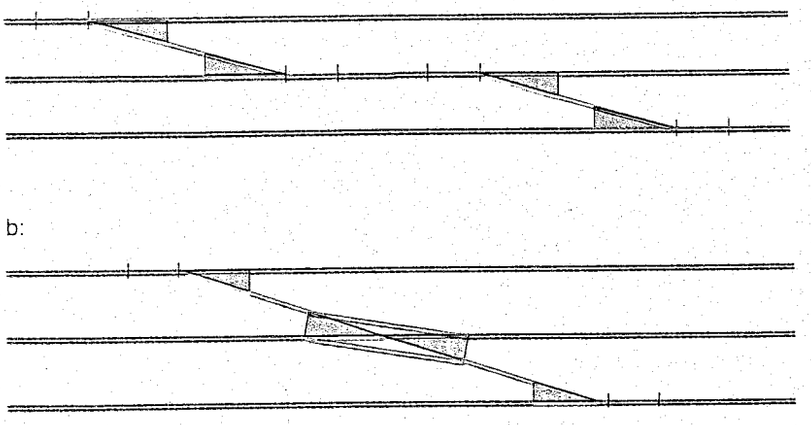


Figure 11.17: Switch connection between more than two parallel tracks

Double cross-overs allow traffic from both directions to change from one route to the other. Figure 11.18 shows two solutions of this type.

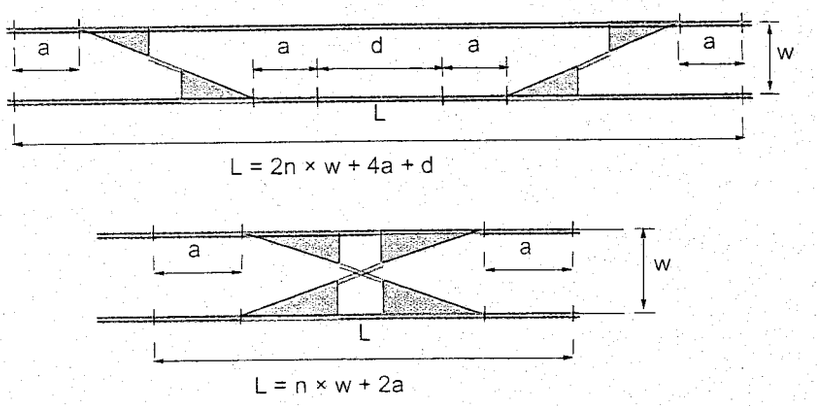
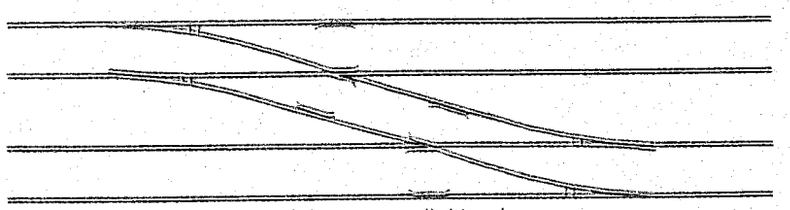
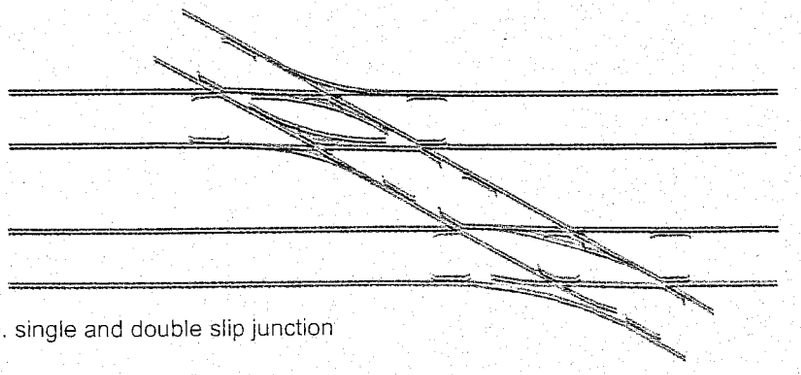


Figure 11.18: Double cross-overs

Figure 11.18 gives a few more examples of crossovers. If two parallel tracks are connected by means of a crossing which joins four turnouts to the main tracks, this produces a scissors cross-over as seen in Figure 11.18c and Figure 11.18d.

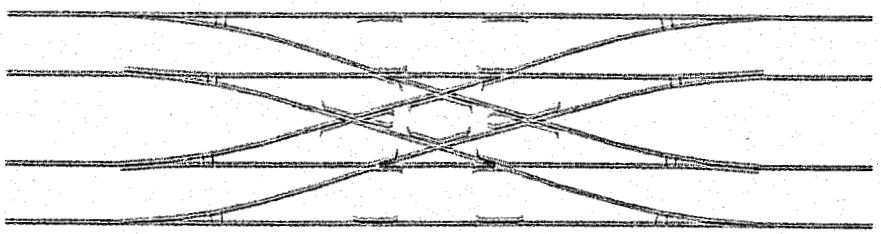


a. two turnouts combining to join two parallel tracks

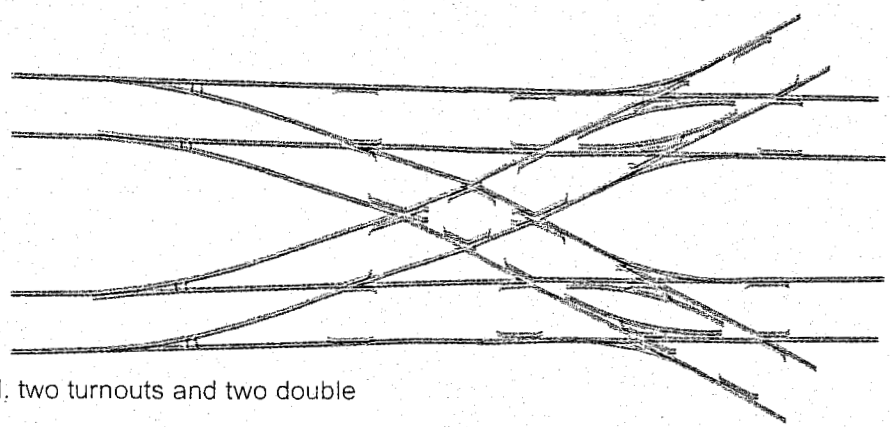


b. single and double slip junction

Figure 11.19: Crossovers



c. two intersecting crossovers with central diamond crossing



d. two turnouts and two double slips with central diamond crossing

Figure 11.19: Crossovers

Figure 11.20 shows an example of the switch and crossing layouts in Utrecht. This shows how complex these situations can be in practice and what immense problems of maintenance can result.

Figure 11.20: Switch layout in Utrecht

11.8 Switch calculation

11.8.1 Relation between curve radius and crossing angle

Under switch calculation we understand here the determination of the most important geometric measures. These are based on the admissible lateral acceleration a = 0.8 m/s² which was mentioned before. The connection between acceleration, speed, and curve radius is:

$$a = \frac{v^2}{R} \tag{11.1}$$

Therefore the maximum speed can be calculated as:

$$V_{max} = 2.79\sqrt{R} \tag{11.2}$$

with V[km/h] and R[m].

In Figure 11.21 a normal right-handed turnout is drawn with a straight crossing. If we assume that the outer curve is tangent with the stock rail at the point of contact, which is approximately correct, it follows from Figure 11.21 that:

$$d = R(1 - \cos\alpha) \tag{11.3}$$

in which:

d = indicates transition point curve to straight track, depending on constructional requirements;
R = radius of curve part closure rail;
α = crossing angle.

Assuming $\alpha << 1$, then $\cos\alpha \approx 1 - \frac{1}{2}\alpha^2$ and $\alpha \approx 1{:}n$, and (11.3) becomes:

$$R = 2dn^2 \tag{11.4}$$

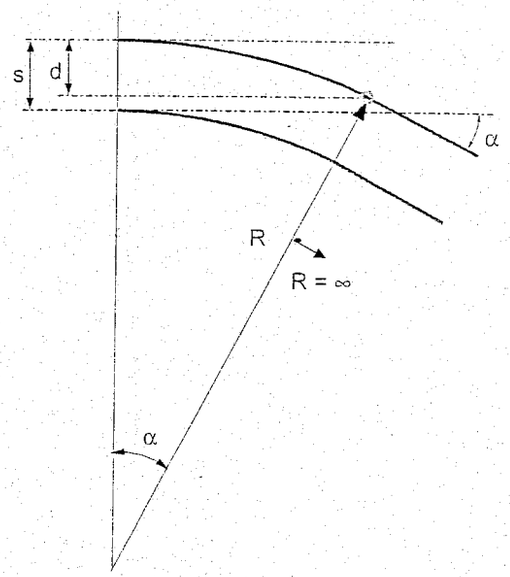


Figure 11.21: Relation between crossing angle and curve radius (straight crossing)

It should be noted that with a curved crossing the value of d is higher than the track gauge s (which is 1.435 m in standard gauge track). In normal situations the value of d in a straight crossing is about 1.25 m, whereas in a curved crossing this value is 1.70 m at most. These values are determined by the track gauge and the distance which is necessary to install the crucible of a thermit weld with a width of 300 mm. Taking this into account, it follows:

for a straight crossing with R in [m]:

$$R \approx 2.5n^2 \tag{11.5}$$

for a curved crossing with R in [m]:

$$R \approx 3.4n^2 \tag{11.6}$$

For the most common cases tables are used of the main dimensions of turnouts based on these calculations. The appropriate value of the maximum speed V_{max} (11.2) is also mentioned.

These tables show that even with small crossing angles the maximum speed is relatively low. A speed of 120 km/h requires a curve radius of about 200 m which corresponds with an angle of ca. 1:25. Such turnouts are difficult to manufacture and the length of the turnout will be rather high.

A symmetrical turnout is used in case of symmetrically diverging tracks. The application in switch connections between parallel tracks is unfavourable because of the creation of S-curve situations. From Figure 11.22 it can be easily deduced that the angle between both tracks is only half-effective with respect to the curve radius. According to (11.6), the relation between radius and angle reads for a symmetrical turnout:

$$R \approx 6.8n^2 \qquad (11.7)$$

This means that regarding equal angle the curve radius of a symmetrical turnout is twice that of a normal turnout. Of course, the maximum speed calculated in (11.2) applies to both tracks. The symmetrical turnouts 1:15 and 1:20, with radii 1200 m and 2000 m, can be utilised with speeds of 100 km/h and 125 km/h respectively.

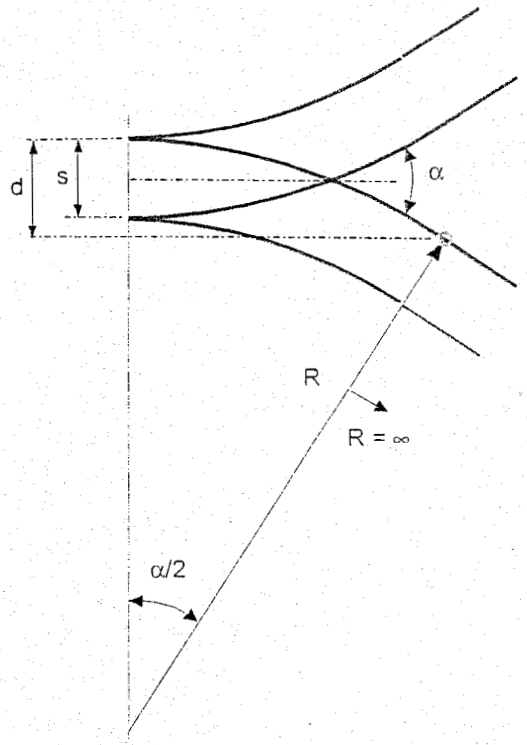


Figure 11.22: Relation between turnout angle and curve radius (symmetric turnout)

11.8.2 Calculation of main dimensions

In Figure 11.23 the double line drawing represents the straight and diverging track of a normal right-handed turnout. As can be seen, the switch blade is retracted over a distance CD with respect to the theoretical contact point C. Moreover, the switch blade is not tangent with the stock rail at point D, but makes a small angle because it is not constructively possible to make a perfect tangent.

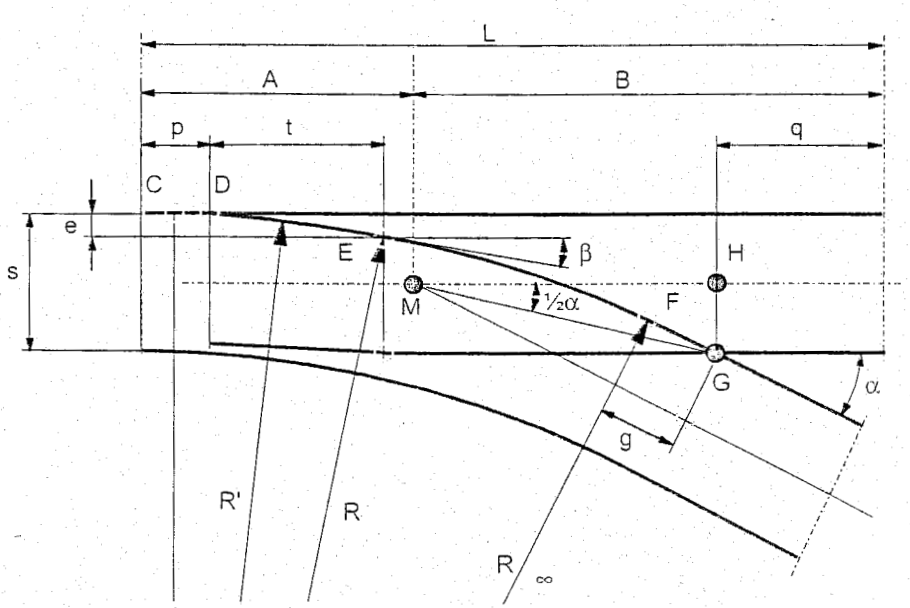


Figure 11.23: Calculation of curve radius and length of normal turnout

Both straight and curved switch blades are used. In case of a curved switch blade, the part a from D to the heel E is curved. In open position there should be enough space over the full length of the switch blade to allow a free passage of the wheel flange. This requirement determines the length of the switch blade.

The advantage of a straight switch blade, apart from its simple manufacture, is that the same design is used for left and right-handed turnouts. A curved switch requires less length and allows a larger curve radius. The full curve runs from E to F and ends at distance g before the theoretical point G of the crossing.

The following quantities depend on the dimensions of existing constructional parts and are supposed to be known:

α = Crossing angle;
β = Angle of tangent line at the heel (point E);
s = Track gauge;
e = Position of heel;
g = Length of straight part before theoretical point G;
p = Distance from front turnout to beginning of switch blade (point D);
q = Distance from beginning of crossing to rear of turnout;
t = Length of switch blade (projection on stock rail).

It follows from Figure 11.23 that:

$$s = e + R(\cos\beta - \cos\alpha) + g\sin\alpha \tag{11.8}$$

from which the curve radius can be determined:

$$R = \frac{s - e - g\sin\alpha}{\cos\beta - \cos\alpha} \tag{11.9}$$

The full length of the turnout can be calculated from:

$$L = p + t + R(\sin\alpha - \sin\beta) + g\cos\alpha + q \tag{11.10}$$

Finally, the location of the mathematical point follows from:

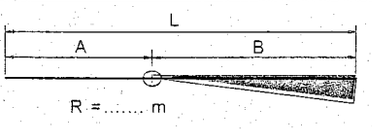
$$MH = \tfrac{1}{2}s\cot g\tfrac{1}{2}\alpha \tag{11.11}$$

The values of the quantities printed in the NS tables follow from the relations:

$$\begin{aligned} B &= MH + q \\ A &= L - B \end{aligned} \tag{11.12}$$

Table 11.1 lists some switch types commonly applied by the NS with their main dimensions.

Common turnouts



angle of intersection	construction	L [mm]	A mm]	B [mm]	R [m]	V [km/h]
1:34.7	UIC54 with movable crossing	99332	36878	62454	2300	140
1:15	UIC54 with curved crossing	47277	21221	26056	725	80
1:15	UIC54 with straight crossing	42706	16550	26056	600	70
1:12	UIC54 with curved crossing	38320	17420	20900	465	60
1:9	UIC54 with curved crossing	32185	14185	18000	260	40

Table 11.1: Survey of some common switches and crossings

11.8.3 Geometrical design of switches and crossings

In order that a switch and crossing layout can be manufactured, assembled, and laid into a track, every part of it must be very accurately specified in dimensional terms, usually on a drawing. Scales of 1:100 are used for this purpose, and 1:50 or even larger scales may be used on occasions. It is undesirable to attempt to scale off from even the largest scale drawing dimensions in order to achieve the necessary accuracy to obtain parts which properly fit together. Hence every line on the drawing must be defined by dimensions, and these must be confirmed by calculation. Only then can it be reasonably certain that the manufacturer will produce components which correspond with the engineer's intentions. It is to this end that it is necessary to define all layouts in geometrical terms.
Until recently, all this work had to be done by individual calculation, and, as a result, a substantial body of formulae has been developed over the years to make the task easier. Much of this technique has been made obsolete by the introduction of Computer Aided Design CAD).
An example of such a program is TURN which is used by the TU Delft in combination with the FEM package ALGOR.

11.9 Production, transport and laying of switches

Switches are prepared at the switch manufacturer. In Figure 11.24 the production of the middle of a cross-over in concrete is shown.

After preparation the switches will be transported to the definitive location either by car, ship or by special switch transport trains.

At the building site the switches can be handled by special heavy rail cranes or cranes from the road. An alternative is the use of mobile gantries on temporary rails or mobile gantries at retractable rail wheels.

After the switch is installed and the several parts are welded together one or more special tamping machines are used. The switch can now be connected to the signalling, heating and operation systems and tested for use.

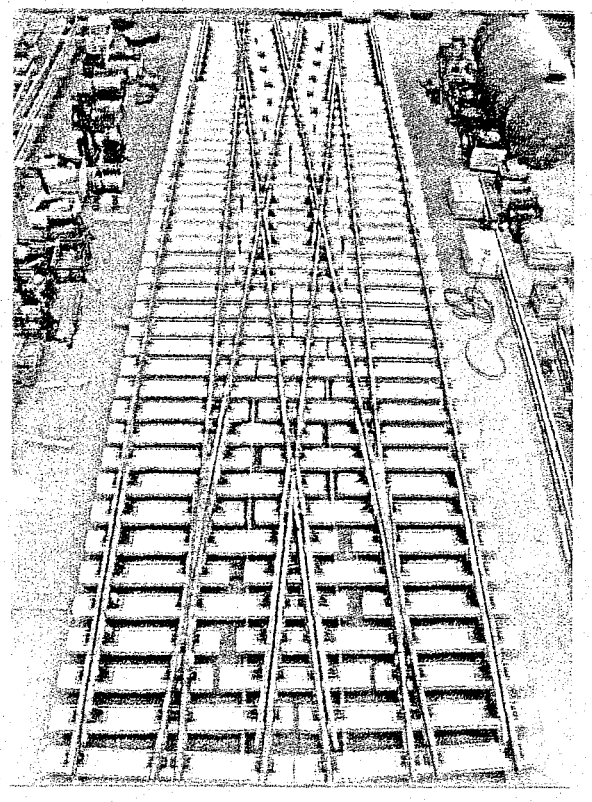


Figure 11.24: Assembling of a cross-over in concrete

Switches for trams and metros

For metros and tramways the principle of a switch is the same as for railway track. However, the following structural differences should be noted:

- The curve in switches for trams and metros can be much lower by the lower speed and bogie distance, so the switch length is shorter
- The switch can be integrated in the pavement of the road
- The railtype can be different by the lower load and speed.
- Sometimes switches for trams have no power or have simple engines; the blade will be moved by the tramwheels.
- The switches are operated from the tram driver by signal, not from an operating centre
- By tramways there are much more crossings at more different angles.

12 TRACK MAINTENANCE AND RENEWAL

12.1 Introduction

Track maintenance means the total process of maintenance and renewal required to ensure that the track meets safety and quality standards at minimum cost. Figure 12.1 gives a schematic summary of the various components, which go to make up the maintenance process. Annual maintenance on the NS network, with its 4500 km of main line tracks, comprises renewal of roughly 140 km of main line, 40 km of secondary tracks and sidings, 1000 km of mechanical tamping, 60 km of ballast cleaning, 10 km of corrective grinding and renewal of 250 switches. In addition to this the track requires spot maintenance on a daily basis.

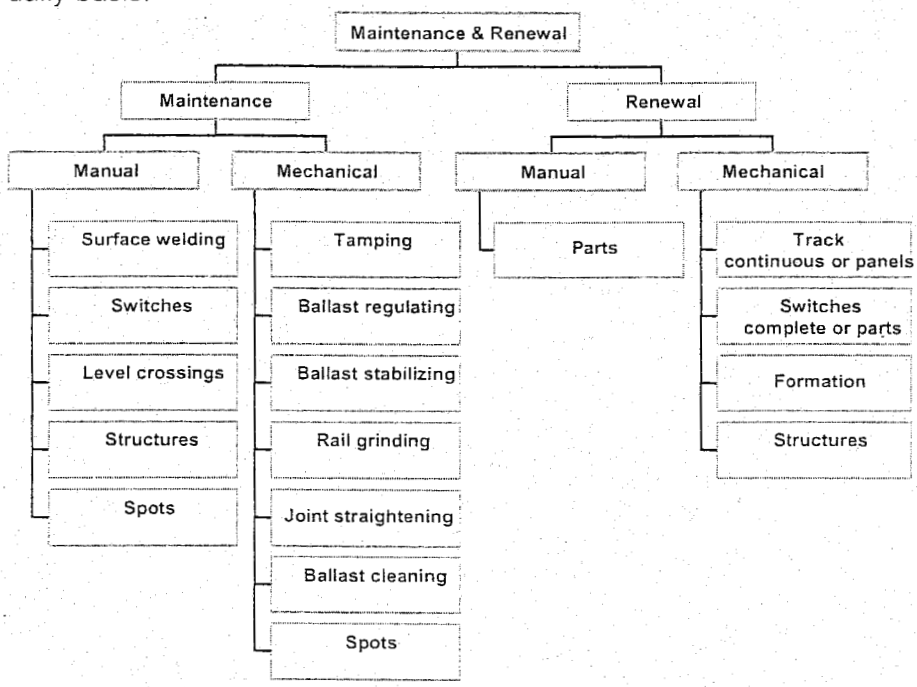


Figure 12.1: Schematic survey of maintenance and renewal process

Maintenance and renewal are in principle scheduled on the basis of control data from measuring systems, visual observation and financial-economic data, bearing in mind local conditions. The measuring systems are discussed in Chapter 16, and how the guidelines are obtained is explained in Chapter 18. This introduction concentrates on visual inspection and safety.

Visual inspection

The purpose of visual inspection is principally to check whether circumstances have arisen which may jeopardize safety of railway traffic. Inspection frequency varies depending on speed limit and daily train tonnage from a few times a week on the most important lines to once a month on the least important lines. Extra inspections are necessary in exceptional circumstances, such as very hot weather. Visual inspection becomes more and more supported by video inspection systems, which detect material faults by photo imaging (Chapter 16).

Safety

The braking distance of trains is much longer than that of cars or trams. Trains cannot be brought to a standstill in time if people or vehicles unexpectedly appear on the track. Similarly, it is not possible to halt traffic temporarily each time work is required on the track. This is why comprehensive stringent safety regulations apply to work within the structure gauge. Firstly the track must always be in a safe condition for approaching trains and secondly the safety of the track maintenance crews must be ensured.

12.2 General maintenance aspects

Track maintenance can be divided into:

- rail geometry;
- track geometry;
- track structures;
- ballast bed;
- level crossings;
- miscellaneous.

Maintenance of track geometry can in turn be subdivided into incidental (spot) maintenance, in other words, repair of local irregularities, and systematic maintenance which is carried out as a matter of course mainly with heavy track maintenance machines. The latter is referred to as mechanized maintenance.
Mechanized maintenance is carried out with:

- tamping machines: to correct level, cant and alignment;
- ballast regulators: to establish correct ballast profile;
- stabilizers: to compact ballast;
- rail-grinding machines: to remove corrugations and grind welds;
- STRAIT: to straighten welds;
- ballast cleaners: to clean ballast bed.

Current practice is to carry out maintenance exclusively according to requirement, in other words when indicated by measured data or observations. So-called maintenance sections, measuring 5 to 10 km each, should as far as possible be treated in their entirety.

12.3 Spot maintenance of track geometry

Incidental maintenance to repair local defects can be carried out manually, supported by small machines. Systematic maintenance is carried out as much as possible by heavy machines. The possibility of track possession plays an important part in this. The most important spot work consists in:

- levelling and tamping using vibrating compactors or tamping tines;
- measured shovel packing;
- spot tamping by special machines;
- rectifying track gauge.

Using jacks to raise the track and then filling the space with ballast material rectify defects in track level. This can be done using manual compactors, vibrating compactors or impact tampers. Often maintenance is required at level crossings or other short sections of track. The most important incidental maintenance consists in rectifying unacceptable twist.

Measured shovel packing is a method, which originated in France. The greatest advantage over tamping is that with measured shovel packing the ballast bed beneath the sleepers is not disturbed. In view of the labour-intensive nature of this method it is only used to a very limited extent. Besides, the fine stone is not suited to the coarse ballast.

An alternative is the hand-held stone blower, shown in Figure 12.2, using accurately determined quantities of special-grade ballast, which are blown under the sleepers. The amount of ballast is measured out on the basis of a recording of the track ahead using void meters and the AUTOGRAPH system [277], which is shown in Figure 12.97. With the aid of inclinometers, AUTOGRAPH measures the absolute level and cant of the track at every sleeper. An on-board minicomputer is used to log data. The on-board computer can be used for simple line side data processing such as the calculation of track quality in terms of standard deviations and profile plotting. Alternatively, the data can be downloaded onto an IBM compatible PC or mainframe-computer for further processing and analysis. By a complementary system comprising another three wheeled measuring unit and mid-chord versine measuring potentiometer, also alignment and gauge are measured.

Figure 12.2: Hand-held stone blower

Isolated local defects should already be corrected when first appearing; otherwise the dynamic forces of train traffic soon will create additional irregularities. Manual methods and simple or lightweight machines have only a limited capacity for durable elimination of such faults. Special heavy weight tamping machines therefore have been developed for spot maintenance (Figure 12.3).

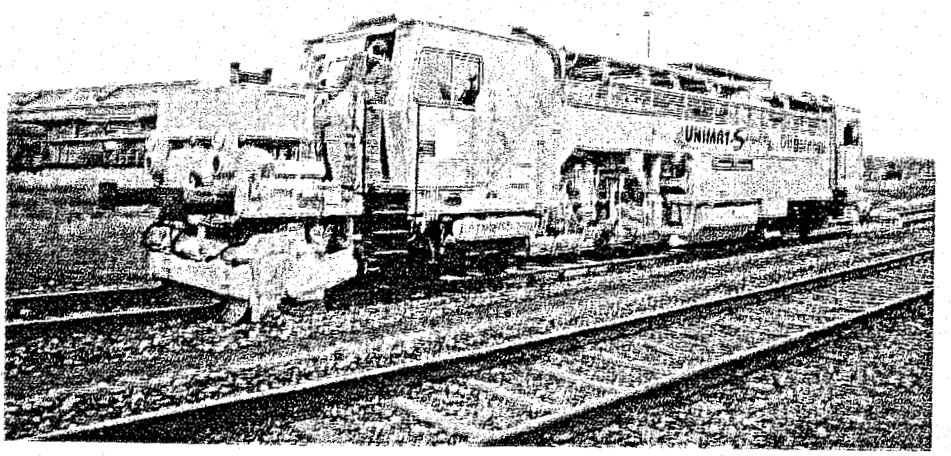


Figure 12.3: Spot maintenance machine as used by German Railway DB

The levelling- and lining system and the on-board geometry calculation computer are designed for spotting single irregularities and calculate correction values which correlate to the adjacent geometry values. Then, heavy-duty tamping units enable a durable correction of spot-fault. The tamping process is adjusted to single or double insertion and normal or high-pressure tamping, as desired by the lifting amount and compactness of ballast at every sleeper. 'Design overlift' to encounter the initial settlement is also possible. The machine has also a high travelling speed (100 km/h), so that the spots, which are scattered over the whole network, can be reached in short time and track possessions can be kept to a minimum.

12.4 Rail grinding and reprofiling

Irregularities in rail geometry can give rise to very high dynamic loads. These geometry defects partly occur during manufacturing of the rails (known as rolling defects), and partly during operation in the form of corrugations, some examples of which are given in Chapter 10. The only remedy for such defects is grinding. Long waves, as in rolling defects, are difficult to remove because a great deal of material has to be removed.

12.4.1 Rail grinding machines

Two grinding principles are applied to rails: rotating stones and stones oscillating longitudinally. In the Plasser GWM 250 rail-grinding machine pictured in Figure 12.4 the stones are placed in a 2.5 m long frame, which oscillates longitudinally; a detail is shown in Figure 12.5. As a result of this movement the rail material is abraded. The stones continually adapt to the mean shape of the railhead and this has a smoothing effect.

Figure 12.4: Rail grinding machine GWM 250

NS use the model GWM 220 in the main track to grind rails after renewal. Such preventive grinding stops or slows down the development of corrugations. The machine is also used in combination with STRAIT for grinding welds.

The system with rotating stones is used in rail-grinding machines, which are hired out worldwide by the Swiss firm Speno. This grinding principle is extremely effective. The stones are placed in units as depicted in Figure 12.6. The units can be pivoted, which enables the grinding action to be concentrated on specific areas of the railhead, permitting a streamlining of the process particularly on flattened or lipped rails. The pivoting principle is sketched in Figure 12.7.

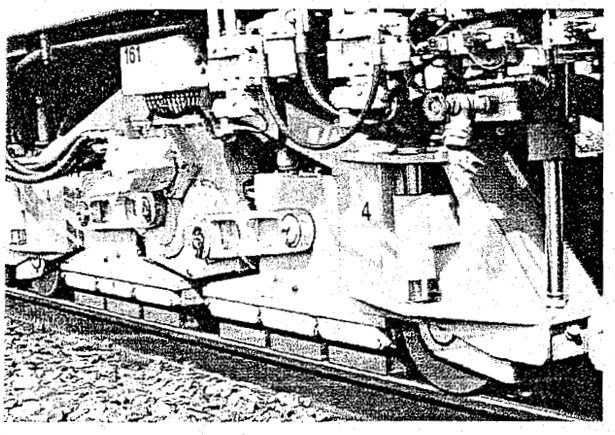


Figure 12.5: Grinding unit with oscillating stones

The stones form an angle such that the rail profile approximates to the form of a polygon. In this way the railhead can be reprofiled, which principally has enormous advantages in the case of plastic deformation and extensive corrugations.
Figure 12.8 shows a rail ground in this manner, the facets made by the various stones being clearly visible on the rail. These smooth areas very soon merge into a continuous profile.

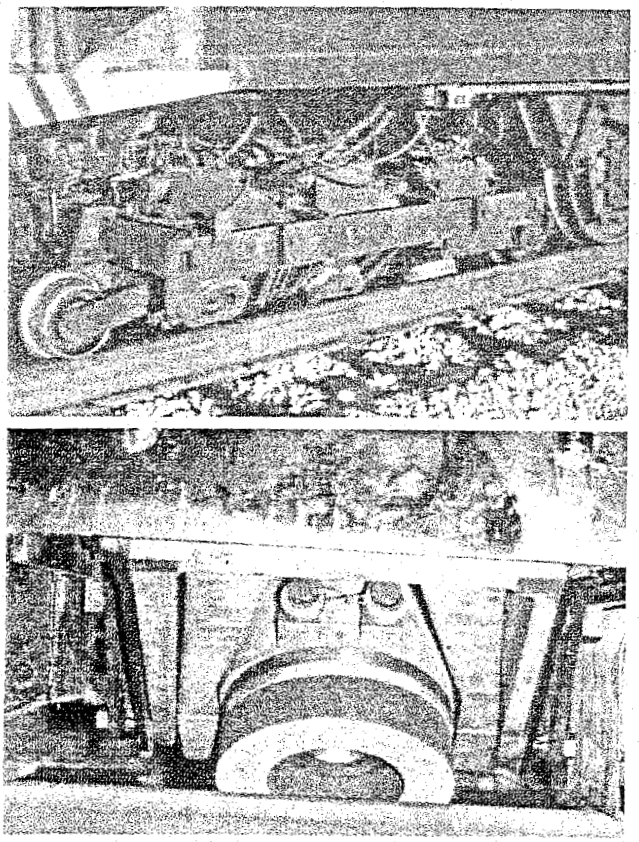


Figure 12.6: Grinding units with rotating stones

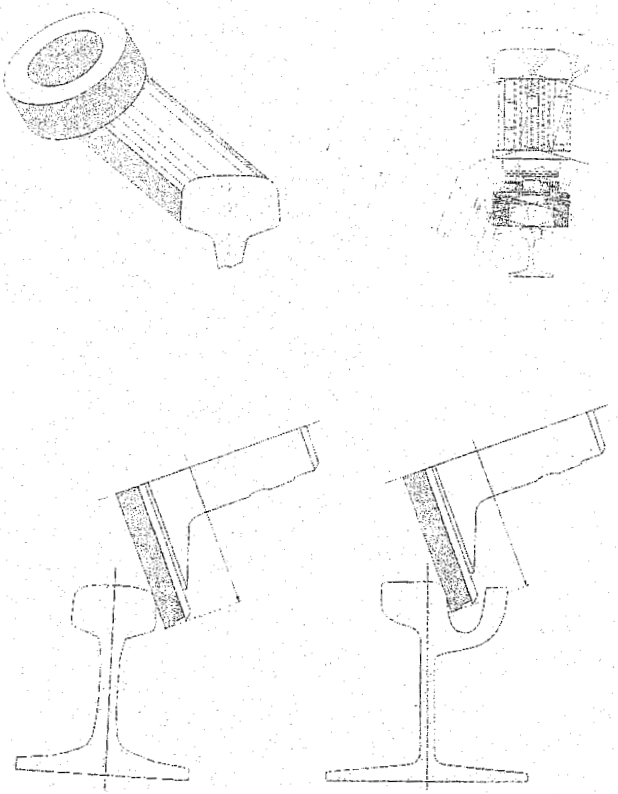


Figure 12.7: Principle of pivoting the Speno grinding units

A feature mainly applied on heavy haul railways is asymmetric grinding by means of which the wheel/rail contact point is shifted towards the inside of the high rail and towards the outside of the low rail. This gives better steering of the wheelset by which flanging is prevented or at least reduced, thus lessening the problem of side wear, severe corrugations and shelling. Figure 12.9 shows the profiles of high and low rail which were ground asymmetrically with the Speno train. The shift of the contact is clearly visible and tallies with the applied principle. For more details on asymmetric grinding please refer to [157].

Figure 12.8: As-ground rail with the different facets clearly visible

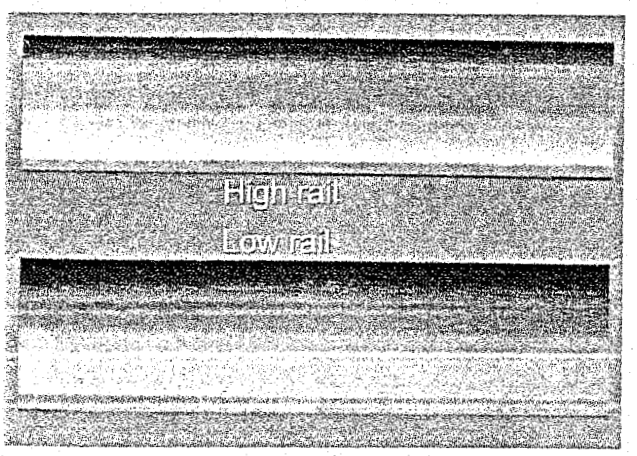


Figure 12.9: Asymmetric ground rail profiles

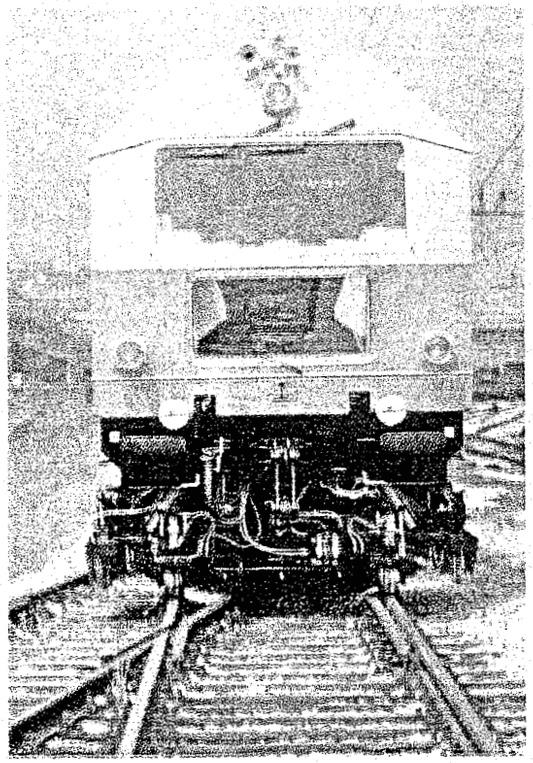


Figure 12.10: Speno RR 16 P/D switch grinder

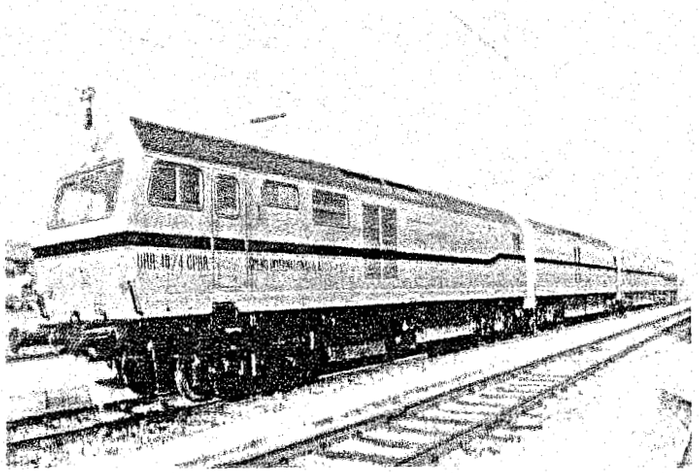


Figure 12.11: Speno 48 stone grinder

Figure 12.10 shows the RR 16 P/D switch grinder. Specially designed grinding trolleys allow full reprofiling of the gauge face of rails close to guardrails or switchblades.
Figure 12.11 presents a view of the 48 stone grinder, which has been delivered to Chinese Railways.

12.4.2 Rail reprofiling machines

If the rail show severe profile deformation (e.g. heavy haul traffic) or if the profile has to be modified, the application of rail planning machines is advantageous because they can remove larger amounts of material in one pass. Figure 12.12 shows a deformed railhead and Figure 12.13 the same after reprofiling.

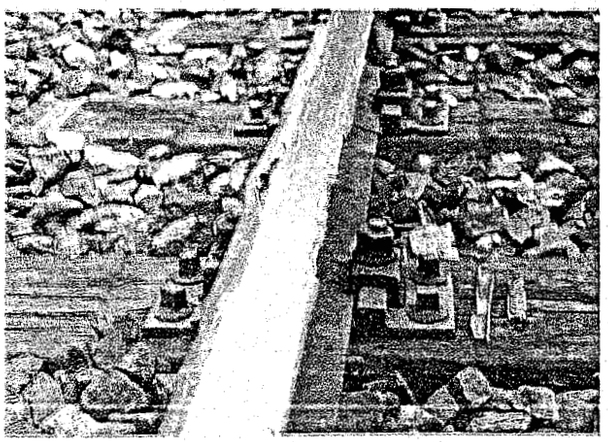


Figure 12.12: Overlap on railhead

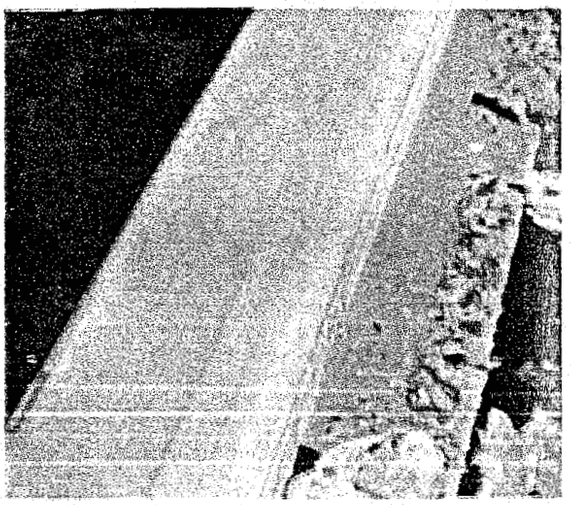


Figure 12.13: Planed rail

The rail planing machine is a four axle heavy weight machine which carries the planning units between the axles and a swarf collector at the rear end (Figure 12.14).

The Planing Unit consists of a tool support and a tool carrier holding a planing cutter. The tool carrier is operated with quick release fastenings (Figure 12.15).

The planing unit is guided along and hydraulically pressed against the rails by vertical and horizontal guiding rollers.

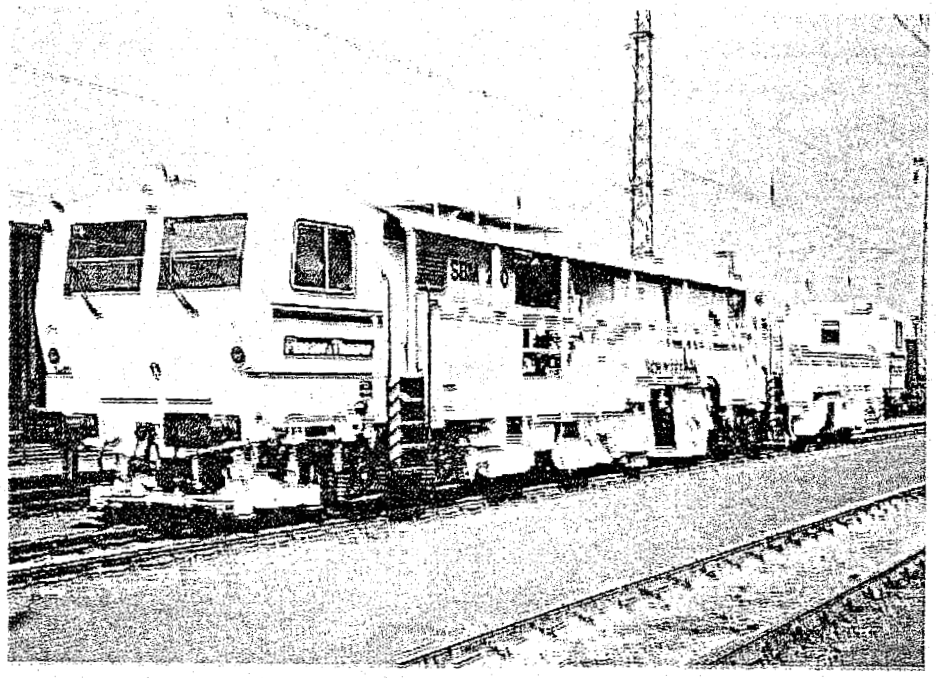


Figure 12.14: Rail planing machine

The number and positioning of the guide rollers ensures that on an uneven rail surface the planing cutter still always performs a straight longitudinal movement. The vertical and horizontal adjustment of the planing cutters is performed hydraulically. The adjustment values are set using a digital potentiometer.

High-pressure jets are positioned in the frame of the planing unit so that it is possible to cool the planed surfaces. This means that the service life of the planing cutters can be well used.

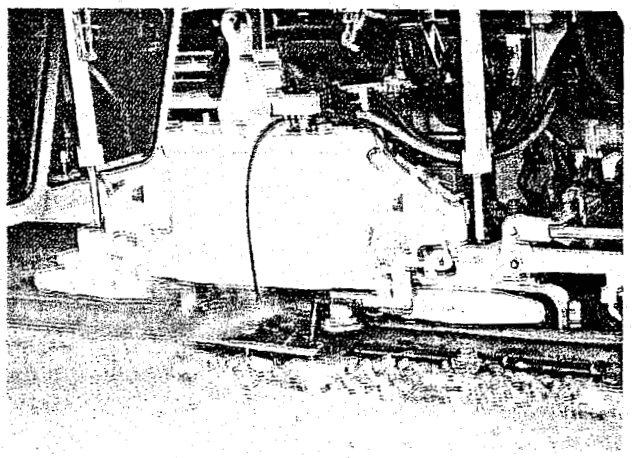


Figure 12.15: Rail planing unit

The re-profiling of the rails by the planing machine is performed in several working passes (Figure 12.16).

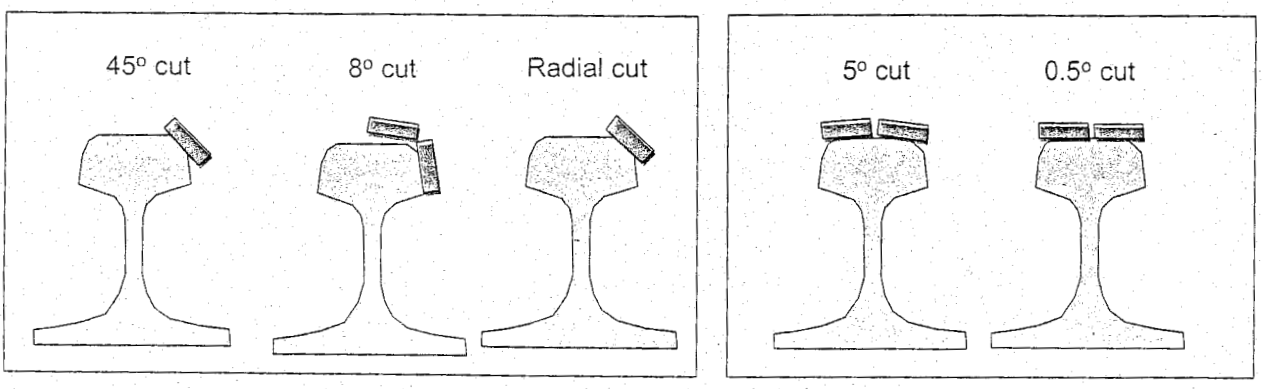


Figure 12.16: Cutting the rail profile in several passes

12.5 Correcting weld geometry

12.5.1 STRAIT principle

Structural changes which occur during welding of rails in the heat-affected zone lead to relatively large spread in the material characteristics and consequently in the relation between force and displacement. What is more, this relation is highly non-linear. Therefore a method has been developed which is entirely independent of profile and material characteristics and based on an iterative straightening principle referred to as STRAIT (Straightening of Rail welds by Automated Iteration Techniques) [81], [82].

Figure 12.17 shows the iteration principle of the mobile system, which is incorporated in a Plasser and Theurer tamping machine. The bending operation is controlled on the basis of two displacement transducers spaced 0.2 m apart, which are connected by a measuring frame 1.20 m long. The difference in the two transducer values represents the step, which is an essential item of information for the subsequent grinding process.
The transducer with the highest displacement value in absolute terms at the start is used for the actual controlling of the iteration process. During the first iteration the weld is loaded almost to the yield point. For a 700 N/mm^2 grade this corresponds to an initial displacement or starting value of 4.5 mm whereas for higher tensile strengths a starting value of 5.5 mm can be used.

Once the starting value has been reached the load is reduced to zero. In the unloaded state STRAIT measures the difference in displacement with respect to the upper limit of the tolerance interval (0.5 mm), this being the maximum overlift which can occur after the straightening operation. This difference in displacement is added to the starting value and forms what can be considered as a new starting value for the second iteration. Next the load is again increased until this displacement has been achieved: this occurs with yield. After the load has been removed the old starting value is again increased by the difference in displacement in relation to the upper limit and the third iteration takes place. This procedure is repeated until displacement of the weld in the unloaded state exceeds the lower limit of the tolerance interval; for the mobile STRAIT system this value is 0.2 mm. As a rule, the process converges in 3 to 4 iterations towards the target value.

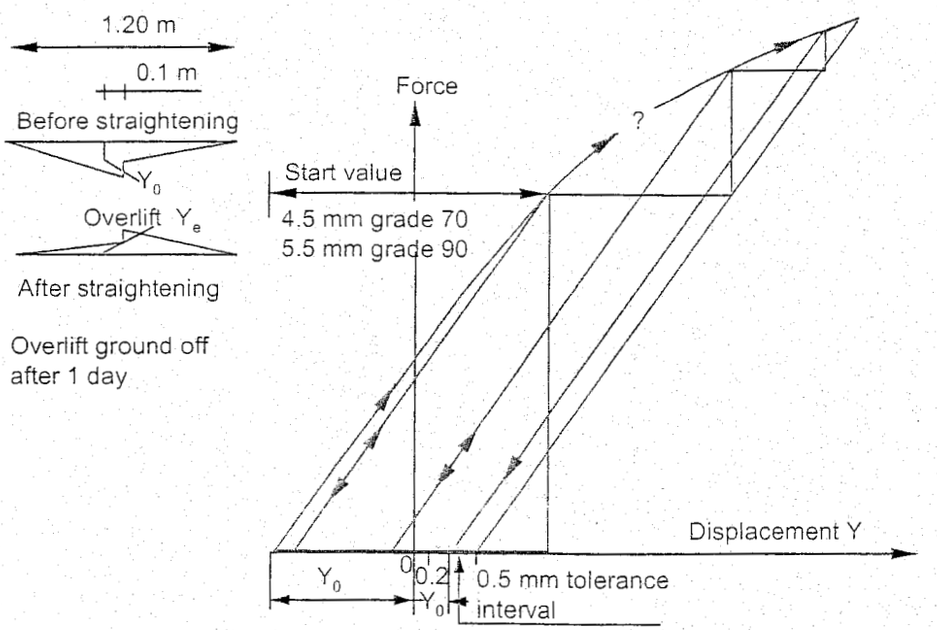


Figure 12.17: STRAIT principle

12.5.2 Mobile weld correction

Figure 12.18 shows the hydraulic straightening device incorporated in a tamping machine and consisting of a block provided at the extremities with a hydraulic cylinder, and a rail head claw in the centre for lifting the weld. This three-point bending system produces a maximum force in the railhead claw of about 600 kN when bending UIC 54/ 900. By using the alternatively available head claw, none of the ballast need to be excavated.

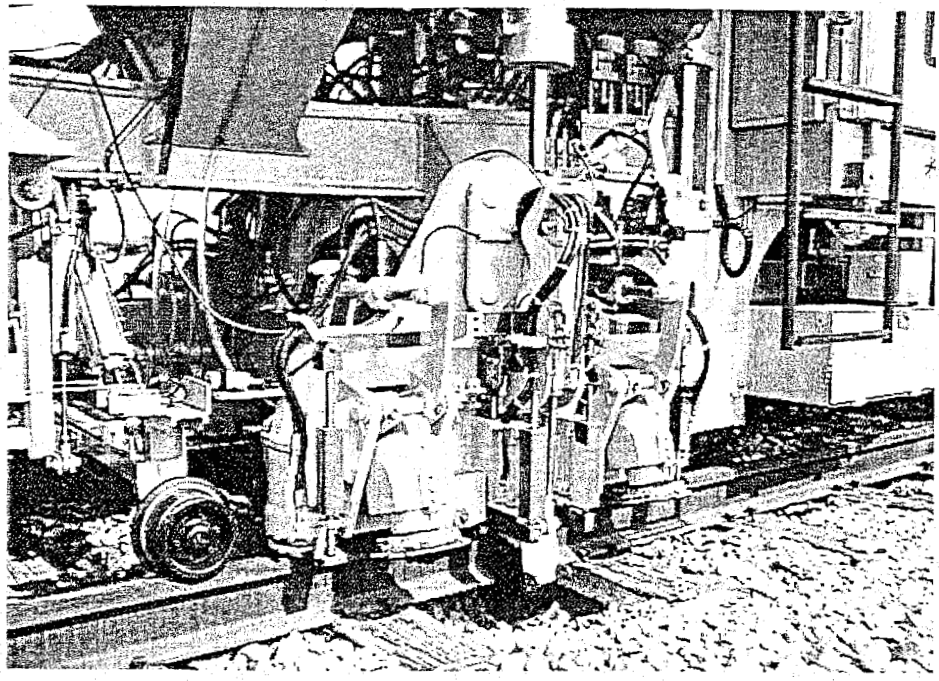


Figure 12.18: Straightening unit on STRAIT machine

Before straightening commences the measuring system must be calibrated. For this purpose the previously discussed electronic gauge with digital display is used.

To ensure durability it is necessary that directly after bending, the sleepers around the weld are tamped; this is done on 8 sleepers. After bending, resulting in an overlift of the weld of some millimetres, the surplus material and particularly the steps are removed by grinding. The Plasser GWM is used for this. To smooth the weld geometry the machine moves back and forth during grinding of a weld at an amplitude of about 1.5 m. With the exception of extreme situations a grinding period of 2 minutes is sufficient.

Correction of track welds must be seen as one coherent process consisting of straightening, tamping the sleepers and grinding. Weld correction of this type, together with preventive grinding of the entire track in three passes is standard practice on NS following renewal. The underlying principle can be summarized as the reduction of dynamic load by removing all geometrical imperfections in rail and weld geometry, which not only slows down development of corrugations but also deterioration of track geometry and track component quality.

The crucial factor in assessing the effectiveness of weld correction is the durability of the result for a given tonnage. NS has corrected thousands of welds and no plastic deformation has since been observed on these welds [82]. Ultrasonic testing has confirmed that the bending process does not give rise to any metallurgical defects.

12.6 Tamping machines

12.6.1 General considerations

Track geometry maintenance until 1945 was manual labour. Patents were applied, however, starting from 1896, proposing mechanical means to replace the strenuous efforts, which kept armies of workers engaged. Seen from today neither the technical principles nor the available steam engines allowed for a technical break-through.

It was not before 1933 when Mr. August Scheuchzer, a Swiss contractor, designed a successful purely mechanical mechanism to carry out mechanized track tamping. Correction of track geometry had still to be done by accompanying gangs manually, with manually operated rail-lifters set in respective distances to hold the track in its improved position until it was fixed by tamping ballast under the sleepers.

With the presentation of the first hydraulic track tamping machine in 1953 by the two Austrians Franz Plasser and Josef Theurer a development started in mechanization and rationalization of track maintenance, which is still ongoing today. From the beginning the aim was to improve work performance (speed and quality) by innovative machine solutions – like the hydraulic, non-synchronous tamping system - and thereby reduce the costs of the process.

The most important improvements of the basic idea are only briefly touched:

- Introduction of and combination with levelling systems;
- Combination with ballast regulators and compactors to a mechanized maintenance train (Figure 12.19);
- Invention of two-sleeper-tamping;
- Combination with lining systems to combined levelling-lining-and tamping-machines;
- Numerous improvements to upgrade work-speed, accuracy and operating comfort;
- Computerization of levelling and lining, including precalculation of lifts and slues after initial measurements;
- Introduction of dynamic track stabilization;
- Invention of the continuous action tampers (Figure 12.20);
- And as the latest step so far the development of the 3-sleeper-tamping mechanism.

Current practice is to maintain track almost exclusively using tamping machines, which correct level, cant and alignment. Irregularities in the track geometry with a wavelength of up to 20 to 30 m are automatically smoothed on the basis of an in-built measuring system. To correct long waves the machines use laser. In curves, external control data is necessary; for this reason facilities such as the on-board computer ALC and special track geometry survey machines are available.

Under certain circumstances special systems such as the pneumatic ballast injector, or "stone blower", developed in England, are used.

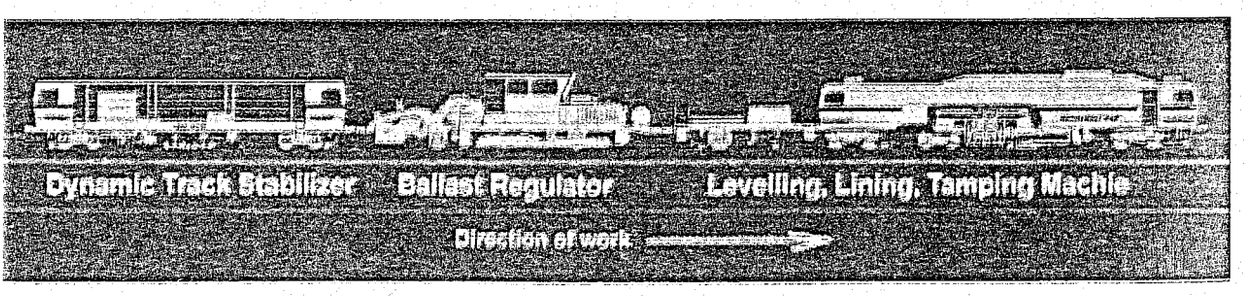


Figure 12.19: Mechanized Track Maintenance Train MDZ

12.6.2 Tamping principle

The tamping machine lifts the track up to the level determined by the measuring system and also positions it laterally; then the ballast under the sleepers is squeezed using so-called tamping tines. Figure 12.21 shows the principle of tamping with a Duomatic machine in which two sleepers are tamped simultaneously. The principle of a single tamping unit is shown in Figure 12.22. The essential features of tamping machines are tamping quality and tamping output. The tamping is non-synchronous with directional vibration. Non-synchronous means that each tamping tine is moved with the same force, independently of the path. This ensures that uniform consolidation is achieved. The ballast bed underneath each sleeper is consolidated equally. The squeeze pressure can be adjusted as required to adapt to the prevailing condition of the ballast bed.

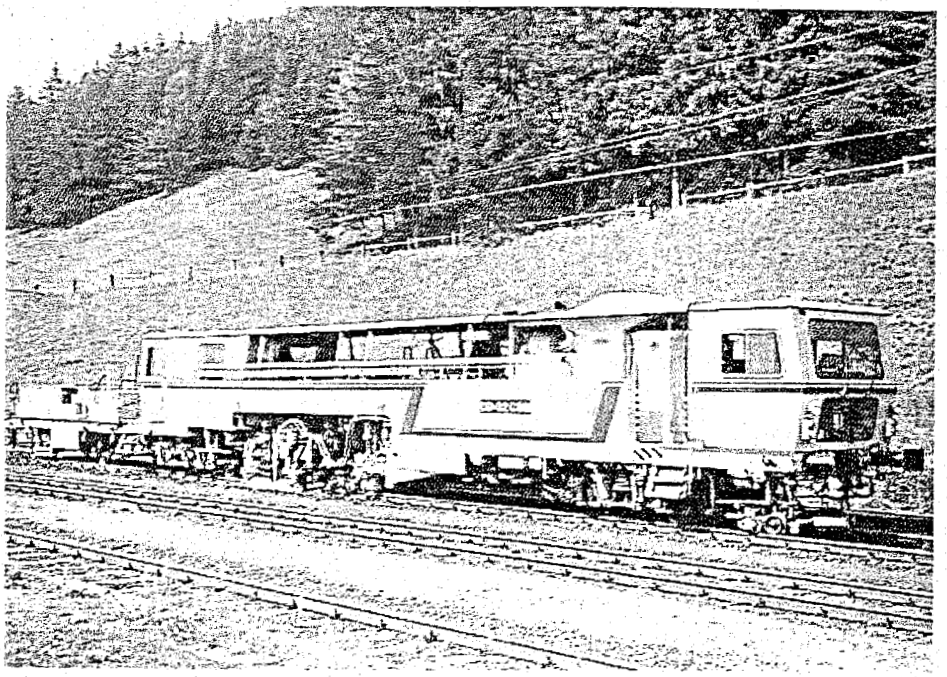


Figure 12.20: Continuous action levelling, lining, tamping machine

Directional vibration means that the vibration acts in the squeeze direction, thus supporting the consolidation effect and not loosening the ballast.

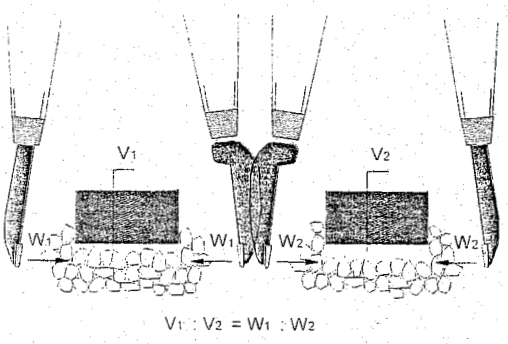


Figure 12.21: Duomatic non-synchronous tamping principle

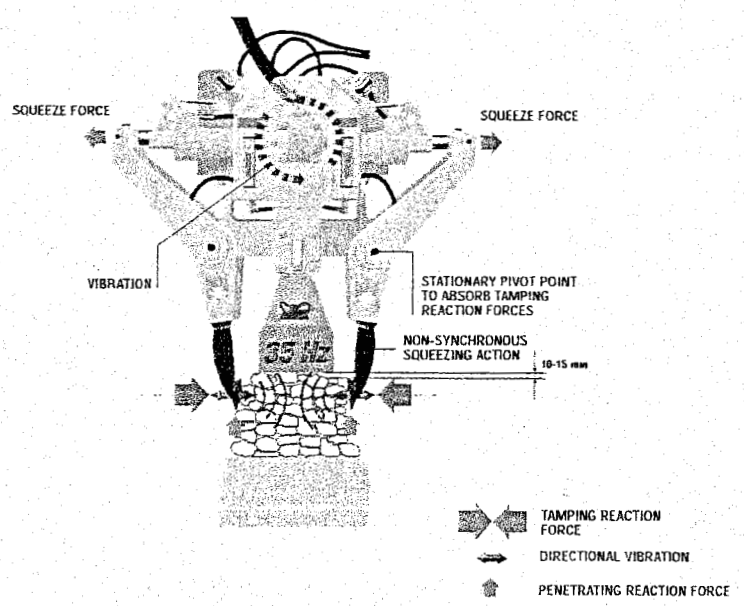


Figure 12.22: Tamping unit

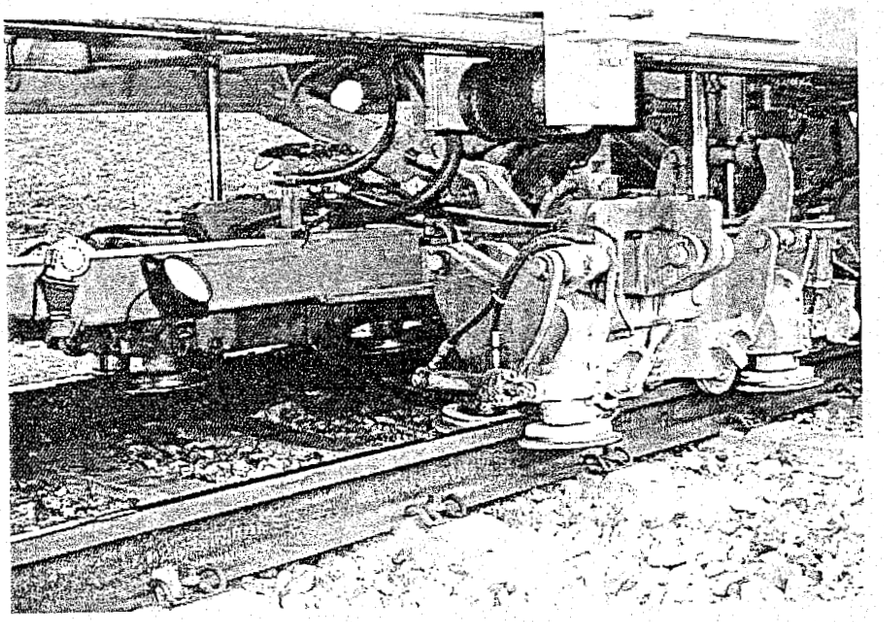


Figure 12.23: Lifting and lining mechanism with roller clamps

Figure 12.23 depicts a detail of the combined lifting and lining mechanism.

A main objective in the development of tamping-machines always was the need for high working output, as characterized by hourly working rates or annual performances. This is a result of the situation within the railways, which not only increased axle-loads and running speeds. The most challenging development with respect to track times for maintenance is the ongoing rise of the number of trains per time unit. Available occupation times were reduced over the years, from "long" occupations for instance from 5 to 2 hours. At the same time rail transport was concentrated gradually onto main trunks with the highest standards in rail technology, but also the highest traffic densities, of course.

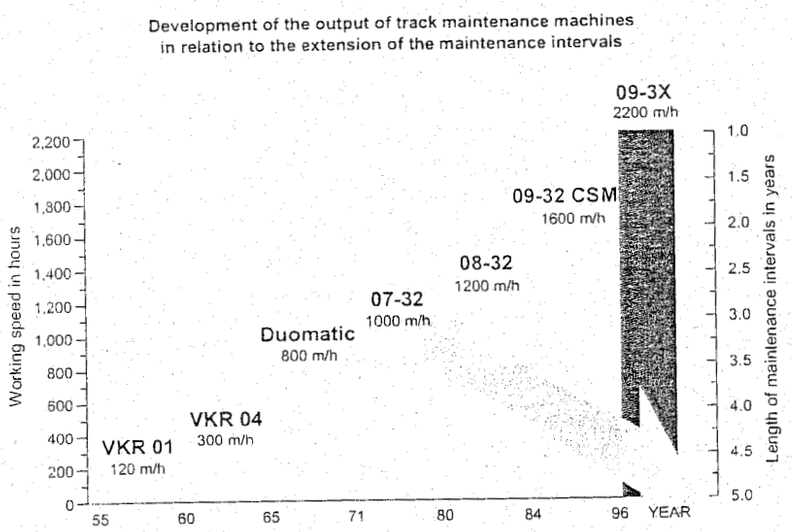


Figure 12.24: Development of working speed

Today a number of brand-new lines are operated in Europe. Their economic justification rests on high transport capacity and undoubted availability. So "availability" of infrastructure is a central matter in modern rail engineering.

All these circumstances pushed up the working speeds over the years considerably. Was it an incredible step forward to tamp with 60 m/h in the early days of mechanization – and thus to replace the exhausting manual labour – so are the standards today at more than 1500 m/h with continuously action two-sleeper-tampers and more than 2100 m/h with the latest 3sleeper-tampers (Figure 12.24). With it goes the shortening of approach times, quick set-up and removal from the site and increase of durability of the work result.

Continuous action tamping

With conventional tamping machines the whole machine has to slow down at each sleeper and then accelerate again. With the new generation of tamping machines of the Plasser and Theurer 09-CSM type the machine moves at constant speed with only the tamping unit jumping from sleeper to sleeper. In this way only 20% of the mass is accelerated and decelerated. This gives a higher output up to 2200 m/h, a high degree of comfort in the operating cabs and low energy costs. The 09-3X Tamping Express is shown in Figure 12.25 and the three sleeper tamping unit in Figure 12.26.

Figure 12.25: Continuous action three-sleeper tamping machine

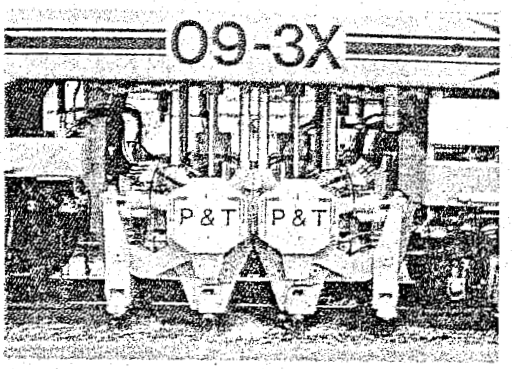


Figure 12.26: Three sleeper tamping unit

Switch and crossing tamping

For the maintenance of switches and crossings special switch tampers such as the Plasser Unimat 08-275 have been designed. This machine is equipped with 2 universal high pressure tamping units, each with 8 side tilting tamping tines arranged in pairs as depicted in Figure 12.27. The tamping units can be displaced laterally to be centred exactly over the area to be tamped. The switch tamper may also be used as a standard plain line tamper.

Heavier designs of switches and crossings due to the use of concrete sleepers and heavy rail profiles demand additional measures for their treatment. When lifting such turnouts in the area of the long sleepers with the standard two rail-lifting unit, the reaction forces on the fastenings are already exceeding their yield strength. Therefore an additional lifting arm was fitted to the machine, by which the turnout rail is simultaneously lifted with the rails of the main track. This additional feature of the Unimat 08-275 3S helps to avoid the development of undue stress on fastenings and sleepers.

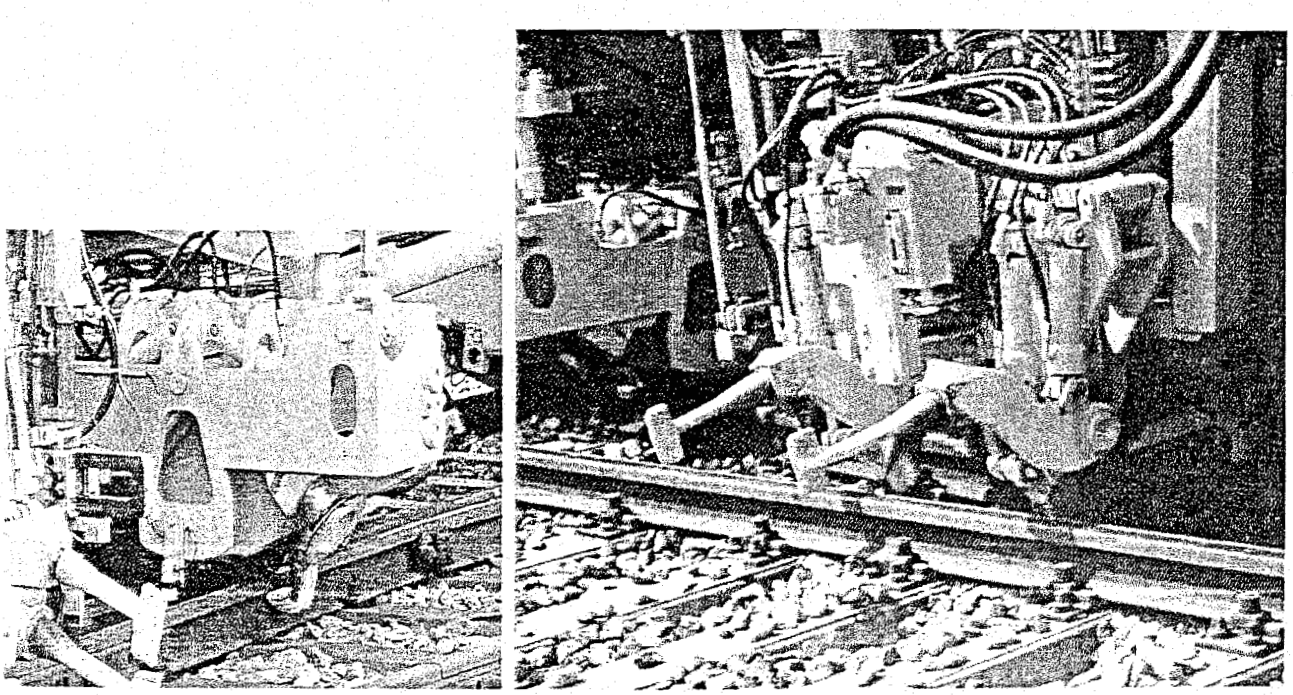


Figure 12.27: Switch lifting, slewing and tamping unit

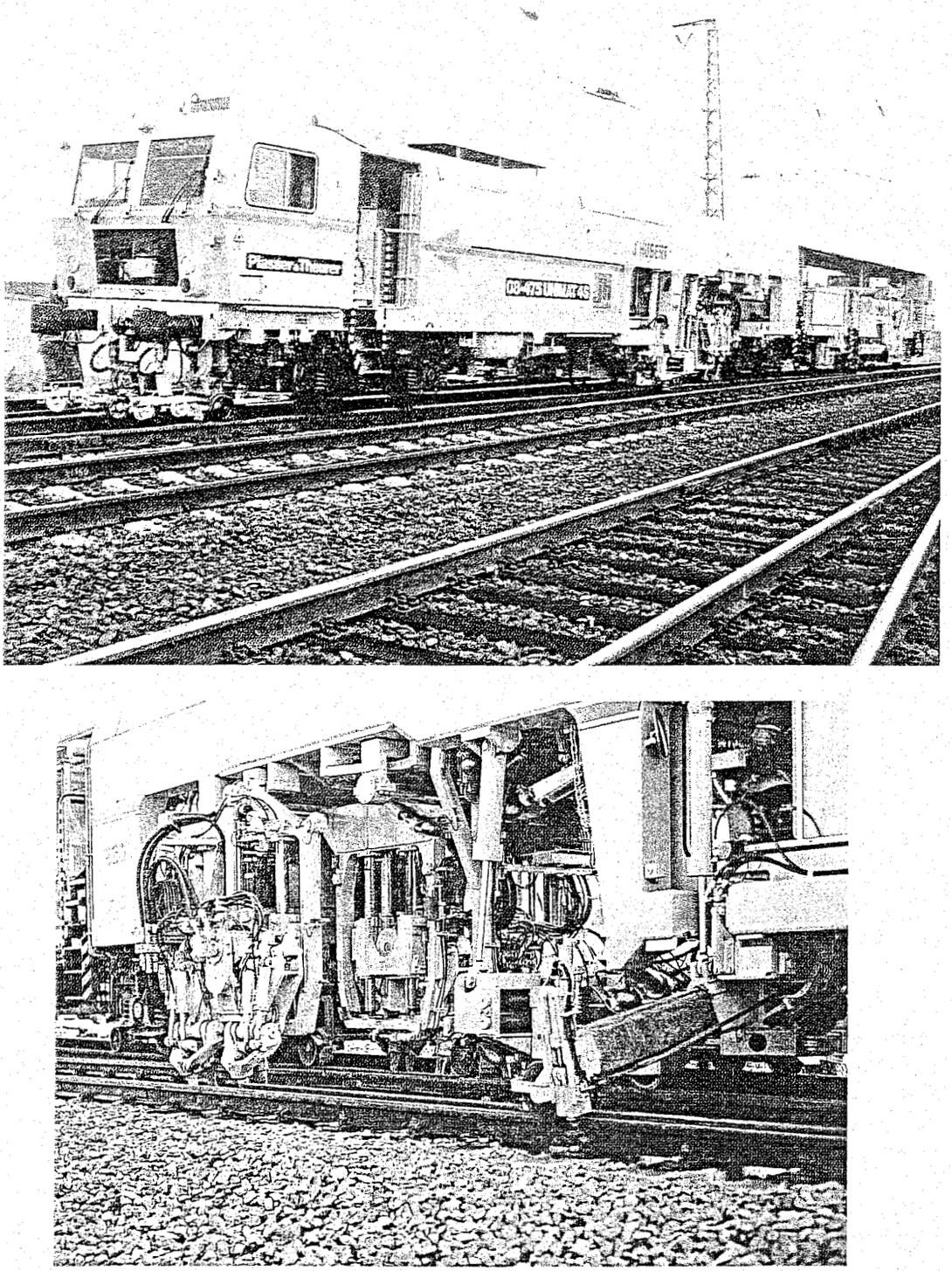


Figure 12.28: Unimat 08-475 4S with four rail tamping and outrigger for three rail lifting

In addition to 3-rail lifting, the 4-rail tamping brought about a further improvement in the quality of S&C maintenance (Figure 12.28). The tamping units of the Unimat 08-475 4S are split into four parts. The outer parts are mounted on telescopic arms so that the tamping tools can reach a distance of 3200 mm from track centre. This enables the main track and the turnout track to be tamped in one go; the whole switch unit is sufficiently supported when carrying out the first tamping pass on the main track. There is no danger that the switch may tilt because the long sleepers have not been tamped over the whole length within the first tamping pass.

12.6.3 Levelling and lining

12.6.3.1 Smoothing principle of modern tamping machines

Level

The tamping/levelling principle for the vertical direction is shown schematically in Figure 12.29. In this three-point measuring system point A is located on the already corrected track. The track is to be given an overall lift equal to the lift height BC. The track at working point D is raised, and the sleeper packed so that it comes to rest on the straight line AB at E. Levelling can be expressed mathematically as:

$$Z_n(x) = \frac{b}{\ell} Z_n(x-a) + \frac{a}{\ell} Z_0(x+b) \qquad (12.1)$$

Here Z_n is the new level and Z_0 the old level. The factor a/ℓ is the reduction factor by which any defects are reduced. Via Fourier transformation the same relation is obtained in the wavelength domain. This relation has the form:

$$FT\{Z_n\} = H(\lambda) FT\{Z_0\} \qquad (12.2)$$

in which:

$$H(\lambda) = \frac{a\xi^b}{\ell - b\xi^{-a}} \qquad (12.3)$$

with: $\xi = e^{\frac{i2\pi}{\lambda}}$; λ = wavelength

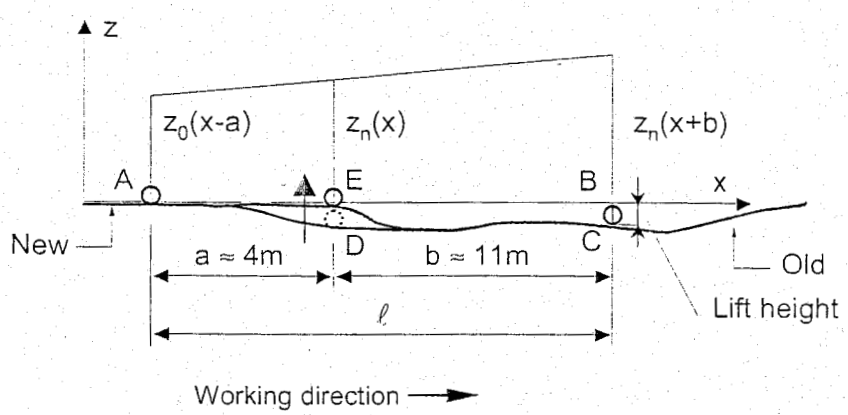


Figure 12.29: Tamping/levelling principle

The transfer function H gives the relationship between the level amplitude after tamping and before tamping as a function of wavelength. The modulus of this function is shown in Figure 12.30 for a Plasser and Theurer 07-32 tamper, where a = 3.88 and b = 10.52 m.

Various tests have been carried out to determine the real transfer function of the result produced by a tamper. This research is described in [71] a representative example of this type of actual transfer function, measured directly after tamping, is likewise presented in Figure 12.30, from which it emerges that the improvement is approximately half that expected on the basis of the theory.

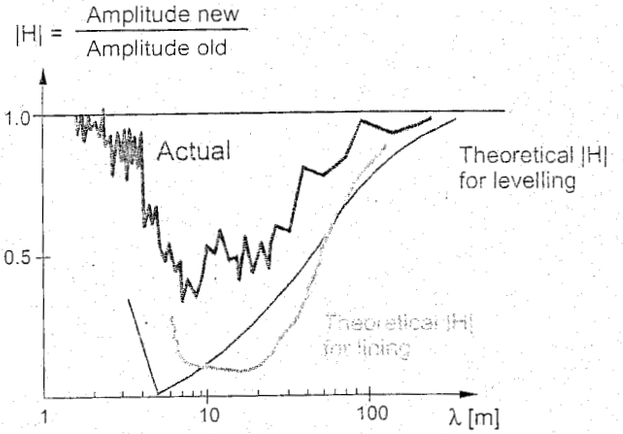


Figure 12.30: Comparison of theoretical and actual transfer function for 07-32 tamper

It should also be pointed out here that proportion of the initial improvement disappears again fairly quickly after tamping without additional stabilization.

Alignment

The lining system of modern track maintenance machines is based on a 4-point measurement, or, if design lining is applied, a 3-point measurement as is the case for vertical profile. In the 4-point system the machine measures three ordinates, Y_1, Y_2, and Y_4, as seen in Figure 12.31, which together determine a second-degree curve by means of which local track alignment is approximated. At work point C the track is aligned so that it comes to lie on this curve. The correct position is verified by means of versines h and H, the quotient of which has a fixed value.

If the machine is being used in the automatic mode the leading point of the long chord D follows the old track geometry whereas points A and B follow the geometry, which has just been corrected (Figure 12.32).

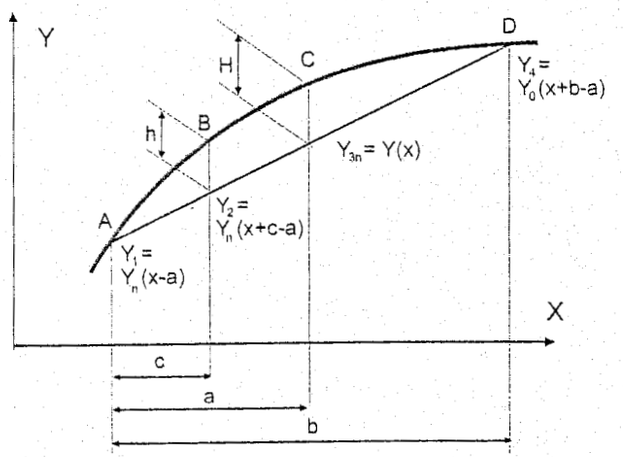


Figure 12.31: 4-point lining principle

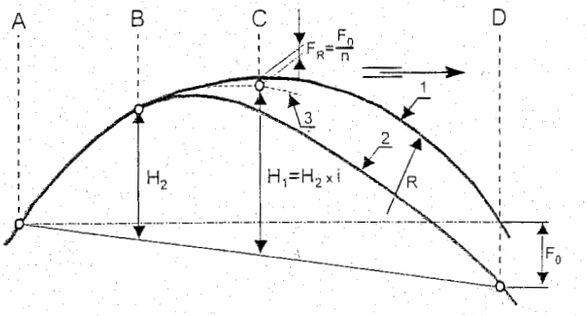


Figure 12.32: 4-point lining system with error F_0 at front end of lining chord

Design lining is also possible with these machines, in which case the new position must be defined beforehand. There are two options for supplying these data: the correction values are entered at D either manually or automatically by the on-board geometry computer (ALC) so that this point follows the ideal geometry. Minor errors in input values are smoothed out by the lining system of the machine.

The error reduction, which in theory can be achieved using the 4-point lining system, will be explained with the aid of Figure 12.31. Points A and B are already corrected points, for which the ordinates are indicated by Y_n (new). The location of the uncorrected point D is indicated by Y (old). If values a, b and c are known, the position of C can be expressed in that of A, B and D. In a similar way as for level the new geometry for alignment is to be expressed in the old according to the recursive relationship below:

$$Y_n(x) = \frac{a}{c}\frac{\alpha-1}{\alpha}Y_n(x-a+c) + \left[\frac{b-a}{b} - \frac{a(b-c)}{bc}\frac{\alpha-1}{\alpha}\right]Y_n(x-a) + \frac{1}{\alpha}\frac{a}{b}Y_0(x+b-a) \qquad (12.4)$$

The transfer function between new and old geometry which can be derived after Fourier transformation reads:

$$H(\lambda) = \frac{\frac{1}{\alpha}\frac{a}{b}\xi^{b-a}}{1-\frac{a}{c}\frac{\alpha-1}{\alpha}\xi^{-a+c}-\frac{(a-c)(a-b)}{bc}\xi^{-a}} \qquad (12.5)$$

This function is also depicted in Figure 12.30 for the Plasser and Theurer 07-32 where a = 9.785 m, b = 20.785 m and c = 5 m.

The 3-point lining principle

As has already been commented, lining is also possible using the previously discussed 3-point system (Figure 12.33). Machines which are equipped with the 4-point system can be switched to this mode. In curves with a radius R the new value at C has to match the versine BC*CD/(2R), which is set manually or computed by an on-board system. In the latter case first a measuring run is carried out and then the versines are calculated.

The transfer functions of the 4-point and 3-point systems differ only slightly so, from this point of view, no preference can be expressed for one system or the other. For practical purposes the 3-point system has the drawback that in curves the actual radius has to be known. Sometimes it is wrongly claimed that using the above-mentioned versine the radius of the track can be adjusted. If the set value of the radius differs from the actual value this will simply result in a displacement of the track over the whole curve length without the magnitude of the original radius being affected. On the other hand the 3-point system has advantages in any design lining application, as only one reference point in the rear has a feed-back into the transfer function.

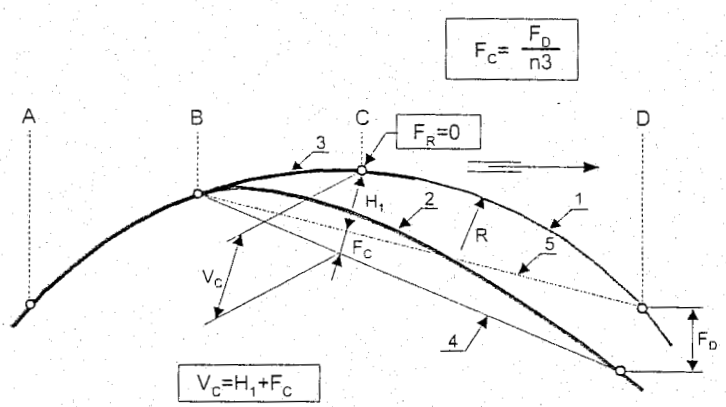


Figure 12.33: 3-point lining system

Laser levelling and lining system for tangent track

To level and line long straight track sections, the front reference points of the levelling and lining system are controlled by a laser beam (Figure 12.34). This system consists of a laser receiving camera, which is mounted on the front tensioning car of the tamping machine, and the laser transmitter on a small trolley which is set up in front of the machine in a distance of up to 300 m.

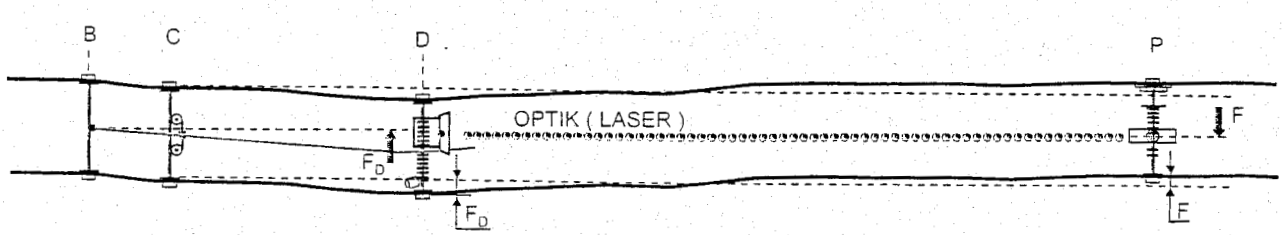


Figure 12.34: Laser alignment system

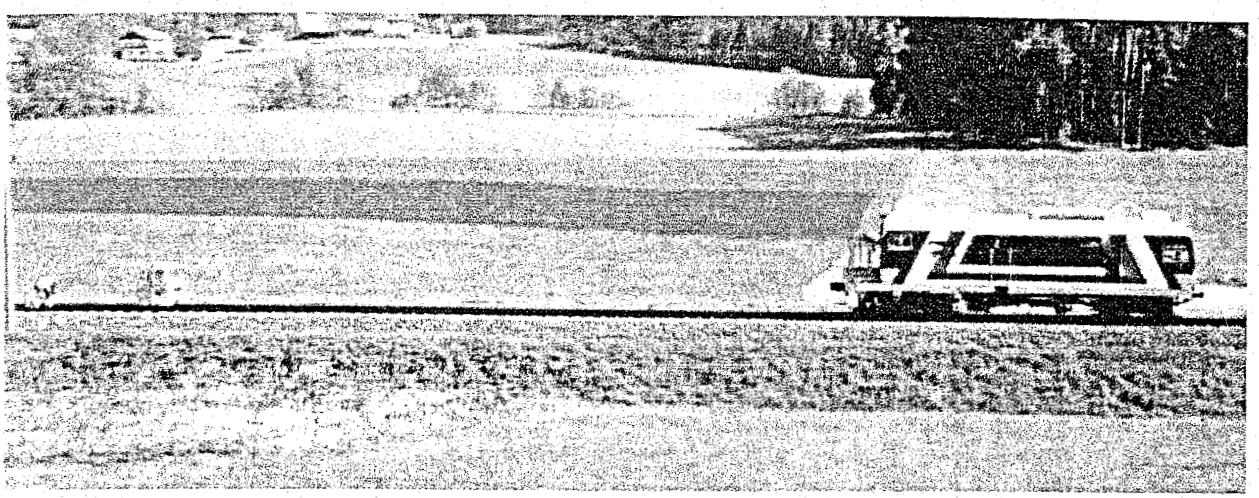


Figure 12.39: Laser satellite and EM-SAT 120

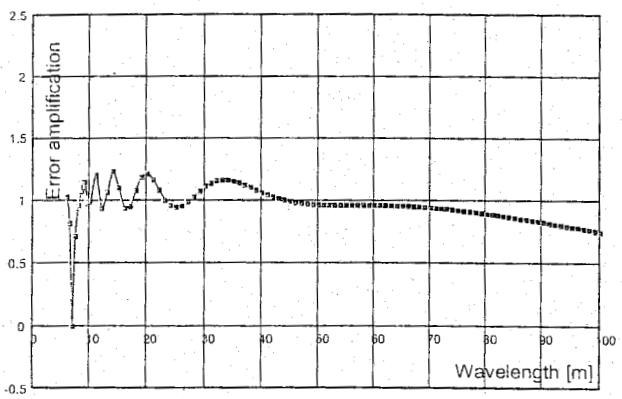


Figure 12.37: Transfer function of lining system after application of factor method

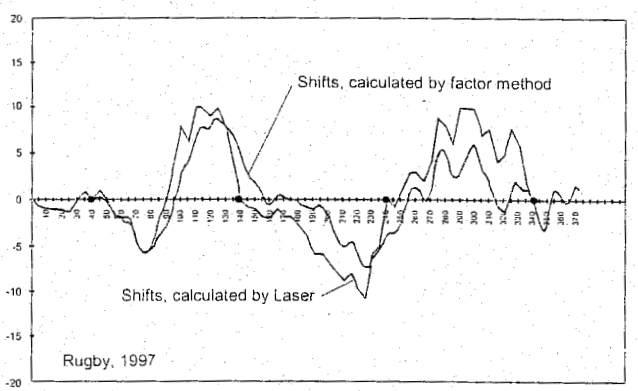


Figure 12.38: Correlation of alignment between laser and factor system

The target data achieved in this way can be stored and filed on hard disk or floppy disk for later application. The memory capacity of the computer permits the storage of target data for several sections of track. When required for maintenance work, the appropriate data only has to be selected and activated.

12.6.5 EM-SAT

The automated, computer-assisted surveying technique EM-SAT (Figure 12.39) measures long wave track faults in longitudinal level and curvature (alignment) and calculates precise control values for the tamping machine. The EM-SAT is used for:

- Planning and determining of tamping and lining sections,
- Surveying prior to tamping work according to guideline 824.3010,
- Surveying for elimination of spot faults using a spot fault tamper,
- Recording of track renewal (verification of the track bed structure in respect of longitudinal level and alignment) before and after the tamping runs;
- Acceptance of new track to identify possible long wave faults

The EM-SAT continuously determines – above all – long wave faults in longitudinal level and curvature (alignment). The actual values obtained are compared to the target values, recorded (analogue print) and transmitted to the ALC of the tamping machine for further use.

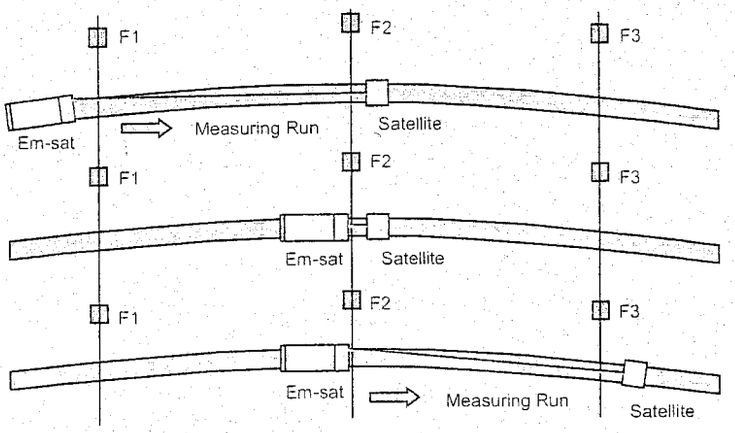


Figure 12.40: Measuring principle of EM-SAT

Measurements are taken in cyclic action (Figure 12.40) with an average output of 2.5 km/h with simultaneous measurement of longitudinal level and alignment. Besides the displacement and lifting values (measuring accuracy < 1 mm), superelevation and track gauge faults can be recorded at the same time. The position of the Laser measuring chord has to be checked against the outside reference points (fixed points, F1, F2, F3,...), deviations from the target position are also corrected by the calculation software. The checks at the fixed points are done either manually or with a non contact fixed point measuring unit.

GPS measurements

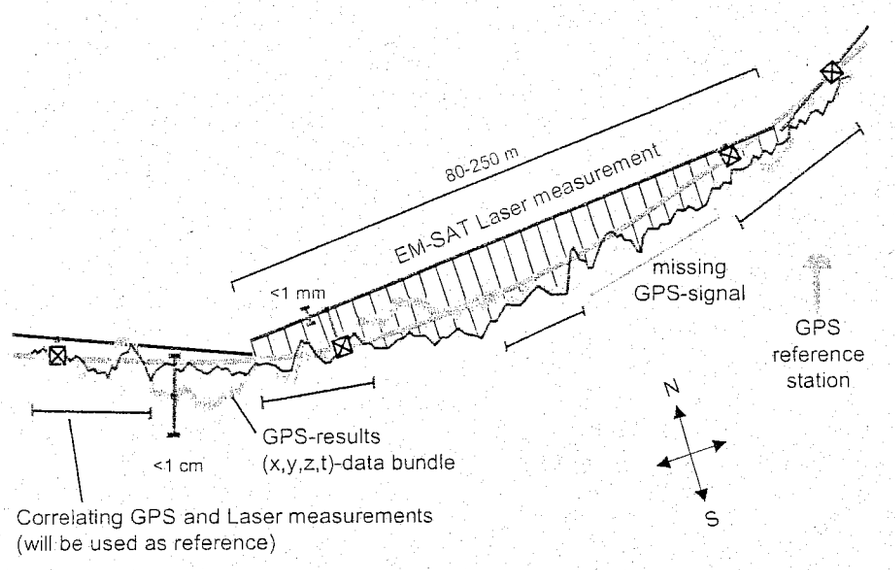


Figure 12.41: GPS-based track measurement

From 1993 to 1995, the Technical University Graz and the Research and Testing Department of Plasser & Theurer studied in a joint research project the possible uses of GPS for track surveying [164]. The study showed, that he highest absolute alignment accuracy achievable with GPS is approximately ± 6 mm. The accuracy in level is worse by a factor between 1.5 and 2. GPS is not precise enough to be used instead of a laser reference chord measuring system. The requirements in terms of track correction values to guide tamping machines are about 1 mm.

With the EM-SAT mechanised measuring of long chords has become reality. The absolute accuracy of EM-SAT measurements is better than 1 mm. To determine the chord position and hence the actual and target track geometry in a co-ordinate system, fixed points are necessary. The drawback of fixed points is their inconstancy (in an absolute sense, they are not "fixed"), they are complicated and expensive in terms of measuring work and maintenance, and they slow down the work progress of mechanised systems.

GPS provides an elegant method of replacing fixed points (Figure 12.41). The achievable accuracy in line of ± 6 mm and ± 9 to 12 mm in level is fully sufficient for these purposes. Simultaneous measurement of the actual track using laser reference chords and GPS makes it possible to transform the highly accurate relative laser reference chord data into absolute co-coordinative data. This also

serves to check the quality of the GPS data. In areas where the measured actual positions of the laser reference chord and GPS are very similar, very reliable GPS data are available which serve as support points for the end points of the chord. Fixed points and their disadvantages could now be a thing of the past.

For high precision GPS measurements (10-20 mm) it is necessary to have fixed GPS reference stations (10- 20 km apart), e.g. in stations or GPS reference networks.

12.7 Stone blowing

12.7.1 General principle

The key to the development of stone blowing was the observation that ballast has a "memory" of the shape to which it had deteriorated prior to each tamping operation. Measurements of the lift/settlement behaviour on a number of Railtrack sites clearly showed that durability of the lift improves when it exceeds 20 - 25 mm (Figure 12.42). This was the reason for developing the stone blowing process which adds stones to the surface of the existing ballast bed, leaving the body of the compact and stable ballast bed undisturbed.

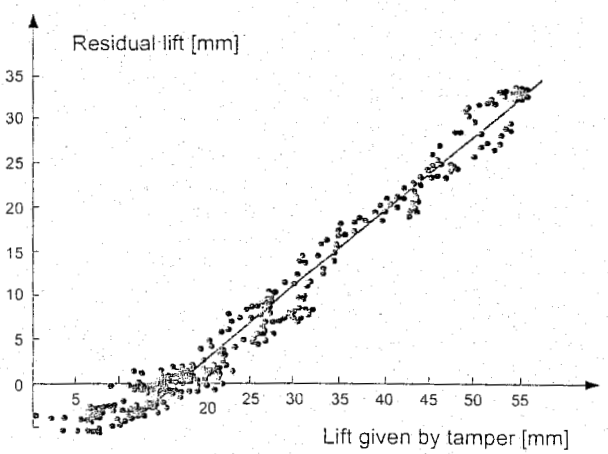


Figure 12.42: Lift settlement characteristic following tamping

The principle is outlined in Figure 12.43. The sleeper is lifted to create a void. The stone blowing tube is driven alongside the sleeper and a measured quantity of stone, derived from the known relationship between added stone and lift achieved, is blown into the void. Subsequently, the tube is withdrawn and the sleeper is lowered onto the added stone. Moreover, according to [182], this can be done at a production rate that is comparable to other maintenance techniques, and the machine can treat areas of track that are otherwise inaccessible for machine maintenance.

In addition to the hand-held version, Railtrack uses the stone blowing machine PBI 84 (Figure 12.44) which was introduced in 1996. In 2001, there were 13 of these machines in operation. The stone blower works in a design mode rather than a smoothing mode. Special measuring equipment is attached to the stone blower to measure the unloaded track profile, which is converted into a loaded profile by superimposing the displacement of a loaded axle relative to a chord created by two unloaded measuring axles.

The frequently heard objection to the principle of stone blowing concerns the injection of relatively small stones under the sleepers. This is usually explained on the grounds that the stone specification calls for a grading between 14 and 20 mm, which is within Railtrack's specification for new ballast. The grading was chosen to avoid any drainage problems and is also larger than the stone size associated with measured shovel packing.

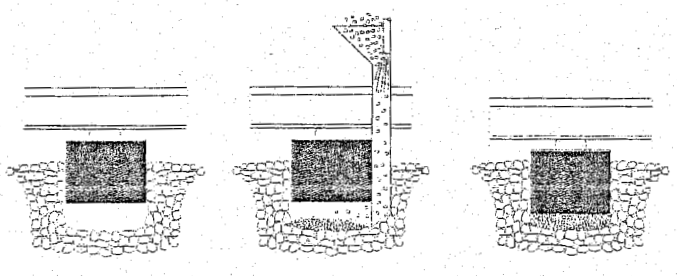


Figure 12.43: Principle of stone blowing

The stone blower requires knowledge of the actual track profile which is free of phase shifts. By using a number of strategically placed unloaded axles, it is possible to collect relevant information in order to reconstruct an accurate true profile for wavelengths of up to 100 m. In addition, a number of independently loaded axles are used to measure the load-deflected profile of the track in relation to this unloaded profile.

Average Output	440 meters/hours (1,400 feet/hour)
Gross Weight	113 tonnes (124 tons) carried on three bogies
Maximum Axle Load	18 tonnes (20tons)
Length	32.2 meters (106 feet)
Travel Speed	100 kph (60 mph) maximum
Measuring Speed	16 kph (10 mph) maximum
Engine	Cummins KTTA19C, 522 kilowatt (700 horsepower) @ 1,950 rpm
Auxiliary Power Unit	Cummins 4B3.9 diesel, 28 kilowatt electrical output
Track Lifting	0 to 80 mm (0 to 3.2 in.)
Track Slewing	0 to 80 mm (0 to 3.2 in.)
Minimum Working Radius	150 meters (500 feet)
Fuel Capacity	4,500 liters (1,200 gallons)
Hydraulic Tank Capacity	950 liters (250 gallons)
Stone Capacity	16tonnes (18tons)
Size	#6-approximately 20 mm (0.8 in.)
LA Abrasion Index	11-19
Stone Delivery	0.5 to 22 kg (1 to 48 lb.) per tie end
Automatic Index Distance	1 tie pair
Index Time	3 seconds
Blowing Time	0.3to 11 seconds

Table 12.1: Basic Stoneblower Data

Positioning is absolutely vital and is achieved by using redundant measuring wheels which keep an independent check on each other. A computer algorithm resolves differences to achieve an accuracy which is better than one sleeper in a 4,000 m site

An inaccuracy resulting in a miscalculation of stone may not be immediately obvious and would only manifest itself as a track geometry irregularity after the passage of significant volumes of traffic. So, an independent system is employed to re-measure the profile in real time during the stone blowing operation.

12.7.3 Stone blowing applications

According to Railtrack tests, and under Railtrack conditions, the durability of stone blowing in comparison with tamping ranges from 2 to 7 times, details of which are explained in Figure 12.45. Stone blowing should not be considered as a substitute for tamping, but rather as a complement. There might be specific circumstances in which stone blowing could be considered more effective. For instance, when a track has been fully consolidated, the finishing correction of level can be achieved by stone blowing, just as the grinding process is the finishing touch to weld straightening [182]. As a result of the increased durability, stone blowing is also claimed to result in a reduction of total track possessions required

According to [182], the largest benefits of stone blowing could be achieved on tracks which required frequent tamping and hence had high maintenance costs. However, it could be argued that such robust benefits were true due to the fact that approximately 20 % of the UK railway network required tamping at more frequent intervals than six months in order to maintain the desired geometrical standard. In other words, these were typically tracks where the ballast was approaching the end of its service life and might otherwise had been scheduled for renewal.

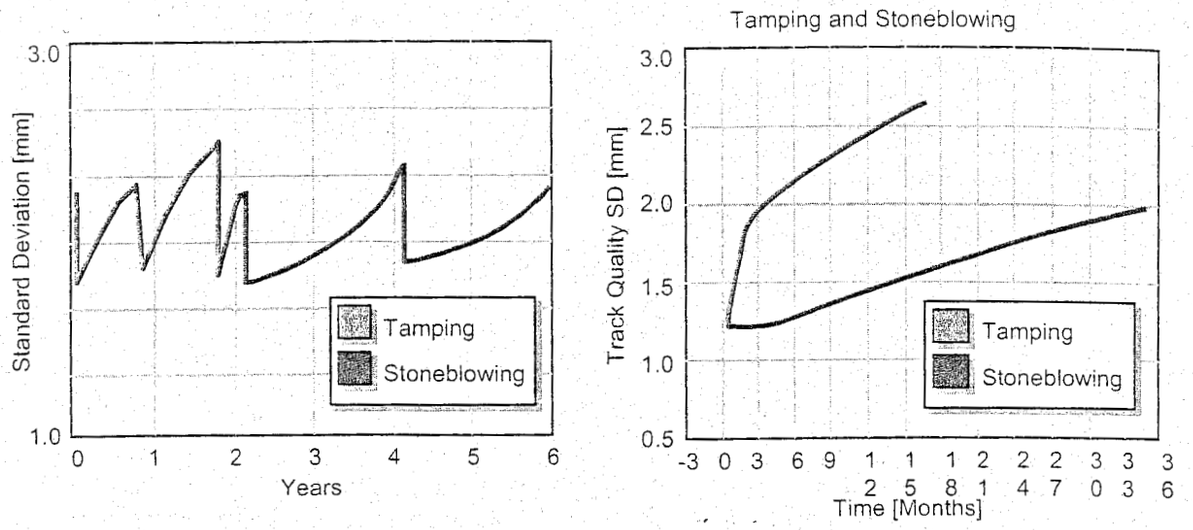


Figure 12.45: Typical post-maintenance behaviour for Stoneblowing and Tamping

Design tamping on longer waves is most effective with design lifts in excess of 25 mm. Stone blowing, on the other hand, is more effective for shorter wavelengths, with lifts in the range of 0 - 30 mm. Thus, the stone blowing process is, to a large extend, complementary to design tamping.

According to [182], track renewal can only be economically justified if the reduction in maintenance costs resulting from a renewal, plus any revenue saved by avoiding speed restrictions, outweighs the amortised cost of the renewal. By introducing a machine which can hold acceptable geometry on track with ballast that was regarded as 'untampable' and was scheduled for renewal, the balance of the maintenance and renewal equation is shifted.

Railtrack also investigated the effect of the maintenance process itself on ballast damage. This investigation showed that tamping at relatively short intervals was one of the main causes of ballast damage (Figure 12.46), each tamping cycle being equivalent to approximately 20 MGT of traffic [182] [183]. Stoneblowing, however, was shown to have only a negligible effect.

The UK contracting company GTRM [196] tried to use stone blowing at some of the active bankslip sites where the top and cross-level went off quickly and which needed to be tamped on a 4-6 week basis. Stone blowing proved not to be capable of lifting the track sufficiently and obtaining considerable corrections to the longitudinal profile and cross-level. Hence, they realised that tamping was more suitable for bankslips. Likewise, they found tamping to be more appropriate for newly reballasted track, ballasted track in reasonable condition, or wooden track. However, GTRM also reported significant track quality improvements on several sites where stoneblowing was used.

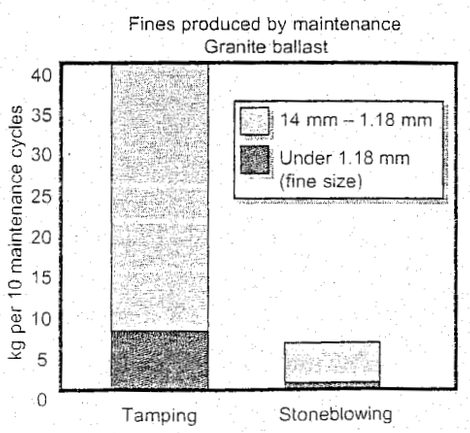


Figure 12.46: Ballast damage caused by tamping & stoneblowing

12.9 Ballast profiling and stabilization

Ballast profiling

The correct ballast cross-section, defined by the permanent way regulations, is of great significance. Above all, it is important for the safety of the track against track buckling. Moreover, any unnecessary spreading of the ballast bed represents a cost factor. To counteract this, it is vital to produce the right ballast cross-section both in the course of track maintenance as well as when placing new ballast. If this work is not carried out regularly, huge quantities of ballast could be lying unused in the track network.

Ballast regulators reshape the ballast bed and should be used in connection with every tamping operation. Standard ballast regulators reshape the ballast bed by several runs backwards and forward. With the development of continuous action tampers it became necessary to redesign also the ballast regulators – "one go" continuous action machines were introduced (Figure 12.51). In addition the machines now can be fitted with large scale hoppers which enable better distribution of the ballast and thereby savings of new ballast (Figure 12.52).

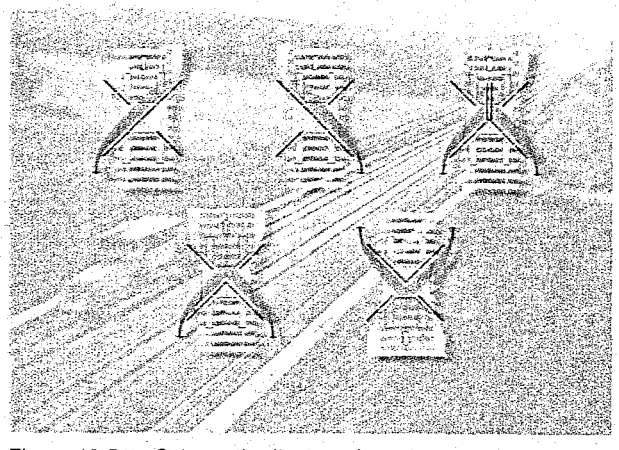


Figure 12.51: Schematic display of continuous action ballast regulation

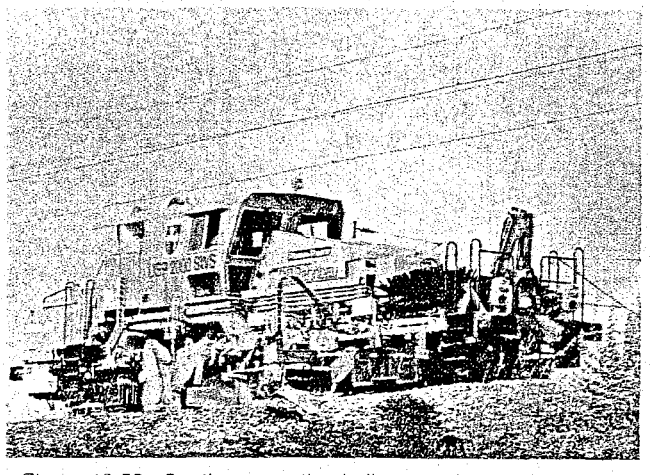


Figure 12.52: Continuous action ballast regulator with ballast hopper

Dynamic track stabilization

Ballast profiling is followed by stabilization with the Dynamic Track Stabilizer. The aim of dynamic stabilization is to achieve an improved anchoring of the track grid in the ballast bed. After dynamic stabilization the condition of the track offers greater operating safety and enables, particularly on newly laid track or after full track maintenance, travel at the maximum line speed. This helps to avoid speed restrictions and to reduce operational hindrances overall.

The Dynamic Track Stabilizer (DTS) shown in Figure 12.53 consists of a 4-axle wagon fitted with a 370 kW diesel engine for driving the excenters and pressure cylinders of the stabilising units (Figure 12.54). With this machine the ballast is compacted via the rails. The method is based on lateral excitation of the track whilst the track is being loaded vertically with a load of approximately 356 kN. The excitation frequency can be adjusted between 0 and 42 Hz, with the force amplitude increasing to a maximum of 320 kN. The vibration transmitted to the ballast lies in the natural frequency range of the ballast and causes the stones to settle closer together within the cavities. The result is a force-free re-arrangement of the ballast stones. The number of cavities is reduced. Instead of individual contact points the, ballast stones have a larger number of contact surfaces and contact edges. Between sleepers and ballast stones too, the sum of contact surfaces rises considerably.

The high energy caused locally in the track by the stabilizer can give rise to ground vibration on the line. Chapter 15 considers this in more detail. Operation on bridges and in tunnels is generally possible, certain operation parameters have to be observed.

12.10 Mechanised track maintenance train

Figure 12.56: Mechanised track maintenance train MDZ 2000

The aim of a complete continuous action, production-line treatment of the track with quality control of the completed work is achieved with the concept of a Mechanised Maintenance Train MDZ consisting of the levelling, lining, tamping machine, the ballast regulator and the dynamic track stabiliser. The last machine of the consist should also be equipped with a track recorder to document the finished job. The MDZ can be composed in different performance categories but it is important that the machines match in working and travelling speed. The standard consist of a MDZ for high capacity lines is shown in Figure 12.19.

By integration of ballast regulation into the tamping machine or the dynamic track stabiliser the MDZ consists only of two machines (Figure 12.56)

12.11 Ballast cleaner

The main tasks of the ballast bed can be defined as follows [292]:

- uniform distribution of the wheelset forces on the subgrade;
- reduction of dynamic stress, caused by dynamic axle loading;
- ability to be maintained and to hold the vertical position achieved during maintenance;
- assurance of the horizontal stability (lateral resistance to displacement).

Figure 12.57 shows, that on polluted ballast, the load distribution function of the ballast bed and the full drainage function of the subgrade must be restored by undercutting-cleaning. The general rule is:

- ballast cleaning becomes appropriate when there are more than 30% of fines of less than 22 mm size in the ballast [63].
- ballast cleaning is absolutely necessary when there is more than 40% pollution.

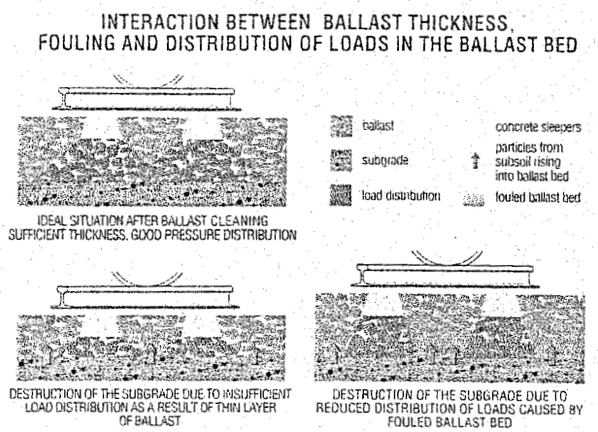


Figure 12.57: Function of the ballast bed

The ballast cleaner, as depicted in Figure 12.58, is used to dig away the ballast to a minimum depth of 25 cm below the sleepers by means of a chain to which "excavating teeth" are attached. The chain transports the spent ballast upwards to a system of vibrating sieves where material smaller than 35 mm is filtered out. The clean coarse material is returned to the track.

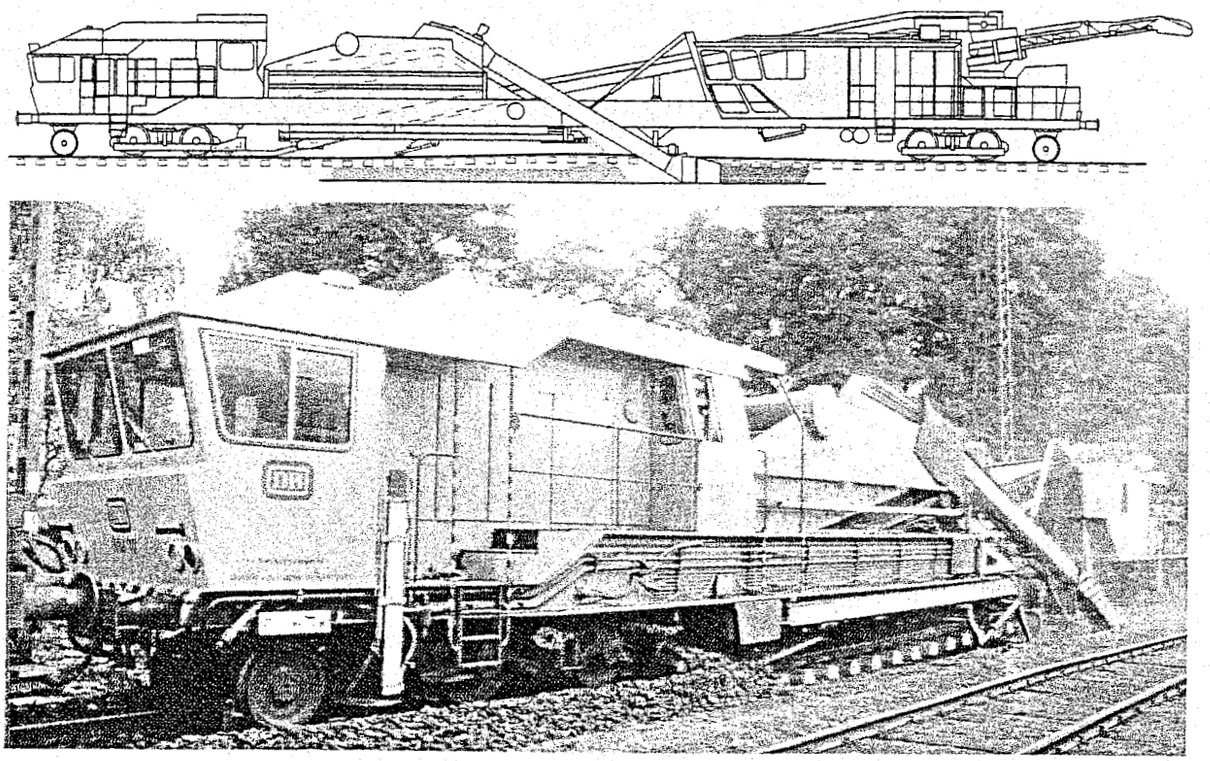


Figure 12.58: Ballast cleaning machine RM 80

During the cleaning operation about 1.5 t of dirty ballast is released per metre of track. Processing of this determines the work rate of the machine. The practice to dump contaminated ballast to the side of the track is not applicable any more today. Waste loading systems which can keep up with the working speed of modern ballast cleaners (up to 1000 m³/hr) are run in front of the machine.

These are special hopper cars as depicted in Figure 12.59 with bottom conveyor belts and transfer conveyors so that each car can load the next one. The cars are loaded with waste one after the other. Starting from the front.
As soon as a number of cars is loaded they are detached from the rest and transported to the waste dump. There they unload themselves via the swayable transfer conveyor bands. The cars which remained with the machine serve as buffer units till the empty ones return. The loading capacity of the hopper units ranges from 65 to 100 tons. The number of cars depends on the anticipated daily output.

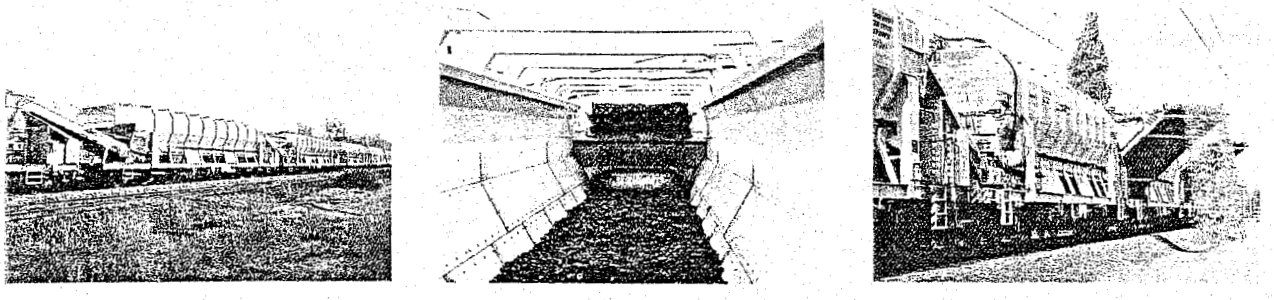


Figure 12.59: Waste loading system MFS (System Knape)

The waste material has to be transported to special dumps for decontamination, deposition or recycling. NS transports this material to a recycling unit where it is processed down to concrete gravel.

12.12 Formation rehabilitation machines

The improvement of sections of track with subsoil problems is absolutely necessary for many reasons. Substructure with insufficient bearing strength is primarily a serious technical problem because it jeopardises the stability of the permanent way to the highest degree. This results in enormous expenditure for the track maintenance and the life cycle of the entire permanent way material would be shortened considerably. Poor quality subsoil can also cause enormous additional costs, both with regard to higher track maintenance costs and the associated operational hindrance costs and also in view of the shortened life cycle of the permanent way material.

The insertion of a formation protective layer (FPL) is an effective and well-proven method to raise the bearing strength of the subsoil and has been used since the early seventies. As a result of this measure there is an enormous reduction of the costs for track maintenance.

The formation protection layer consists of a special sand-gravel mixture which has to fulfil the filter rule of Terzaghi, which determines the size of the voids and the frost criteria of Casagrande which determines the allowed fraction of fine grain < 0.02 mm and the demanded non-uniformity of the material [102], [246] (Figure 12.60). The layer may be reinforced by a geosynthetic underneath.

The material excavated during formation rehabilitation contains valuable raw material which can be utilised for the track improvement. The re-use of this recycled material enables a reduction of the rehabilitation costs by reducing the expenditure for transport and the quantity of new material required.

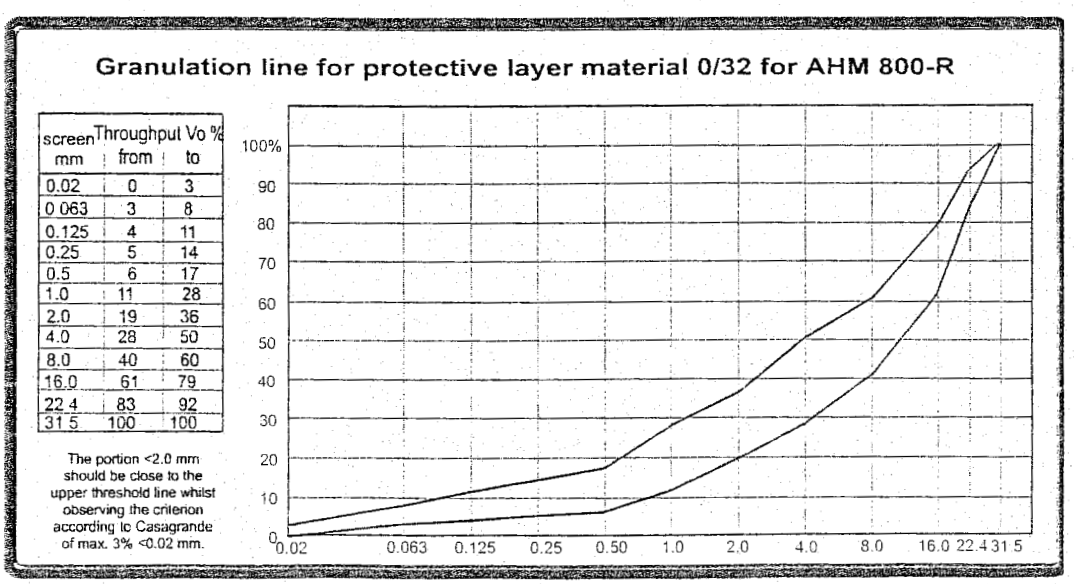


Figure 12.60: ÖBB specification of granulation of the protective layer material

Blanketing machine with ballast rehabilitation

Manual insertion of sand blankets is not only costly, also the compactness and uniformity of the blanket cannot be maintained. Insertion of blankets by road construction equipment needs to close down the track during the rehabilitation process and again in many cases the quality of the finished product is questionable. Different machines and machine systems are available, which can insert sand blankets, geosynthetics or other protection layers under the track, in track possessions, without the necessity to dismantle the track.

The AHM 800 R is a formation rehabilitation machine which uses re-cycled ballast for the formation protective layer (Figure 12.61). The AHM 800 R has two independent excavation devices. The smaller front excavating chain picks up the top layer of the old ballast bed (20 to 25 cm). The material is freed of small metal parts then an impact crusher breaks the ballast stones to a size of 0 to 35 mm. The crushed ballast is mixed with water plus additional FPL material in a mixing plant and prepared for installation.The second, larger excavation unit removes the remaining old ballast and the top layer of the old subgrade. Earth compactors smooth the remaining surface. If required, a geotextile or a fabric layer can be rolled over this or styrofoam slabs or geogrids can be laid (Figure 12.62 and Figure 12.63).

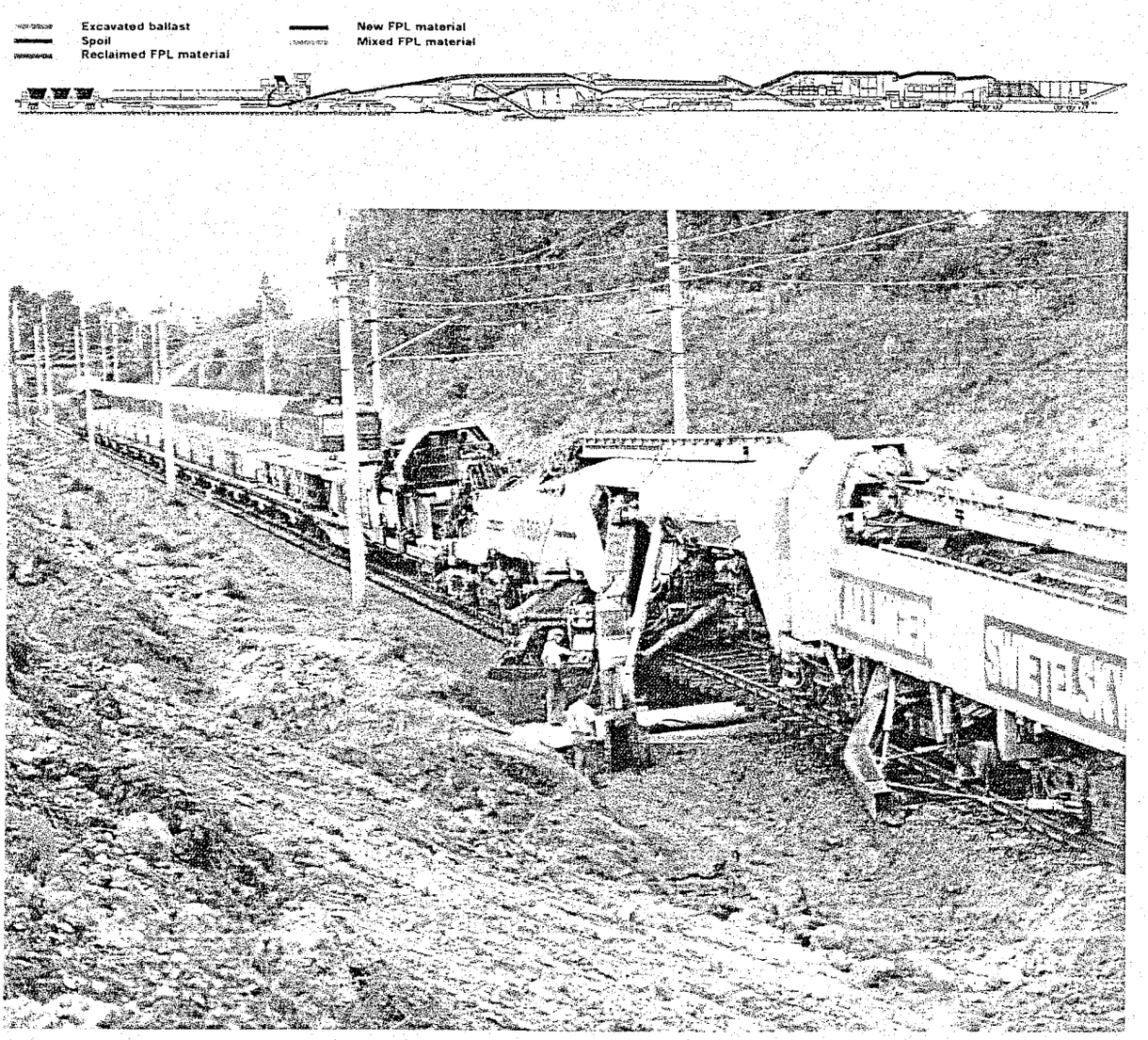


Figure 12.61: Formation rehabilitation machine AHM 800 R

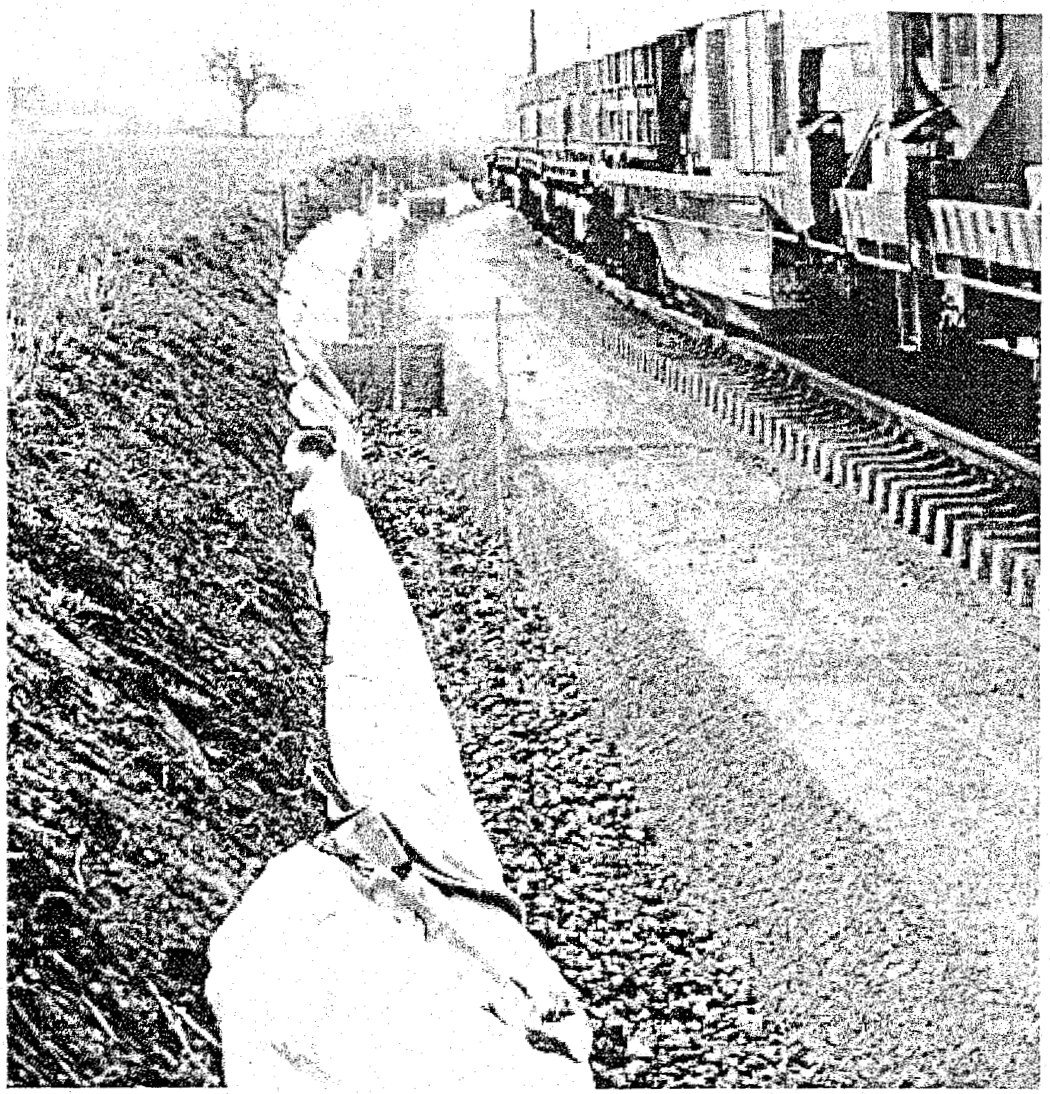


Figure 12.62: Example of tracks after formation rehabilitation with geotextile

A slewing conveyor belt inserts and distributes the sand/gravel mixture across the entire width of the track. A grading unit and six consecutive vibrator plates compact the sand-gravel mixture to produce the finished formation protective layer. The humidity of the FPL material is measured regularly. If necessary, more water is added using a spraying unit.

In this way a formation protective layer (FPL) up to a thickness of 50 cm can be installed in one operation. The track skeleton is laid on the finished FPL. All materials are transported on conveyor belts within the machine. The spoil lifted by the big excavating chain is loaded onto the material conveyor and hopper units pushed by the machine. The additional FPL material is supplied by containers which are unloaded by the gantry crane into the collecting hopper. One gantry crane ensures the machine is supplied from these containers.

Formation rehabilitation with integrated ballast cleaning

Another formation rehabilitation machine of the latest design, the RPM 2002 (Figure 12.64) was developed specially for the requirements of German Railway: the upper layer is removed separately and cleaned within the machine. The stones are sharpened in a crushing plant. After the formation protective layer has been inserted and compacted, the reclaimed ballast is distributed in the track together with the new ballast. A tamping unit in the rear part of the machine enables line speeds of 70 km/hr immediately behind the machine. For cleaning of wet and clogged ballast the star screen had been developed (Figure 12.65) where rotating cams peel the dirt from the stones.

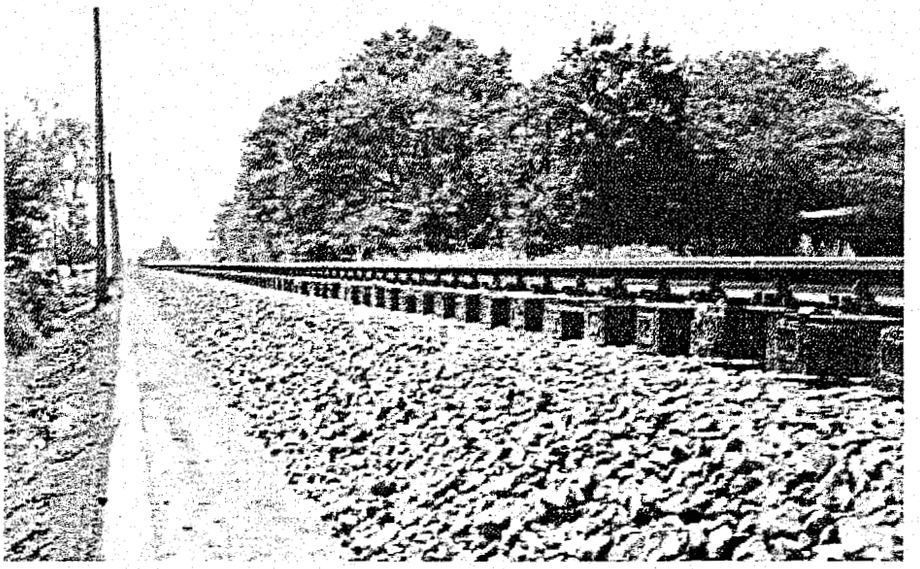


Figure 12.63: Example of tracks after formation rehabilitation without geotextile

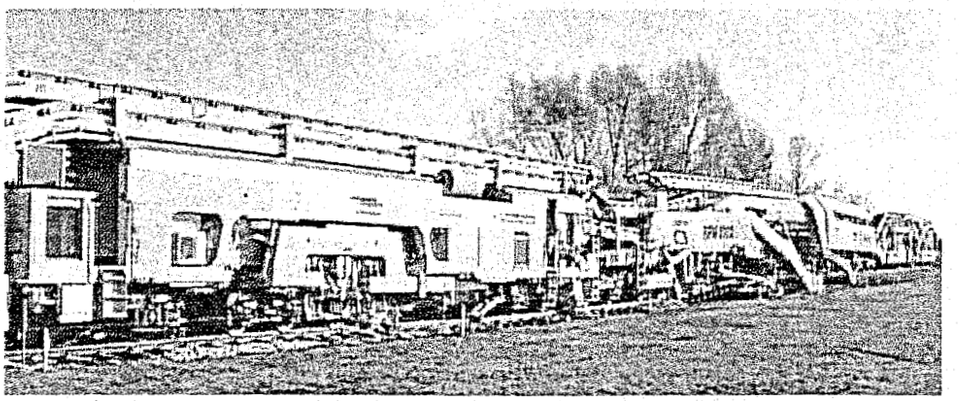


Figure 12.64: Formation rehabilitation machine RPM 2002

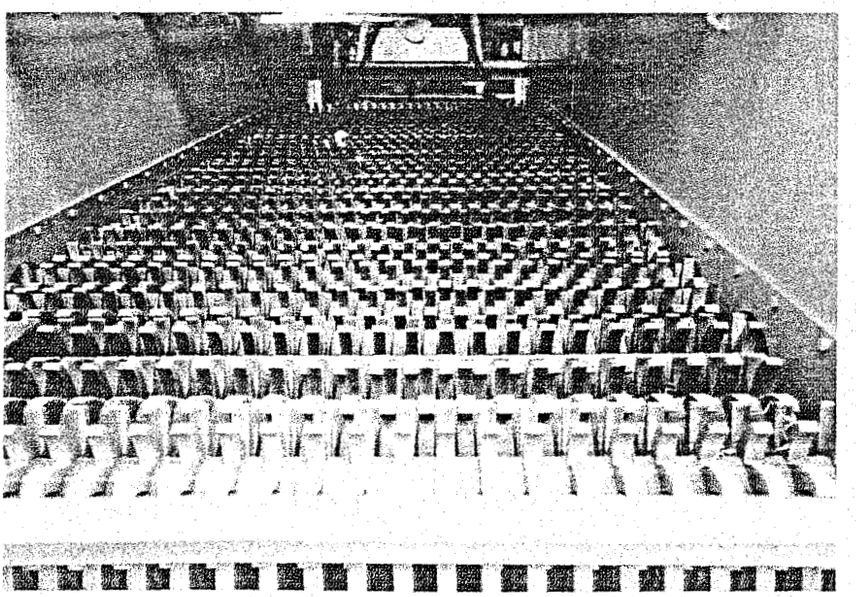


Figure 12.65: Star screen (system Wiebe)

12.13 High temperatures

On CWR track, changes in length are not possible which means that when there are changes in temperature stresses arise in the rails. During maintenance the high temperatures and related compressive stresses are of prime importance. This is principally the case for maintenance during which the lateral resistance of the ballast bed is reduced substantially. This is only restored slowly as trains pass over the track. Although the use of ballast stabilizers directly behind tampers and ballast cleaners clearly improves the situation, safety considerations somewhat limit this possibility if rail temperatures are, or are likely to be, on the high side.

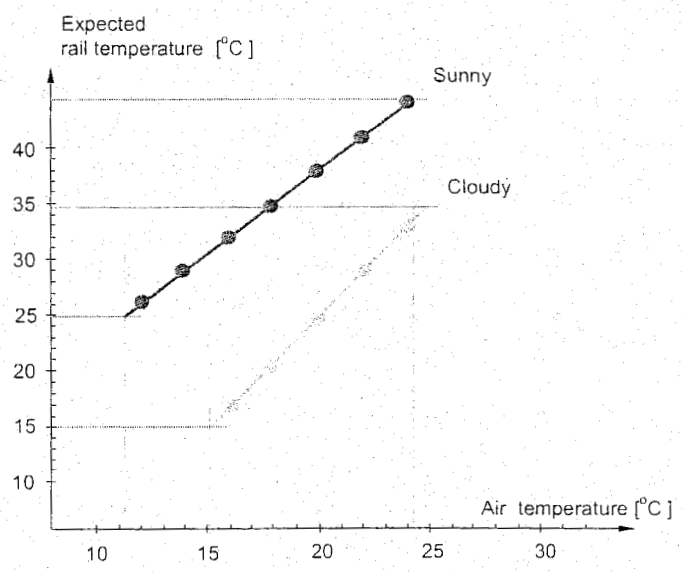


Figure 12.66: Relationship between air temperature and rail temperature

Rail temperatures differ enormously from air temperatures especially in sunny weather and are difficult to predict. On the basis of an extensive measuring programme the relationship between air temperature and rail temperature, shown in Figure 12.66, has been determined for cloudy and sunny weather.

If on CWR track the rail temperature is greater than 35° C loosening of the ballast over a distance of more than 2.50 m is not permitted[195]. This applies to ballast renewal, ballast cleaning, levelling and tamping, not forgetting renewal of the sleepers. If these operations take place during the night at lower temperatures this poses no obstacle provided tamping is followed by the DTS and a speed limit is imposed.

12.14 Maintenance of the track structure

The introduction of continuous welded rail, elastic fastenings and better quality control has over the last few decades greatly reduced the need for maintenance.

Rails

The most important type of rail maintenance consists in removing defects. This can be achieved by using a short length of rail or, alternatively, by using a wide-gap thermit weld (SkV 75), as long as the defect is limited to a distance of 60 mm.

Another category is formed by deformations of the running surface, such as corrugations, burring and plastic flow. These can be remedied by grinding.

Resurfacing of rails is a method used in the case of local defects such as minor horizontal cracks in the rail head, battered rail heads, rolling seams, flaking on the rail head and wheel burns. The latter are caused by slipping of the driving wheels during acceleration.

Sleepers

Sleeper maintenance is essentially a matter of plugging wooden sleepers at bolt holes with warped threads, mainly near switches. In all other cases of defect the sleeper is replaced.

Switch maintenance

Switches require notably more maintenance than plain track. The aspects below are considered the most important on NS:

- track gauge:
  prescribed track gauges allow for a tolerance of maximum + 5 and - 2 mm but the track gauge at switches and crossings may never be less than 1435 mm;

- switch opening:
  the specified switch opening at a guard rail is 41 mm. The maximum permissible tolerance is + 2 mm. Large deviations from this are rectified by means of shims and blocks;

- switches and crossings:
  thorough inspection is necessary to detect vertical and lateral wear, point attack, shelling, fracturing and cracking. These are repaired by grinding, resurfacing or total replacement;

- pair of switches:
  the wear tolerance on a pair of switches must never be exceeded. There are gauges for measuring whether the wear on the switch blade and stock rail is still within the permissible limits. Any unacceptable deviations can be rectified by grinding or by replacing the pair of switches.

12.15 General observations on track renewal

There are in fact no precise criteria for track renewal because of the number of factors involved. It is important to make maximum use of unbiased information. This subject is dealt with at some length in Chapter 18. A recapitulation of various factors which deserve consideration when preparing a track renewal plan will suffice here. For instance:

- data on rail fractures in relation to traffic tonnage borne;

- rail wear;

- quality of fastenings;

- quality of sleepers, especially in respect of cracks and deterioration of support surfaces;

- current geometry and trends based on track recording car information;

- development of maintenance costs;

- benefit and use of removed materials;

A number of principles can be used as a basis for track renewal. In the case of partial track renewal those track components which are worn are replaced. This inevitably results in variable quality and is therefore only considered for branch lines in Europe and for lines mainly carrying heavy freight traffic, as in the United States.

Almost everywhere in Europe the tendency is to aim for renewal of complete sections of track. For optimum cost-effective use of the structural components the service lives of the individual parts have to be adapted to one another in whole number ratios, for example 1:1 or 1:2. Local conditions, such as formation, ballast bed and curves, can affect the ratio and therefore the choice of track system. To promote homogeneous track quality it is advisable that whole sections of 5 to 10 km, called maintenance sections, be renewed at a time.

At NS the decision for complete track renewal or partial renewal and for the time of execution is

based on life-cycle calculations by the net present value method [24]. In principle, the categories of cost which are considered are:

- investments, initial investments as well as those during the lifetime;
- operating cost: cost for maintenance and renewal and spot repair;
- cost for non-availability of track for traffic.

12.16 Manual track renewal

Manual track renewal causes virtually no obstruction to normal operations. Only on the sections where the sleepers are being renewed there is a restriction to 75 % of the speed limit. This was partly the reason why NS still used manual track renewal on a large scale. Another reason was that the ballastbed on which many tracks in Holland were founded consisted of gravel. Working gravel by hand is a reasonable option. Lastly the type of sleeper is important when selecting the laying system: concrete sleepers can only be laid mechanically because of their weight.

The use of CWR track imposes conditions on the working method; at the point of transition to CWR track the ballast resistance must be sufficiently great and must not be disturbed again too soon after renewal. Furthermore, there must be no irregularities in alignment. This is why the general rule is that sleepers are renewed first and then rails.

Renewal of sleepers

If the ballast bed is made up of gravel, ballast chips or crushed gravel are unloaded before the sleepers are renewed. The point of this is that this broken material mixes with the gravel when the sleepers are being replaced, thus improving the ballast bed and increasing the ballast resistance. The new sleepers and fastenings are unloaded next to the track under renewal.

The sleepers are replaced during the day one by one. After the ballast has been removed from around the sleeper, the sleeper is pulled out to the side of the track and then the new sleeper is installed.

Renewal of rails

The new 180 m-long rails are unloaded next to the sleepers and during the day are welded together to form lengths which can be replaced in a single operation. This length depends on. the length of possessed track and can vary from 540 m to 1080 m. Replacement of rails proceeds as follows:

- the fastenings are loosened and the old rails moved towards the inside of the track;
- the new rails are positioned in the sleeper supports and thermit welded to the rails which have already been renewed. Every 6 m rollers are placed under the rails to be renewed. With the aid of hydraulic tensioning equipment the rails are lengthened until they reach a length which corresponds to a neutral rail temperature of 25° C;
- once the rails have reached the required length the rollers are removed and the rails are fastened onto the sleepers;
- during the day the old rails are sawn into lengths measuring 180 m and during the night are removed by the special trains which supplied the new rails;
- night-time is also the time when the old sleepers and fastenings are loaded for transportation;
- finishing treatment to the ballast bed is either mechanical or manual.

12.17 Mechanical track renewal

12.17.1 Introduction

Even when other branches of industry had long since replaced heavy manual labour by machinery, renewal of track was still to a large extent a manual operation. The reason for this is straightforward. Very intensive operating generally means that it is not possible to take track out of service for an adequate length of time for renewal purposes, but this is an essential precondition for the use of heavy machinery.

The use of concrete sleepers and the need for improvements to the ballast bed by renewal or cleaning point increasingly to mechanization. In choosing from the systems available a number of factors play a role, such as necessary investment, required output, possibilities of track possession and storage space.

12.17.2 Track possession

As previously mentioned, track possession is particularly difficult on NS and has to be restricted mainly to night-time, especially at weekends. This has not only technical and organizational, but also social consequences. Weekend night shifts in particular pose increasing problems as regards staffing. In this respect a plan has been developed covering several years which makes provision for night working during the week. During the week the maximum track possession period at night is rarely more than 7 hours and sometimes is no more than 5½ hours. Usually night freight traffic can use adjacent track or be diverted; this is virtually out of the question for passenger traffic, although sometimes the last passenger trains of the evening are replaced by buses. The amount of time which the track is possessed has an enormous influence on output. For instance, with the track out of service for 5½ hours the output-amounts to 480 m and for 7 hours this figure is 700 m.

12.17.3 Gantry crane method

In many countries this is the oldest known system in which the new track is pre-assembled on a construction site. During possession of the line, the gantry cranes are then used to lift the old sections of track and place the new ones. These cranes run on temporary rails which are laid on either side of the track.

There are 2 variants of the gantry crane method. In the first one assembled track sections are installed. In the second version, which is mainly concerned with CWR track, only the sleepers are laid and directly afterwards the rails are placed onto them. The latter method is used most widely on NS.

Originally the choice was to lay pre-assembled sections only, mainly because this system kept the track out of service for a minimum length of time. In view of the fact that the track has to be of the CWR type and too many thermit welds are undesirable it is necessary to install temporary rails which are later replaced by the definitive long-welded rails.

The variant in which sleepers are laid and continuous welded rail fastened directly onto them is more attractive (Figure 12.67). In this so-called continuous method the quality of the work is clearly better and costs less. For this reason attempts are made to create situations where track possession will allow this continuous method to be used.

12.17.4 Track section method

The track sections are pre-assembled by laying the sleepers on an assembly track and installing temporary 30 m-long rails. The prefabricated sections are raised by gantry cranes and loaded onto the section train, each train having a capacity for maximum 900 m of track.

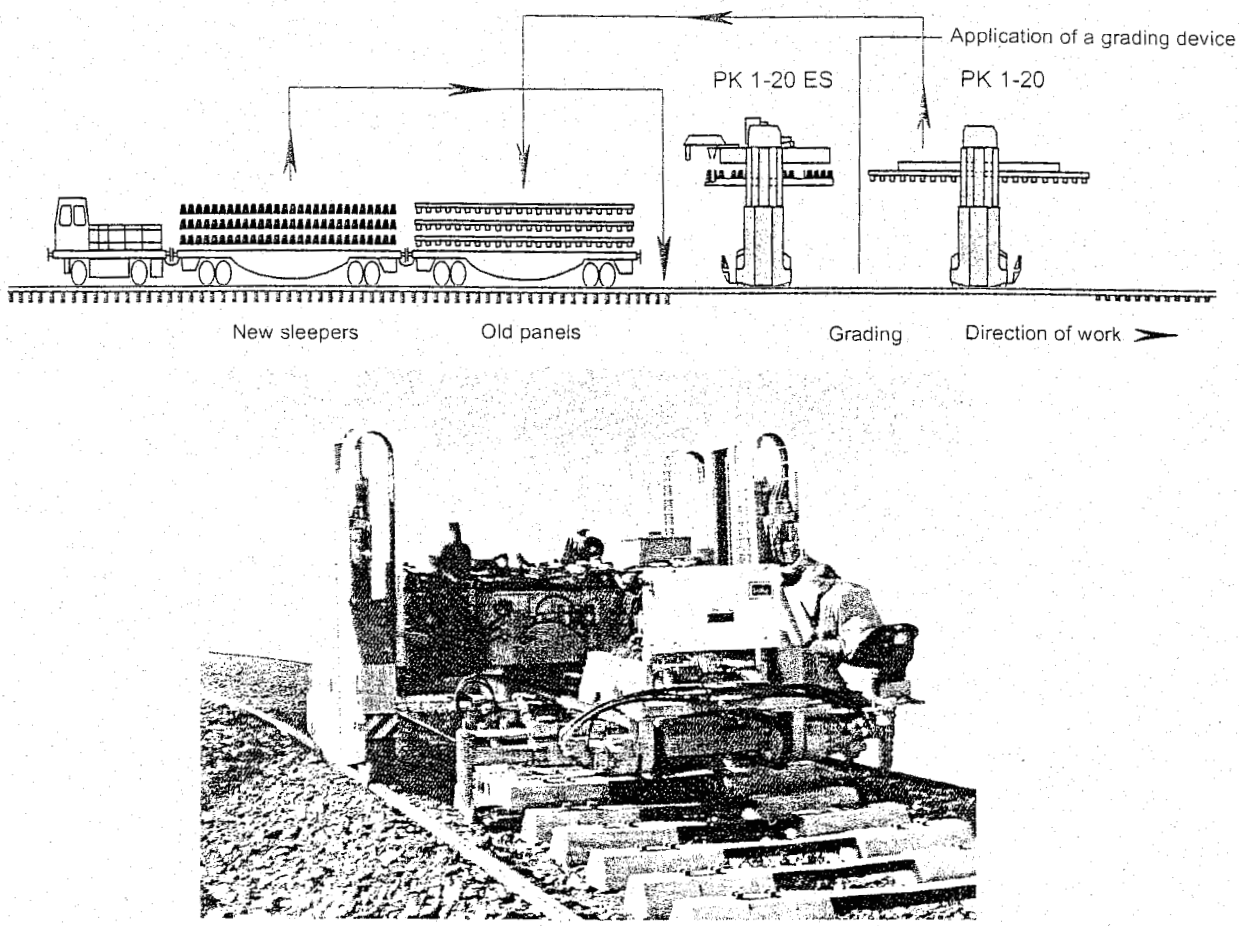


Figure 12.67: Continuous track relaying with gantry cranes

The work on the track can be split into preparation, track section renewal, rail renewal and finishing work. The task begins with the unloading of 180 m-long rails into trenches specially dug in the shoulder of the ballast bed. The rails are positioned at exactly the right height and for the sake of the gantry crane set at a gauge of 3.31 m. Experience has shown that the rails need no support. By filling out with ballast sufficient stability is achieved. For welding the rails are supported on blocks on either side of the weld.

Section renewal on NS takes place at night during track possession [105]. A number of work trains depart in the following order:

- train with track sections, also including a staff coach, an equipment vehicle, a wagon carrying the cranes and a wagon carrying a bulldozer;
- ballast train;

At the location designated for renewal the gantry cranes rise from the wagon and run on the guide rails. The cranes, provided with special rail clamps, lift the first section of old track, move back along the rails and deposit the track section on the empty wagon of the section train. A second section of track is then removed.

The bulldozer is unloaded in the "gap" which has been created and begins digging out the ballast. Since the bulldozer blade covers the total trench width and is fitted with guide rollers supported on the new rails it is possible to dig accurately to the required depth.

When the trench has been dug the gantry cranes take a new track section from the wagons and lay it in the trench. Then the following cycle is continuously repeated: the gantry cranes remove a track

section, deposit it on the emptied wagon, take a new track section to the trench and position it. This cycle lasts 7 to 10 minutes.

As soon as the number of renewed sections permits, the ballast train is unloaded. Following this the track is levelled and aligned using a tamper-liner.

After a few weeks the temporary 30 m rails are replaced by the final continuous welded rails. Finally the ballast is profiled mechanically and the track is once more levelled and aligned.

The major problem of renewal according to the section method is with jointed track which is temporarily laid on the new sleepers. The joints of the temporary rails are often of such poor quality that the sleepers crack around these joints and once the final rails have been installed have to be replaced. For this reason a method is being sought in which the new continuous welded rails can be installed directly. This continuous method is discussed in the following section.

12.17.5 Continuous method

Although in principle the continuous method does not require a construction site, existing tracks are used for storing the new sleepers and for dismantling the old track sections. The procedure for track renewal during possessions at night is summarized in Table 12.2. The following points are worth noting in comparison with the pre-assembled track section method:

1.	Remove and load old sections
2.	Excavate ballast
3.	Place sleepers
4.	Adjust sleepers and place rails
5.	Place rollers under rails
6.	Adjust rails to correct length (stretching/heating) and fix fastenings
7.	Discharge ballast, tamp and stabilize track

Table 12.2: Summary of NS continuous relaying process

The sleepers are loaded crosswise on special wagons, with special attention being paid to securing the sides.

The new sleepers are lifted in groups with the help of a lifting bar suspended from two gantry cranes and are laid on the track.

A rail installation machine positions the sleepers and then places the new welded rails on the sleepers. By using long outriggers this machine can run on the rails just laid.

The rails are placed direct onto rollers behind the rail installation machine. Subsequent tensioning or heating ensures that the rails are lengthened so that once they have been fastened to the sleepers they are stress-free at 25 °C. An output of about 400 m track during a period of 5 hours and about 540 m during a period of 6 hours can be achieved.
Before the track is brought back into service the entire ballast profile is installed, followed by mechanical tamping and stabilizing. During the next night tamping and stabilizing is carried out once more.

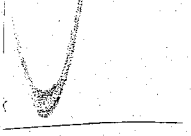
The photographs in Figure 12.68 to Figure 12.86 show various stages during renewal according to the continuous method.

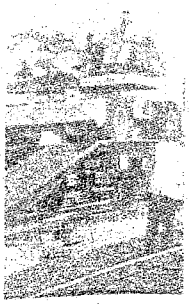
As has already been mentioned, it is of utmost importance that the track quality is raised to as high a level as possible immediately after renewal. If this does not take place at the welding depot, all welds on NS are straightened and ground and at the same time the entire track undergoes preventive grinding in three passes. This is carried out as soon as possible after renewal and certainly no later than after one year. At the same time as straightening and grinding the track is tamped once again.

Figure 12.6
descends t

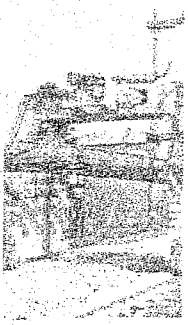
Figure 12.7
train;

Figure 12.72:
concrete bloc

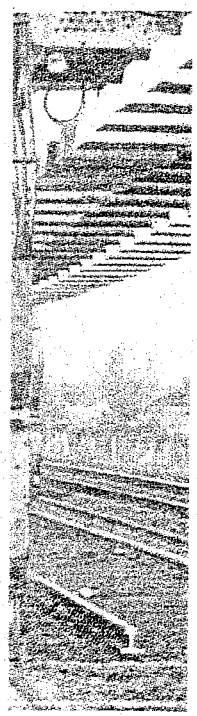




ain after arrival at
o track;



al of old section; lo



nuous method; loa

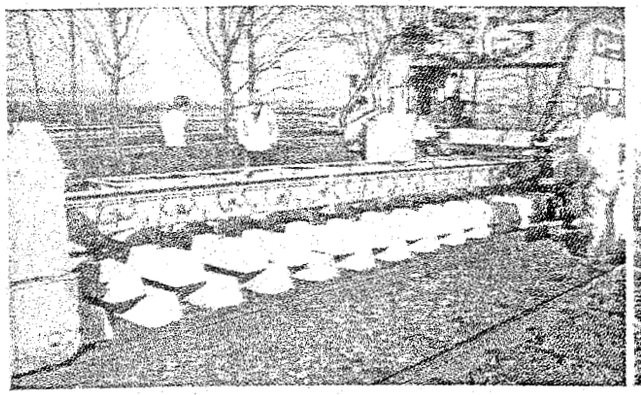


Figure 12.74: Unloading sleepers "one by one"; 2 x 15 m = 30 m;

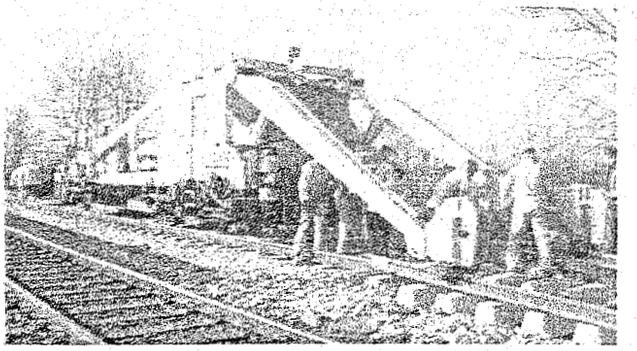


Figure 12.75: Machine places rail on sleepers;

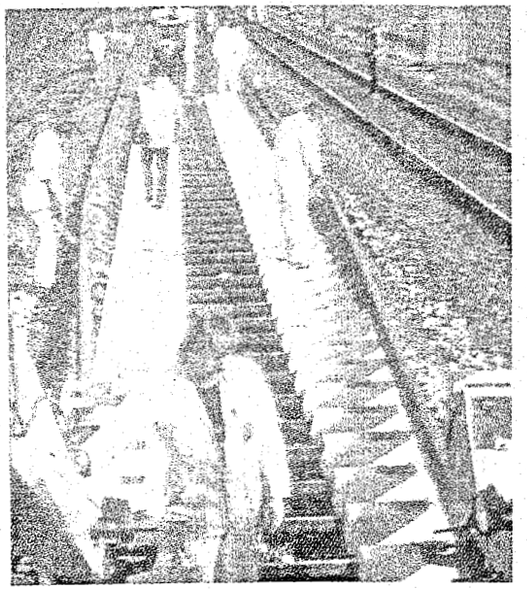


Figure 12.76: From temporary rail to final rail;

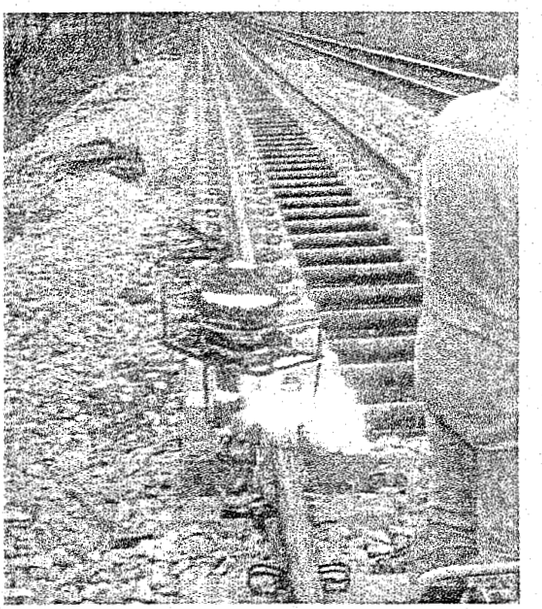


Figure 12.77: Thermit weld at beginning of renewal;

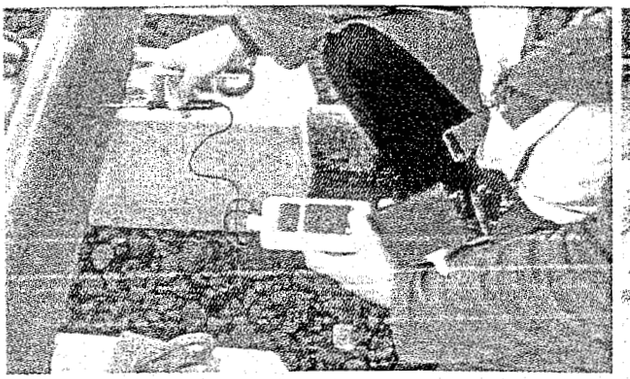


Figure 12.78: Measuring rail temperature;

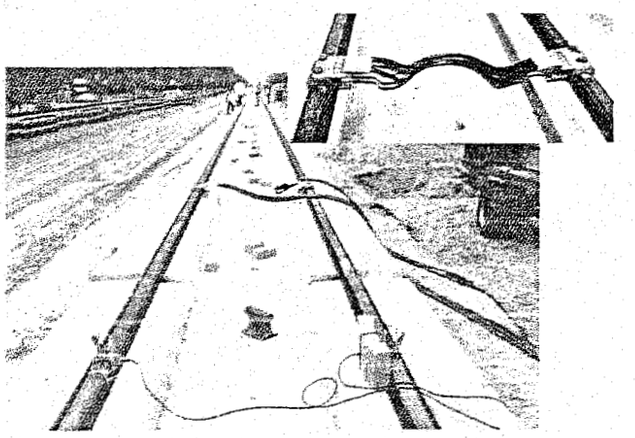


Figure 12.79: Heating rail;

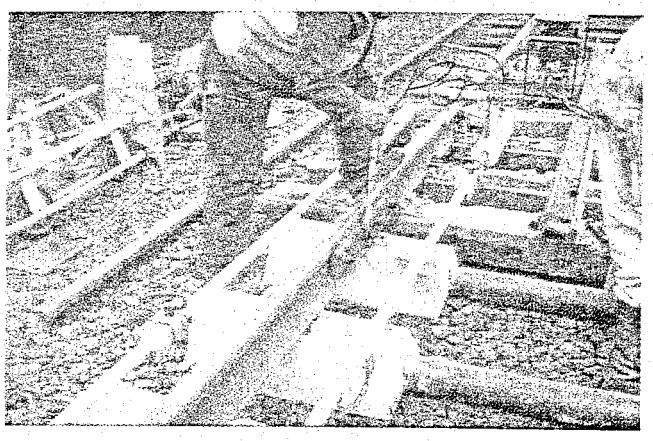


Figure 12.80: Connecting renewed track to old track; hydraulic stretching device;

Figure 12.81: Fixing DE spring clips;

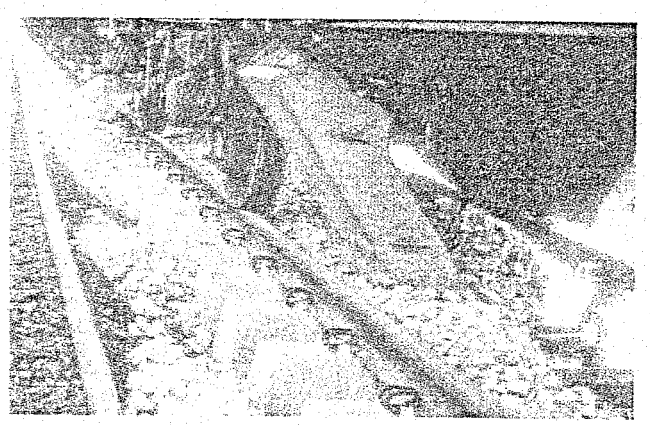


Figure 12.82: Discharging ballast;

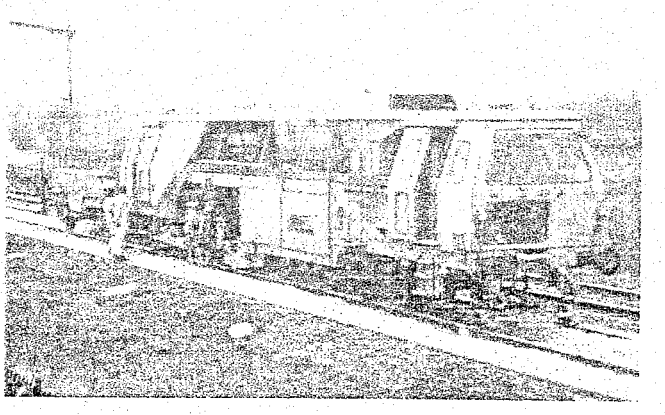


Figure 12.83: Tamping track using Mainliner;

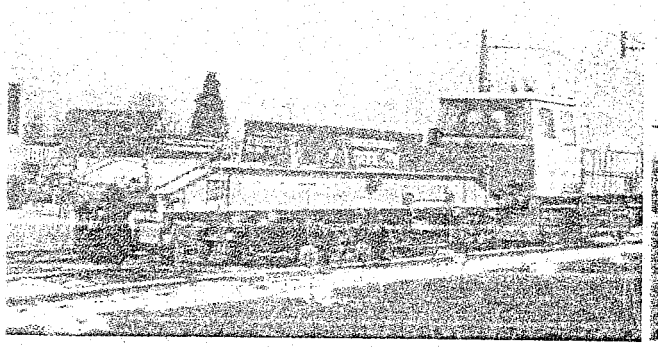


Figure 12.84: Stabilizing the track;

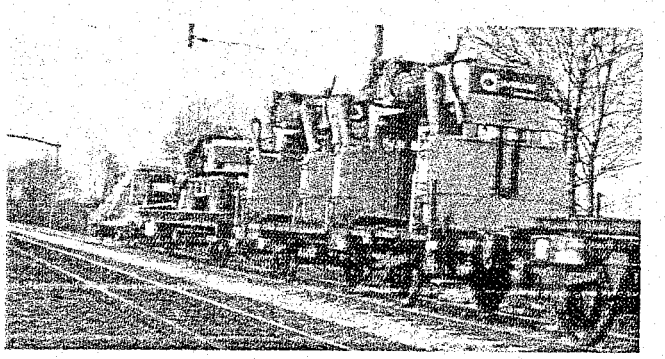


Figure 12.85: Departure of work/section train;

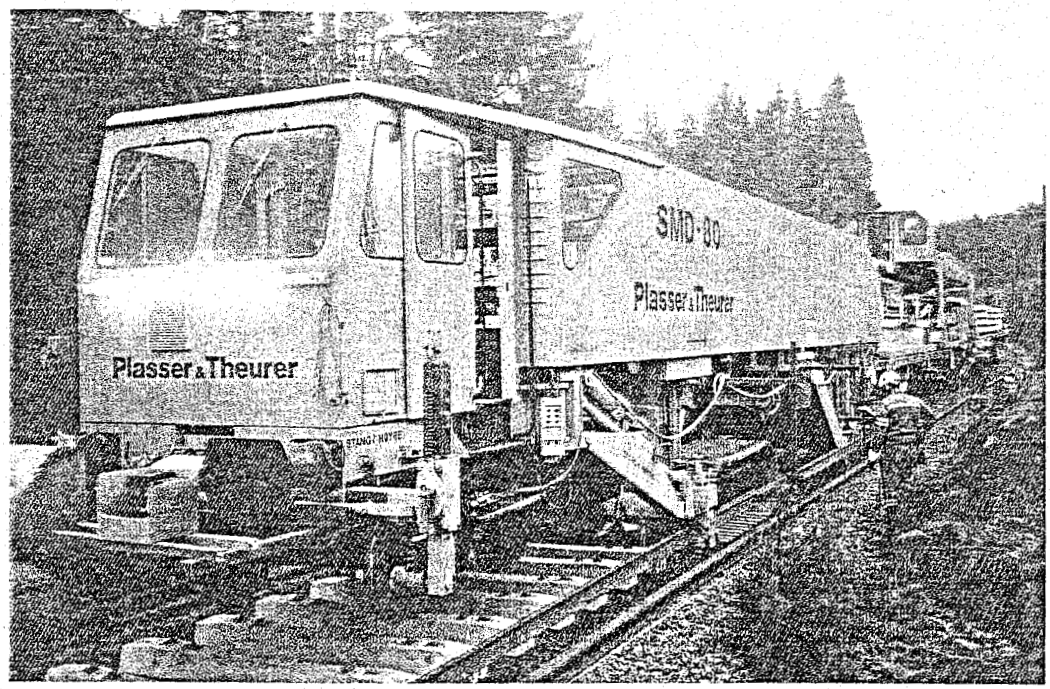


Figure 12.88: SMD 80 track relaying and laying train

The various principles used in the different Plasser & Theurer machines are presented in Figure 12.89.

12.18 Switch renewal

Prefabrication

Traditionally switches and crossings are assembled on a construction site. After laying out the switch sleepers firstly the pair of switches is put into place and then the other switch components and rails. Once the geometry has been corrected, first the straight section and then the curved rail is fixed in place. It is most important that the correct track gauge and switch opening are installed. Another method is to prefabricate switches near the place of installation on a temporary assembly site. This can be a platform beside the track or a section of track.

Modern switches and crossings are high technology units with a high degree of precision. The best method therefore is to assemble them in the factory and to transport the pre-assembled units on special transport cars to the site. Figure 12.90 shows such a wagon loaded with the newly developed "plug in" switch of VAE.

Installing switches

More and more switches are being installed by means of cranes. This is a fast, useful approach especially for short switches. The important thing is digging the trench. Use of laser control on the bulldozers guarantees an extremely accurate excavation depth.

For precise installation of large and heavy switches and crossings special switch laying machines have been developed, which can handle complete units with the necessary care and high work speed (Figure 12.91).

The use of welded switches means that attention must also be paid to the stress distribution in the switches and in the adjacent tracks. The switch itself is not heated or stretched but the adjacent track is. Care is taken to ensure that with CWR switches at least 100 m of track on either side is also continuous welded. This prevents undesirable displacements and uneven stresses in the switches.

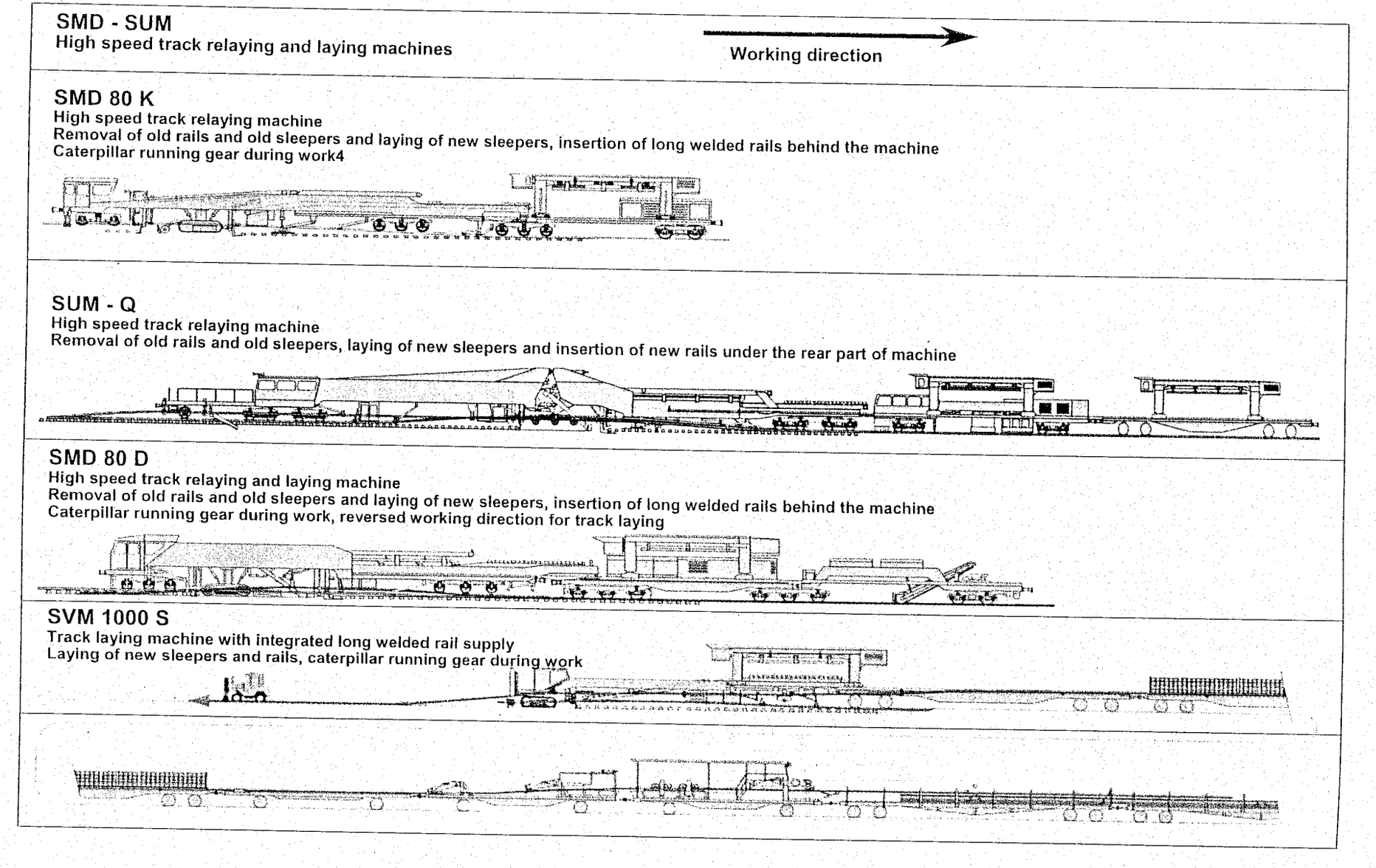


Figure 12.89: Various principles applied for renewal trains

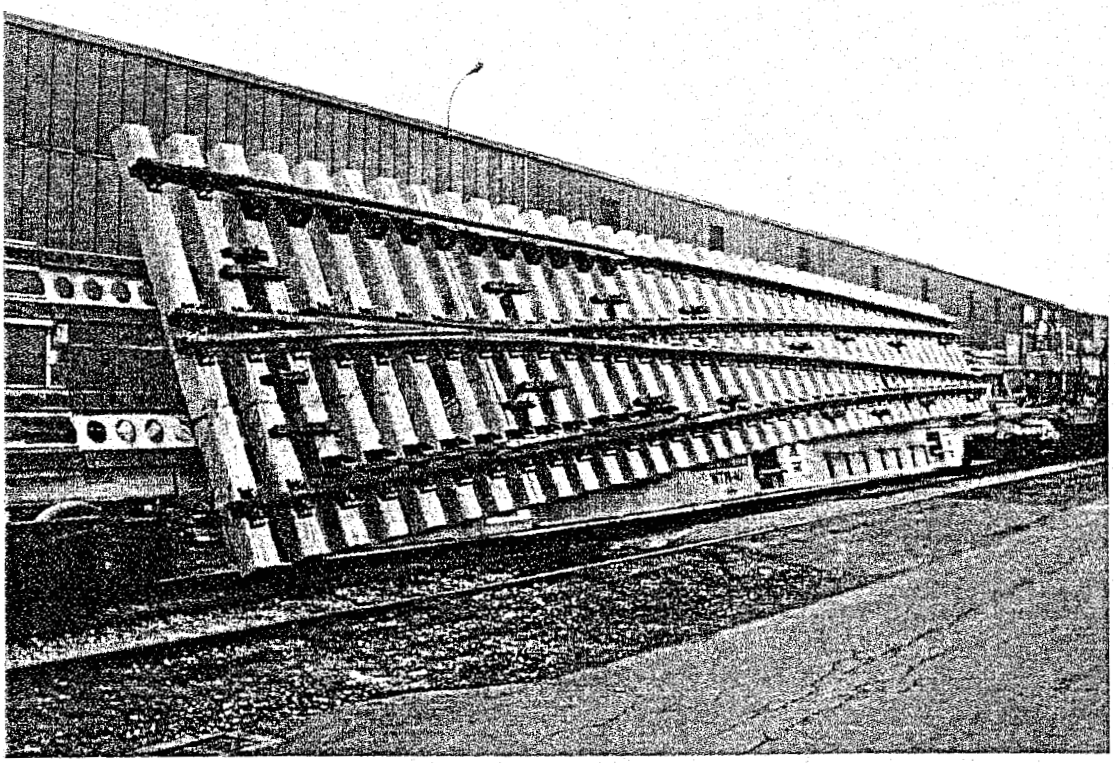


Figure 12.90: "Plug in" crossing on special transport vehicle

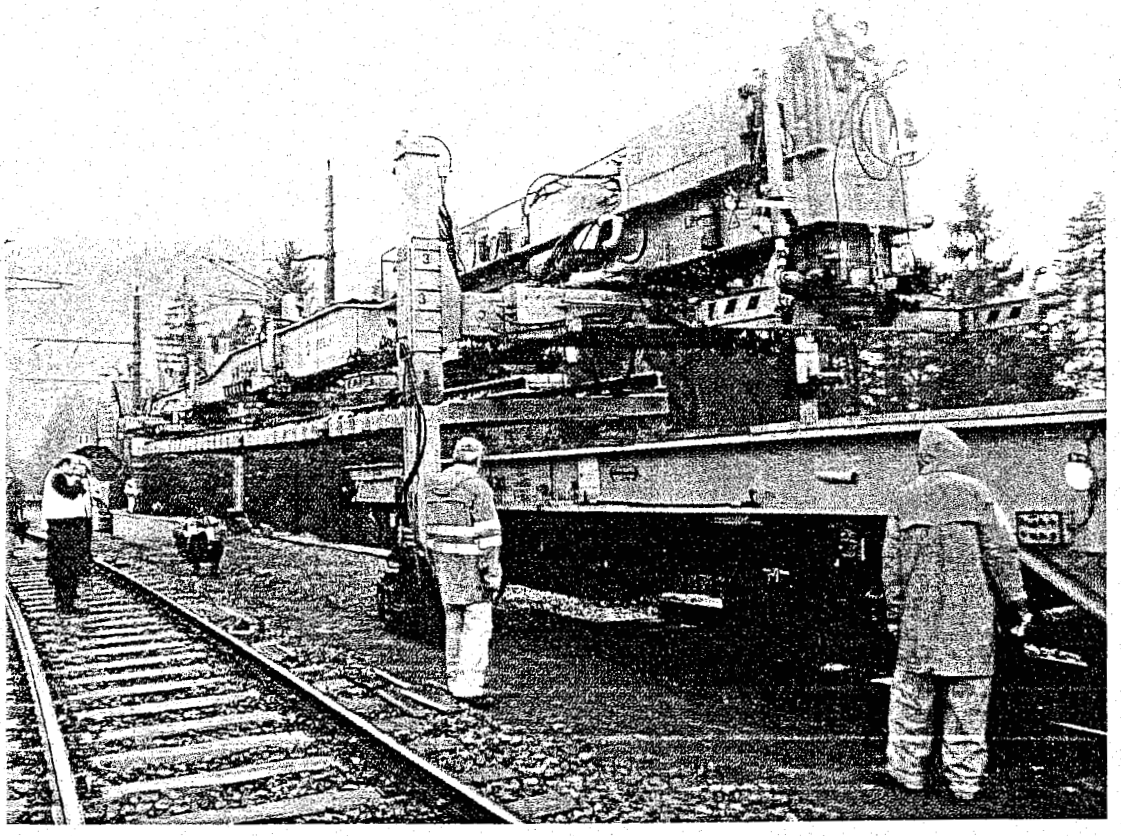


Figure 12.91: WM 500 switch installation machine

12.19 Track laying

12.19.1 General considerations

This section looks at the topic of track laying. A basic distinction can be made between the following two methods for laying new track:

- "Overhead" construction. This is the method used when there is no track at all to start with. All materials are transported via the track which is under construction. Transportation of rails longer than 30 m is cumbersome. This means that the rails have to be welded together once in the track or have to be replaced later by continuous welded rails. Lately methods have been developed to install also long welded rails overhead.

- Construction "from the side". This method is applied if a track already exists. The materials can then be transported via this track and discharged sideways onto the right location. Long rails of the order of 200 m then pose no problem.

12.19.2 Track construction trains

There are a few companies worldwide specializing in the manufacture of track maintenance and construction machines. The Austrian Plasser & Theurer and the Swiss Matisa in particular export to all parts of the world. They have developed a variety of construction trains for laying track.

In all cases the system is designed to lay or renew track at high speed. The design of the machines is fully geared towards laying CWR track. Individual sleepers are placed via a conveyor system. Rails are installed on the wagons for the gantry cranes, which convey a wheeled pallet, to run on. Such machines have an output of 600 m per hour.

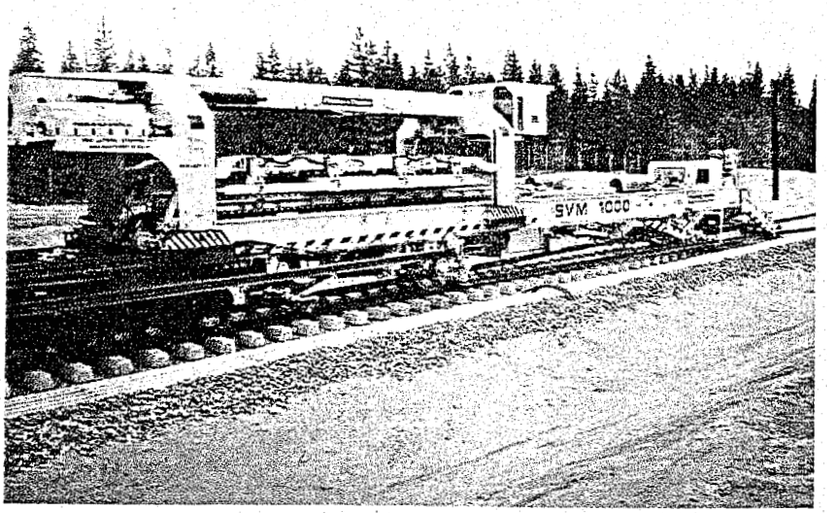


Figure 12.92: SVM 1000 S track laying machine

Figure 12.92 shows the Plasser & Theurer track laying machine SVM 1000 S. The principle is annotated in Figure 12.89. The SVM 1000 S consists of the actual laying machine, the sleeper-transporting wagons, one or more gantry cranes which run on the wagons, a long-rail conveying unit and a rail loading train.

The long rails are pulled from the rail loading train by the rail conveying unit and pushed over guide rollers to the front of the laying machine. There the ends are gripped by a tractor and pulled further forward till a complete rail length is in front of the laying machine. Then the rails are taken over by the rail guiding clamps of the machine.

Running on the sleeper wagons, a gantry crane picks up the sleepers and carries them to the transfer and storage platform of the sleeper laying machine. In the case of longer sleeper trains, this transport is done by several gantry cranes working together in relay.

Using conveyor chains the sleepers are taken from the storage platform to the sleeper laying unit which automatically lays them in the correct position.

The new rails which had been pulled into the front the machine are guided into the baseplates by hydraulic roller tongs at the rear end of the machine. The machine rests on caterpillar tracks at the front while the remaining part of the laying train runs on the newly laid track.

12.19.3 Platow system

In Russia use has been made for a considerable number of years of the Platow system, pictured in Figure 12.93, which is, also suitable for track renewal.

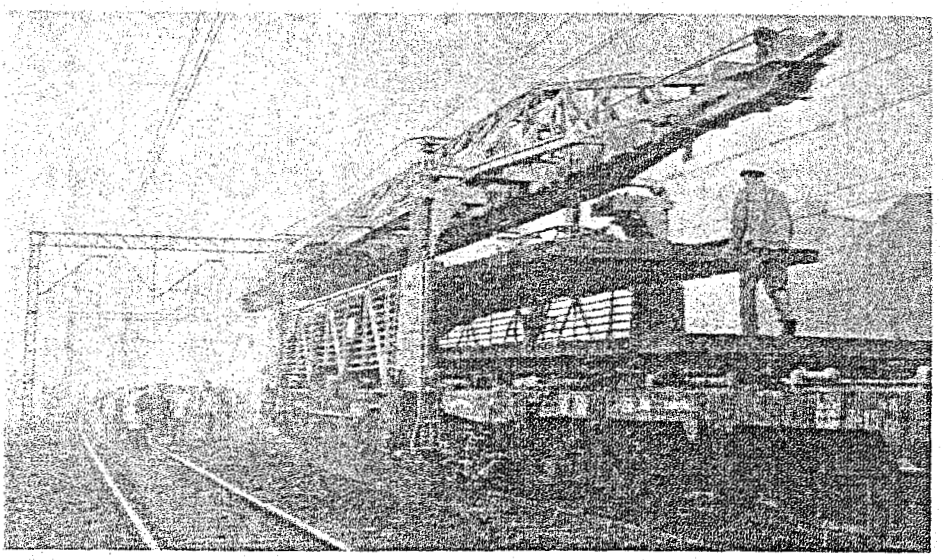


Figure 12.93: Platow system for track laying

Pre-assembled track sections are loaded on special wagons over which the sections can be transported quickly, using winches, to the head of the train where a heavy crane takes the track sections from the wagon and lays them on the ballast bed. The cranes were manufactured in fairly large numbers by the Kaluga works for various countries in Eastern Europe and for developing countries.

Cranes in use today weigh 78 t and can lay track sections with concrete sleepers in lengths of up to 25 m. Frequently the rails are subsequently replaced by new continuous welded rails measuring 210 m in length. The output of this system is very high, covering 810 m per hour, or, in other words, 1 minute 50 seconds per track section. It goes without saying that at such a fast rate all other operations have to be adapted. In this respect the supply of ballast and accuracy of laying pose the major problems.

12.19.4 TGV tracks

For the construction of the Paris-Lyon TGV line exceptionally high requirements were specified regarding accuracy and care. To meet these requirements two different French systems were applied.

In the Drouard method a temporary track is laid over a distance of 10 km. The sleepers and 288 m lengths of rail are conveyed via this track. The materials are discharged laterally; some of the ballast is also transported on this track. Of the 30 cm of ballast which eventually comes to lie under the sleepers, 20 cm is in place before the track is laid.

Once the track is laid the rest of the ballast is brought to the new track, which is raised by 10 cm and tamped, and the rails are connected using thermit welds. After 10 km of track has been laid the temporary track is disassembled and re-placed over the next 10 km.

In the Desquennes and Giral method use of a temporary track is dispensed with. The 288 m-long rails are moved along the length of the prepared line by means of caterpillars. The rails are welded together to form lengths of about 1200 m.

The rails are spaced 3.10 m apart and are not supported by any special means. The front bogie of the huge machine which brings the sleepers from the work train to the line and which places the rails on the sleepers runs on these rails. The rear bogie of the machine runs on the track which has just been laid.

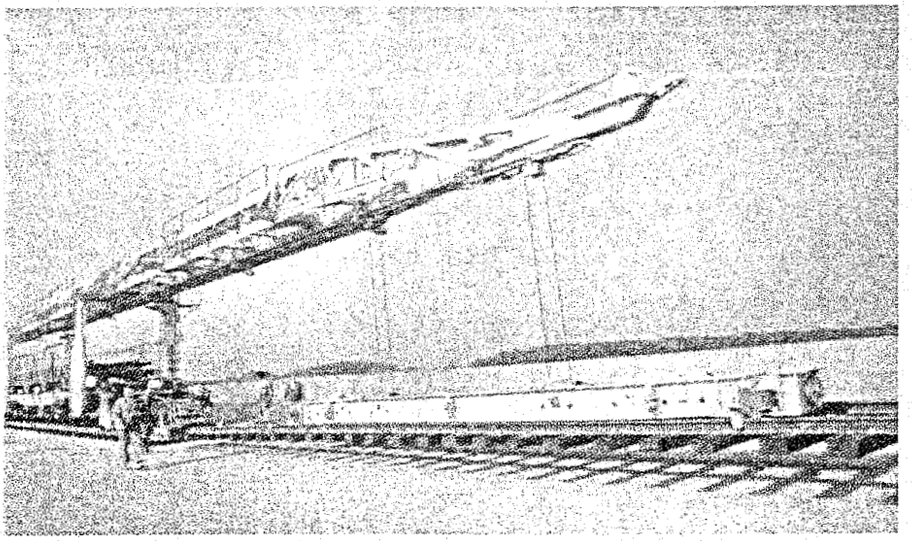


Figure 12.94: Laying of auxiliary track for TGV-A

The work train carrying the sleepers also runs on this track. Gantry cranes are used to convey the sleepers above the train. An output of 300 m/h is possible with this system. The ballast is conveyed entirely along the new section of track.

The track laying system used for the TGV Atlantique involved the laying of an auxiliary track with a tilting panel launching beam as shown in Figure 12.94.

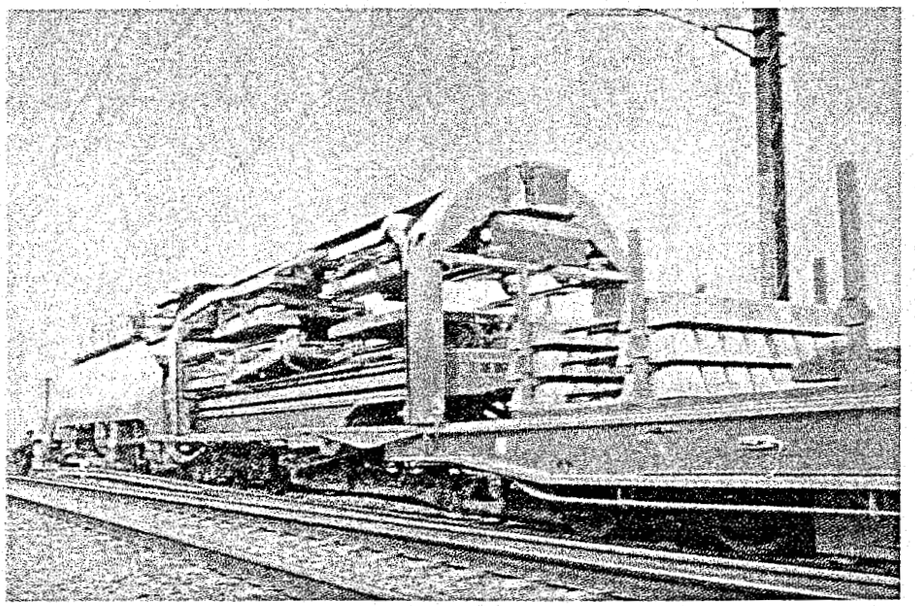


Figure 12.95: TGV-A sleeper supply

The materials are transported via this track. Gantry cranes which run on the new rails placed next to the auxiliary track convey the sleepers to their correct position. Figure 12.95 and Figure 12.96 show how this is done. Ballasting and track adjustment are carried out in eight successive passes - five tamping operations each with maximum lifting of 80 mm, one tamping and levelling operation, one final tamping operation, and one "maintenance" tamping operation after stabilizing and before the track is put into service.

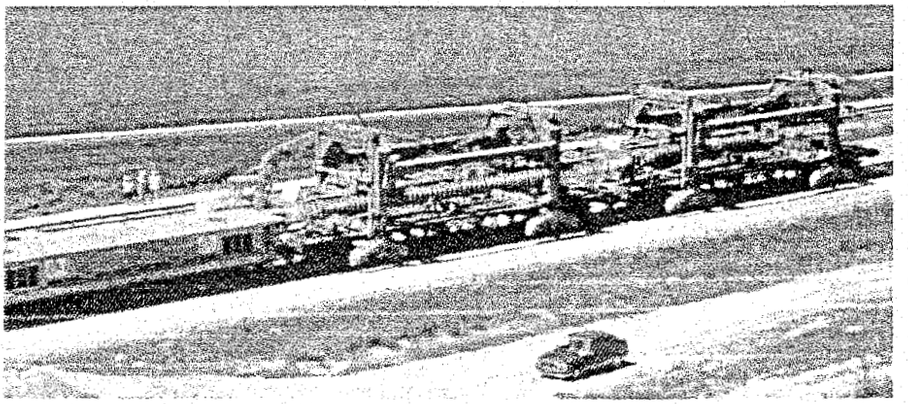


Figure 12.96: TGV-A: sleeper laying

The last three operations are carried out using a continuous-action tamper equipped with lasers for lining and levelling, and assisted by an on-board computer which controls the adjustment of the curves by using the laser beams.

12.20 Deterioration of Track Geometry

12.20.1 Introduction

Under the influence of dynamic track loads the track geometry deteriorates. The mechanism governing this phenomenon is rather complex. If a track is freshly tamped it is well-known that directly afterwards relatively large settlements occur. If every point of the track were to settle by the same amount no irregularities would develop. However, these settlements are often far from uniform due to inhomogeneities in support conditions, track structure and load distribution. This results in differential settlements which lead to the development of irregularities in the wavebands experienced by the rolling stock.

Many investigations have been carried out on the fundamentals of the deterioration mechanism and the possibilities of controlling this phenomenon via existing or improved maintenance methods. Many of these investigations were coordinated by the Office for Research and Experiments (ORE)'. This chapter mainly refers to the work carried out by committee D 161, which completed its work in 1988. The first stage of this work consisted in studying existing data from the sites used by ORE D 117 and also data available from the normal track recordings of a number of administrations. This new examination showed that the factors governing the rate of deterioration were not obvious and that the unknown factors in the track were the most important in determining both the average quality and the rate of deterioration. Based on the experience gained from this examination, new test sites were specially selected to examine the various aspects governing the rate of deterioration.

12.20.2 Historical records

The original tests on the rate of deterioration of track geometry were carried out by ORE D 117. This committee attempted to analyse the effect of different types of traffic, track construction and maintenance machine on the quality of the track and its rate of deterioration. While the results of these tests on the effects of these variations were not very conclusive, the committee nonetheless felt that the quality of the track on relaying was the most important factor.

Some of the measurements taken on the D 117 sites by SJ were continued so that a much longer time history of the deterioration rate was available, and also so that the effect of track maintenance could be evaluated. BR, SBB and SNCF had been recording track geometry on much of their networks on a regular basis for several years. Some analysis was carried out on a selection of these records with the aim of acquiring a greater understanding of the variation in geometry during a maintenance cycle. The main conclusions from this work were:

a. both the vertical quality and alignment deteriorate linearly with tonnage or time between maintenance operations after the first initial settlement. This trend is not always the case for sections with high deterioration rates;

b. the rate of deterioration is very different from section to section even for apparently identical sections carrying the same traffic;

c. neither type of traffic nor track construction could be statistically shown to have a marked effect on the quality or on the rate of deterioration;

d. the rate of deterioration appears to be a constant parameter for a section of track regardless of the quality achieved by the maintenance machine;

e. in general tamping machines improves the quality of a section of track to a more or less constant value.

In addition to examining geometry data, reference [218] also considers some of the fundamental questions as regards vehicle/track interaction and the effect this has on the track geometry. The track

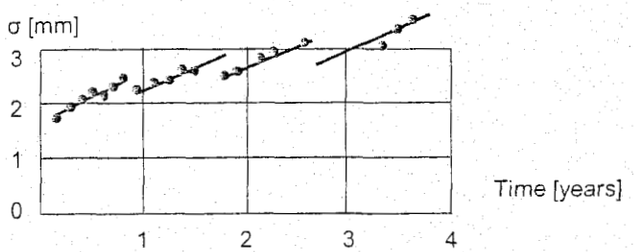
12.20.4 Deterioration rates of geometry

The deterioration rates of quality indices are calculated as a function of traffic in mm/MGT or of time in mm/year. Without including quick settlement and rapid deterioration of track immediately after tamping, the deterioration rate generally displays a linear trend between two maintenance operations as shown in Figure 12.97.

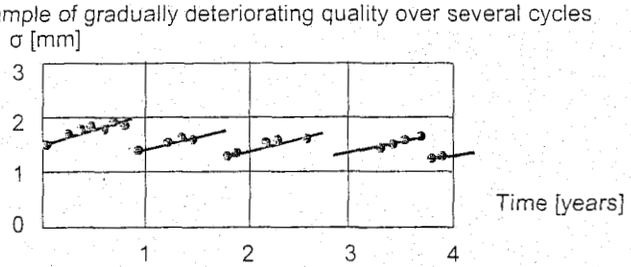
	BR	SBB	SJ	SNCF	NS
Longitudinal level	2.0	1.1	1.6	0.7-1.4	1.1
Alignment	-	0.3	0.6	-	0.8

Table 12.3: Mean deterioration in mm/100 MGT

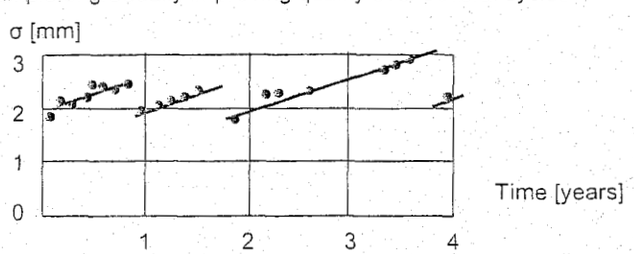
BR has carried out a large number of investigations concerning track deterioration. A number of representative examples are displayed in Figure 12.100. BR established that the average quality over several maintenance cycles remains stable in 65% of the cases but tends to increase slowly in 33% of the cases (Figure 12.100.a) and improve marginally for the remaining 2% (Figure 12.100.b).



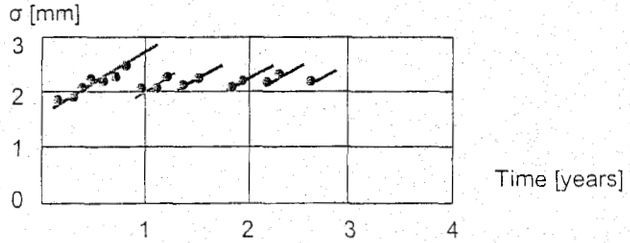
a . Example of gradually deteriorating quality over several cycles



b . Example of gradually improving quality over several cycles



c . Example of a longer than normal maintenance cycle



d . Example of a shorter than normal maintenance cycle

Figure 12.100: Various trends of track deterioration observed on BR

The following table, obtained from reference [88], gives the mean rates of deterioration observed on various networks and expressed in standard deviation as measured by the NS car in mm/100 MGT.

The scatter in deterioration rate values is very large: for longitudinal level the maxima are about 10 mm/100 MGT, with current mean values from 1 to 2 mm/100 MGT, and for alignment the maxima are only about 2 mm/100 MGT. For this reason the longitudinal level indices, which vary more quickly, are preferred to the alignment indices when a decision has to be made on levelling/lining work.

The scatter is also very large between short zones situated very close together. Figure 12.101 shows two zones on the same line, of the same age and construction carrying the same traffic, a 0.7 mm/100 MGT slope being obtained in the first zone, and a 2.0 mm/100 MGT slope in the second.

D 161 attempted to study factors which could influence the variations in deterioration rate under different track and traffic conditions for the longitudinal level using multiple linear regressions, but these were unsuccessful.

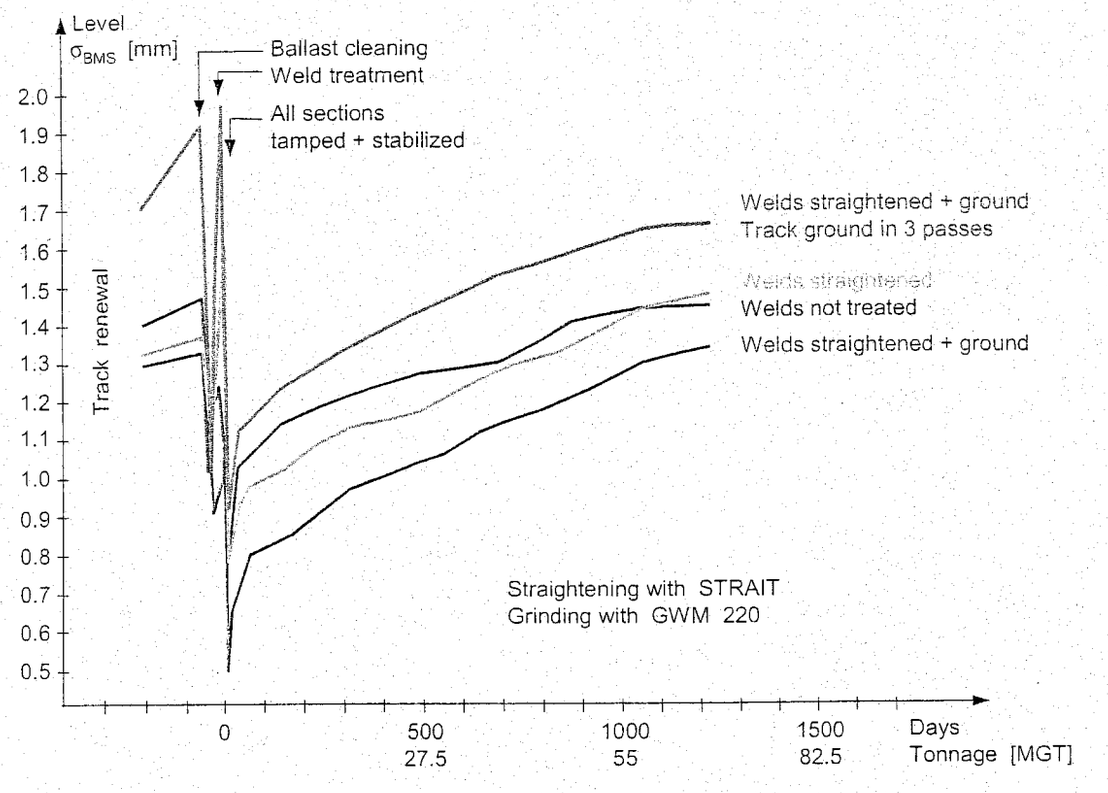


Figure 12.103: Average deterioration in level for 4 test sections of 1.5 km on NS

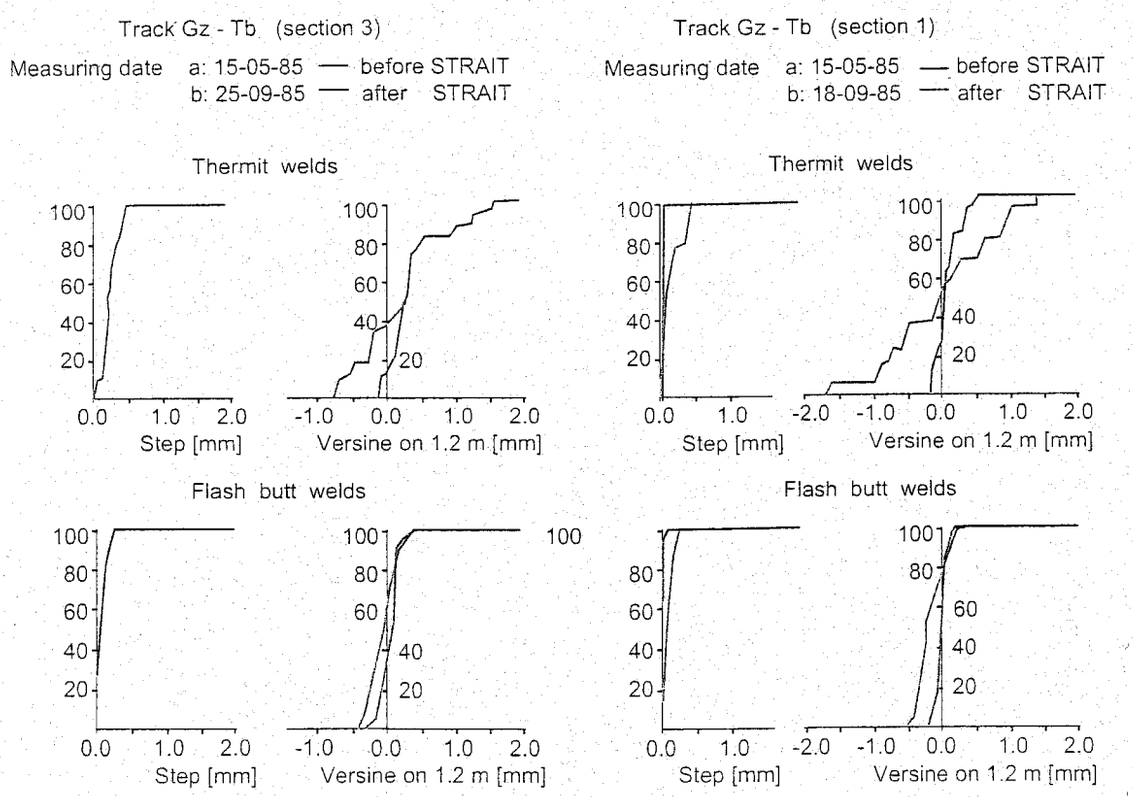


Figure 12.104: STRAIT method: improvement in weld geometry

However, recordings made with an electronic straight edge over 1.20 m, displayed in Figure 12.104, does show the efficacy of improvement operations on the weld quality. This is confirmed by the weld signal produced by BMS-2, presented in Figure 12.105. The welds in the first two sections hardly give any response at all, while the opposite is true in sections 3 and 4. This tallies with the number of cracked sleepers. From this information it would seem that straightening and grinding is effective in reducing cracking of concrete sleepers, but straightening without grinding is substantially less effective.

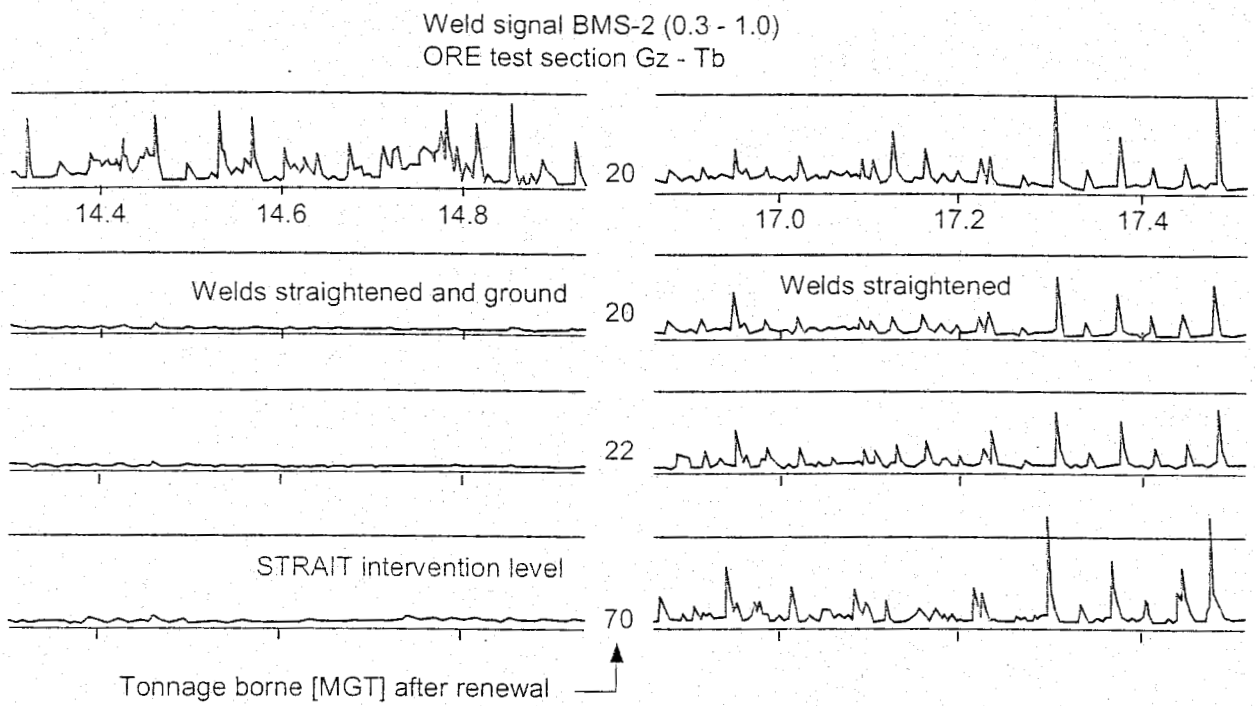


Figure 12.105: Weld signal BMS-2 on ORE D 161 test section Gz-Tb.

12.20.7 Development of corrugation

The axle box accelerations recorded on the D 161 Gilze-Tilburg test section are presented in Figure 12.106 and show the effect of preventive grinding on the rate of development of corrugations. The preventively ground section 1 hardly displays any corrugations after 50 MGT, whereas some rails in the other sections, with 70 MGT after renewal, are already close to the grinding intervention level.

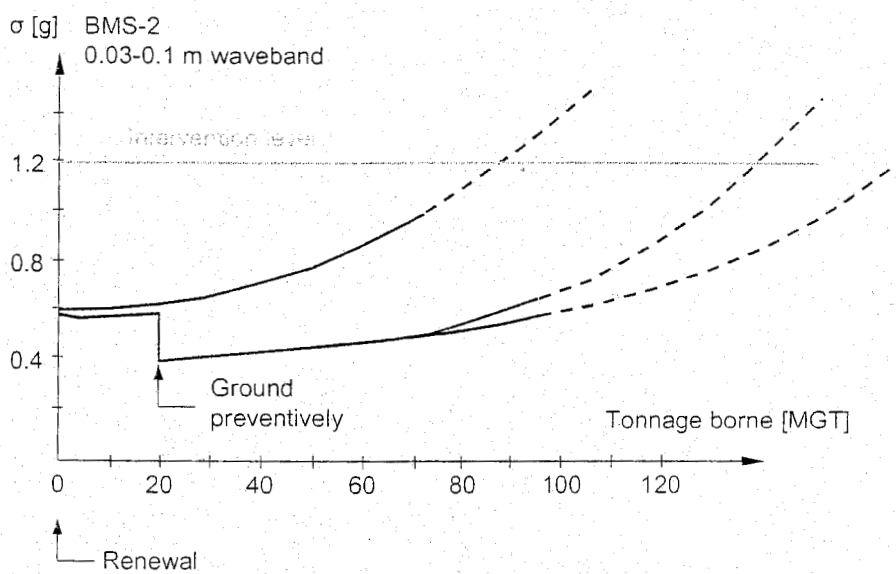


Figure 12.106: Effect of preventive grinding on corrugation development

12.20.8 Effect of stone blowing

A selective method for correction of level is the addition of stone under the sleeper on top of the consolidated ballast bed. BR tests generally showed higher durability of the geometry after blowing than after tamping.

The example in Figure 12.107 shows that the rate of deterioration of track treated by blowing is quite different from that treated by tamping only. Immediately after stone blowing deterioration becomes very slow. The level previously reached before intervention is found only after a period corresponding to at least two maintenance cycles of tamping on the example presented.

For further details please refer to 12.7.

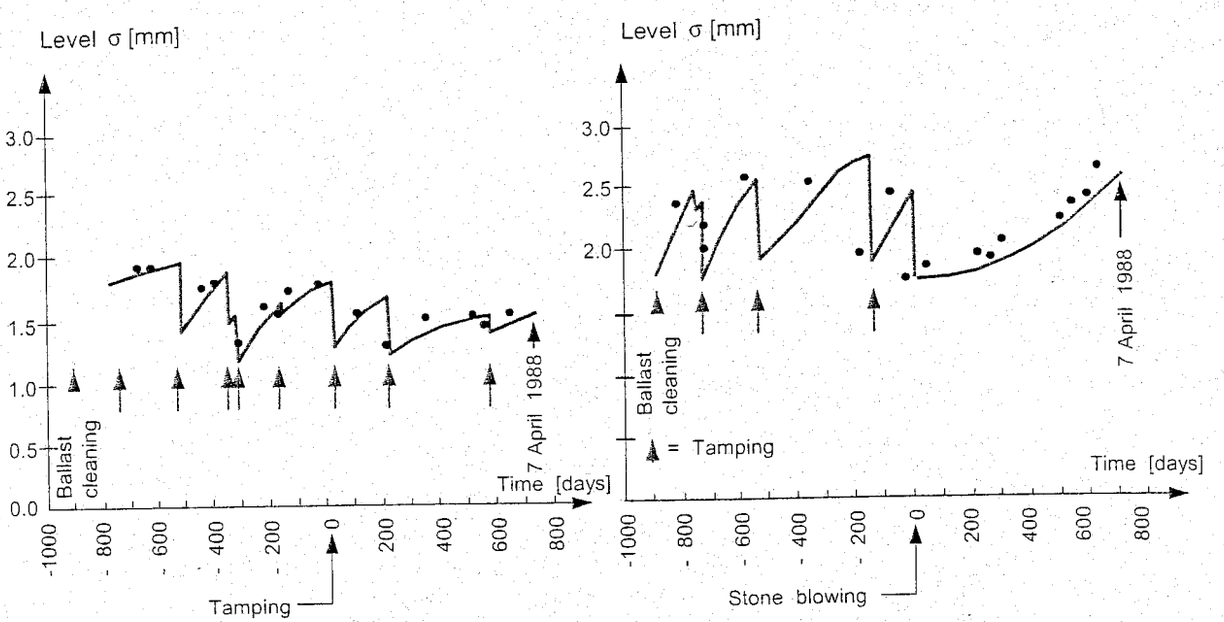


Figure 12.107: Results of stone blowing test.

12.20.9 Development of lateral track resistance

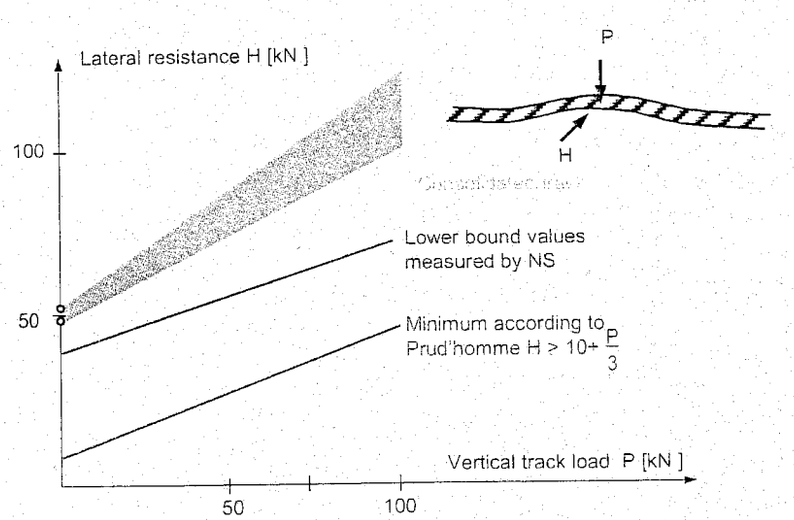


Figure 12.108: Lateral track resistance versus vertical track load measured by NS.

Tamping operations generally result in an improvement in track geometry and a decrease in lateral track resistance. As the accumulated tonnage grows the geometry deteriorates and the lateral resistance improves. After complete track renewal or a ballast cleaning operation the lateral resistance drops so much that, without special measures, a speed restriction is required. The duration of the speed restriction is in fact determined by the rate at which the traffic restores the lateral resistance. At the end of the seventies NS performed a series of tests with an adapted tamping machine, as described in [73], in which lateral resistance as a function of vertical track load was determined on the Eindhoven-Venlo line [68]. The results are summarized in Figure 12.108.

For a good quality consolidated track the lateral force at which the track starts to displace, if no vertical load is applied, is about 50 kN. If a vertical track load of 100 kN is applied this value becomes 100-130 kN. The difference between NP 46 and UIC 54 rail profiles falls within this range. In general the scatter is rather large. For this reason Figure 12.108 also indicates the minimum values found on the Eindhoven-Venlo line operated for a trial period at 160 km/h. These forces are substantially higher than the minimum values required according to the Prud'homme criterion. The lower-bound values measured on the Eindhoven-Venlo line may be regarded as an indication for terminating speed restrictions.

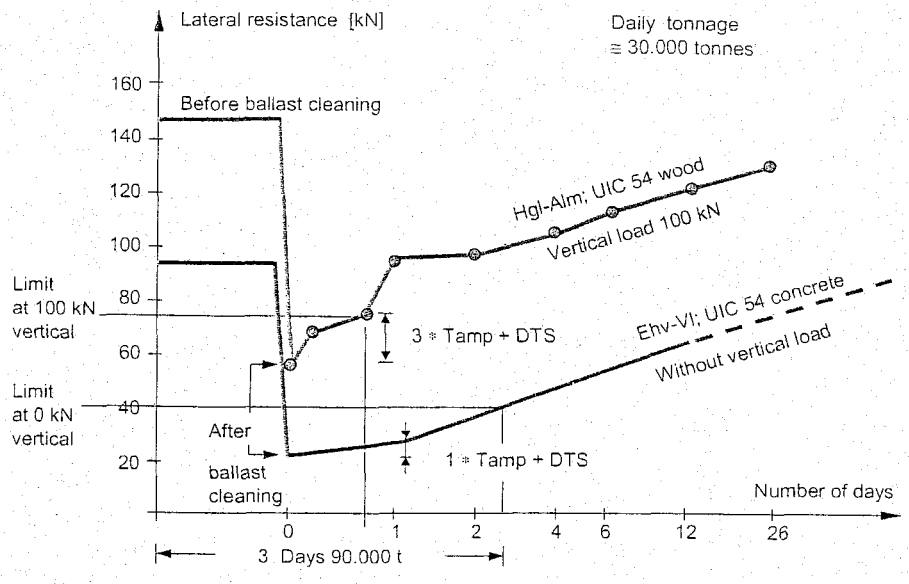


Figure 12.109: Development of lateral resistance as a function of tonnage measured by NS.

With the above-mentioned tamping machine lateral resistance measurements were carried out after ballast cleaning, without a vertical track load on the Eindhoven-Venlo line, and with a vertical track load of 100 kN on the Hengelo-Almelo line. After ballast cleaning the track was tamped and then stabilized with the DTS. The results are presented in Figure 12.109. The effect of stabilizing can be clearly seen. The lower-bound values, indicated in Figure 12.108, are attained within 3 days of carrying train traffic, which corresponds to a cumulative tonnage of 100.000 tonnes.

In the eighties DB had been carrying out similar tests with a second generation DTS of the 62 N type, which is heavier than the one used on NS. These tests were especially aimed at investigating lateral resistance following track renewal and the possibilities of restoring the loss of resistance via application of the DTS. The test results have been published in [148]. Figure 12.110 shows the mean lateral resistance of individual sleepers as a function of tonnage borne for sections with and without application of the DTS. The measurement results show that DTS increases lateral resistance with an amount equivalent to at least 100.000 tonnes of accumulated traffic. On the basis of these findings DB had concluded that after total track renewal the track may be operated immediately at a speed of 250 km/h, provided the DTS has been applied. This will of course require thorough attention at the transition point between old track and renewed track, as in practice problems normally occur at this location.

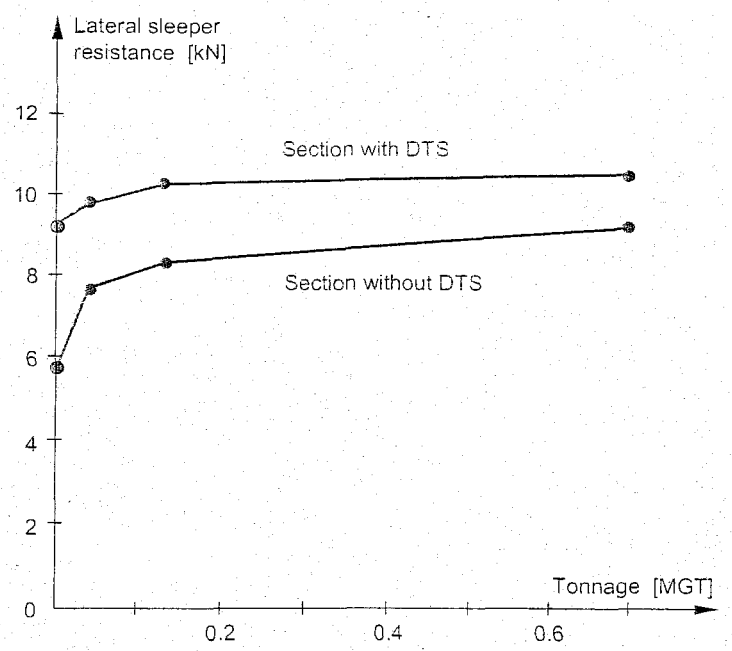


Figure 12.110: Increase in lateral sleeper resistance versus tonnage with and without DTS measured by DB

13 NUMERICAL OPTIMIZATION OF RAILWAY TRACK

13.1 Introduction

Designing is a complex process that includes several stages starting with making sketches and ending with a ready to use product. In each stage of the designing process a number of decisions are to be made, so that designing can be considered a decision making process. In this chapter we will show how mathematical methods such as numerical optimization can help in a decision making process.

The design process commences with defining the requirements of a product. This means that a design has to perform a certain task and has to satisfy certain criteria. Using modern numerical methods, such as e.g. finite element method, the complex behaviour of a design under various loading situations can be simulated. Such simulations can help to estimate the performance of a design and, moreover, reduce the number of expensive prototypes and laboratory tests. Tremendous progress in computer technology has enormously increased the possibilities to numerically simulate complex systems. Modern numerical models used for analysis of static and dynamic behaviour of a railway track have been discussed in the previous chapters.

When the static and dynamic behaviour of the design has been analysed, the next step is to optimise it. Design optimization generally means improving the system's performance during the working cycles while keeping a number of manufacturing, operational, and failure conditions as well as cost limitations in mind. For example, to optimise a railway track one can think of reducing noise produced by a moving train, improving passengers' comfort, or reducing maintenance costs.

In the case of traditional design of technical systems optimization is carried out in a primitive way by modifying design parameters and repeated numerical analyses. The modifications are mostly based on the designer's experience and possibly also on information about the sensitivity of the system's performance to changes in the design parameters. However, it is a time consuming process and, moreover, success cannot be guaranteed.

The most systematic way to improve the design is to use numerical optimization techniques. Combined with advanced numerical simulation analysis, these techniques search for an optimal design based on which possible prototypes can be built. In the forthcoming section a theory and application of structural optimization of railway engineering will be presented. Starting with a brief introduction to numerical optimization, some practical aspects of using an optimization theory will be discussed in Section 13.2.

All optimazation problems have been solved using a modern optimization technique called Multipoint Approximations based on the Response Surface fitting (MARS) method which is briefly discussed in Section 13.3.

Then, three railway engineering applications are presented in Section 13.4 and further. The first one deals with optimization of embedded rail structure. The other two applications are so-called inverse problems in which an optimization technique is used to determine some of the system's parameters. One problem concerns determining ballast lateral resistance parameters based on measurements obtained using a tamping machine. Another problem deals with identifying the dynamic properties of the elastic compound of an embedded rail structure using an hammer excitation test.

13.2 Elements of structural optimization

13.2.1 General optimization problem

Once the behaviour of the mechanical system has been analysed, an optimization problem can be formulated. Firstly, a group of parameters is to be chosen which can be varied to improve the system. Such parameters are called *design variables* and they are denoted by a vector

$$\boldsymbol{x} = (x_1, \ldots, x_N)^T \tag{13.1}$$

The components of the vector $\boldsymbol{x}$ represent various parameters of the structure, such as geometry, material, stiffness, and damping properties, which can be varied to improve the performance of the structure. The process of optimization also implies that there is some *objective function* $F_0(\boldsymbol{x})$ that can be improved which provides a basis to choose between alternative acceptable designs.

The formulation of an optimization problem be restricted by i.e. limitations imposed on various behavioural characteristics of the system. These limitations can be introduced as *inequality constraints* and/ or *equality constraints*. The constraints that impose lower and upper limits on the values of design variables are called *side limits*. Since the equality constraints can always be replaced by two inequality constraints only the problems with inequality constraints are considered. Depending on the problem under consideration, the objective and constraint functions can describe various structural and dynamic response quantities of the system such as weight, reaction forces, stresses, natural frequencies, displacements, velocities, accelerations etc. Cost, maintenance, and safety requirements can be used in the formulation of the optimization problem as well.

Once design variables, objective functions, and constraint functions have been defined an optimization problem can be stated:
Minimise

$$F_0(\boldsymbol{x}) \ , \ \boldsymbol{x} \in R^N \tag{13.2}$$

subject to the constraints

$$F_j(\boldsymbol{x}) \leq 1 \quad (j = 1, \ldots, M) \tag{13.3}$$

and the side limits

$$A_i \leq x_i \leq B_i \quad (i = 1, \ldots, N) \ . \tag{13.4}$$

It is assumed that the optimization problem is a minimisation rather than a maximisation problem. If a formulated problem involves maximising a function, it is always possible to minimise its reciprocal value. Each of the constraints (13.3) is a dimensionless function of the design variables $\boldsymbol{x}$, which can be obtained by normalising the corresponding response quantity, namely

$$F_j(\boldsymbol{x}) = \frac{\underline{F}(\boldsymbol{x})}{\underline{F}_{max}} \leq 1 \tag{13.5}$$

where $\underline{F}(\boldsymbol{x})$ is a function representing a response quantity of the mechanical system and $\underline{F}_{max}$ is the maximum allowable value of the response.
A proper definition of an optimization problem is very important. For example, omitting even one important constraint can result in an unrealistic optimal design.

13.2.2 Solution process

To solve the optimization problem (13.2)-(13.4), various *mathematical programming methods* can be used. Some of them, referred to as *zero-order* methods, only require values of the objective and constraint functions. These methods are easy to use and less sensitive to numerical noise. On the other hand, first order and higher order methods which make use of the first order and higher derivatives of the functions with respect to the design variables, provide better convergence characteristics of the optimization process. They do, however, require additional effort to develop sensitivity analysis and, generally, are more affected by the presence of numerical noise.

Most of the optimization methods are numerical search techniques. Starting from the initial design, they move by making relatively small steps in the direction of improving the value of the objective function and/or reducing the degree of violation of the constraints. The search process is usually terminated when no improvements of the objective function can be made without violating the constraints.

Solving the optimization problem is an iterative process which involves multiple evaluations of the objective and constraint functions (13.2)-(13.3) in the points of the design variable space assigned by the optimization method. Starting from an initial design, an optimiser searches for a better design by assigning variable space to points in the design. The values of the constraint and objective functions are to be evaluated by the response analysis software in this space. Based on the obtained function values, a new design point is assigned and submitted to evaluation by response analyses software. The process is repeated iteratively until the optimization criteria have been satisfied. Since values of the functions can be obtained using one of the numerical methods, the optimization process implies intensive exchange of information between the optimiser and response analysis software.

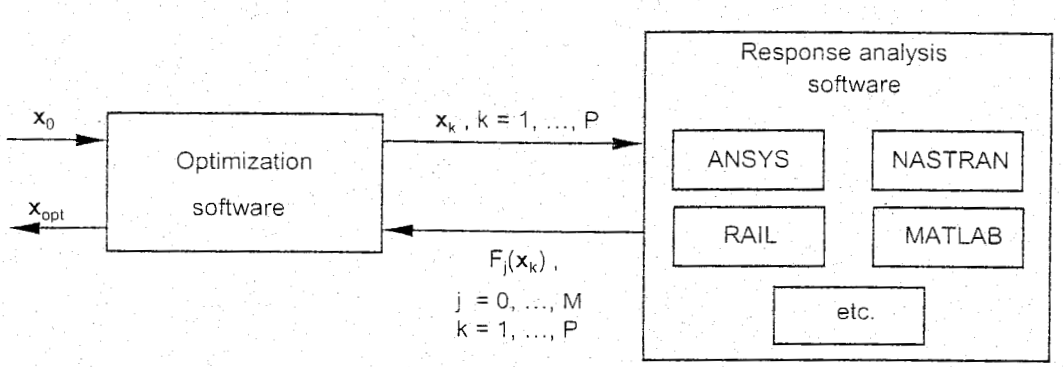


Figure 13.1: Computer implementation of optimization method using external link between optimiser and response analysis software

A possible computer implementation of an optimization method is shown in Figure 13.1. The response analysis software, in this case, is externally linked to an optimiser. The advantage of linking the optimiser in this way is that various response analysis software can be used to obtain the values of the constraint and objective functions. This makes an optimization software problem independent, i.e. it can be used for various optimization problems. The optimization software IMOPT (Chapter 13.3) used here employs an external link approach. Alternatively, response analyses can be internally linked to an optimiser as e.g. in some commercial codes which have their own optimiser.

13.2.3 Approximation concept

The majority of structural optimization problems (13.2)-(13.4) have the following characteristic features:

- The objective and constraint functions are implicit functions of design variables;

- To calculate values of these functions means to use numerical response analysis techniques which usually involve a large amount of computer time;

- Function values and (or) their derivatives often contain numerical noise, i.e. they can only be estimated with a finite accuracy.

During the optimization process the values of the objective and constraint functions have to be calculated many times using one of the numerical methods. If the computational cost of the evaluation of the response quantities is low or moderate, one of the conventional methods of mathematical programming can be employed to solve the optimization problem (13.2)-(13.4). However, in the case of a complex system such calculations may be time consuming and even prohibitive. This problem has been mitigated in the mid-seventies by introducing approximation concepts [245]. According to the approximation concepts, the original functions are replaced by approximate ones which are computationally less expensive. Thus, in each iteration step the original optimization problem (13.2)-(13.4) is replaced by a succession of more simple approximated subproblems:

Minimise

$$\tilde{F}_0^k(\boldsymbol{x}) , \qquad \boldsymbol{x} \in R^N \tag{13.6}$$

subject to

$$\tilde{F}_j^k(\boldsymbol{x}) \le 1 \qquad (j = 1, \ldots, M) \tag{13.7}$$

and

$$A_i^k \le x_i \le B_i^k , \quad A_i^k \ge A_i , \quad B_i^k \le B_i \qquad (i = 1, \ldots, N) \tag{13.8}$$

where the superscript k is the number of the iteration step, $\tilde{F}_j^k(\boldsymbol{x})$ is the approximation of the original function $F_j(\boldsymbol{x})$, and A_i^k and B_i^k are *move limits* defining the range of applicability of the approximations (local, global or mid-range). Since the functions (13.6)-(13.7) are chosen to be simple, computationally inexpensive and noiseless, any conventional method of mathematical programming can be used to solve the problem (13.6)-(13.8). The obtained point $\boldsymbol{x}^k$ is then chosen as starting point for the $(k+1)$-th step and the optimization problem (13.6)-(13.8), reformulated with new functions $\tilde{F}_j^{k+1}(\boldsymbol{x})$ $(j = 0, \ldots, M)$ and move limits A_i^{k+1} and B_i^{k+1} $(i = 1, \ldots, N)$, is to be solved. The process is repeated until the convergence criteria are satisfied.

Depending on their range of applicability, the approximations can be classified as *local* (valid in the vicinity of a current point in the design variable space), *global* (in the whole region defined by side limits), and *mid-range* (somewhere in between). Local function approximations are based on the exact match of function values and their derivatives at one point in the design variable space using the first-order or second-order Taylor series.

Most optimization methods based on an approximation concept have a typical structure. Thus, one can recognize four main parts in any optimization method based on approximation concepts:

- Developing approximation functions;

- Solving the approximated optimization problem (13.6)-(13.8);

- Strategy for changing the move limits;

- Convergence criteria.

Each of these parts is very important for the success of the optimization process. They are briefly discussed in the following section in which the optimization method used in numerical examples is described.

13.3 MARS method

A technique that uses Multipoint Approximations based on Response Surface fitting (MARS) has been chosen [173], [274] as the general optimization method suitable for railway engineering applications. The method is based on the approximation concepts (13.6)-(13.8). The approximations used in this method belong to the mid-range class (see previous section). They are based on information about the original functions obtained at several points in the design variable space. Thus, the construction of the approximations also includes the problem of choosing the points in the space of the design variables. These and other aspects are briefly discussed below.

To build the approximation functions the methods of regression analysis [25] are used. According to each approximation $\tilde{F}$ in (13.6)-(13.7) is defined as a function of the design variables $\boldsymbol{x}$ and tuning parameters $\boldsymbol{a}$ (for brevity the indices k and j will be omitted). To determine the components of vector $\boldsymbol{a}$ the following weighted least-squares minimisation problem is to be solved [173], [274]:

Find vector $\boldsymbol{a}$ that minimises

$$G(\boldsymbol{a}) \equiv \sum_{p=1}^{P} \{ w_p [F(\boldsymbol{x}_p) - \tilde{F}(\boldsymbol{x}_p, \boldsymbol{a})]^2 \} \rightarrow min . \tag{13.9}$$

Here $F(\boldsymbol{x}_p)$ is the value of the original function from (13.6)-(13.7) evaluated at the point of the design parameters space $\boldsymbol{x}_p$, and P is the total number of such plan points; w_p is a weight factor that characterises the relative contribution of the information about the original function at the point $\boldsymbol{x}_p$. It should be noted that if the design sensitivities of the original functions are available, they can easily be used in the building of approximations. The function G in (13.9) then reads

$$G(\boldsymbol{a}) \equiv \sum_{p=1}^{P} w_p \left([F(\boldsymbol{x}_p) - \tilde{F}(\boldsymbol{x}_p, \boldsymbol{a})]^2 + \gamma \delta_p \sum_{i=1}^{N} [F(\boldsymbol{x}_p)_{,i} - \tilde{F}(\boldsymbol{x}_p, \boldsymbol{a})_{,i}]^2 \right), \delta_p = 1 / \left(\sum_{i=1}^{N} (F(\boldsymbol{x}_p)_{,i})^2 \right) \tag{13.10}$$

where δ_p is a normalising coefficient and $0 < \gamma < 1$ reflects the weight of the sensitivity information compared to the information on the function values [173].

Selecting the points belonging to the plan of numerical experiments in (13.9)-(13.10) is a separate problem. The minimum number of points needed to build approximations (13.6)-(13.7) is equal to the number of design variables plus one. The most simple way to choose is by taking a step from a base point in each coordinate direction, possibly including the points of the previous iterations belonging to the current search subregion and its neighbourhood [173] as shown in Figure 13.2a. However, if a problem of domain-dependent calculability arises, which means that at some points of the design variable space the functions cannot be evaluated, a random plan of experiments (Figure 13.2b) is a good remedy [173], [178].

Choosing an approximation function $\tilde{F}$ is important for the success of optimization. The basic requirements for such a function are:

- It must depend on the same design variables as the original function;
- It has to contain a number of tuning parameters to be defined using the general (non-linear) least-squares method;
- It must be simple enough to be used in numerous repeated calculations;

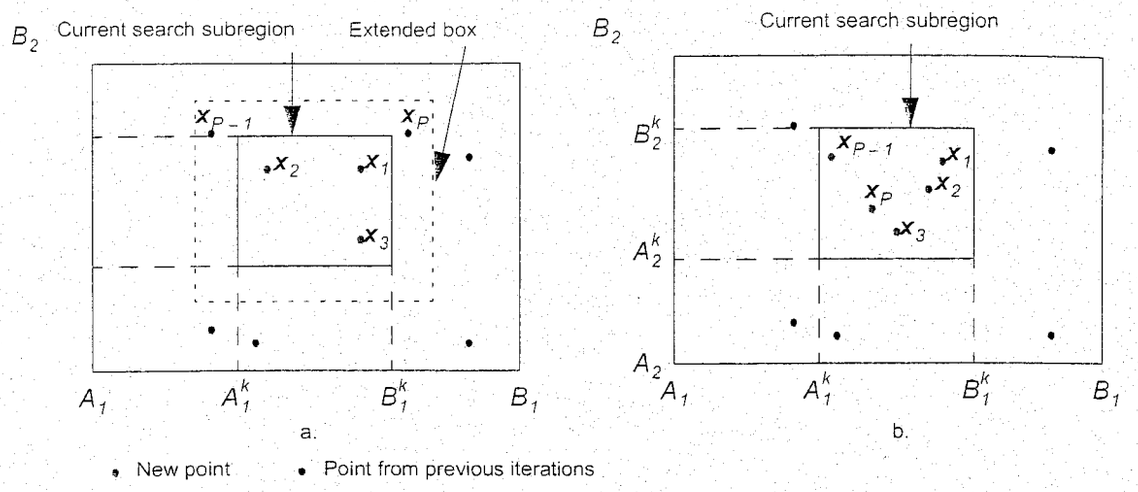


Figure 13.2: Extended (a) and random (b) plan of experiments in MARS method

– It should not contain a considerable level of numerical noise in order to not cause convergence problems in the optimization process.

Simple yet quite efficient approximations are intrinsically linear (with respect to the tuning parameters) models, namely linear and multiplicative models:

$$\tilde{F}(\boldsymbol{x}, \boldsymbol{a}) = a_0 + \sum_{i=1}^{P} a_i x_i \quad \text{and} \quad \tilde{F}(\boldsymbol{x}, \boldsymbol{a}) = a_0 \prod_{i=1}^{P} (x_i)^{a_i}. \tag{13.11}$$

These models have been successfully applied to various design optimization problems [173], [175], [178]. For details on approximations using the MARS method we refer to [173].

The obtained approximation functions are used in the formulation of the optimization problem (13.6)-(13.8). In order to solve this problem any conventional method of non-linear mathematical problem can be used. In MARS a Sequential Quadratic Programming (SQP) method has been chosen [229]. The solution of the problem is considered a starting point for the next iteration. The move limits are changed depending on the quality of approximation and location of the optimal solution in the previous step. The main rules of the strategy to change the move limits employed in MARS are:

– If the approximating functions do not adequately represent the original ones in the current optimum point, which means that the search subregion is larger than the range of applicability of the current approximations, the move limits (13.8) are changed to reduce the size of the search subregion;

– If the approximations are good and the solution to the optimization problem (13.6)-(13.7) is an internal point of the search subregion, which means it could be considered as an approximation of the solution of the original optimization problem (13.2)-(13.4), the search subregion is reduced;

– If the current optimum point belongs to the boundary of the search subregion (at least one of the move limits is active) and the approximations are good, the size of the subregion is not changed for the next iteration.

An example of changing move limits is when a solution of the approximation problem on k-th step $\boldsymbol{x}^k$ is located on the boundary and inside of the search subregion (adequate approximations are assumed), see Figure 13.3.

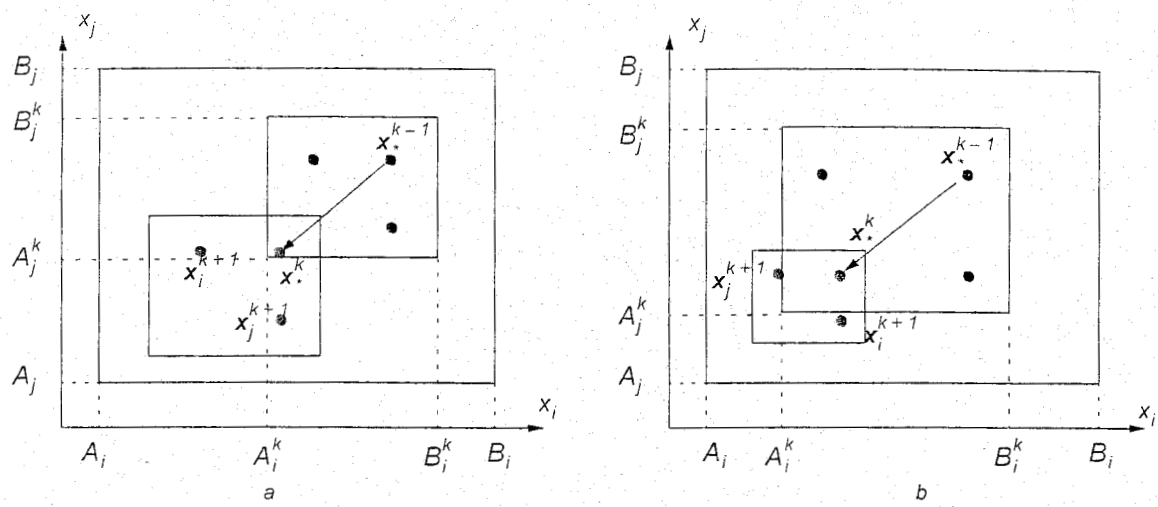


Figure 13.3: Changing of move limits on boundary (a) and inside (b). Plan of experiments without using derivatives

The iteration process is terminated and the solution of the problem (13.6)-(13.8) is considered a good approximation of the solution of the original problem (13.2)-(13.4) if:

- Approximations are good;
- None of the move limits is active (the solution is located inside of the search subregion);
- The search subregion is small enough.

More information on the MARS method and its recent developments can be found in [173], [275].

13.4 Optimal design of embedded rail structure (ERS)

13.4.1 Introduction

In this Section a procedure used for optimal design of a railway track is presented which has been developed at Delft University of Technology. The procedure contains all necessary parts of the design process, such as:

- Numerical modelling of track behaviour;
- Verification and calibration of a numerical model using results of laboratory tests;
- Optimization.

Here, the procedure is applied to the optimal design of an Embedded Rail Structure. The criteria regarding optimal design are based on the requirements to railway tracks relating to the design's cost efficiency, minimum noise emission, and minimum wear of rails and wheels. The design variables comprise the material and geometry properties of ERS, namely elastic properties and volume of compound, shape of rails, and size of troughs.The static and dynamic behaviour of ERS is analysed using the finite element models built using ANSYS and RAIL which have been described in Chapter 5 and Chapter 6. In the subsequent sections the following topics will be discussed:

- Requirements for optimum design of ERS and its numerical interpretation;
- Choice of optimal design criteria;
- Formulation of an optimization problem;

13.4.3 Optimization problem

As mentioned above, in order to use a numerical optimization method a problem should be stated with the following form (13.2)-(13.4). To optimise ERS, eight geometry parameters and two material parameters of compound have been chosen as the design variables $\boldsymbol{x}$. They are shown in Figure 13.9. Note that the variables x_5 and x_6 represent the filling level of compound relative to the total height of the rail. The lower and upper bounds on the design variables are collected in Table 13.1.

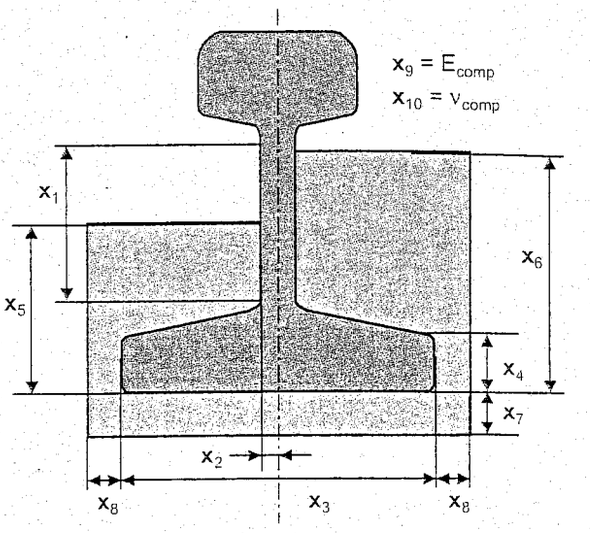


Figure 13.9: Design variables of the optimization problem

In order to achieve an optimum design of ERS, the cost, acoustic noise (13.12), and maintenance (wheel-rail wear) effort should be minimal. Therefore, the corresponding response quantities of ERS, such as the amount of compound, area of open surface, resonant frequency, and contact forces respectively, are i.e.

$$F_C \equiv V \rightarrow min\,, \qquad F_N \rightarrow min\,, \qquad F_M \equiv W \rightarrow min \tag{13.17}$$

The requirements that prevent damaging of ERS, train derailment (13.13), and buckling of rails (13.16) have been used as the constraints:

$$F_1(\boldsymbol{x}) \equiv \frac{u_{x,2}}{u_{x,\,allow}} \leq 1\,,\; F_2(\boldsymbol{x}) \equiv \frac{\sigma}{\sigma_{allow}} \leq 1\,,\; F_3(\boldsymbol{x}) \equiv \frac{P^{*}_{bckl}}{P_{bckl}} \leq 1 \tag{13.18}$$

Thus, the optimization problem (13.17)-(13.18) has three objectives. Such problems are characterised by an objective conflict, i.e. when the values of one objective are improved at least one other objective deteriorates. Regarding this problem an improvement of the acoustic properties can be achieved by stiffening the compound, thereby increasing the resonant frequency of ERS. On the other hand, the stiffer the compound the higher the wheel-rail contact forces. And this would entail that the maintenance characteristics of ERS become worse.

Design variable	Lower bound	Upper bound	Initial value	Units
x_1	8.9	85.0	80.0	mm
x_2	8.0	36.0	8.0	mm
x_3	36.0	70.0	70.0	mm
x_4	15.0	30.0	25.0	mm
x_5	0.100	0.750	0.471	-
x_6	0.100	0.900	0.669	-
x_7	5.0	15.0	10.0	mm
x_8	15.0	60.0	15.0	mm
x_9	1.0	10.0	5.6	MPa
x_{10}	0.40	0.46	0.46	-

Table 13.1: Side limits and initial values of design parameters

A typical approach to solving a multi-objective problem is to convert it to a single objective one [10]. One way to do this is by considering the most important function as the only objective and imposing limits on the other objective functions (treat them as the constraints) [67]. The other way is to make a composition of the objective functions while assigning weights (preferences) that reflect the relative importance of objectives for each function. A solution to a composition problem is a compromise solution since it depends on weight factors. An entire set of compromised solutions is called the Pareto set. In practice, a representative subset of the Pareto set is to be determined so that the person in charge (a decision maker) can choose the most appropriate solution [255]

The latter approach has been used here to solve the optimization problem (13.17)-(13.18). Thus, the following objective function has been constructed:

$$F_0 \equiv \left\{ w_C \frac{F_C}{F_{C.ut}} + w_N \frac{F_N}{F_{N.ut}} + w_M \frac{F_M}{F_{M.ut}} \right\} \rightarrow min \qquad (13.19)$$

in which the weight coefficients w_C, w_N, w_M are the preference factors ($w_C + w_N + w_M = 1$) that reflect the relative importance of compound, noise, and maintenance reduction in the final design of ERS. $F_{C.ut}$, $F_{N.ut}$, $F_{M.ut}$ are the normalising coefficients since the objective functions have different units.

Preference factors can reflect different points of view on optimum design of ERS (point of view of society, investor, maintenance contractor, etc.). Based on results of optimization with different sets of preference factors, a decision about a final design can be made. Here, each normalising coefficient is chosen as an optimal value of an objective function of a corresponding single criterion optimization problem (13.17)-(13.18), i.e. a problem with one criterion from (13.17) and a full set of constraints (13.18). These solutions are also called *utopian* solutions since they cannot be obtained during a combined multi criteria optimization.

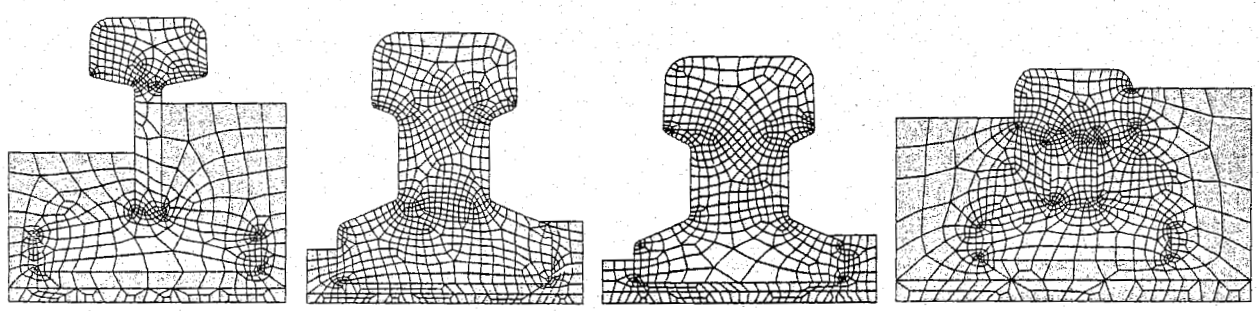


Figure 13.10: Initial design of ERS with conventional rail (a) and results of single criterion optimization (b - compound, c - noise, d - maintenance)

The initial design and results of the single criterion optimization are given in Figure 13.10. The numerical results are collected in Table 13.2. From these results it can be seen that the optimum designs in the single criterion optimization problem with respect to the cost and noise (Figure 13.10a and Figure 13.10b) are close to each other (Figure 13.10c). This can be explained by the fact that the compound and noise objectives do not conflict that much. Both designs have acoustic properties that are quite good (improved response frequency 350 Hz and 380 Hz respectively and the open surface A is almost reduced twice) while they are poorly with respect to maintenance properties (high contact forces 14.1 kN and 15.1 kN).

During the third single criterion optimization (Table 13.2) the properties related to maintenance of ERS have been improved dramatically (contact forces of 17.5 kN for the initial design and 10.0 kN for the epitomised one). However, the acoustic properties are much worse (noise objective 1.09) than in the case of the optimal designs of the two first single optimizations (noise objective 0.54 and 0.46 respectively). Moreover, the amount of compound in the optimum design of the maintenance optimization has become even larger (19.973 dm^3/m) than the one in the initial design (13.536 dm^3/m).
The small amount of compound above the rail (Figure 13.10b and Figure 13.10c) is sufficient to prevent upward buckling since the stiffness of the rail, which is important for determining the buckling force according to the chosen analytical criteria (13.15)-(13.16), has increased. A more accurate calculation of the buckling force using the non-linear numerical analysis could be useful in future optimization.

	Initial design	Single criterion optimization		
		Compound	Noise	Maintenance
Design variables				
x_1[mm]	80.0	40.9	35.4	39.4
x_2[mm]	8.0	22.4	23.2	13.7
x_3[mm]	70.0	52.5	50.1	58.6
x_4[mm]	25.0	29.9	19.5	23.6
x_5	0.471	0.136	0.100	0.748
x_6	0.669	0.257	0.200	0.900
x_7[mm]	10.0	5.0	5.0	12.8
x_8[mm]	15.0	15.0	15.0	48.1
x_9[MPa]	5.60	9.44	9.99	4.13
x_{10}	0.46	0.43	0.49	0.41
Constraints				
σ [MPa]	1.07	1.15	2.24	0.91
$u_{x,2}$ [mm]	1.83	1.08	1.38	1.61
P_{bckl} [MN]	12.5	1.7	1.1	5.4
Objectives				
F_C [dm^3/m]	13.536	2.103	1.806	19.973
A [dm^2/m]	134.8	74.3	67.5	120.7
f_r [Hz]	250	350	380	190
F_N	1.12	0.54	0.46	1.09
F_M [kN]	17.5	14.1	15.1	10.0

Table 13.2: Numerical results of single criterion optimization

To investigate the sensitivity of an optimal solution to the choice of preference factors, several multi criteria optimization problems with different sets of the preference coefficients have been solved. In the first optimization, it is assumed that all the objectives are equally important, i.e. $w_C = w_N = w_M$. The optimum design is shown in Figure 13.11b. The numerical results are collected in Table 13.3.

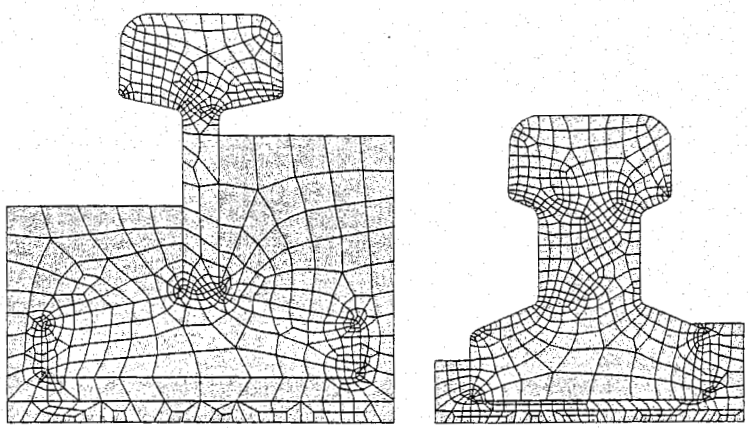


Figure 13.11: Initial design of ERS with conventional rail (a) and result of multi criteria optimization with equal preference (b)

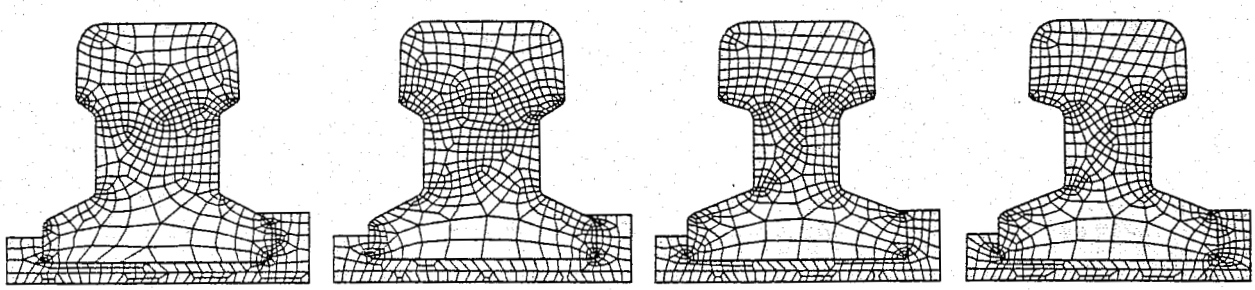


Figure 13.12: Geometrical results of multi criteria optimization, variation of preferences for noise (w_N) and maintenance (w_M)

Figure 13.12 shows the results of optimization problems with varied preference factors related to the reduction of the maintenance effort and the noise (see Table 13.3) while considering the cost objective less important (the corresponding preference coefficient is constant $w_C = 0.1$). The numerical results collected in Table 13.3 clearly reflect the effect of the chosen preferences for the optimum design. From this Table it can be seen that the stiffness of compound and resonant frequency of ERS increase (from 5.39 MPa to 9.25 and from 290 Hz to 380 Hz respectively) as the noise objective becomes more important (w_N increases from 0.1 to 0.8). On the other hand, the contact forces increase as well (from 11.2 kN to 13.8 kN) as the maintenance becomes less important (w_M decreases from 0.8 to 0.1). From Figure 13.12 one can see the tendency of the rail shape of the optimum design to change as the maintenance objective becomes less important (from a to d). It should also be noted that the only difference between the two last designs (c and d) is the thickness of the compound layer under the rail (see Table 13.3).

In another sequence of optimization problems, the preference factor for the noise objective was constant $w_N = 0.1$ while the preferences for the other two objectives were varied. The ERS shape results from optimization are given in Figure 13.13 and the numerical results are collected in Table 13.3. Again the stiffness of compound (and therefore the vertical stiffness of the structure) increases as the reduction of maintenance becomes less important (w_M reduces).

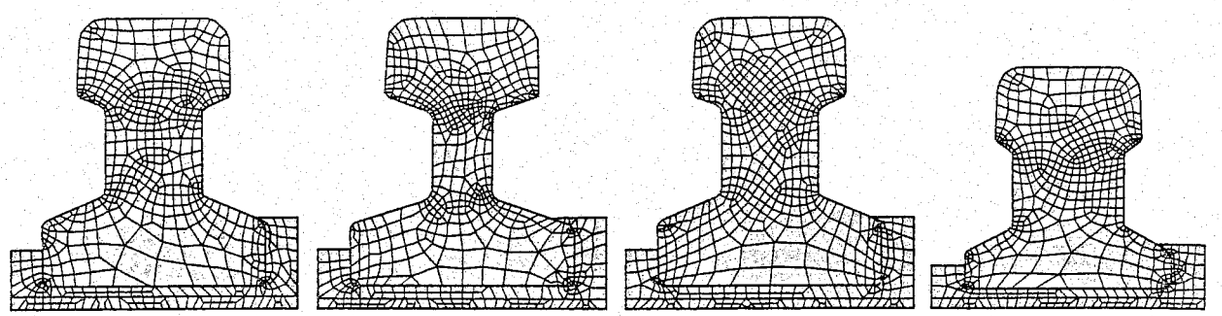


Figure 13.13: Geometrical results of multi criteria optimization, variation of preferences for compound (w_C), and maintenance (w_M)

In Figure 13.12 and Figure 13.13 one can see the remarkable difference in the rail shape of the optimum design when the compound objective is more important than the noise objective (compare Figure 13.12 b,c,d and Figure 13.13 a,b,c). Although the compound and noise objectives do not conflict very much, the reduction of noise has more influence on the changing compound stiffness than the reduction of the compound volume as can be observed from (compare the results for $w_C = 0.6$, $w_N = 0.1$, $w_M = 0.3$ and for $w_C = 0.1$, $w_N = 0.6$, $w_M = 0.3$). Since the vertical stiffness of ERS is important in order to reduce noise, the E modulus of compound increases relatively fast during the optimization. Therefore, the stiffness of compound and the bending stiffness of the rail of the optimal ERS design are higher when noise reduction is considered to be more important than reducing the amount of compound, i.e. $w_N > w_C$ (compare Figure 13.12c with Figure 13.13b, and also the corresponding rows in Table 13.3).

The contribution of the objective functions to the optimal design of ERS is investigated in Figure 13.15 in which the sensitivity of the objectives to variation in the preference factors is shown. The values of the objective functions are normalised using the maximum value of the corresponding objective function obtained in both optimization sets. From this figure the conflicting nature of the optimization problem can clearly be seen. Other compromised solutions can be found by assigning different preference factors for the objectives and performing a similar optimization.

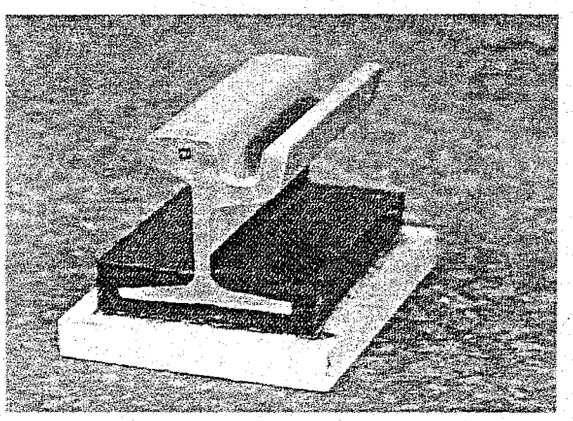


Figure 13.14: ERS design used for tram tracks (ALH Rail Coatings Ltd.)

Finally, one practical design of ERS used for tram tracks is given in Figure 13.14. Comparing the compound shape of the design to the designs shown in Figure 13.12 and Figure 13.13, it can be concluded that the results obtained from the numerical optimization are very close to a realistic design of ERS, and that numerical optimization can help to better understand and improve the performance of Embedded Rail Structure.

	Initial design	Multi criteria optimization							
		Wc=0.3 Wn=0.3 Wm=0.3	Wc=0.1 Wn=0.1 Wm=0.8	Wc=0.1 Wn=0.4 Wm=0.5	Wc=0.1 Wn=0.6 Wm=0.3	Wc=0.1 Wn=0.8 Wm=0.1	Wc=0.3 Wn=0.1 Wm=0.6	Wc=0.6 Wn=0.1 Wm=0.3	Wc=0.8 Wn=0.1 Wm=0.1
Design variables									
x_1[mm]	80.0	41.2	36.4	33.8	35.7	35.7	41.9	40.9	40.9
x_2[mm]	8.0	22.3	27.2	25.0	19.6	19.6	21.8	14.3	22.4
x_3[mm]	70.0	52.7	51.0	49.4	50.2	50.2	53.0	52.5	52.5
x_4[mm]	25.0	29.7	17.7	15.9	20.1	20.1	30.0	29.9	29.9
x_5	0.471	0.140	0.102	0.100	0.100	0.100	0.146	0.137	0.137
x_6	0.669	0.261	0.201	0.183	0.203	0.203	0.269	0.258	0.257
x_7[mm]	10.0	5.0	5.0	5.3	5.3	5.0	5.0	5.0	5.0
x_8[mm]	15.0	15.0	16.4	15.0	15.0	15.0	15.0	15.0	15.0
x_9[MPa]	5.60	6.86	5.39	7.16	9.25	9.25	4.82	7.94	9.4
x_{10}	0.46	0.46	0.44	0.44	0.43	0.43	0.50	0.43	0.43
Constraints									
σ [MPa]	1.07	1.14	1.09	1.14	1.18	1.14	1.14	1.19	1.19
$u_{x,2}$ [mm]	1.83	1.44	1.72	1.30	1.06	1.03	1.99	1.29	1.08
P_{bckl} [MN]	12.5	1.7	1.2	1.2	1.4	1.4	1.5	1.7	1.7
Objectives									
F_C [dm^3/m]	13.536	2.126	1.911	1.816	1.886	1.820	2.172	2.110	2.103
A [dm^2/m]	134.8	74.9	67.5	65.3	69.3	69.2	76.1	77.6	74.3
f_r [Hz]	250	310	290	320	350	380	270	340	350
F_N	1.12	0.57	0.53	0.49	0.50	0.48	0.62	0.57	0.54
F_M [kN]	17.5	12.3	11.2	12.4	13.5	13.8	11.6	12.7	14.0

Table 13.3: Numerical results of multi criteria optimization

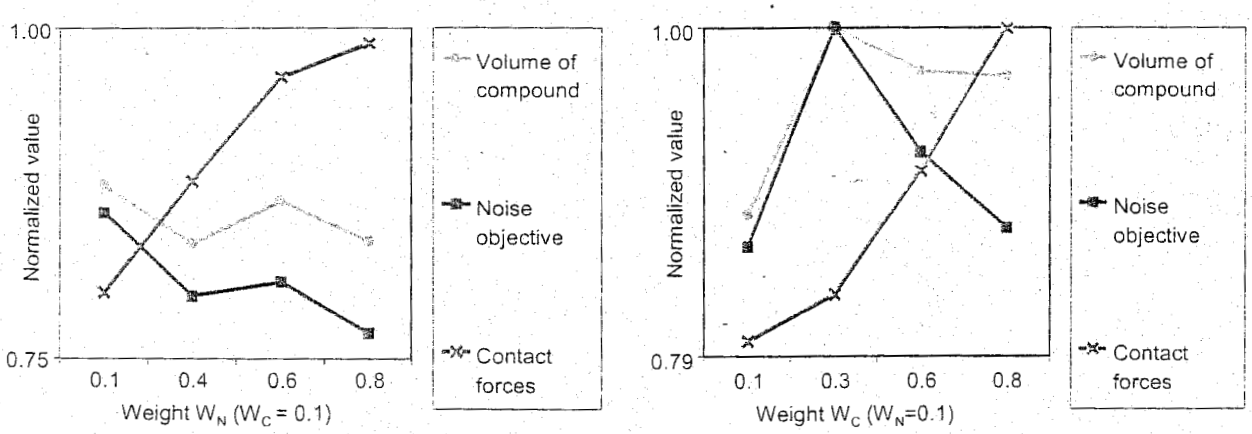


Figure 13.15: Sensitivity of objectives to variation of preference factors

- Formulation of a multi criteria optimization problem;
- Obtaining results.

A typical ERS consists of a continuous reinforced concrete slab that rests on a concrete stabilised road bed, which in turn is placed on a sand base as shown in Figure 13.4. Two (conventional) rails are embedded in visco-elastic compound (polyurethane mixture) for which two troughs are made in the slab.

Obviously, the design of ERS shown in Figure 13.4 is not optimal. For example, the use of conventional rails (such as UIC54) which are relatively high requires a high stiffness compound (to ensure sufficient lateral resistance of ERS in order to prevent train derailment). An elastic strip under each rail is therefore a necessary part of a structure, although it complicates an ERS design. Moreover, the high filling level of compound (Figure 13.4) increases the costs of a structure since compound is quite expensive material. On the other hand, the amount of compound in ERS can be reduced by using smaller rails. Moreover, the elastic strips under the rails can be eliminated by adjusting the stiffness properties of the compound.
Recently, an ERS with a new rail profile SA42 (Figure 13.5) was developed within the framework of the STV-project (Stiller TreinVerkeer: Silent Train Traffic). It has been experimentally demonstrated that such rail can reduce the level of acoustic noise by 5 dB(A). This result also shows that the efficiency of using conventional rails in ERS is rather arguable, and that the performance of ERS can be improved.

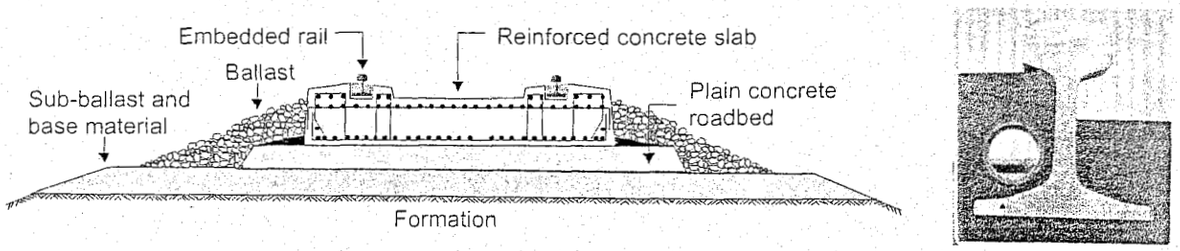


Figure 13.4: Embedded Rail Structure (ERS)

In this Section, an ERS design is improved using a numerical optimization technique based on the MARS method described in Section 13.3. First, requirements for an optimum design have been formulated and transformed to equivalent response quantities of numerical models of ERS. Using these responses, a multi criteria optimization problem has been formulated. To solve the problem, it has been transformed to a single objective problem using a preference function approach. The optimization procedure and results are discussed below.

13.4.2 Requirements for optimum design of ERS

An optimization process commences with defining an optimum design. In this stage a decision has to be made about the requirements for the best performing railway track. Such requirements are then transformed to equivalent response quantities of a track structure. The response quantities - evaluated numerically - are used in the formulation of an optimization problem (13.2)-(13.4) to solve which MARS method (Section 13.3) is used.

Numerical models

Here the response quantities relating to cost efficiency, acoustic properties, and maintenance effort are considered of importance to the optimum performance of ERS. To estimate the performance of an ERS design, static and dynamic models have been developed. The static response quantities such as stresses and displacements of an embedded rail structure under various loading conditions have been obtained using a general purpose finite element package ANSYS.
The 2-D and 3-D FE models of ERS are shown in Figure 13.5. Before these models were included in the optimization process, they were verified by comparing the results of laboratory tests and finite element calculations [175].

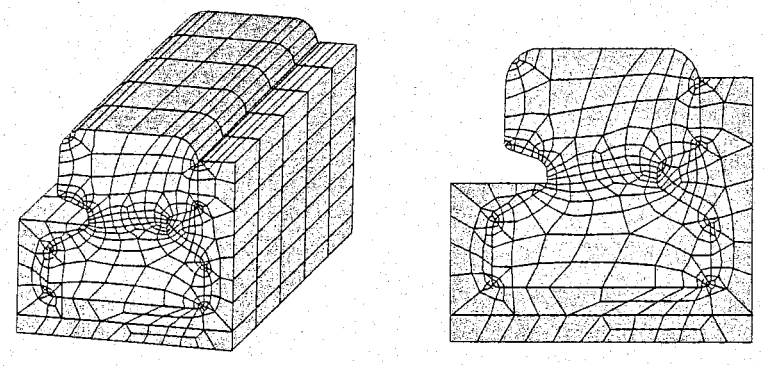


Figure 13.5: 3-D (a) and 2-D (b) finite element models of ERS with SA42 rail

It should be noted that the elastic strip under the rail, which is a common part of existing designs, is now replaced by elastic compound. In [175] it is demonstrated that the same behaviour of a structure can be achieved by only using a compound with adjusted E modulus and Poisson ratio.

Three loading cases have been considered to obtain the static response quantities of a structure for assessment of ERS design (Figure 13.8). The dynamic responses of ERS have been obtained using a finite element program RAIL that is described in Section 6.9. The numerical mode of ERS, built using RAIL, is shown in Figure 13.6. Here, the application of RAIL focuses on two aspects, namely acoustic noise produced by a track and wheel-rail wear.

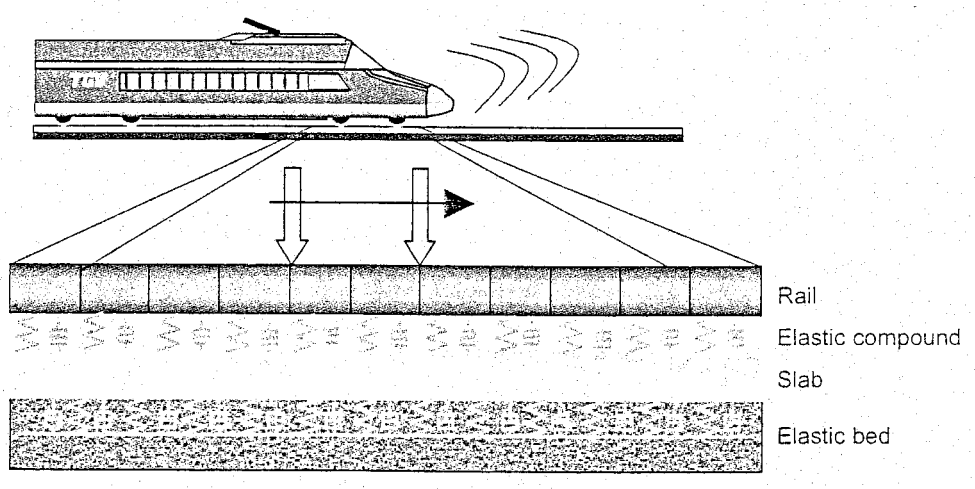


Figure 13.6: RAIL model of ERS (moving train loading case)

In order to ensure that the static and dynamic models describe the behaviour of the same ERS, they have been coupled to each other by adjusting geometrical properties such as cross-sectional moment of inertia, etc. of the rails in the dynamic analysis based on the parameters of the static model. Also, the static and dynamic vertical stiffness of ERS has been correlated. To determine the static stiffness of a track the vertical load F_y has been applied at the top of the rail head as shown in Figure 13.8a. The static (K_{stat}) and dynamic (K_{dyn}) vertical stiffness are then calculated as $K_{stat} = F_y / u_{y,1}$ and $K_{dyn} = 2K_{stat}$ ($u_{y,1}$ is the vertical displacement of the rail corresponding to this loading case).

Minimum cost requirement

The cost efficiency of ERS can be estimated by the amount of elastic compound used in it (Figure 13.7b). To reduce the costs, the *volume of elastic compound V* [dm^3/m] used in ERS should be *minimal*.

Minimum acoustic noise requirement

The '*open surface*' of ERS, i.e. the surface of the rail and the elastic compound that is in contact with open air (Figure 13.7c), defines the level of the *acoustic noise*. The noise radiating from the ERS structure increases in proportion to the size of the open surface. Thus, in order to reduce the noise produced by a structure, the open surface of ERS A [dm^2/m] should be minimised.

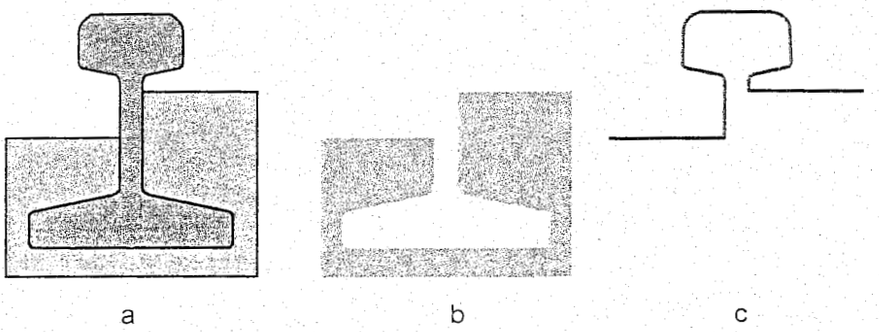


Figure 13.7: Model of ERS (a) to estimate volume of compound (b) and area of open surface (c) used in optimization

An optimum railway track should possess good dynamic properties. Large vibrations on a track can diminish passengers' comfort and increase maintenance effort of vehicles, rails, and other structures. To avoid the harmful track vibrations, resonant frequencies of ERS should not coincide with any of the vehicle resonant frequencies. One way to achieve that is to shift the *resonant frequencies* of a track as far as possible from those of a vehicle. In other words, the first principal resonant frequency of ERS in vertical direction should be maximised. It should be noted that the acoustic characteristics of ERS (i.e. the ability of a track to damp vibrations before it starts to radiate them via its open surface) improve as the first resonant frequency is increased [175]. The resonant frequency f_r is obtained by applying an impulse load to ERS and performing the frequency response analysis (Figure 13.6).

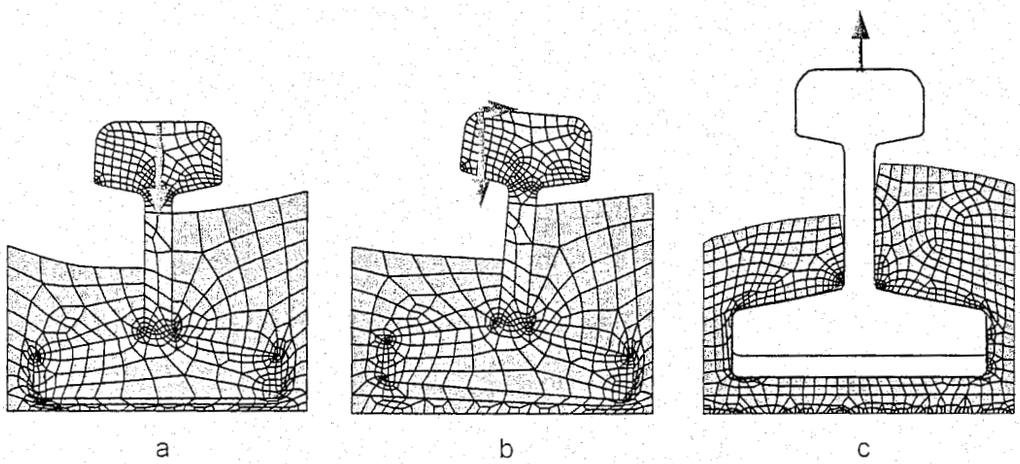


Figure 13.8: Loading cases for static analysis (ANSYS) of ERS

Since the level of acoustic noise radiating from ERS depends on both the resonant frequency f_r and the area of the open surface, the following formula has been used to estimate the acoustic noise F_N:

$$F_N = A\frac{f_{max} - f_r}{f_{max} - f_{min}}, \tag{13.12}$$

in which f_{min} = 100 Hz and f_{max} = 1000 Hz are the lower and upper boundary of the resonant frequency.

Maintenance requirements

Wear of rails defines to a large extent the maintenance costs of a track structure. Likewise, wear of wheels defines the maintenance costs of a vehicle. Here the wheel-rail wear is estimated by analysing wheel-rail contact forces. The standard deviation of the contact forces should be minimal in order to reduce the wheel-rail wear [169]. The standard deviation of the wheel-rail contact forces was obtained from the 0.8 sec simulation of the TGV train moving at 90 m/s (324 km/h).

Safety requirements

Two safety requirements concern the lateral displacements of the rails and strength of the compound. Under a specific angular loading condition, the lateral displacements should not exceed a prescribed limit in order to avoid the gauge from widening and, ultimately, train derailment. Therefore, a special loading case has been considered (Figure 13.8b) in which an angular concentrated load has been applied to the rail (at 22 degrees relative to the vertical). For an optimum design, the lateral displacements $u_{x.2}$ and maximum (Von Misses) stress σ_{max} should be below their maximum allowable values, i.e.

$$u_{x.2} \le u_{x.allow} \text{ and } \sigma_{max} \le \sigma_{allow} \tag{13.13}$$

Here, the maximum allowable lateral displacement $u_{x.allow} = 0.002m$ has been chosen. Based on experimental data, the following approximation of the maximum tensile stress has been used at which cracking of compound occurs (the approximation is valid for compound with E module between 1MPa and 10 MPa):

$$\sigma_{lens.max} = 0.09E + 0.8[MPa] \tag{13.14}$$

The maximum allowable stress in (13.13) is then taken as 90% of the maximum tensile stress (13.14).

Another requirement concerns upward buckling of a rail. This can happen if there is no (or not sufficient) adhesion between the rail and compound. The vertical (upward) stiffness is then reduced significantly and only depends on the shape of the rail. For a safe design of ERS, the minimum buckling load should not be lower than the prescribed value P^*_{bckl}:

$$P_{bckl} \ge P^*_{bckl} \tag{13.15}$$

Here the minimum allowable buckling force $P^*_{bckl} = 1MN$ has been used that corresponds to a situation in which very high normal forces in rails are caused by temperature variation and train braking. The buckling force P_{bckl} can be estimated analytically by modelling ERS as a beam on an elastic foundation and using the energy approach [270] that reads

$$P_{bckl} = 2\sqrt{EIK_{up}} \tag{13.16}$$

where EI is the flexure of the rail and K_{up} is the upward stiffness of ERS per unit length. To estimate the upward stiffness of ERS, the model has been adjusted so that the rail and compound have no common nodes and there is no friction between the rail and compound (Figure 13.8c). The vertical displacements of the rail are obtained by applying the vertical upward load $F_{y.3}$ and performing the non-linear contact analysis. The vertical displacement of the rail $u_{y.3}$ is then used for the stiffness evaluation as $K_{up} = F_{y.3}/u_{y.3}$. For details on the model data used for the static and dynamic analyses, we refer to [175].

13.4.4 Remarks and conclusions

The MARS method has been applied to the optimal design of an Embedded Rail Structure. The optimization criteria are formulated using the requirements relating to cost efficiency and safety of railway tracks, minimum noise emission, and minimum wear of wheels and rails. Both the static and dynamic behaviour of ERS has been analysed using finite element numerical models.

The multi-objective optimization problem was solved by transforming it into a single objective problem which had been composed from the original single criterion problems using the preference factor approach. The preference factors representing the relative importance of each objective function during the optimization process can reflect different points of view on optimum design (point of view of society, investor, maintenance contractor, etc.). To prevent numerical difficulties, the value of the objective functions were normalised using the results of the single optimizations (utopian solutions) that were performed beforehand.

To demonstrate the efficiency of the presented procedure, several optimal designs of ERS were obtained using different sets of the preference coefficients. Other designs can be obtained by performing the optimization with different preferences for the optimum design. This can help to make a decision on a final design of ERS.

13.5 Determination of ballast lateral resistance using optimization technique

13.5.1 Introduction

In this Section we describe a procedure to analyse the lateral resistance of a railway track making use of a numerical optimization technique. The procedure is part of a method used to determine train speed limits on newly laid tracks or tracks after renewal (see Chapter 12). The problem that arises identifying ballast parameters belongs to a class of so-called *inverse problems* in which parameters of an unknown system (object) are to be determined based on the system's responses (object). By minimising the difference between the responses of a system and its numerical model the system's parameters can be determined.

Ballast parameters and safe speed limits of trains are analysed using the same 3-D finite element model described in Chapter 7. Track responses in the shape of load/displacements diagrams are obtained by shifting a real track in lateral direction using a tamping machine. The measurements are also simulated using the FE model of a track. Based on the experimental data and the results of numerical simulation, ballast parameters are identified using an optimization technique (a Sequential Quadratic Programming method). The procedure is described below.
Sufficient lateral strength of a railway track is important for safe track operations. It allows the track to resist high lateral forces which can be caused by a moving train and/or temperature variation. The lateral strength of a track is to a large extent defined by the lateral resistance of ballast. In a newly laid track or track after full maintenance the ballast particles are not well enough consolidated. Therefore the lateral resistance of such a track is poor (see Chapter 7). In order to prevent track buckling, the lateral forces occurring during operations on newly laid tracks or tracks after maintenance should be restricted. Usually, this is achieved by temporarily restricting train speeds on renewed sections of such track. As the number of train passes grows, the ballast becomes more consolidated and its lateral resistance increases. As a result, the restriction of the speed limits can be lifted.

Obviously, the introduction of temporary speed restrictions deteriorates the operational efficiency of a railway track. One way to reduce the operational hindrance is to employ dynamic stabilising directly after track maintenance. However, it is still necessary to know whether the imposed speed limitations are safe or not. In order to apply proper speed limits, knowledge about the lateral resistance of a track and its interpretation are required to determine safe speed limits [176].

13.5.2 Measuring the lateral resistance of track

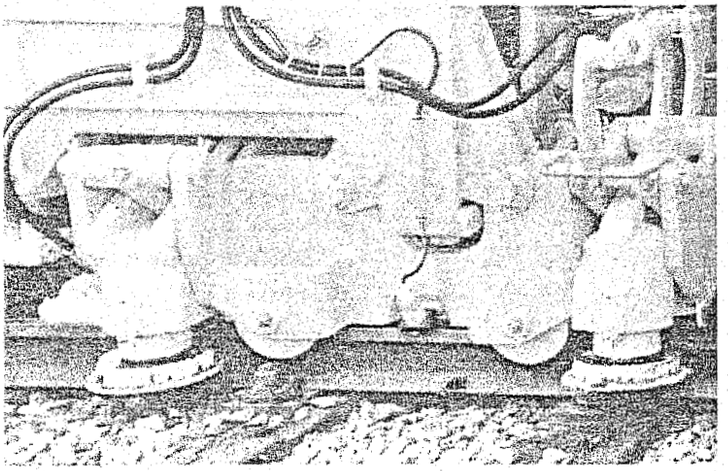


Figure 13.18: Measuring track lateral resistance using tamping machine

A technique used to measure the lateral resistance of a track employed by the Nederlandse Spoorwegen (Netherlands Railways) is presented here. The technique makes use of a tamping machine, which, instead of lifting the track, shifts the track frame in the lateral direction. To amplify the effect of vertical loads on the lateral resistance of a track, an additional vertical load can be applied as well. The resulting lateral displacements of the frame are measured and recorded. Finally, the applied lateral force and resulting lateral displacements of the track are combined in a force-displacement diagram.

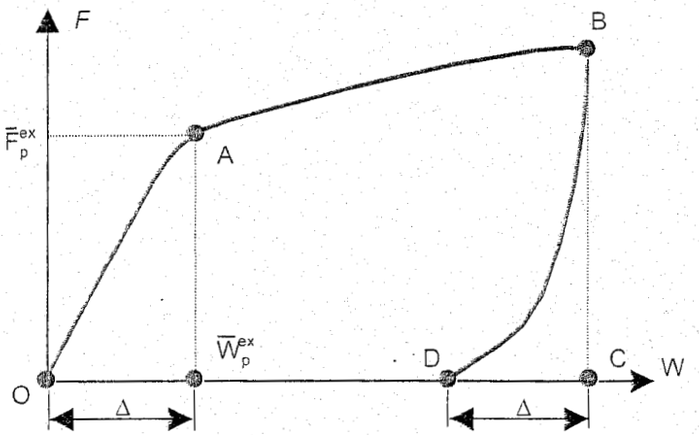


Figure 13.19: Lateral force-displacement diagram of track measured using tamping machine (not to scale)

A typical force-displacement diagram is shown in Figure 13.19. The applied lateral force is slowly increased which results in movement of the track frame (path O-A-B in Figure 13.19). As the lateral displacements achieve a prescribed maximum value (point B in Figure 13.19), shifting grips are released so that the track frame moves back (path B-C-D in Figure 13.19). It should be noted that the maximum displacements of the track during the test are relatively small (max. 5mm). This means that the residual displacements of the track are not large and the original track geometry can easily be restored. The elastic characteristics of the track can be obtained by considering the residual displacements of the track (path O-D in Figure 13.19).

The maximum elastic displacements $\bar{W}_p^{ex}$ of the track are then estimated by the displacements of the track after the lateral load has been released (path B-D in Figure 13.19), i.e. $\bar{W}_p^{ex} = \Delta$. The maximum elastic force $\bar{F}_p^{ex}$ can be obtained from the force-displacement diagram (Figure 13.19) as the load which corresponds to the displacement $\bar{W}_p^{ex}$. The elastic parameters of the track $\bar{F}_p^{ex}$ and $\bar{W}_p^{ex}$ measured using a tamping machine will be used here for ballast parameter identification.

13.5.3 Ballast parameter identification

General parameter identification problem

In a typical parameter identification problem there is an object (or process) to be investigated and a numerical model that can describe the behaviour of such an object. An investigated object is considered to be like a black box that produces responses depending on the values of certain input parameters $\boldsymbol{X}$. A numerical model is characterised by the same parameters $\boldsymbol{X}$ as the investigated object as well as by certain tuning parameters $\boldsymbol{a} = [a_1, \ldots, a_N]$. The goal of an identification problem is to find a set of parameters $\boldsymbol{a}$ that will allow the obtained numerical model to most accurately describe the object's behaviour. In other words, the difference between the object and model responses for the same set of input parameters should be minimal. Thus, a general parameter identification problem can be formulated as the following minimisation problem:

Minimise

$$G(\boldsymbol{a}) \equiv \sum_{1}^{P} \left[\Phi(\boldsymbol{X}_p) - \tilde{\Phi}(\boldsymbol{X}_p, \boldsymbol{a}) \right]^2 \rightarrow min \tag{13.20}$$

subject to the constraints

$$A_i \leq a_i \leq B_i, \quad i=1, \ldots, N \tag{13.21}$$

$\boldsymbol{a} = [a_1, \ldots, a_N]$ is the vector of the tuning parameters of the numerical model chosen as the design variables that can be varied during the optimization process;

in which:

Φ is the set of responses obtained from experiments with the object;
$\tilde{\Phi}$ is the set of responses derived from the numerical model;
A_i and B_i are the upper and lower boundaries of the tuning parameter;
$\boldsymbol{X}_p$ is the p-th set of input parameters of the original object;
P is the total number of such sets.

The problem's solution (13.20)-(13.21) $\boldsymbol{a}^*$ can be found using a conventional method of mathematical programming.

Numerical model

The 3-D Finite Element (FE) model of a railway track used here is shown in Figure 13.20 (which is identical to Figure 7.43). The model can be used for stability analysis of various track structures. The model parameters are described in Chapter 7. Ballast behaviour in lateral direction is characterised by four parameters F_p, W_p, F_l, and W_l related to elastic/plastic limits, and by the friction coefficient $tan\varphi$ between ballast and sleepers.

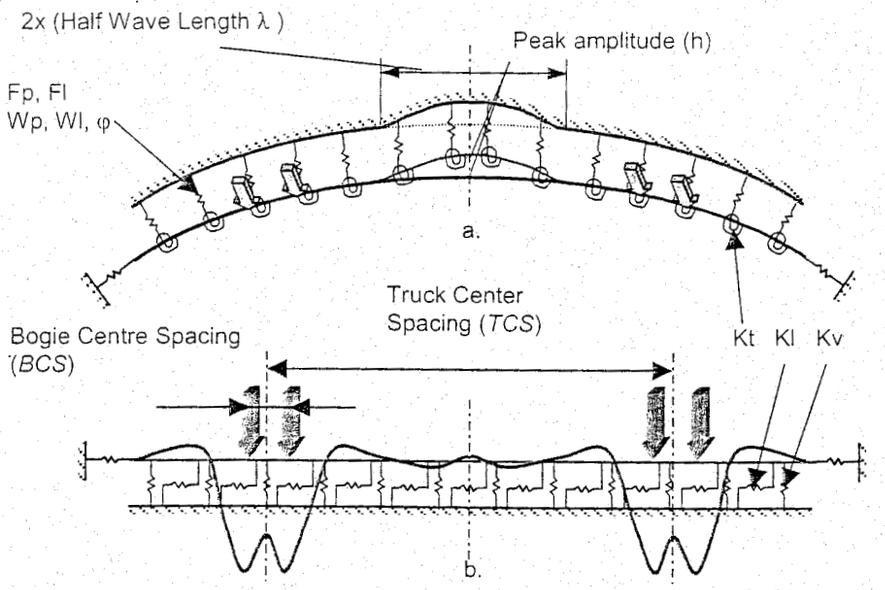


Figure 13.20: FE model of track used for identifying ballast parameters

An example of simulation of the lateral behaviour of a track using a bilinear ballast model is given in Figure 13.21. In this example, a track is loaded by the lateral force that uniformly increases from *0* to *17 kN*. Additionally, the vertical load $F_v = 100kN$ is applied to a track (the same application point is used as with the lateral load). The simulated lateral displacements of the track plotted for the load application point are shown in Figure 13.21b. The corresponding ballast behaviour in the lateral direction (bilinear model) is shown in Figure 13.21a. From these figures it can be observed that plastic deformations of the track $\overline{W}_p$ occur when the ballast begins to yield (W_p), i.e. $\overline{W}_p = W_p = 2mm$.

Thus, if the ballast parameters are known, the numerical model (Figure 13.20) can be used to simulate track behaviour under various loadings. On the other hand, if the lateral behaviour of an entire track is known, one can try to solve an inverse problem in order to determine the parameters related to the lateral ballast resistance. The procedure used to determine ballast parameters using an optimization technique is described below.

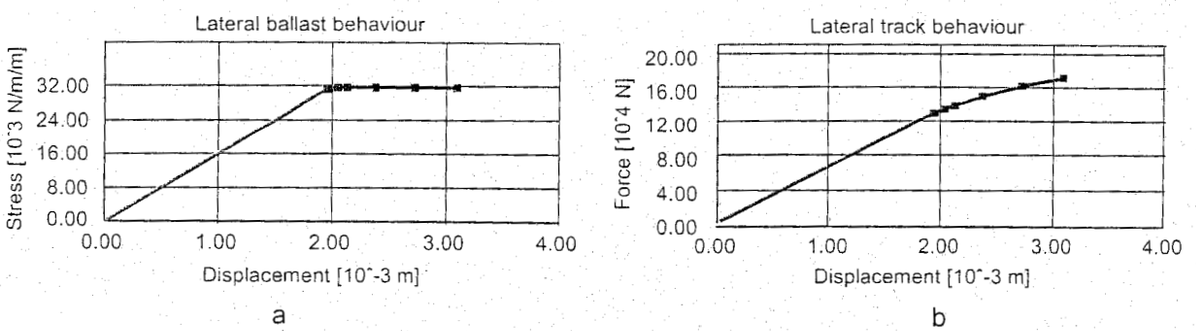
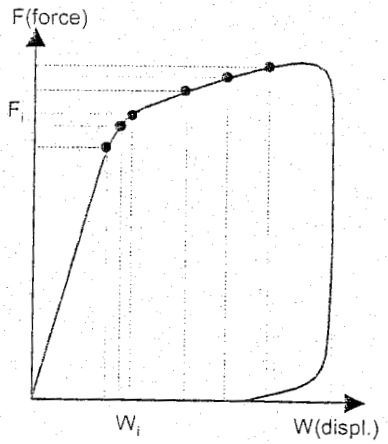


Figure 13.21: Lateral behaviour of ballast (a) and entire track (b) obtained from numerical simulation

Identification of ballast parameters

The general method for model parameter identification described above has been applied to determine the ballast parameters F_p, W_p, F_l, W_l, and $tan\varphi$. According to this method, one can try to vary the parameters of the numerical model in order to minimise the difference between the measurement data and the results of numerical simulation (Figure 13.22).



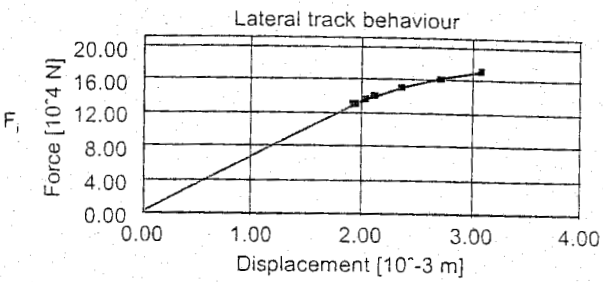


Figure 13.22: Ballast parameter identification: minimisation differences between measurement results and numerical simulation responses

Since the parameters characterising the elastic properties of the ballast (F_p and W_p) do not affect the plastic behaviour of the track and in order to speed up the optimization process, the parameters F_p and W_p can be determined first, apart from the other parameters. Moreover, due to a small shift of the track during measurements and, as a result, the lack of experimental data on plastic deformation of a track, it is difficult to obtain the values of the parameters F_l and W_l which characterise the plastic behaviour of ballast. This is why determination of these parameters is not discussed here. However, if the experimental data is available the same technique can be applied to determine the parameters F_l and W_l as for the elastic parameters.

If $\bar{F}_p^{ex}$ and $\bar{W}_p^{ex}$ are measured elastic limits (force and displacement) of a track (Figure 13.19), then, using the methodology of the previous subsection, one can try to find the values of the ballast parameters F_p and W_p. This should be done in such a way that the simulated and measured lateral displacement of a track ($\bar{W}_p$ and $\bar{W}_p^{ex}$) in a reference point caused by the limit force $\bar{F}_p^{ex}$ are as close as possible. Thus, the following minimisation problem similar to the problem (16.20)-(16.21) is to be solved:

Minimise

$$G(\boldsymbol{a}) \equiv [\Phi(\boldsymbol{X}_p) - \tilde{\Phi}(\boldsymbol{X}_p, \boldsymbol{a})]^2 \rightarrow min \tag{13.22}$$

in which

$$\boldsymbol{a} = [F_p, W_p], \; \boldsymbol{X} = [\bar{F}_p^{ex}], \; \Phi(\boldsymbol{X}) = [\bar{W}_p^{ex}], \; \tilde{\Phi}(\boldsymbol{X}) = [\bar{W}_p]. \tag{13.23}$$

Here, the vector of the responses of the track structure Φ (and consequently of the model $\tilde{\Phi}$) consists of the displacement of only one point on the track $\bar{W}_p^{ex}$ (and model displacement $\bar{W}_p$), namely the application point of the lateral force s. If displacements of other points of the track are available they can be taken into account as well. Because of the linear relation between the force and displacement of the track, the displacements corresponding to one loading step $s = \bar{F}_p^{ex}$ are considered (i.e. $P = 1$ in (13.20). When non-linear behaviour of a track is analysed, more loading steps (components of the vector $\boldsymbol{X}$) should be used.

Since the track displacements are obtained performing a static linear analysis, the optimization problem (13.22)-(13.23) can be solved relatively quickly. To speed up the optimization it can also be assumed that yielding of the entire track occurs when the ballast begins to yield, i.e. $W_p = \bar{W}_p = \bar{W}_p^{ex}$. The vector of the design variables of the problem (13.22)-(13.23) then only has one component, namely $\boldsymbol{a} = [F_p]$. If no extra vertical load is applied, the obtained optimal parameters $a = [F_p^* W_p^*]$ are the elastic limits of ballast, i.e. $F_p = F_p^*$ and $W_p = W_p^*$. In case the measurement data have been obtained with extra vertical load, the obtained optimal parameter F_p^* is considered as

the maximum yielding force $s_{max} = F_p^*$ defined as $s_{max} = F_p - s_v tan\varphi$ (see Chapter 7). Based on F_p^* the peak strength of ballast can be calculated as follows:

$$F_p = F_p^* + s_v tan\varphi \qquad (13.24)$$

in which s_v is the vertical force that acts on each lateral spring element caused by vertical load F_v (see definition of the spring element for ballast lateral behaviour in Chapter 7. The corresponding limit deformation W_p is evaluated as follows

$$W_p = \frac{F_p}{K_{el}}, \quad K_{el} = \frac{F_p^*}{W_p^*} \qquad (13.25)$$

It should be noted that the rail temperature during the measurements can also be taken into account in the ballast parameter identification problem.

The minimisation problem (13.22)-(13.23) has been solved using an optimization technique based on a Sequential Quadratic Programming (SQP) method [229]. The method is iterative. It is based on successive linearisations of P non-linear differences in (13.22), combining a first order trust region method with a local method which uses approximate second order information. The first order derivatives are approximated by finite differences. The choice for the SQP method in order to solve the problem (13.22)-(13.23) is explained by the fact that the function G was relatively smooth and, therefore, the SQP method did not encounter any difficulties. If the objective function G is strongly non-linear, another optimization method is to be used, e.g. the MARS method (see example in Chapter 13.6).

Verification of parameter identification procedure

Experimental data measured on a ballast track with wooden sleepers have been used [193] to verify the parameter identification procedure. The data was obtained employing a tamping machine. During measurements a neutral rail temperature was assumed, whereas the friction coefficient tanφ = 1.1 was used for the wooden sleepers. Other parameters of the model which have been used to determine ballast parameters can be found in [174].

Since no extra vertical load was applied to the track structure during the measurements ($F_v = s_v = 0N$), the values of the parameters F_p^* and W_p^* obtained from the solution to the minimisation problem (13.22)-(13.23) have been used as the ballast parameters (see above), viz. $F_p = F_p^* = 44680\ N/m$ and $W_p = W_p^* = 0.0018\ m$ (Table 13.4). Ballast (lateral spring element) and track lateral behaviour are shown in Figure 13.24 using the obtained ballast parameters when no extra vertical load is applied.

Input			Output			
$\bar{F}_p^{ex}$	$\bar{W}_p^{ex}$	T_{exp}	$F_p'[N/m/m']$	$F_p[N/m]$	$W_p[m]$	$K_{el}[N/m/m]$
64800	0.0018	0	12471	44680	0.0018	6928333

Table 13.4: Results of identification problem using measurements without extra vertical load

To verify the results of identification, the lateral resistance data of the same track were used measured with an extra vertical load $F_v = 100kN$ (applied at the application point of the lateral force [193]). The measurements with the extra vertical load have been simulated using the model with the obtained ballast parameters F_p and W_p. The results of the simulation are shown in Figure 13.23.

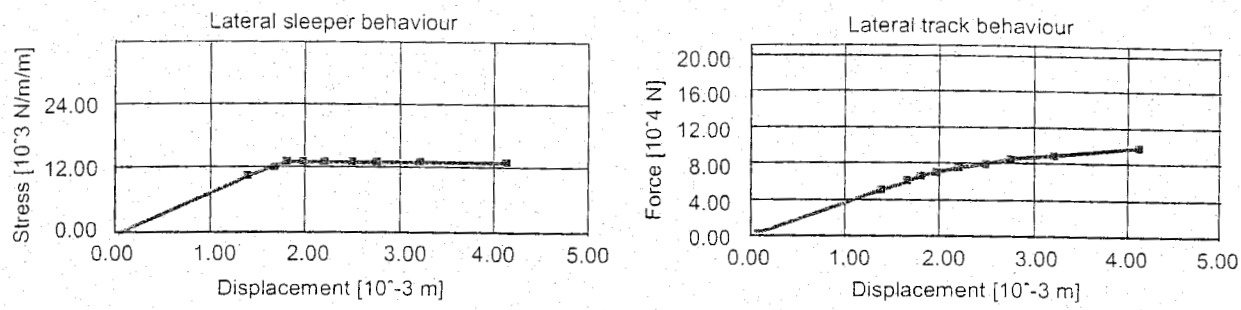


Figure 13.23: Model obtained using measurement data without vertical load; case with no vertical load

Comparing these results with the experimental data, it can be seen that the simulated and measured (average) values of the maximum elastic force, $\bar{F}_p = 132kN$ in Figure 13.24 and $\bar{F}_p^{ex} = 132kN$ in [193] respectively, are very close. From this comparison it can be concluded that the numerical model provides correct results. The verification procedure is summarised in Figure 13.25. It should be noted that the measurement results regarding extra vertical force can also be used to determine the friction coefficient between the ballast and sleeper. The verification shows that a numerical model of which the parameters are obtained based on measurement data can very well describe the behaviour of the track and, therefore, can be used for simulations under various loadings. The numerical model can, for example, be used to determine the maximum allowable buckling temperature of the rails (see the procedure used to determine temporary speed limits on tracks after renovation in Chapter 7).

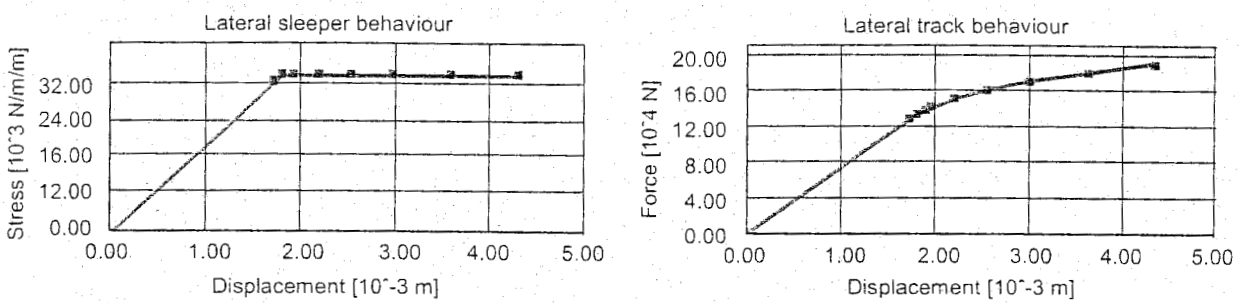


Figure 13.24: Model obtained using measurement data without vertical load; case with vertical load 100kN

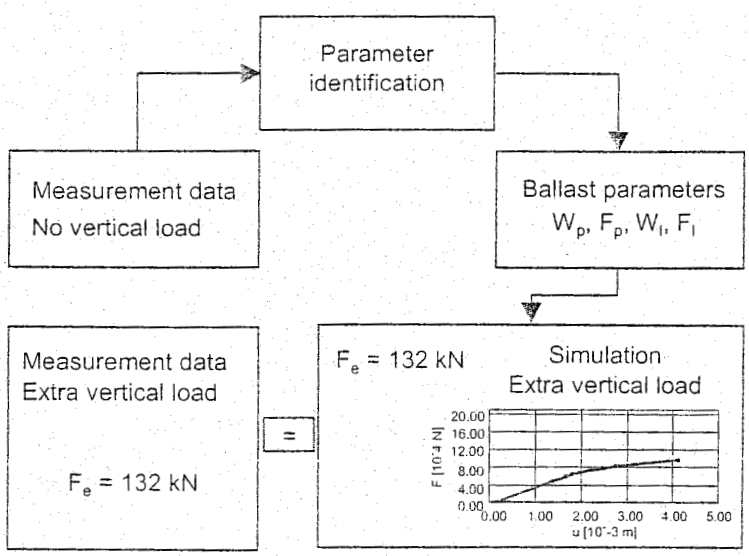


Figure 13.25: Flow-chart of verification of identification procedure

13.5.4 Conclusions

A procedure has been presented to determine the lateral ballast resistance based on measurement results with a real track. The ballast parameter identification is performed using an optimization technique. The procedure has been verified using the measurement results with and without extra vertical load.
The ballast parameter identification and speed limit determination techniques have been implemented in the software BLATRES that can be installed on existing tamping machines.

13.6 Identification of dynamic properties of railway track

13.6.1 Introduction

In this Chapter a methodology is presented in order to assess the dynamic characteristics of a railway track (an Embedded Rail Structure) using an optimization technique. First, the dynamic responses of a track are obtained using a hammer excitation technique. This is achieved by applying a vertical impulse load to a rail and measuring the resulting accelerations. The measured data are then analysed using the Frequency Response Function (FRF). The hammer test is also simulated in the finite element sofware RAIL, which can be used to analyse the dynamic behaviour of a track under various loadings. To match the simulation and measurement data, a numerical optimization technique has been used called a Multipoint Approximation based on Response Surface Method (MARS). The parameters of the model related to the material properties of the elastic compound (polyurethane and cork mixture) in which the rails are embedded were varied during the identification process. The numerical model with the obtained parameters of compound can then be used to estimate the performance characteristics of the railway track. The compound parameters were compared with the ones obtained using a single degree of freedom mass-spring system.

Safe and efficient railway operations require properly designed rail infrastructure, which means that track quality should meet certain criteria both before and during operation. One of the most important characteristics of a track is its vertical stiffness. Under a high frequency loading the vertical stiffness of a track defines the level of track vibrations and wheel-rail contact forces. These ultimately define the emission of acoustic noise and efforts to maintain rail and wheels. Additionally, large variations in track elasticity can cause serious damage to the track which can result in problems regarding the operational safety of the track. Insufficient vertical stiffness of track can yield rail fatigue problems, whereas track which is too stiff may have high noise emission and wear of wheels and rails as a result. This is why assessing the vertical stiffness of a track is extremely important in order to perform track maintenance on time.

Recently, a method to estimate the dynamic characteristics of a track has been developed within the Dynatrack project [169] at Delft University of Technology (see Chapter 14). The procedure consists of three major parts, namely measurement of dynamic responses of a track, numerical modelling of track behaviour, and parameter identification of track components. The dynamic characteristics of a track are obtained using a Hammer Excitation Test (HET) according to which a vertical impulse load is applied to the top of a rail. The resulting track vibrations recorded within the 40 to 2000 Hz range are the most representative in order to predict noise emission and wheel-rail interaction forces. The dynamic properties of a track are then estimated by analysing the Frequency Response Functions based on the measurement data.

In the present study the dynamic properties of a special type of Embedded Rail Structure (ERS) are investigated. ERS is a ballastless type of track that has a number of advantages compared to a classical (ballasted) track, the most important of which is low maintenance cost. This is especially the case if high-speed operations arise. The low maintenance effort results in high availability of the track which means that the track is only unavailable for operations during a limited period of time. Apart from that, the height and weight of ballastless track is lower in comparison with traditional track. This is important for bridges and tunnels. The typical kind of ERS applied up to now consists of a continu-

ous reinforced concrete slab that rests on a concrete stabilised roadbed which in turn is placed on a sand base (Figure 13.4). Rails in such a structure are continuously supported by an elastic compound made up of a mixture of polyurethane and cork. One of the advantages of ERS is that its behaviour can be easily described by a numerical model. More information on Embedded Rail Structure and other ballestless tracks can be found in Chapter 9.

In order to analyse the dynamic behaviour of a track caused by various loadings such as a moving train or temperature variation, a Finite Element model has been developed at DUT [91]. Various track structures can be described by the model which has been implemented in computer software called RAIL. A general description of the software is given in Chapter 6. The accuracy of track responses obtained using the numerical model strongly depends on the values of the model parameters that are related to track components. Some of these parameters, such as rail geometry, have been described accurately, other parameters can only be obtained experimentally.

Here, the material properties of compound are determined by matching the excitation hammer measurement data to the results of numerical simulation. The identification of the material parameters is formulated as a minimisation problem. To solve this problem an optimization technique has been used called a Multipoint Approximation based on Response Surface Method (MARS) as presented in Section 13.3. This method has been specifically developed for problems which involve multiple response analyses and (time consuming) simulations. To verify this method, the material properties of compound are also obtained using a single degree of freedom mass-spring system with damping. In the subsequent sections the following topics are presented:

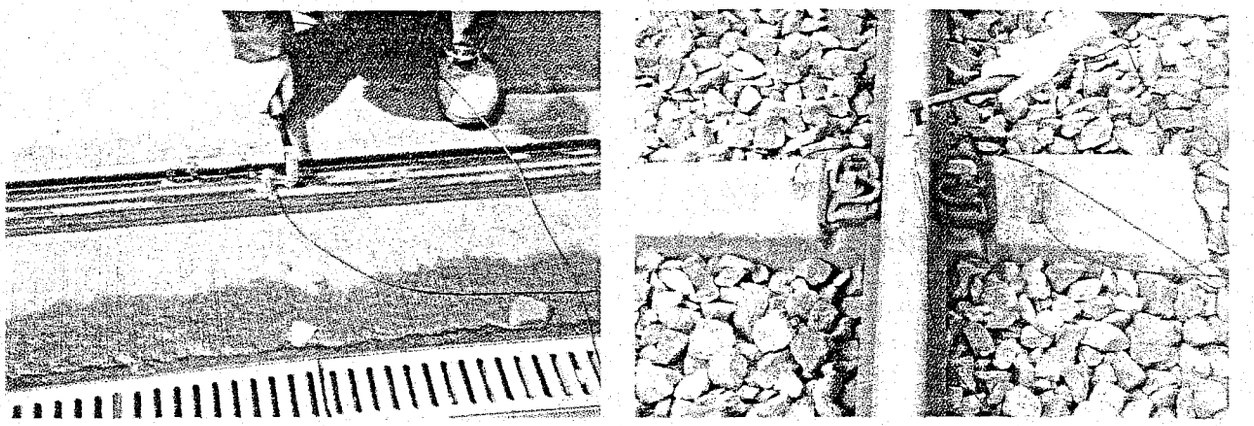
- Hammer Excitation Test (discussed briefly below);
- Identification problem;
- Numerical results.

13.6.2 Hammer excitation test

It is essential to measure track performance in a continuous manner (such as e.g. using a High Speed Deflectograph [232]) in order to obtain information on the overall quality of a track as well as to determine the location of 'weak spots'. Usually, such an overall view provides sufficient ground to perform maintenance on certain sections of a track. However, sometimes weak spots and, in extreme cases, an entire track require closer examination, e.g. if track problems occur repeatedly. For that purpose a spot measuring method has been developed such as the Hammer Excitation Test [167], [169]. Similar to the High Speed Deflectograph, the Hammer Excitation Test concentrates on track elasticity in vertical direction, but now subject to higher frequency loading. This is strongly related to noise, vibrations, and wheel-rail interaction forces. The technique is described below.
The major part of track performance that can be assessed by the Hammer Excitation Test is the dynamic (vibration) behaviour of track structures (Figure 13.26). The overall performance of a track is strongly defined by the characteristics of its components. Therefore, the following properties of a track are important in the performance assessment:

- Dynamic properties such as stiffness and damping (e.g. rail and concrete plate pads, and ballast);
- Inertia properties such as mass, bending, and shear stiffness (e.g. rails and sleepers);
- Geometrical properties such as structural dimensions (e.g. support spacing and support length).

The analysis of the dynamic behaviour of a track includes measuring the applied load as well as the resulting vibration of a track. A vertical (impulse) load applied to the top of the rail (Figure 13.26) produces track vibrations in the frequency range between 30 Hz and 3 kHz, depending on the type of impulse load. The resulting vibrations are recorded using accelerometers which have an application range between 10 Hz and 10 kHz. The recorded signals are then sampled at 20 kHz and converted



a b

Figure 13.26: Hammer Excitation Test on classical track (a) and Embedded Rail Structure (b)

by means of Fast Fourier Transform (FFT) into Frequency Response Functions (FRF), while assuming the linear behaviour of a track.

The inertance frequency response function $H_{AF}[kg^{-1}]$ relates the applied force (input) F and the recorded accelerations of a track (output) A and reads:

$$H_{AF}(f) = \frac{S_{AF}(f)}{S_{FF}(f)} \tag{13.26}$$

in which

$S_{AF}[mN/s]$ is the complex cross-spectrum of accelerations (output);

$S_{FF}[N^2s]$ is the power spectrum of the impulse force (input);

$f[Hz]$ is the frequency.

Sometimes, it is also important to analyse the receptance frequency response function $H_{XF}[m/N]$. Similar to the inertance, it relates the applied force (input) F and the resulting displacement (output) X as:

$$H_{XF}(f) = \frac{S_{XF}(f)}{S_{FF}(f)} = \frac{1}{(2\pi f)^2}\frac{S_{AF}(f)}{S_{FF}(f)}, \quad S_{XF}(f) = (2\pi f)^{-2} S_{AF}(f) \tag{13.27}$$

in which $S_{XF}[mNs]$ is the complex cross-spectrum of displacements (output). From (13.26) and (13.27) the following relation between the inertance and receptance FRF can be obtained:

$$H_{AF}(f) = (2\pi f)^2 H_{XF}(f) \tag{13.28}$$

One of the advantages of FRF is that it is not sensitive to the magnitude and duration of an impulse load. Moreover, in [169] it is shown that the recording signals obtained using different types of hammers are very close, except the extremely low and extremely high frequencies which are not considered here.

13.6.3 Numerical model

In order to better understand and interpret measurement results, a numerical model that is able to describe the various kinds of behaviour of a railway track is necessary. The choice of parameters for such a model is crucial for the assessment and prediction of track performance. Once determined on

technique called Multipoint Approximation based on Response Surface Method (MARS), presented in Section 13.3.

13.6.5 Numerical results

The above-mentioned procedure has been applied to determine the material parameters of compound used in ERS. Thus, the components of the vector of the design variables $x = [K_c\ C_c\ \rho_c]$ in the problem (13.29) are respectively the stiffness, damping, and density of the compound. The initial values and lower and upper boundaries of the design variables are collected in Table 13.5. The inertance FRF $\tilde{H}_{AF}$ in (13.29) has been obtained using the numerical model implemented in RAIL (Figure 13.27). The frequency range from 50Hz to 170Hz was considered since the FRF in this range are mostly defined by the properties of compound. The MARS method has been used to solve the unconstrained minimisation problem (13.29), whereas the multiplicative model has been used to approximate the objective function D.

Design variables	Lower boundary	Upper boundary	Initial value	Optimal value	Units
K_c	10.0	100.0	16.0	26.3	MN/m/m'
C_c	0.0	50.0	15.0	5.2	kNs/m/m'
ρ_c	0.0	6.0	0.1	2.1	kg/m'

Table 13.5: Design variable of track properties identification problem

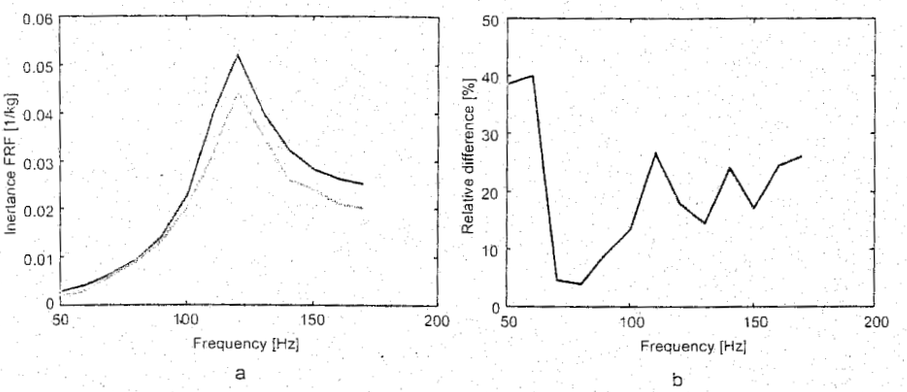


Figure 13.28: Optimal solution using MARS method: measured (green line) and simulated (red line) inertance FRF (a); relative difference between measured and simulated FRF (b)

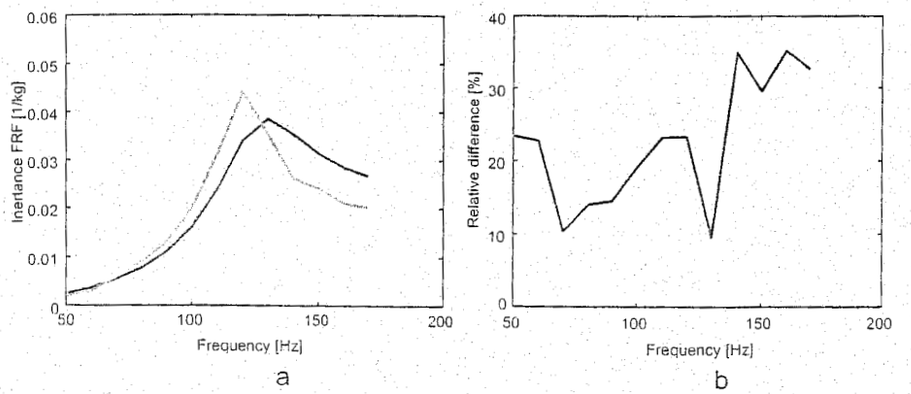


Figure 13.29: Optimal solution using curve-fit and 1DOF mass-spring system: measured (green line) and simulated (red line) inertance FRF (a); relative difference between measured and simulated FRF (b)

An optimum solution (Table 13.5) was found after 20 iterations. The measured and simulated FRF as well as the relative difference between the FRF corresponding to the obtained parameters of compound, are shown in Figure 13.28. To estimate the efficiency of the identification procedure, the compound parameters have been obtained using a single DOF mass-spring system with damping [169] as well. The mass of the system representing the rail mass was kept constant, whereas the stiffness and damping parameters were obtained using the curve-fit function of MATLAB. The obtained parameters were then transformed into the stiffness K_c and damping C_c of the RAIL model (Figure 13.27). The corresponding FRF and their relative differences are shown in Figure 13.29. From Figure 13.28 and Figure 13.29 it can be observed that the model obtained by the MARS method describes the behaviour of the real track more accurately, especially in the neighbourhood of the resonant frequency. The overall difference between the measured and simulated FRF is also smaller for the model obtained using the MARS method (compare the area under graphs in Figure 13.28 and Figure 13.29). The results of the more accurate calculations with these two models are shown in Figure 13.30. Again, the model obtained using the MARS method approximates the measured data more accurately, especially the phase angle between 70Hz and 170Hz. The track model obtained using the MARS method fits better due to the compound mass variation during the identification. Still, some discrepancy exists between the measured and simulated FRF (Figure 13.28) which can be explained by the insufficient accuracy of the track model. The introduction of non-linear elements in the track model could improve how the results fit.

13.6.6 Conclusions

A procedure has been presented in order to assess the dynamic properties of the components of a railway track. The procedure includes measuring the dynamic responses of the track (Hammer Excitation Test), numerical modelling of track behaviour (RAIL), and parameter identification. The procedure has been applied to a ballastless track, namely an Embedded Rail Structure.

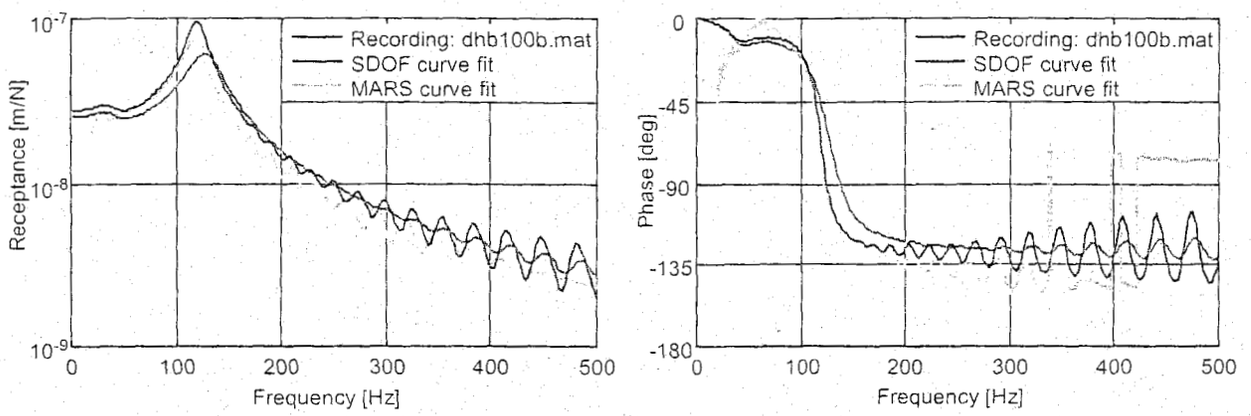


Figure 13.30: Amplitude and phase of FRF using measured data (green line), MARS curve fitting (blue line), and 1DOF curve fitting (red line)

The material parameters of elastic compound have been determined based on measurement data using the MARS optimization method. The track model which includes the obtained parameters of compound can quite adequately describe the behaviour of a real track. The results have shown that the MARS method can effectively be used for parameter identification problems.

The measured and simulated data could fit more accurately if non-linear elements would be introduced in the numerical model of a track.

14 TESTING AND ACCEPTANCE

14.1 Introduction

Railway track is composed out of many different components which have already been introduced in Chapter 5, Chapter 8, and Chapter 9. All components have very specific functions and are designed to fulfil their functions in the best possible way and for the longest possible period of time.
Product development in railway engineering has been highly influenced by empirism and conservatism; a consequence of the prevailing safety requirements, low costs for labour and materials, and availability. During the last decades, techniques have rapidly been developed which allow the production of components according to better specifications and higher quality standards. New materials, mechanisation, automation, production as well as test devices have been introduced and used in railway practice.

Besides the component properties which can be controlled during the manufacturing process, the system properties of the assembled track structures should be controlled as well and should meet quality standards. This puts emphasis on the quality control of construction and maintenance activities carried out by contractors.

This chapter will focus on recent developments in component and structural testing, quality assessment of track components, and structures leading to acceptance. The owner of the track will then be enabled to provide the train operator with safe, durable, and cost-efficient infrastructure.

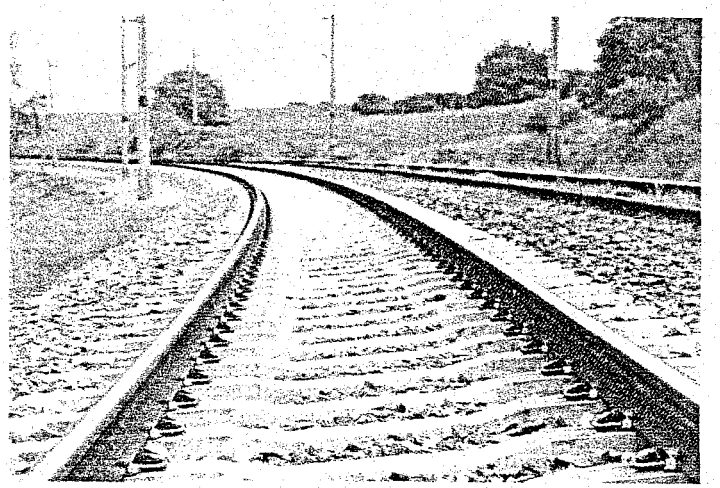


Figure 14.1: New developments in railway track require testing before acceptance

14.2 Component testing and acceptance

14.2.1 Mechanical properties

Track components are supposed to have specific mechanical properties that enable the track to support and guide railway vehicles. In Table 14.1 a number of track components and properties are listed. These properties are arbitrarily categorised and show their most important features. Mechanical properties are not necessarily the most important ones, but they reflect the principle of considering the track as a mechanical system subjected to vehicle loading.

	Elasticity	Strength	Stability	Durability
Rail profile		X		X
Fastening system	X		X	X
Sleepers		X	X	X
Ballast	X		X	X
Slabs		X		X
Track support systems	X		X	X

Table 14.1: Overview of the most important track properties of each component

- Track elasticity is dominated by the stiffness properties of some of the track components such as rail pads and ballast. Damping, also known as energy loss or imaginary stiffness, is very relevant regarding low, medium, or high frequency loading (See 14.2.2).

- Track strength depends on the robustness of the design, the quantities and the qualities of the used materials. This is especially the case regarding rails (and welds), concrete sleepers, and slabs (See 14.2.3).

- Track stability is provided by the rigid frameworks of sleepers and rails, but also by the good resistance of sleepers in ballast (See 14.2.4).

- Finally, for long term performance durability and resistance against fatigue are of course primary requirements for track components. The environmental and loading conditions of a particular track section, however, heavily influence the durability requirements.

14.2.2 Elasticity properties

Track elasticity is responsible for the gradual distribution of forces on the track over the underlying substructure. Discrepancies in or loss of the track elasticity may lead to damage of track and vehicle. In order to determine track elasticity, some test procedures for laboratory specimen have been developed over the years, featuring, in general, static and low frequency loading. The principle of determining several types of track elasticity is illustrated below.

The first example concerns an assembly of rail, rail pad, and fastener installed on a concrete sleeper. In ballasted track structures with concrete sleepers, static and low frequency elasticity is mainly offered by the ballast underneath the sleepers and by the substructure. Wooden sleepers contribute considerably to static and low frequency track elasticity [153]. Concrete sleepers and slab tracks hardly do. They are more sensitive to medium to high frequency loading than wooden sleepers. This necessitates rail pads to meet specific high frequency or impact load filtering requirements. Internationally standardised tests, such as described in [29] and [31], concentrate on determining the static, low frequency and impact (high frequency) stiffness.

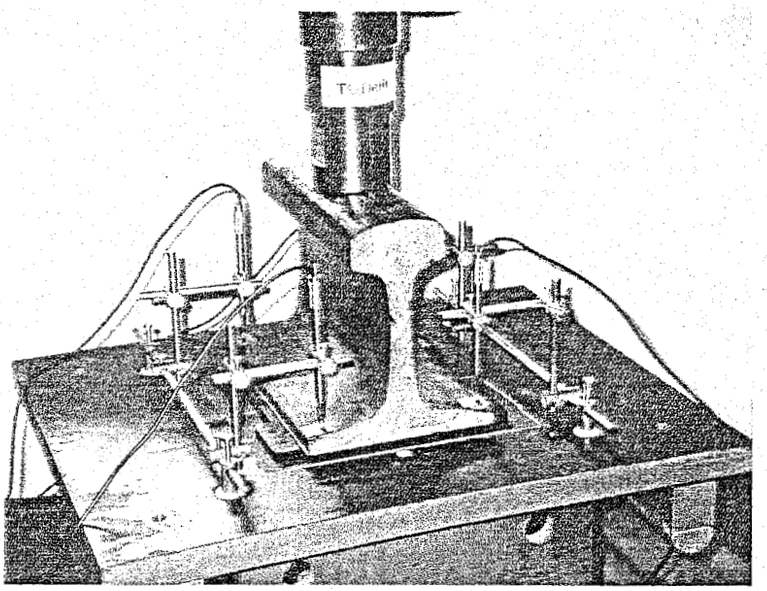


Figure 14.2: Test setup for railpad stiffness tests

The static, low frequency or impact stiffness of rail pads should be determined in reasonable loading ranges. In discretely supported tracks, the maximum load per rail pad highly depends on the axle configuration of the vehicle, the rail type, and support spacing. In general, the maximum load per rail pad is approximately 40% of a wheel load. The (secant) stiffness is often referred to as the typical and useful elasticity value of rail pads. It is important to state the loading range and loading rate explicitly as rail pad materials and rail pad shapes may have non-linear stress dependent features.

In Figure 14.2 a typical configuration is displayed and in Figure 14.3 a typical result of a rail pad loading test.

This rail pad has remarkably linear stiffness properties. In Table 14.2 some static stiffness values of rail pads for concrete sleepers are shown.

In slab track structures, baseplate pads are very commonly applied as a substitute for elasticity which is normally provided by the ballast and the substructure.

These baseplate pads have lower stiffness values than railpads. The lower stiffness is compensated for by applying softer materials and thicker pads.

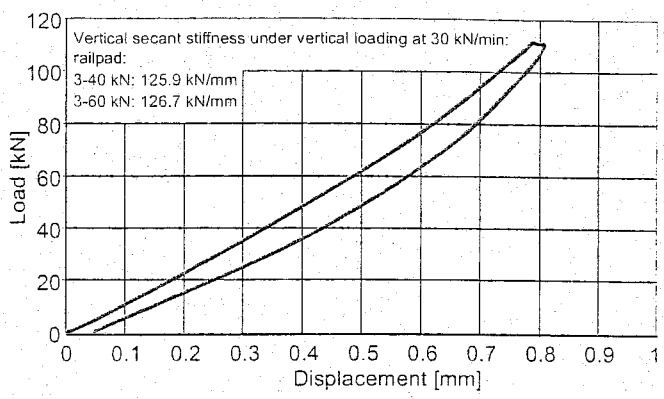


Figure 14.3: Typical railpad stiffness determination graph

Figure 14.5 shows a typical result of a loading test on a baseplate pad, which is shown in Figure 14.4. Generally speaking, softer materials have less linear stiffness properties under static loading, but also under low and medium frequency loading and temperature loading. Stationary loads on the baseplate pad introduce further deformation, as can be noticed at the top of the loading curve in this figure. This material property is called creep.

Supplier	Pandrol [187]	Tiflex [266]	Tiflex [170]	Vossloh [285]	Vossloh [264]
Applied by	BR	GVBA	NS	DB	DB
Material code	4548	Fc9	Fc9	Zw692	Zw687
Material	Studded elastomere	Rubber bonded cork	Rubber bonded cork	Plastic	Plastic
Thickness [mm]	10.0	6.5	4.5	6.0	6.0
Loading rate [kN/min]	50	30	30	30	50
Loading range [kN]	1-80	3-40	3-40	2-50	2-40
(secant) Stiffness [kN/mm]	65	100	125	400	970

Table 14.2: Static stiffness properties of railpads

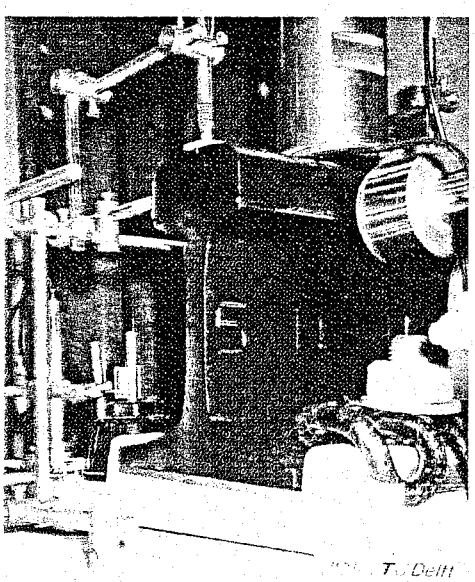


Figure 14.4: Test setup for railpad stiffness tests

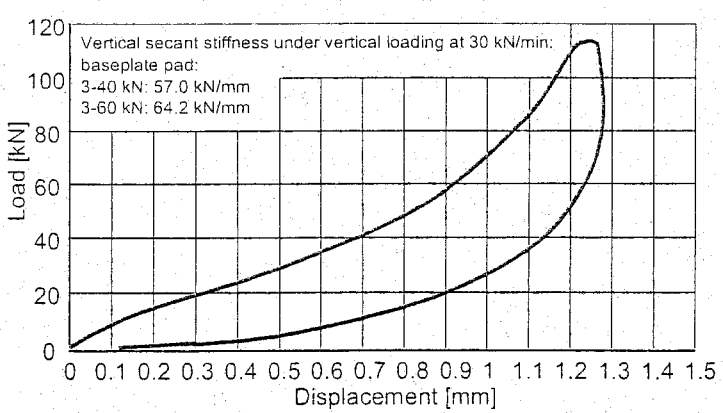


Figure 14.5: Typical baseplate pad stiffness determination graph

The stiffness values of this and other supports have been collected in Table 14.3.

Supplier	Tiflex [170]	Vossloh [284]	Vossloh [285]
Applied by	NS	GVBA	DB
Material code	Fc9 / Fc6	Zw1403 / Zwp40U	Zw692 / Zwp104
Material	Rubber bonded cork	Plastic / Rubber	Plastic / Elastomere
Thickness [mm]	4.5 / 10	5.0 / 27	5.0 / 10
Loading rate [kN/min]	30	30	30
Loading range [kN]	3-40	2-32	2-50
(secant) Stiffness [kN/mm]	39.2	37	22.5

Table 14.3: Static stiffness for slab track structure supports

According to European Standards [29], the stiffness of rail pads and single rail support for slab track is classified as follows.

Pad classification	'soft'	'medium'	'hard'
Pad stiffness	<80	80-150	>150

Support classification	'soft'	'soft-medium'	'medium-hard'	'hard'
Support stiffness	<30	30-50	50-70	>70

Table 14.4: Static stiffness classification in [kN/mm] according to [29]

In the DB standards for slab track [46] the vertical stiffness of a single rail support (spacing 0.65 m) is 22.5 kN/mm +/- 2.5 kN/mm. In above-mentioned general European standards, the DB requirements for slab track supports are thus classified as soft.

One major difference between European standards and e.g. national standards (DB, Railtrack, AREMA, etc.) is that the first only categorise the different supports and pads and assign testing methods, while the second accurately state both the testing method and the required results for acceptance.

In Chapter 8 some ballast bed stiffness values have been stated, so that the equivalent support stiffness of ballasted track can be calculated for comparison. Laboratory or in-situ tests and requirements on ballast stiffness are, however, rare. Static support stiffness of ballasted track varies between 20 and 40 kN/mm per single rail support. Compared to slab track support stiffness this covers the 'soft' end of Table 14.4.

Low frequency loading (<40 Hz) can more accurately represent the elastic behaviour of the track due to the passing wheels of a travelling train than static loading tests can. In Figure 14.6 some loops of low frequency loading of a single support are shown. These loops allow determination of stiffness and damping (energy loss) values per frequency. There is a clear difference between the loading and the unloading branches of each graph. Energy loss is the ratio of the closed loop area over the total area underneath the loading part of the graph. Energy loss for most of the single rail supports varies between 5 and 15 %. This level of energy loss makes these single rail supports to heavily damped structures in mechanical classification. The energy loss of rail pads is in general lower than 10 %, while ballast and other parts of the substructure have extraordinary large energy loss properties.

The low frequency stiffness of a single rail support is nearly always higher than its static stiffness. It is important to state the loading levels and loading frequency when mentioning and comparing these stiffness values. In DB standards [46] the allowable low frequency stiffnesses of single rail supports on slab track are stated explicitly. They increase from static (20-25 kN/mm) to low frequency at 40 Hz (20-40 kN/mm), while the loading conditions are specified in [47]. For realistic temperature conditions +20 % upper and -20 % lower boundaries apply.

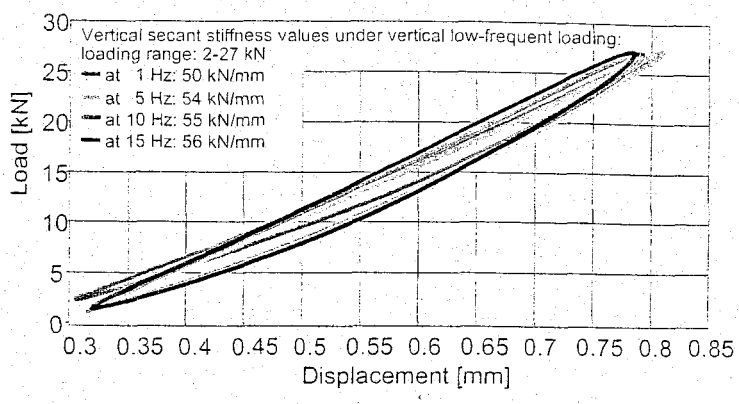


Figure 14.6: Dynamics loops while testing rail supports at different frequencies

Fatigue properties due to re-peated loading should not allow more than a 25 % increase of stiffness according to [29]. More about fatigue testing is mentioned in 14.2.5.
In and on civil structures like bridges and tunnels, the energy absorption of the track structure might be insufficient. If this is the case, the civil structure, track structures, and environment will suffer from damage, vibrations, and noise. In order to abate these problems, extra elasticity (resilience) is added to the track or the civil structure, e.g. by means of super elastic supports or ballast mats.

For additional elasticity and energy absorption, porous but durable materials are used, see Chapter 8. In ballast mats either holes or bonded cork particles are embedded in natural or synthetic rubber products with thickness between 20 and 40 mm. Ballast mats are installed directly on the concrete surface of the civil structure and protect ballast stones from crushing. Typical stiffness properties of some ballast mats are collected in (Table 14.5). They convert the track superstructure into a suspended mass spring system.

Supplier	Phoenix [227]	Tiflex [297]	Clouth [36]
Material code	U33-01	Fc250/Fc101	USM1000
Material	Synthetic and natural rubber	Rubber bonded cork	Natural rubbers
Mat thickness [mm]	33	24 / 6	30
Static secant bed modulus [N/mm^3] between 0.02 and 0.10 N/mm^2	0.015	0.025	0.019
Low frequency bed modulus at 10 Hz [N/mm^3]	0.018	–	0.014
Oscillation amplitude [mm]	0.075	–	
Preload [N/mm^2]	0.03	–	0.03

Table 14.5: Overview of stiffness properties of different ballast mats

Likewise, additional elastic supports (helical springs, absorbers, rubber bearing plates) or mats are installed as bearings underneath and along slabs. A final suggestion is to only insulate the sleepers or blocks on which the rails are fastened, from slab or ballast by means of sleeper boots, block encasings or undersleeper pads. Examples are shown in Chapter 8 and Chapter 9. All elastic components installed in the track – even in switches – are, of course, subject to testing and acceptance. It is, however, impossible to pay attention to each individual testing procedure and the acceptance limits at this point.

14.2.3 Strength properties

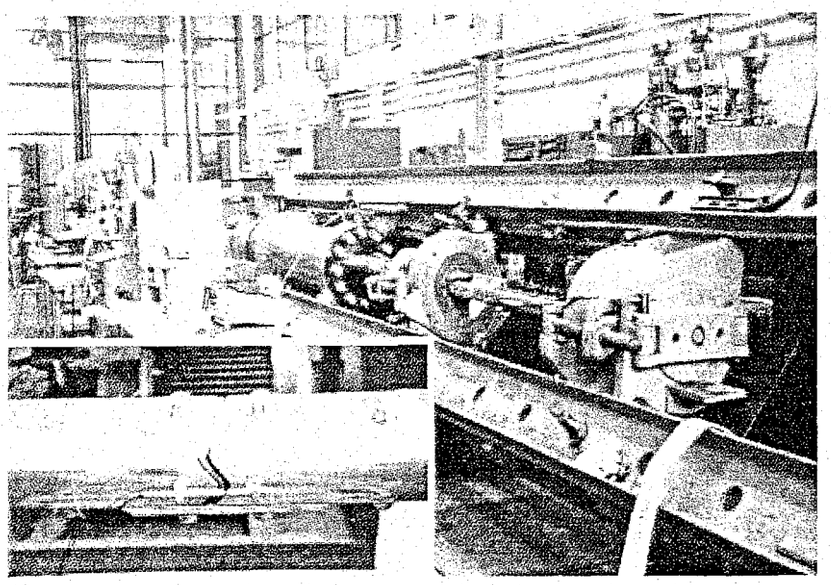


Figure 14.7: Maximum tension force test at insulated rail joints

As track structures are composed out of several different components, all components and materials should be part of the analysis of the overall track strength. The strength properties of rails are related to material quality and shape. In railway applications it is important to consider the strength of the rails in relation to repetitive loading and wear and, therefore, in fact, to fatigue and durability. Chapter 10 gives more information about the quality assessment of rails. In [32] internationally standardised testing methods for rails are listed which cover some strength properties. Rail strength is determined based on material samples.

For rails, joints, and welds, however, strength and fatigue testing is generally prescribed in national standards. Though it is certain that most of the joints and welds are exposed to worse loading than the unmodified rails, strength is determined in a comparable way and with comparable minimum values. In [188] longitudinal tension tests are described for glued (insulated) rail joints. The tension forces can run up to over 1500 kN before the joints crack. For acceptance, minimum force levels of 1200 kN are required. In Figure 14.7 the test set-up is shown.

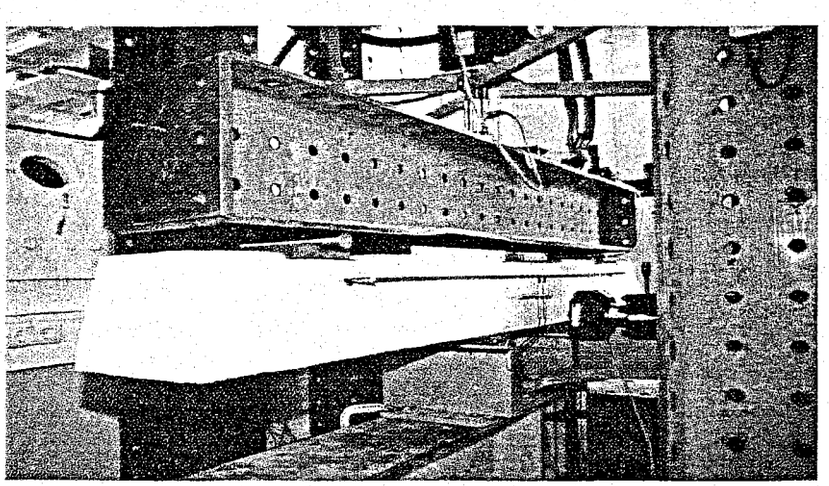


Figure 14.8: Reinforced monoblock sleeper strength test

Comparable features can be distinguished for testing the strength of sleepers as for testing rail joints. For example, the most likely failure mechanism found in monoblock concrete sleepers is cracking in the middle on top or around the housing of the fastening system. In Figure 14.8 one of these two tests is shown. Here too, national (company) standards and testing methods, in general, dominate over the European standard [30]. In this standard, sleeper strength is determined by means of forcedly cracking the sleeper by applying bending stresses in a four-point bending test. The present pre-stressed monoblock sleepers have sufficient resistance to reduce the crack width, though tests have shown that reinforced sleepers can perform equally. Tests have been conducted on reinforced sleepers suitable for light rail purposes [307]. Crack width depends on the level of tension in the reinforcing rods under loading. The tension will steadily increase under the influence of dynamic loading. Results from those tests showed that crack width could be kept within the restrictions under static and fatigue loading.

14.2.4 Stability properties

Track stability guarantees correct track positioning even under severe loading conditions. The most interesting reason why track requires sufficient stability is that it will buckle due to high longitudinal forces in the rails. To overcome this, the framework of rails and sleepers should be sufficiently rigid and embedded in a stabilised ballast bed. Rigidity in the framework is achieved by means of longitudinal and torsional resistance in the fasteners. In [29] tests are listed to determine these resistances. Longitudinal resistance in assembled tracks varies between 8 and 20 kN per rail (per set of fasteners). The torsional resistance of most of the present fasteners is between 30 and 70 kNm/rad [282] and between 100 and 250 kNm/rad per meter of an assembled track framework.

In order to conserve the fixation between rail and sleeper in ballasted track, the longitudinal resistance should be at least 12 kN. Lower or extremely low resistance values are applied in case of interaction with civil structures like bridges. Track stability should then be provided in alternative ways. Lower longitudinal resistance also applies for slab tracks in which case stability of the track framework is almost guaranteed. In [46] a minimum longitudinal resistance of 5 kN per rail (per set of fasteners) is required. In Figure 14.9 a test is shown for determining the longitudinal resistance of a specific embedded rail section in slab track (see Section 9.8), while in Figure 14.10 a test set-up is shown for standard fastening systems mounted on a sleeper.

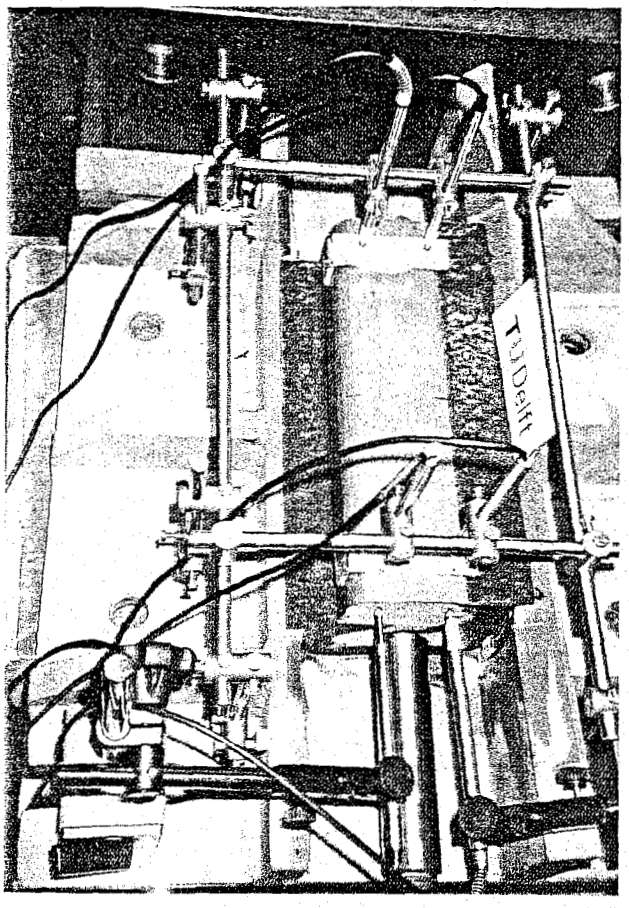


Figure 14.9: Longitudinal resistance test of an embedded rail section

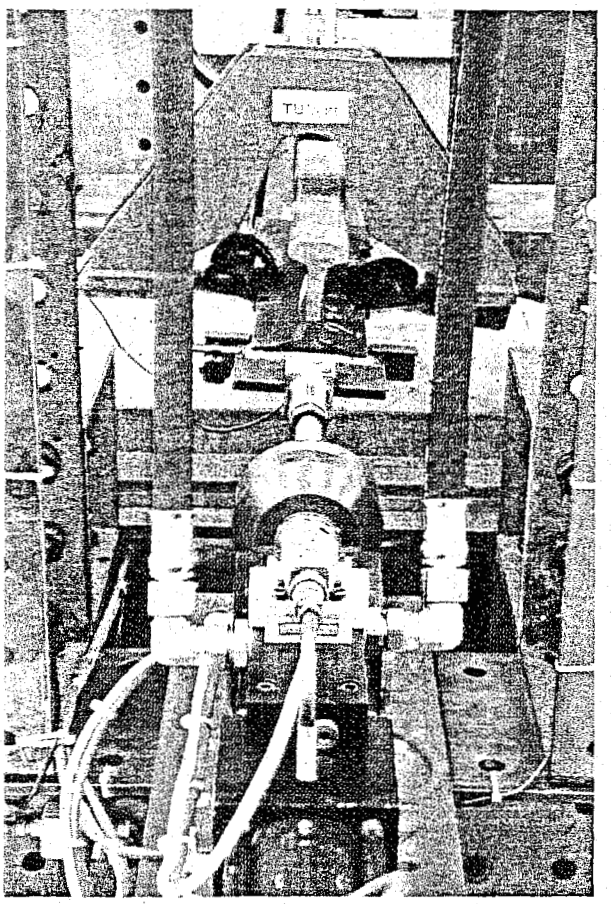


Figure 14.10: Longitudinal resistance test of a fastening system mounted on sleeper

The properties of ballast beds regarding stability have been investigated by many, both theoretically and empirically, in the field and in laboratories. In [282] the results of lateral stability of a ballast bed are reported as a function of vertical track loading and longitudinal forces. Lateral ballast resistance should be determined for track panels instead of single sleepers, as the co-operation of neighbouring sleepers is not negligible. Lateral track resistance in ballast is divided into elastic stiffness and plastic friction.

Purely elastic lateral displacement is in general limited to 2 to 5 mm. The maximum lateral reaction force of ballast per sleeper in unloaded condition is 7 to 11 kN. For displacements above the maximum elastic value, the lateral reaction forces of ballast caused by friction amount to 5 to 10 kN per sleeper. All lateral ballast behaviour depends on the ballast condition, in particular the level of compactness and consolidation, see Chapter 7. A new sleeper design can be adopted to increase track stability as well as improving the lateral resistance of the ballast bed .

14.2.5 Durability and fatigue properties

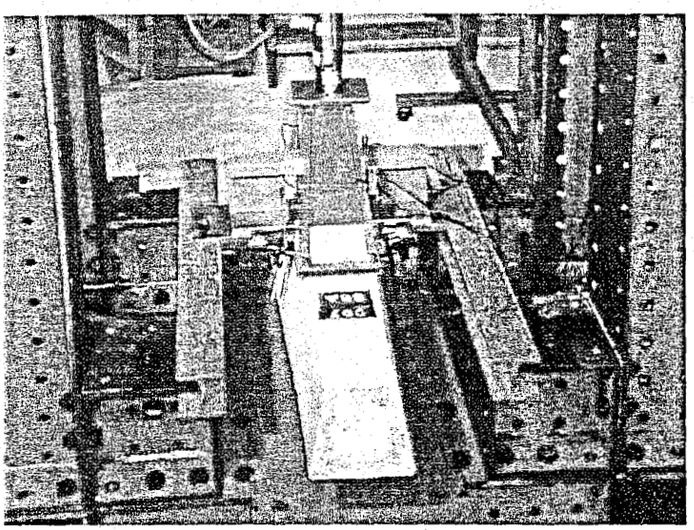


Figure 14.11: Clamping force test of fastening system on sleeper

Durability and fatigue apply to all components that are installed in track structures. In a component's lifetime millions of tonnes will pass, while track geometry and component dimensions and behaviour have to stay within limits.

Most of the components or assemblies are, therefore, tested for approval (in the laboratory and in the field) before being accepted and widely applied.

Two examples of durability and fatigue testing are presented here: testing the clamping force of fasteners and the stiffness properties of embedded rail.

The clamping force of rail fasteners can decrease due to the effects of repetitive loading, causing fatigue in the clips as well as irreversible compression of the rail pad. Repetitive loading tests as described in [29] can be carried out (Figure 14.11). In Table 14.6 the decrease in clamping force is illustrated.

Supplier	Pandrol [187]	Pandrol [187]	Vossloh [32]
Applied by	BR	BR	DB
Fastener code	e1809	e2039	skl1
After 1000 cycles [kN]	12.5	17.4	18.9
After 3e6 cycles [kN]	12.1	16.1	16.2

Table 14.6: Clamping force decrease due to repetitive loading

Considering the durability and fatigue properties of embedded rail structures, see Chapter 9, research has been performed and reported in [132] and [168]. In embedded rail structures, the viscoelastic compound in which the rail is embedded provides elasticity, fixation, and the track's stability. Therefore, durability of this material is extremely important. Durability is tested on resistance against ageing, temperatures, hydration, chemicals, and electric current.

In Table 14.7 the fatigue properties regarding track elasticity are listed for a number of typical embedded rail track structures. UIC60 and UIC54 concern main line applications, while the S41 structure is dedicated to light rail. SA42 is a special low noise rail design with future light rail application which was developed and tested in the Netherlands in 1998 [22] (see Section 9.8.3). Though the testing conditions differ (loading ranges, sample length, loading rate), an increase of static track stiffness is observed. The static stiffness values amount to asymptotic values for further repeated loading.

Rail profile	UIC60 [132]	UIC54 [58]	S41 [133]	SA42 [168]
Applied by	DB	NS	Bremen	-
Elastic compound code	VA60	VA60	VA40	VA90
Elastic strip code	100	102	102	fc6sp
Loading rate [kN/min]	-	30	-	30
Loading range [kN]	0-51	0-?	0-36	2-30
Section length [m]	0.6	0.5	0.44	0.25
Initial static stiffness [kN/mm/m']	43.5 (angular)	~52.1 (vert.)	21.8 (angular)	50.0 (vert.)
Fatigue load angle	22	22	22	22
Fatigue load level	171%	170%	133%	170 %
Number of load repetitions	1 million	1 million	1.3 million	1 million
Initial static stiffness [kN/mm/m']	45.2 (vert.)		22.9 (angular)	62.7 (vert.)
Final static stiffness [kN/mm/m']	53.1 (vert.)		21.8 (angular)	71.8 (vert.)

Table 14.7: Static stiffness of different embedded rail sections due to fatigue loading

The standards for fatigue and durability testing of discrete fastening systems and embedded rail differ per application. The loading condition – in particular the lateral loading on the rail - very much influences the durability and stability of the track. Gauge widening is a common but unsafe observation in fatigue tests. Moreover, the performance of components has to be checked by means of visual inspection of cracks and correct in-place positioning.

For the composition of a realistic fatigue and durability test program, it is further important to consider the loading time and amplitudes carefully. Some systems contain materials with visco-elastic, non-linear, and temperature sensitive properties. This can make it necessary to perform tests under conditioned circumstances, such as shown in Figure 14.12. The loading frequencies of the fatigue tests will never be increased to such an extent that unrealistic system behaviour will start to dominate.

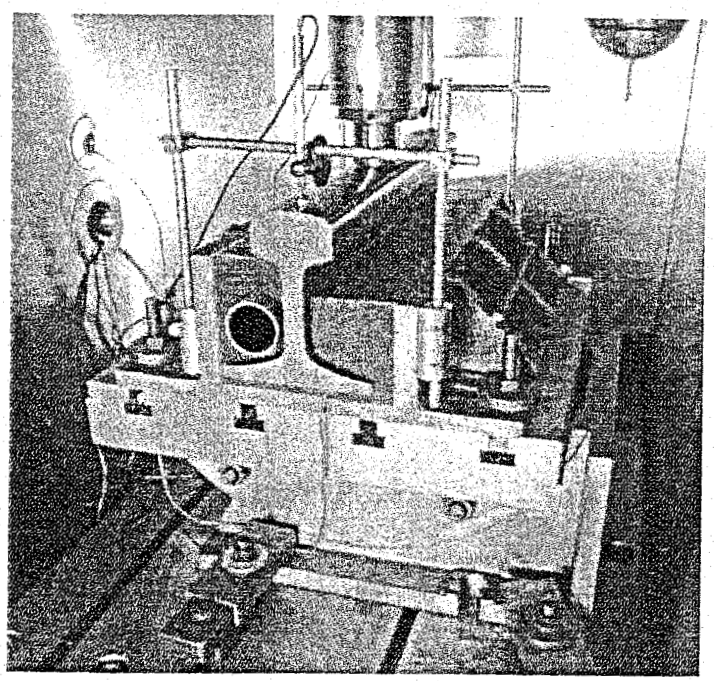


Figure 14.12: Testing of embedded rail sample under conditioned circumstances

14.2.6 Specific component properties

In many countries, the railway track is part of the power supply and signalling system. This implies that the track needs to meet specific electrical requirements, in particular the fasteners and sleepers. Plastic insulators, pads, and the housings of clips and screws assist in separating both rails installed on one sleeper. Insulated rail welds play an important role in signalling systems based on block possessions by trains.

Assessing electrical properties requires testing long-term durability, for example resistance against rust, and measuring rail to rail insulation values. Spraying salt water on fasteners and sleepers for several days or weeks is a commonly applied test. The isolation values still need to be met and the replacement of components still has to be possible.

Assessing acoustics and vibrations is becoming more and more an established part of railway engineering. Dynamic stiffness and damping, both on component and on track level, are tested for high frequency vibrations and absorption properties at different levels of loading. Though extremely important for environmental issues, design guidelines and acceptance standards hardly exist regarding these properties.

Individual track support properties can be of use in estimating noise and vibration radiation conjunction with advanced calculation models, like TWINS [265]. For acceptance, however, most dynamic properties are still evaluated by comparing them to the properties of existing components, and by recording the performance of new and existing track structures. In Figure 14.13 the noise and vibration measurement results of some (super) elastic track support systems are shown with reductions up to several decibels.

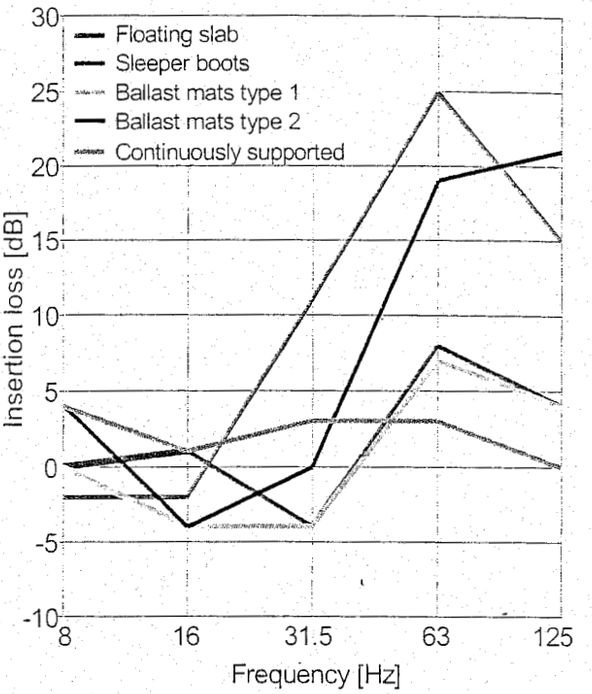


Figure 14.13: Influence of several (super) elastic track support systems on vibration reduction

Not only dynamic properties, but also shapes and dimensions can be altered to improve performance. In [137] the influence of the size of an embedded rail is investigated regarding noise and vibration properties. This has led to the development of the low-noise rail as reported in [22].

If the results are insufficient, noise reduction can be achieved by means of general solutions such as slab track covering tiles, see Figure 14.14 [286], or by a very porous asphalt layer (Figure 14.15). More about the field of structural testing and acceptance will be discussed in the next section.

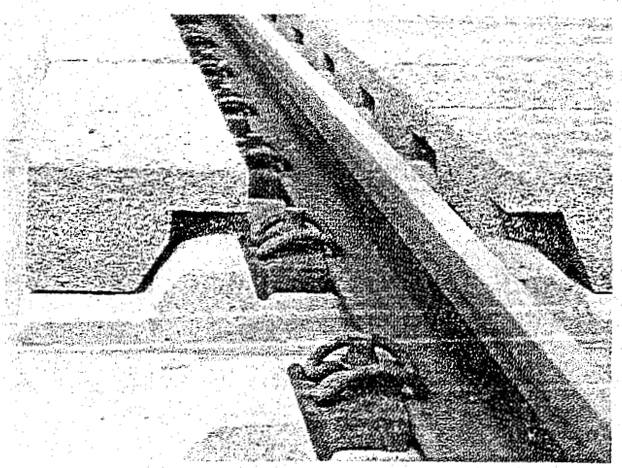


Figure 14.14: Slab track covering tiles

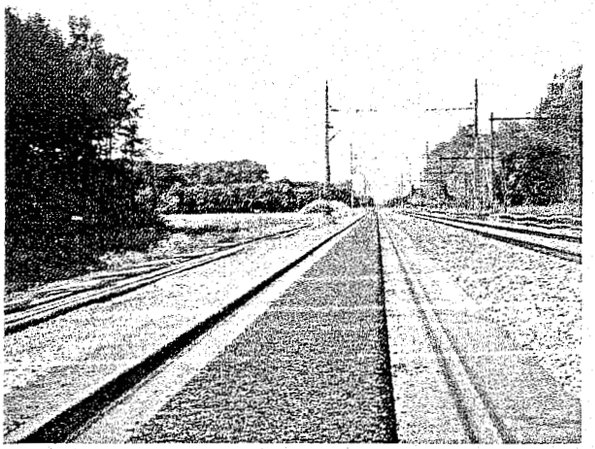


Figure 14.15: Asphalt layer covering slab track structure with embedded rails

14.3 Structural testing and acceptance

The performance of track structures is not only related to the performance of the components. The connections between these two influence the properties of the components and contribute to the properties of the track. This makes it necessary to test and assess track structures as well.

The by far most well known type of track structure test is the performance test: the installed track structure is monitored for visual structural damage and geometrical imperfections. Moreover, this test yields a realistic type of loading, but raises problems regarding safety, duration, and intensive inspections. Acceptance standards are hardly available and far from uniform. In this chapter, some examples of structural testing and acceptance are outlined which can contribute to assessing the quality and the safety of track structures.

Quality and safety of the track structure, which can only be established by means of testing and determining indicative parameters, is based on:

- Average, extremes, and deviation in the properties;
- Tendencies in the average properties over time.

Directly from the moment of constructing a track section deviation in properties exists, while the use of the track causes additional deviation in properties along the track section as well as over time. The indicative parameters should fall within limits which are clearly formulated. In principal, safety has more stringent requirements for acceptance than quality, although these requirements do not apply to every single item the requirements regarding quality apply to. It is for several reasons acceptable that limit values vary.

In this section some examples of testing and acceptance that particularly apply to structural testing will be demonstrated. First, testing for noise and vibration – which is typically related to place – is introduced, followed by passenger comfort (ride quality) – which is related to variance over place. Secondly, the properties of the track in these situations are points of interest as well as, thirdly, the interaction of the track with its surroundings and with the vehicle is. It is, therefore, important to determine the properties of the track in an unbiased way.

14.3.1 Noise and vibration testing of track structures

A common problem that arises from railways is hindrance from noise and vibration (Chapter 15). The abatement of this hindrance has been considered seriously since trains entered urban areas and people were forced to live closer to railway lines. In the beginning, the abatement took place by means of trial-and-error and legislation was not so strict as it is nowadays either.

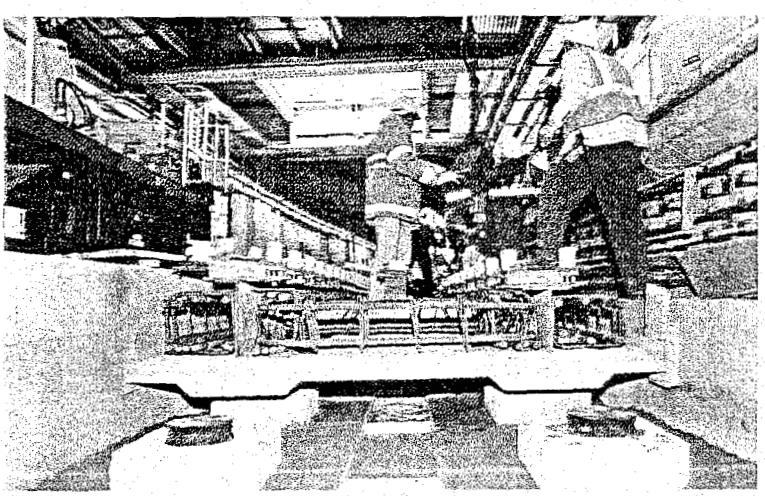


Figure 14.16: Installation of floating slab (photo courtesy of Tiflex/EC Dixon)

Thanks to experience gathered over the years, improved testing facilities, and computer models predicting the noise and vibration properties of tracks, annoyance preventive measures have more and more become part of track construction and track renewal plans. From time to time they are even the main reason. In all those cases, alternation of track structures and properties is suggested to reduce the amount of hindrance.

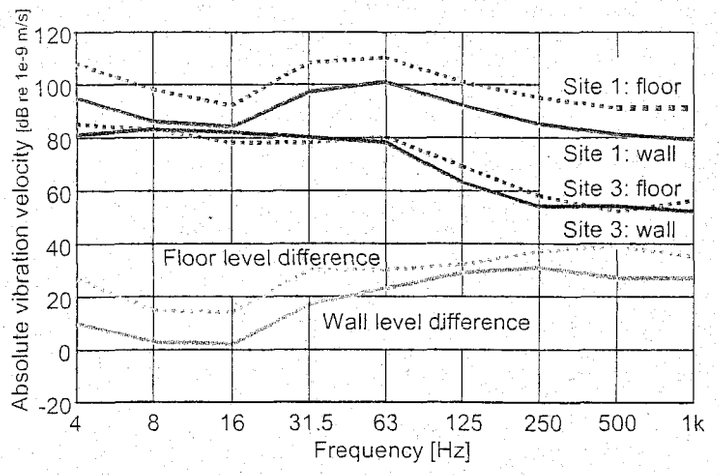


Figure 14.17: Vibration levels of tunnel floor and wall in track mounted directly on tunnel floor (site 1) or on floating slab (site 3)

In order to reduce transmission of vibration energy, damped mass-spring systems are often applied. In [228] and [241] the application of elastically supported slab track (floating slab) is considered. Steel helical springs, elastomere bearers, and full-plane layers (Figure 14.16) separate these concrete slabs and the track from the concrete tunnel floor or bridge deck.

The effectiveness of these measures is often expressed in relation to the vibration levels of initial plain structures. Well-defined points, such as tunnel ceilings or walls, bridge piles, building floors, or walls, are then essential for comparison. Figure 14.17 shows the reducing influence of elastomere supports underneath concrete slabs of 30 m length on the floor and wall vibration levels of a tunnel in Rotterdam (Netherlands) due to metro traffic. Figure 14.13 already showed the results of comparable measurements carried out on several track structures in Austria [228] and Germany [227] and [228]. Here, vibration levels of floating slab, ballast mats, and undersleeper pads are compared to plain ballasted track with comparable traffic.

Floating slabs show higher effectiveness in reducing transmitted vibration energy than ballast mats and undersleeper pads. This is a consequence of the new lowest vertical natural frequency of the system: 4 to 20 Hz for floating slabs, 20 to 60 Hz for ballast mats, and 40 to 100 Hz for undersleeper pads. Additional investigations in some metro systems further show that reduction can be achieved within the rail support systems. The new lowest vertical natural frequencies are then 80 to 140 Hz for block support systems and 120 to 200 Hz for baseplate pad systems.

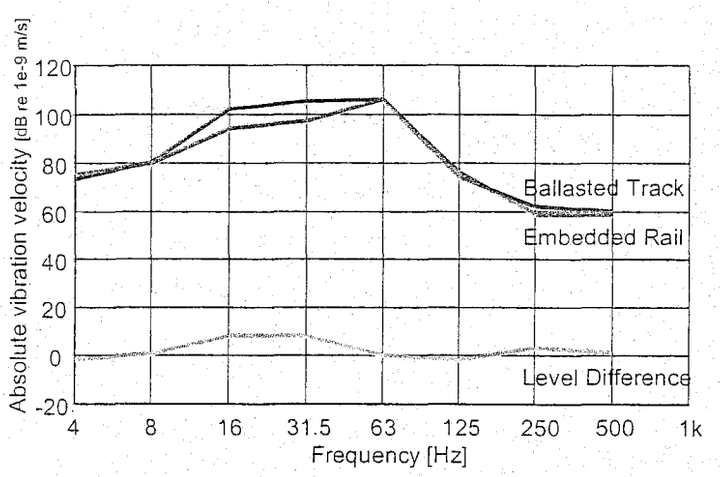


Figure 14.18: Vibration levels of ballasted and slab track recorded at 8m from track centre

In [155], new continuous supports have replaced the existing discrete supports in a metro line in Germany. The effectiveness of this was investigated in order to reduce the vibration energy at the support passing frequency, i.e. 30 Hz when riding at approximately 70 km/h on 0.65m spaced supports. By spreading out track elasticity, the recorded levels due to the forced vibration around 30 Hz are lower. This indicates that the load distribution property of the adapted track structure is rather effective. Thus, continuously supported track lowers the excitation mechanism, so that a reduction of the transmitted energy is achieved in a wider frequency range than only at the support passing frequency.

In [48], the measurement results of a continuously supported embedded rail installed in prefab concrete slabs on a roadbed are reported. The results show the differences in vibration energy transmission between a ballasted track and an embedded rail structure under passenger trains at a track section of Netherlands Railways, recorded at 8m perpendicular to the track (Figure 14.18).

In the beginning of this section we saw that the application of floating slabs, ballast mats, under sleeper pads, baseplate pads, and block supports contributed to lowering the transmitted vibration energy. These applications also introduced and shifted the lowest vertical natural frequency of the track. The continuously supported track structure eliminates many of the higher order natural frequencies in discretely supported track, and easily dominates the lowest vertical natural frequency of discretely supported track as described by the single mass spring system in Chapter 6.

Over the years it has become clear that a decisive design principle to exactly assess the effectiveness of vibration and noise reduction is difficult to construct. Local circumstances (soil properties, track alignment, track and vehicle quality, distance to buildings) make it very complicated to judge improvements and measures at the track in advance. Experiences in comparable situations and professional skill will help to certify that the testing and recording of track during service falls within the limits that are listed for human and structural perception of vibration and noise (e.g.[51], [55]).

14.3.2 Passenger comfort and ride quality

Assessing track quality from the point of view of vehicles and – more important for revenues and safety – of passengers is very important. Track quality should, therefore, be divided in two parts: namely, considering the vehicle and the passenger.

For the passenger, quality assessment is based on comfortable transport which is achieved by means of limiting the low frequency vibrations which passengers are exposed to. The limitations are collected in [93] and [134] and concern 3D average, maximum, and standard deviation values. The suspension systems of the vehicle, the track geometry, and the design speed determine the comfort levels.

In [283], a method is presented in which typical geometrical track elements contribute to improving the comfort levels of passengers in high-speed line vehicles. Especially the transition curves in the horizontal plane are improved and connected in such a way, as introduced by [236] and shown in Figure 14.19, that jerk and impact loads are reduced. The accelerations to which the passengers are exposed are significantly lower.

Track geometry can be measured and controlled with devices and equipment such as presented in Chapter 16. On the other hand, specific testing equipment for passenger comfort exists and can be implemented in any vehicle so that (relatively) weak spots in the track network can be identified from the passenger's perception. This can be part of a permanent testing and monitoring program to assure quality, and can, in principle, be used on any type of railway network.

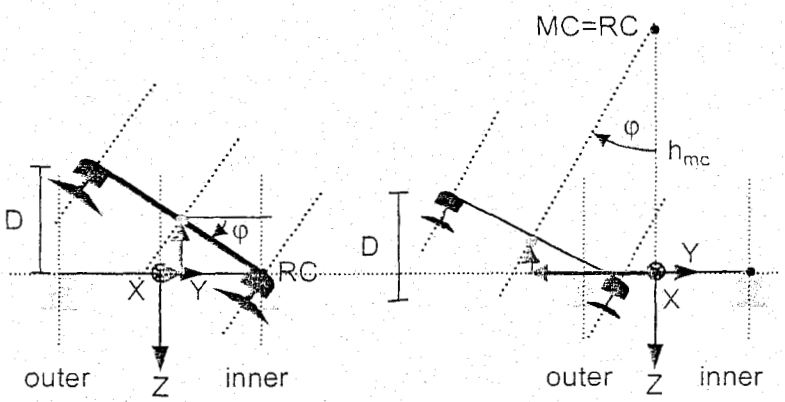


Figure 14.19: Modified track geometry in curves

Analysing vibration levels recorded at certain positions within the vehicle (e.g. at bogies or axle boxes) can identify weak spots in track geometry or in track structures. This is important as high vibration and acceleration levels can be harmful to the vehicle. Just like passenger comfort, ride quality for vehicles is also measured and categorised with acceleration values for average, maximum, and standard deviation.

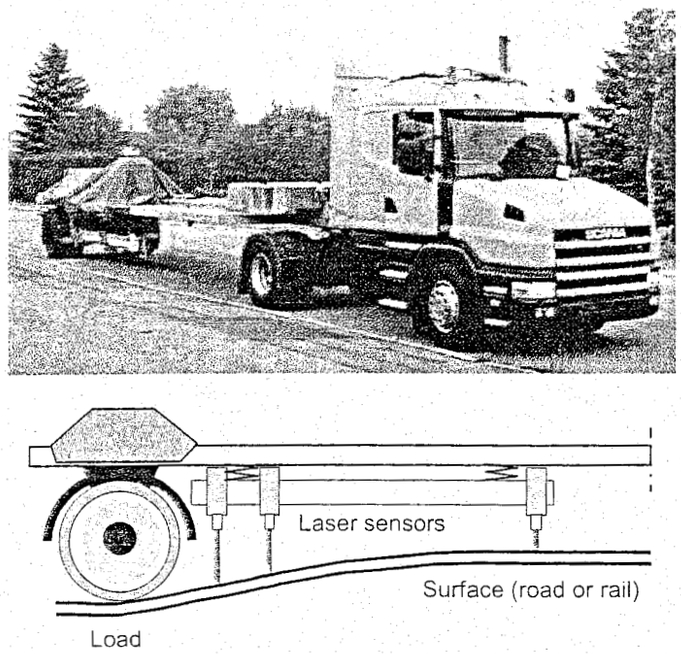


Figure 14.20: High-speed deflectograph for road application

By paying closer attention to e.g vibration recordings, various types of track defects in geometry and structure can be detected. Very commonly observed problems in track structures are the loss of track support (bouncing sleepers) and the stiffening of supports (compacted and pulverised ballast) and transitions close to level-crossings and abutments. The device which is capable of detecting these kinds of problems in track support stiffness contributes to listing maintenance spots, determining the precise type of maintenance, and, thus, to saving on maintenance costs.

One example of such a device is given in [232]. This device measures the velocity of the rail deflection close to an instrumented wheel. After analysing data, track properties can be calculated on nearly single rail support level. The so-called high-speed deflectograph will be able to perform such testing on railway tracks as if has been successfully introduced on road structures (see Figure 14.20).

14.3.3 Dynamic properties of track structures

There are numerous components that contribute to the performance of railway track. The stiffness of track structures can vary in time and place, due to climate, construction, maintenance procedures, and tonnage, and can give rise to impact loading and increased noise and vibration radiation near the track. This requires a closer observation of the dynamic track properties that are associated with these problems.

For track design as well as for track performance, control of dynamic track properties is important. Chapter 6 has already presented several analytic and finite element models, some including vehicles. In principle, the dynamic properties of vehicles and track can be considered separately in transfer functions. For track structures, the transfer function can be expressed as a relation between (wheel) force input and (rail) deflection output, such as indicated for a single beam model by equation (6.39).

$$H_{F \to r}(f) = \frac{Y_r(f)}{Y_F(f)} \tag{14.1}$$

For vehicles, comparable transfer functions can be composed which relate (wheel) force input to (passenger or vehicle) body acceleration.

$$H_{F \to \ddot{b}}(f) = \frac{Y_{\ddot{b}}(f)}{Y_F(f)} \tag{14.2}$$

When considering the force input in both above-mentioned equations equally, the transfer function from rail deflection to body acceleration shows up easily. This proves that dynamic track properties do influence the vibration behaviour of the vehicle on it.

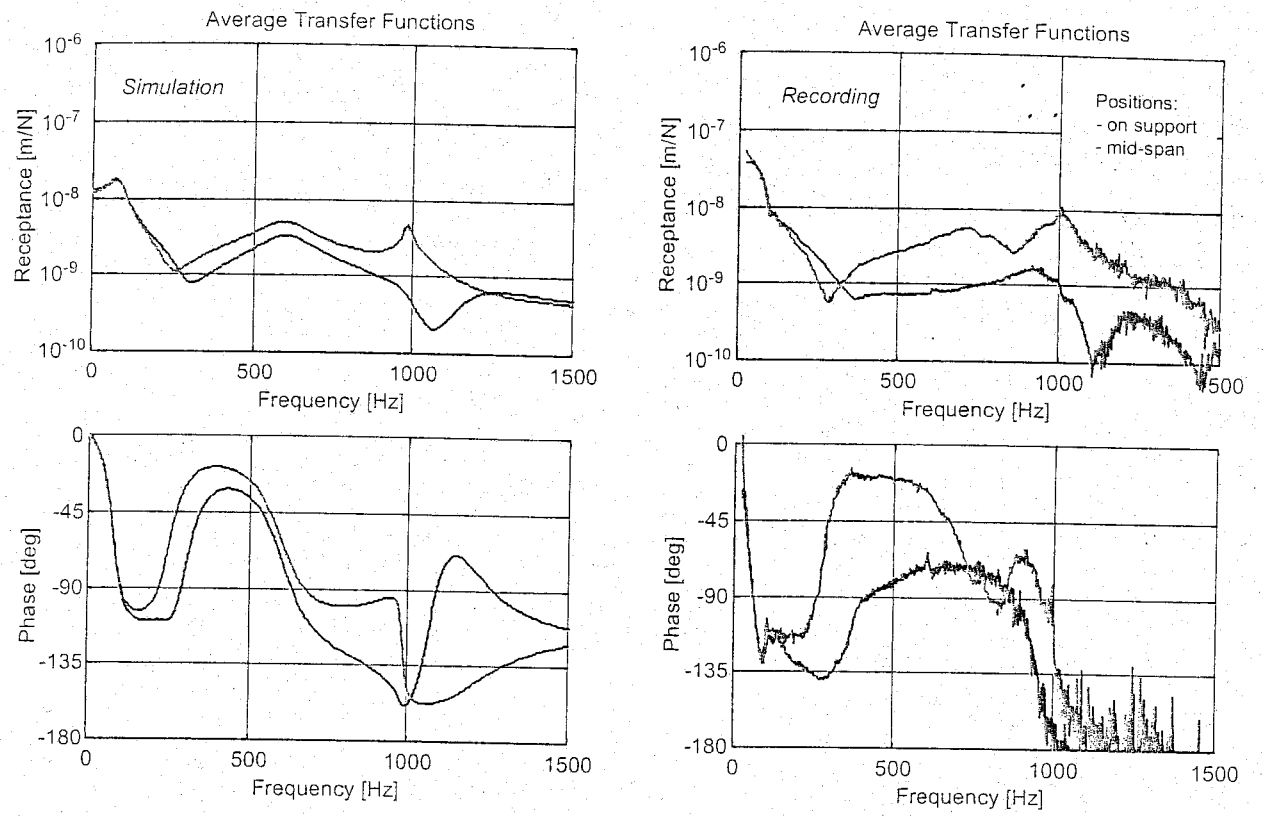


Figure 14.26: (b) Simulated and recorded transfer functions of ballasted track structure.

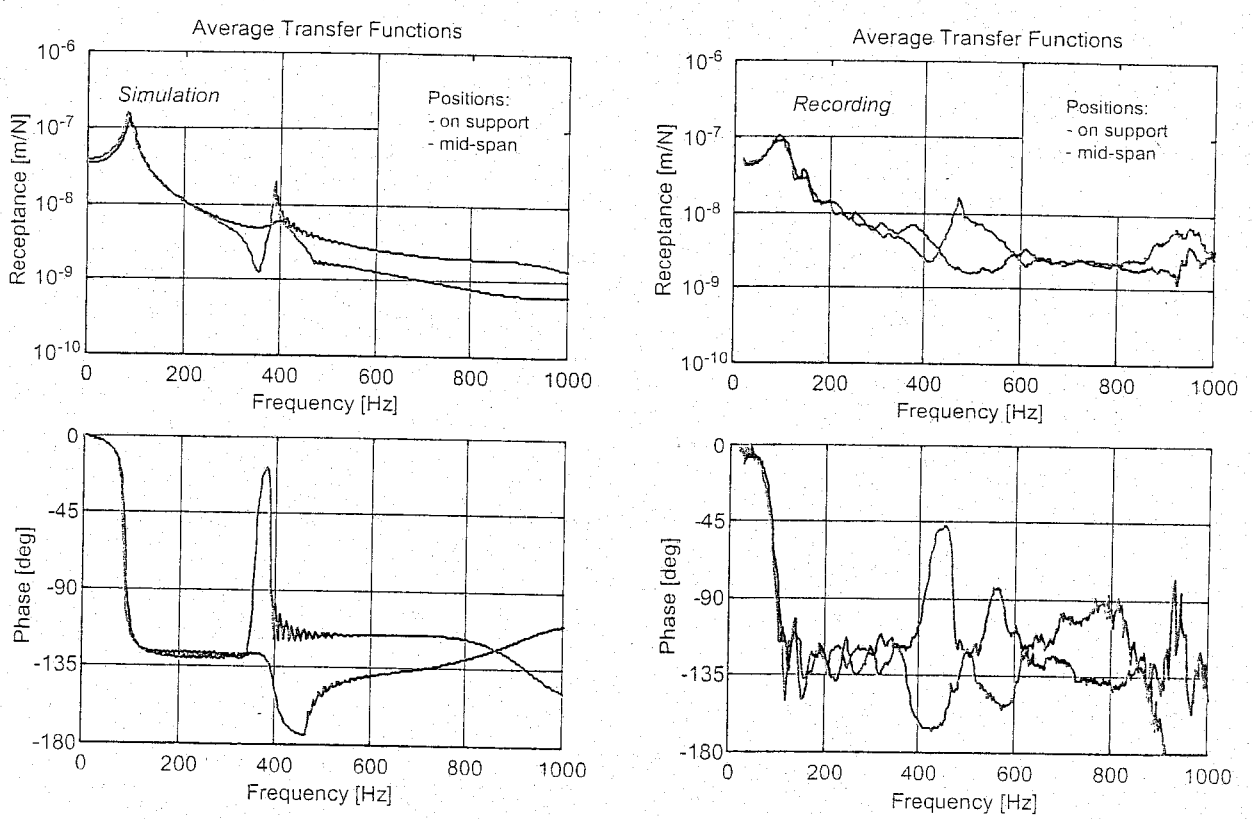


Figure 14.26: (c) Simulated and recorded transfer functions of baseplate slab track structure

15 NOISE AND VIBRATION

15.1 Introduction

In recent years, rail transport systems have increasingly received complaints from people living alongside lines and above underground lines. The disturbance is usually caused by the direct emission of noise or vibration from the railway, but sometimes noise in buildings is produced by the walls vibrating which is referred to as re-radiated noise.

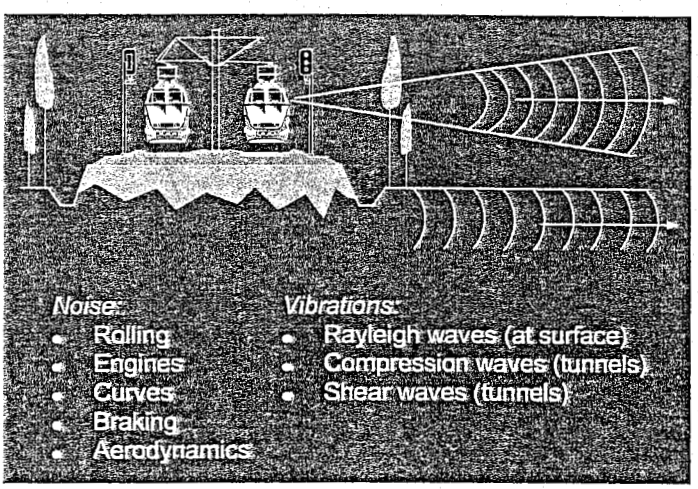


Figure 15.1: Noise and vibration radiation principle

Vibrations and structure-borne noise mainly occur at lower frequencies below 50 Hz. At higher frequencies these vibrations attenuate increasingly rapidly. The energy at higher frequencies is radiated as noise mainly through the wheels and the rails. Roughly speaking, vibrations and structure-borne noise occur in the frequency range 0 - 100 Hz and noise between 30 - 2000 Hz. The principle of noise and vibration radiation is illustrated in Figure 15.1.

15.2 Some definitions

As the dynamic range of the spectral values is rather large, they are mostly represented on a logarithmic scale expressed in dB, according to:

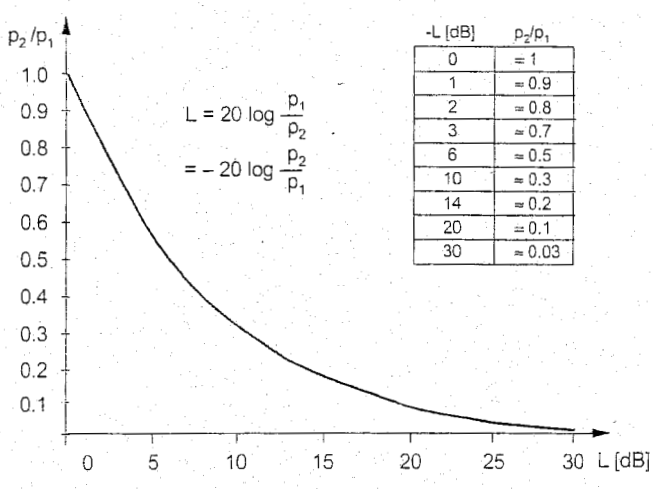
$$L[dB] = 20\log\frac{p_1}{p_2} \qquad (15.1)$$

Figure 15.2 shows the relationship between the linear scale and the dB scale.

Power spectral values are mostly calculated as root mean square (rms), or effective value, per 1/3 octave band. In the low frequencies associated with vibrations, accelerations are often integrated to produce velocities.

When expressing them as dB values a reference value should be given. Noise values are normally expressed relatively to a reference value 2 10^{-5} N/m^2.

The human perception of noise is characterised by the filter shown in Figure 15.3. This A-filter, as it is referred to, removes almost all contributions below 200 Hz. Noise levels are normally presented after having been A-weighted. This is indicated in dB(A). Depending on the application, spectra can either be presented as A-weighted or not.



-L [dB]	p_2/p_1
0	= 1
1	≈ 0.9
2	≈ 0.8
3	≈ 0.7
6	≈ 0.5
10	≈ 0.3
14	≈ 0.2
20	≈ 0.1
30	≈ 0.03

Figure 15.2: Relationship between dB scale and linear scale

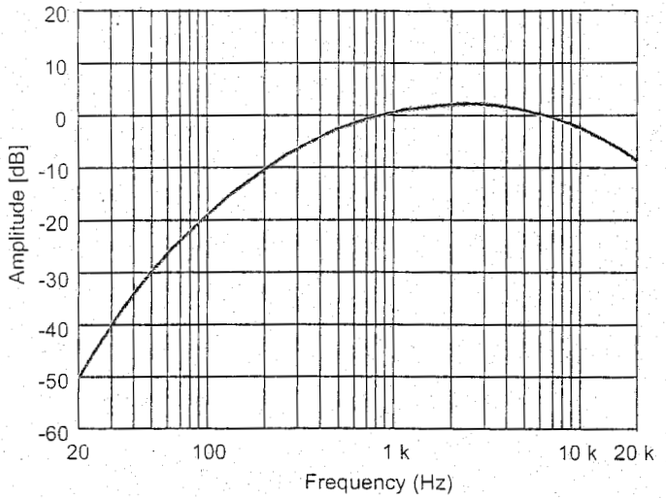


Figure 15.3: Transfer function of A-weighting filter

The equivalent noise energy level L_{eq}, expressed as dB(A), is found by integrating the squared noise pressure signal as follows:

$$L_{eq} = 10\log\frac{1}{T}\int_0^T\left[\frac{p_A(t))}{p_0}\right]^2 dt \qquad (15.2)$$

in which $p_0 = 2\ 10^{-5}$ N/m^2, $p_A(t)$ is the A-weighted noise signal and T is the integration time.

Basically different from the root mean square value in (15.2) is the sound exposure level SEL which represents the equivalent energy of a phenomenon during a reference time $T_r = 1$ s, according to:

$$SEL = 10\log\frac{1}{T_r}\int_0^T\left[\frac{p_A(t))}{p_0}\right]^2 dt \qquad (15.3)$$

The difference between both expressions is the reference time T_r over which the energy is averaged. In (15.2) $T_r = T$, whereas in (15.3) $T_r = 1$ s.

15.3 Ground vibrations

15.3.1 Introduction

Vibrations are caused by the large forces between wheels and rails. These forces fluctuate in response to wheel and rail roughness over a wide range of frequencies. In addition, the distributions of axle loads in a train also produces a force excitation as it passes a fixed point. The latter effect leads to excitation at frequencies which correspond to the vehicle passing frequency and its harmonics, whereas forces due to the wheel or rail roughness have a periodicity determined by the wavelength of the roughness and the vehicle running speed.

The combined vehicle/track system is complex and has many natural frequencies. When one of the excitation frequencies corresponds to a natural frequency of the system there is a particularly strong vibration. A strong vibration also ensues if an excitation frequency or natural frequency of the system corresponds to the sleeper passing frequency. Whereas natural frequencies of the system are independent of running speed, the sleeper-passing frequency increases in proportion to the running speed. As the speed rises, a coincidence of frequencies may occur and the vibrations will reach a peak and then diminish. Thus, vibrations do not increase steadily with speed and a speed reduction may sometimes make the vibrations worse.

Vibrations propagate from the track through the ground by means of compression waves, shear waves and surface (Rayleigh) waves. For each type of wave the energy becomes less as distance from the source increases. This is due to geometric dispersion and energy absorption in the ground. The lowest frequencies are the least damped.

The main lines in the open radiate vibrations by means of Rayleigh waves according to the mechanism sketched in Figure 15.4.

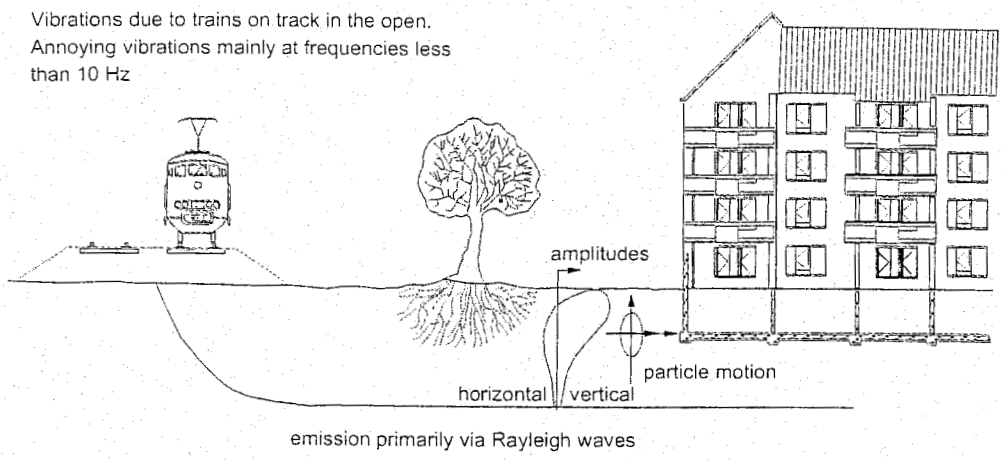


Figure 15.4: Vibration propagation by underground railways

In the case of underground railways the vibration energy is mainly transmitted in the frequency band 30 - 150 Hz through compression and shear waves as indicated in Figure 15.5. These vibrations, which most commonly have an energy peak in the region of 50 Hz, are noticeable as a rumbling noise due to re-radiated sound from walls and ceilings which typically vibrate at 5 - 25 Hz.

Buildings near main line railways are usually at such a distance that the vibrations due to compression and shear waves fade away while the surface waves remain.

The most critical conditions occur when the predominant frequency of ground vibration coincides with the natural frequency of a building. Although the fundamental frequency for lateral vibrations is in the order of 1 -10 Hz, components may have higher natural frequencies.

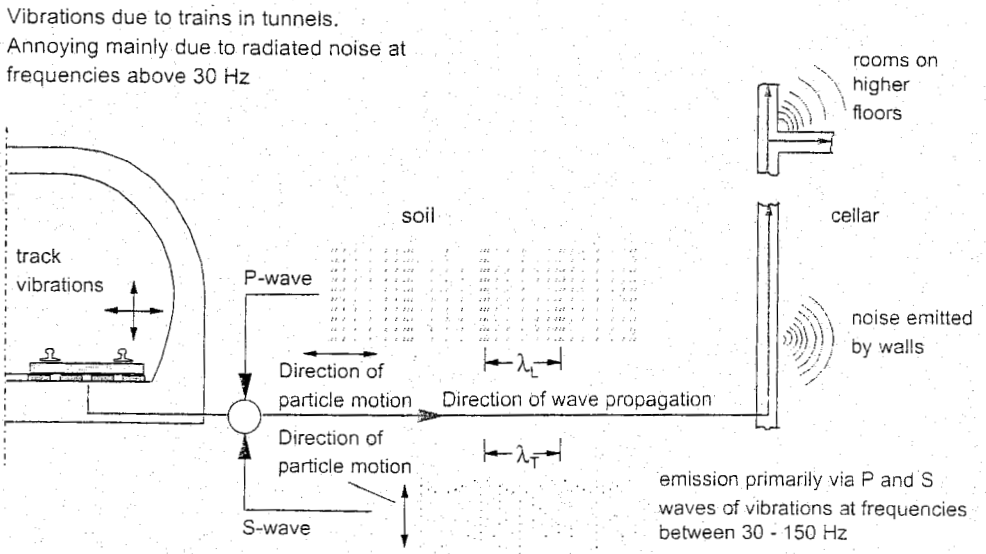


Figure 15.5: Vibration propagation by underground railways

15.3.2 Wave propagation in soils

Vibration energy is transmitted through the soil as body and surface waves. A wave is characterised by the particle motion u, the wave propagation speed c and the wavelength λ. A plane wave propagating in the x-direction at a speed c can be described by the expression:

$$u = Ae^{i\frac{2\pi}{\lambda}(x-ct)} \tag{15.4}$$

For a fundamental discussion of wave theory please refer to literature.

If the load excitation is of a steady-state character, the frequency of the particle motion corresponds to the excitation frequency f. The wavelength λ is determined by the excitation frequency f and the wave propagation speed c according to the relationship:

$$f = \frac{c}{\lambda} \tag{15.5}$$

In the case of transient loads, the frequency f will correspond to the frequency at which the layered nature of the soil structure favours vibration transmission.

The particle motion of compression waves is in the direction of the wave propagation, whereas that of shear waves is in a plane perpendicular to it. In the case of Rayleigh waves occurring in the vicinity of a free surface, soil particles follow a retrograde elliptic motion (counter-clockwise) as indicated in Figure 15.6. At the surface the vertical displacement is about 1.5 times the horizontal displacement.

Rayleigh waves are often compared to water waves. There is one fundamental difference though, for in the case of waves travelling to the right, the water particles move clockwise (prograde) in deep water according to a circular motion, while in shallow water the particles move according to an elliptic motion in which the horizontal displacement is greater than the vertical displacement. With Rayleigh waves and water waves the particles follow a closed path. Therefore, this motion is called orbital motion. It should be emphasized that waves do not transport any mass, just energy.

For shear waves (or transverse waves) the propagation speed equals:

$$c_T = \sqrt{\frac{G}{\rho}} \tag{15.6}$$

in which G is the shear modulus and ρ is the density of the soil.

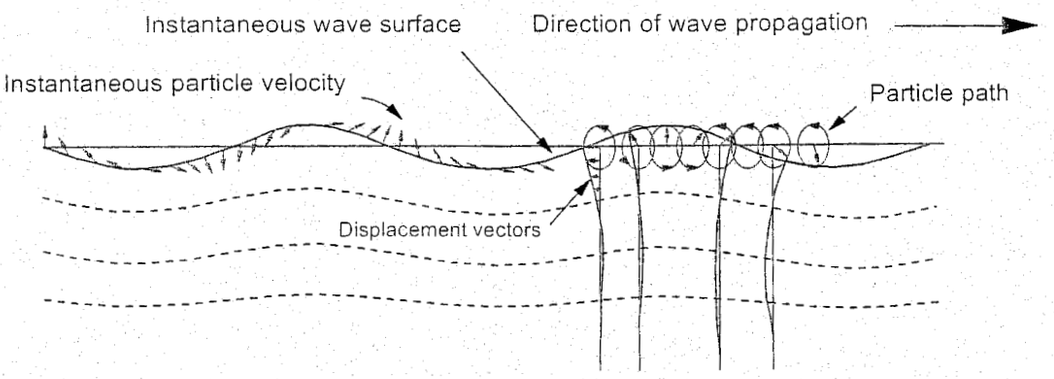


Figure 15.6: Schematic drawing of Raleigh surface waves

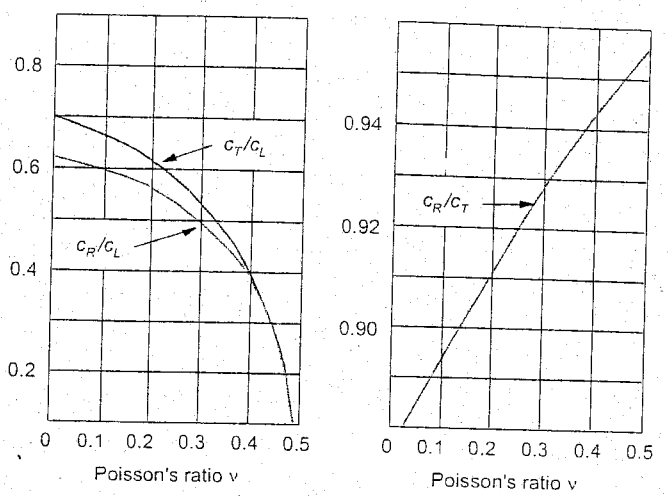


Figure 15.7: Wave speed dependency on Poisson's ratio

Ratios c_R/c_L, c_R/c_T, and c_T/c_L, where c_R = Raleigh wave speed, c_l = longitudinal wave speed, c_R = transverse wave speed

Figure 15.7 shows the longitudinal wave speed c_L and the Rayleigh wave speed c_R in relation to the transverse wave speed c_T as a function of Poisson's ratio ν. Rayleigh waves are just slightly slower than shear waves, while compression waves can have significantly higher velocities.

Figure 15.8 shows the basic features of the wave fields generated by a vertically vibrating source at the surface of a half space. The distance to each wave front is drawn in proportion to the wave velocity for a medium with ν = 0.25.

The major part of the vibration is transmitted as surface waves (67%), while shear waves and compression waves account for 26% and 7% respectively. Surface waves are, therefore, usually of great practical importance. The vibration energy gradually decreases with increasing distance from the source because the energy spreads over an increasingly large area. In addition to this geometrical damping, vibration energy is dissipated as a result of material damping. The amplitude of Rayleigh waves decreases with distance according to:

$$A_2 = A_1 \sqrt{\frac{R_1}{R_2}} e^{-\alpha(R_2 - R_1)} \qquad (15.7)$$

in which A_1 and A_2 are the vibration amplitudes at two points located at distances of R_1 and R_2 respectively from the vibration source. In this so-called Barkan formula the absorption coefficient α is a measure of the material damping. The absorption coefficient depends mainly on the soil type and

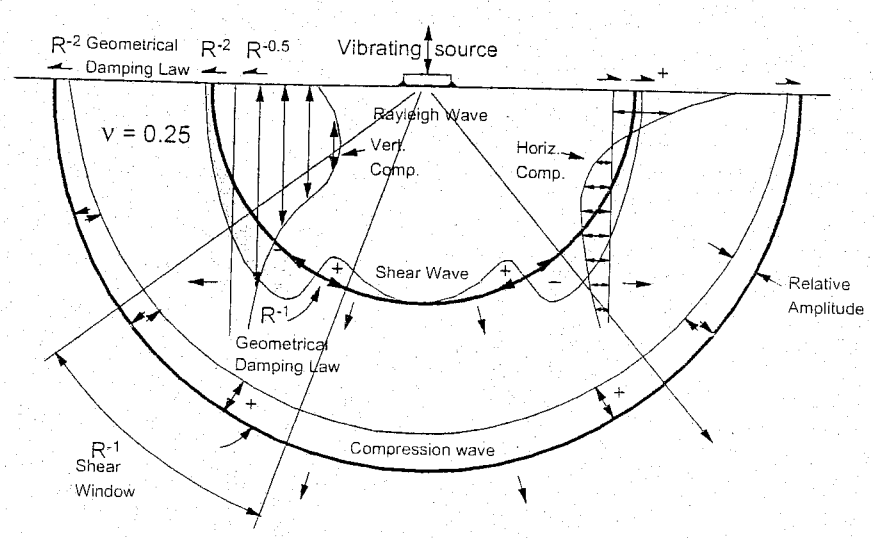


Figure 15.8: Displacement waves in an elastic half-space

the degree of water saturation and varies from 0.01 - 0.1 m^{-1} for saturated loose soils and dense soils respectively. The amplitudes of body waves attenuate much more rapidly, as the geometrical damping on the interior is proportional to R^{-1}, and along the surface is equal to R^{-2}.

Figure 15.9 shows horizontal and vertical vibration amplitude as a function of depth for Rayleigh waves. Obviously, these waves attenuate rapidly with depth. The energy propagates almost completely through a layer extending down to a depth of one wavelength.

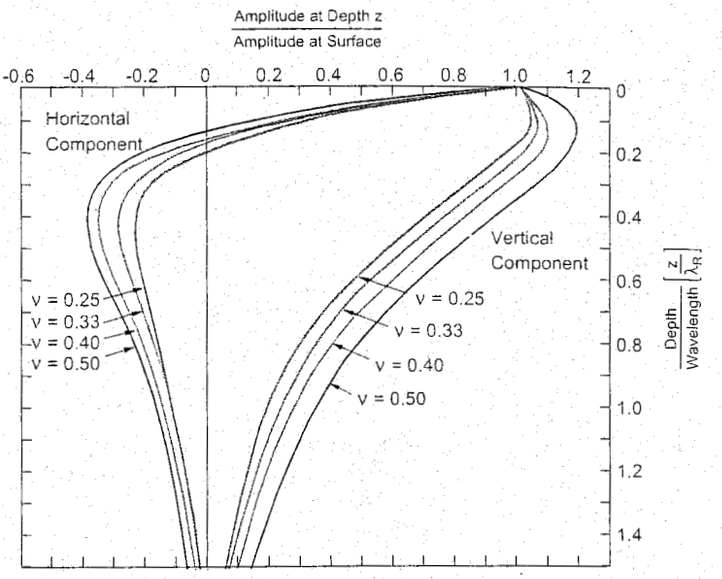


Figure 15.9: Displacement versus depth for Raleigh waves

The generation of waves during passage of a train can be considered to be a more or less steady-state process. The energy radiated through the waves depends on the transfer function between train load and track displacement.
This frequency-related function depends on the track and soil properties. The track models discussed in Chapter 6, or the approximate solutions presented in [235], can be used to determine these frequency response characteristics.

For most railway vibration problems on plain tracks in the open, the predominant frequencies of the load spectra are of the order of 5 to 50 Hz. As Rayleigh wave velocities vary between approximately 50 and 250 m/s, the wavelength of Rayleigh waves is expected to be in a range from 5 to 50 m.

In natural soil deposits, which are often stratified, resonance effects can occur. The lowest natural frequency f of a homogeneous soil with a thickness h can be determined from:

$$f_n = \frac{c_T}{4h} \tag{15.8}$$

in which c_T is the shear wave propagation velocity. For example, a clay layer of 5 m thickness plus a shear wave velocity of 100 m/s has a resonance frequency of 5 Hz. One of the constraints regarding high speeds is that the running speed should be well below the wave propagation velocity to avoid undesirable interference phenomena

An example of calculating critical train speed effects using the program SPOOR has already been discussed in Chapter 6.

15.3.3 Human perception

The human body is susceptible to the following vibration frequencies:

0.1 - 0.2 Hz: resonance of the organ of balance, resulting in phenomena characteristic of seasickness;
4 - 8 Hz: resonance of the contents of abdominal cavity and thoracic cavity;
30 - 80 Hz: resonance of eyes in the eye sockets, resulting in loss of focus.

The audibility limit lies at a frequency of approximately 20 Hz [7].

Figure 15.10 shows the bounds between vibration and structure-borne noise perception. The human body is able to perceive vibrations with a displacement amplitude of 0.001 mm; fingertips are about 50 times more sensitive.

It is generally assumed that amplitudes of 0.2 mm at 5 Hz do not cause any damage to buildings. As this value is about 200 times the perception limit, it is not so incredible that such vibrations lead to complaints.

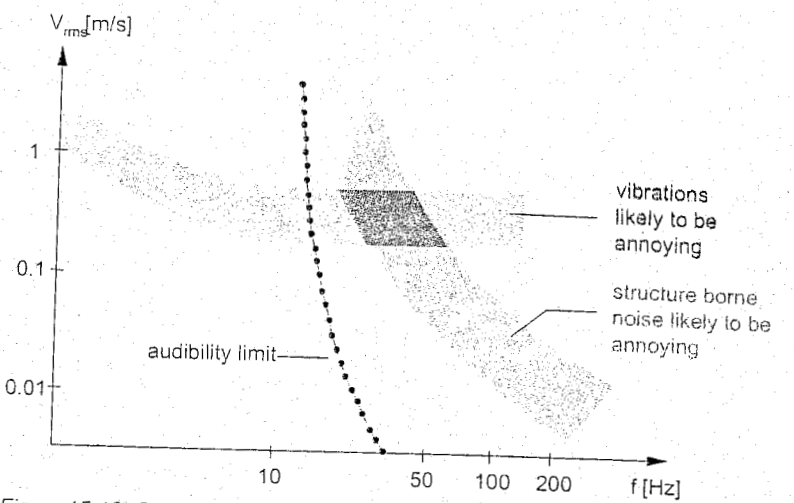


Figure 15.10: Global classification of annoyance due to vibrations and structure-borne noise

Vibrations can be expressed in terms of displacement, velocity, or acceleration. The common units are velocity in mm/s.

Figure 15.11 shows an example from the German DIN standard 4150, with curves representing equal perception of vibrations by human beings. Identical curves are used in ISO standard 2631, VDI standard 2057 and BSI standard 6472.

The equal perception value EP is defined as:

$$EP = \frac{0.18f}{\sqrt{1+\left[\frac{f}{5.6}\right]^2}} v_{rms}(f) \quad (15.9)$$

Furthermore:

$$v_{peak} = \sqrt{2}\, v_{rms} \quad (15.10)$$

Displacement d, velocity v and acceleration a are related as follows:

$$a = 2\pi f v = (2\pi f)^2 d \quad (15.11)$$

When interpreting the EP curves, one should take into account the fact that the essential factor in assessing vibrations is the force: acceleration times effective mass, or velocity times the integrated effective mass. The effective mass decreases with frequency and tends towards zero. Once integrated, the effective mass, which more or less reflects the EP curves, tends towards a constant value.

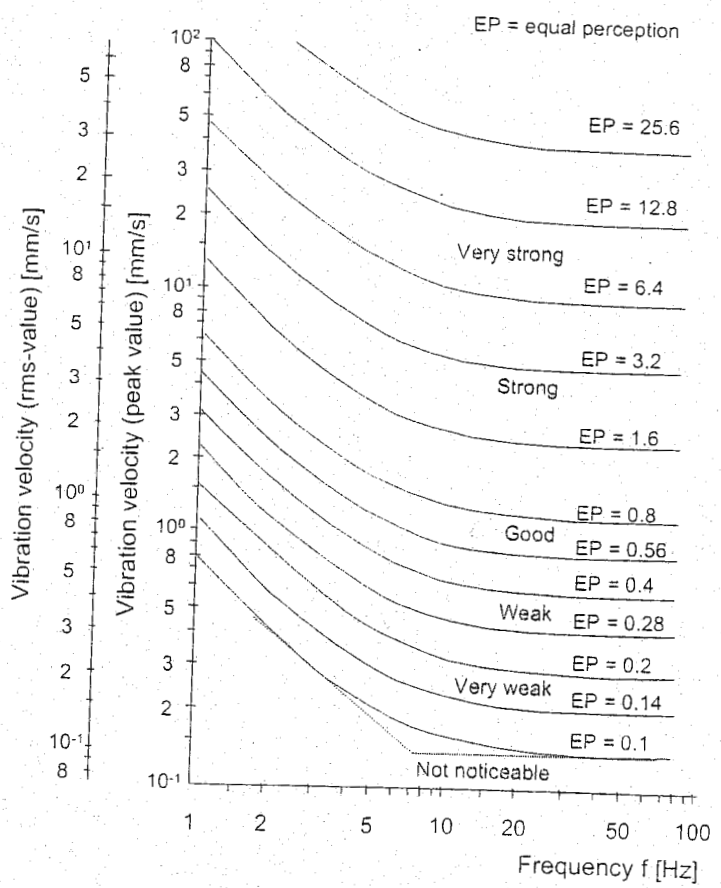


Figure 15.11: Curves of equal vibration perception

The relationship between linear scale and dB scale is represented in Figure Figure 15.12. The reference values are not internationally standardised.

For example, v_0 was 10^{-8} m/s in Figure 15.12, while ISO prefers 10^{-9} m/s and German literature often refers to 5.10^{-8} m/s. The vibration levels can be classified globally as shown in Table 15.1 .

From the point of view of those who live near the track, vibrations should preferably be kept below a velocity amplitude of 1 mm/s. Most of the complaints investigated by NS concerned vibrations with an amplitude of less than 2 mm/s [118].

The largest vibration amplitude at 10 m distance from the track centre estimated by NS from statistical data was 7 mm/s. For public buildings a value of 8 mm/s, measured at the foundation, is normally taken as the limit. This is in accordance with DIN standard 4150 which says that the possibility of damage to buildings is excluded below this value [7].

	equation	reference value
acceleration	$L = 20 \log_{10}(a/a_0)$ dB	$a_0 = 10^{-5}$ m/s²
velocity	$L = 20 \log_{10}(v/v_0)$ dB	$v_0 = 10^{-8}$ m/s
displacement	$L = 20 \log_{10}(d/d_0)$ dB	$d_0 = 10^{-11}$ m

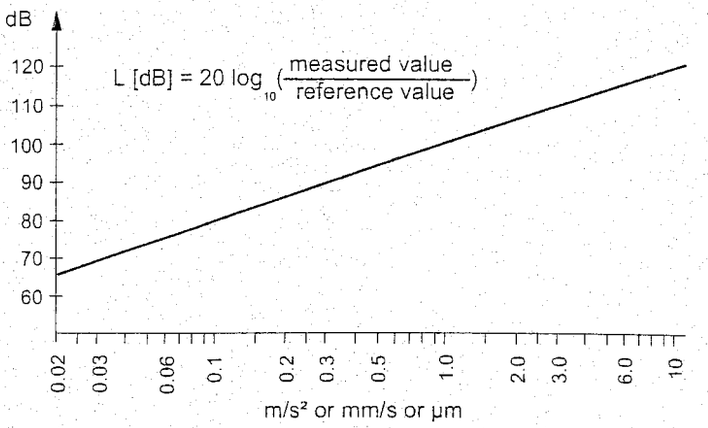


Figure 15.12: Relationship between linear scale and dB scale for acceleration a, velocity v, and displacement d

EP value	LEP [dB] $v_0 = 5.10^{-8}$ m/s	LEP [dB] $v_0 = 10^{-8}$ m/s	LEP [dB] $v_0 = 10^{-9}$ m/s	Classification of perceptibility
< 0.1	< 66	< 80	< 100	not noticeable
0.10 - 0.25	66 - 74	80 - 88	100 - 108	very weak
0.25 - 0.63	74 - 82	88 - 96	108 - 116	weak
0.63 - 1.60	82 - 90	96 - 104	116 - 124	good
1.60 - 4.00	90 - 98	104 - 112	124 - 132	strong
> 4.00	> 98	> 112	> 132	very strong

Table 15.1: Classifications of vibration levels

Tests carried out by ORE D151 [98] showed that masonry walls that had been subjected to a 14 mm/s peak vibration at 13 Hz over a period of 60 days did not undergo any substantial damage.

15.3.4 Measured vibrations

A number of investigations on vibration transmission carried out by ORE D151 are described in [215]. Figure 15.13 shows acceleration spectra measured on the first floor of a building during the passage of a train at 60 km/h. In the test three wheel tread conditions were considered: newly re-profiled wheels, wheels worn in service over 150.000 km, and wheels worn in service after 150.000 km with flats developed by skidding, i.e. emergency braking with suppression of wheel-slide protection.

According to Figure 15.13, the measured accelerations are in the frequency band 30 - 250 Hz. Wheel treads worn in service seem to show a significant increase in vibration, in the order of 5 to 10 dB, compared to newly profiled wheels. Wheel flats produce a further increase in vibration, of the order of 5 dB, at frequencies above 60 Hz. The wheel flat energy is also transmitted to the buildings at higher frequencies and produces an acoustic effect which increases the noise level by about 5 dB(A).

An increase in ground vibration may also be expected from irregularities in the rail, in particular from joints and rolling defects. Because of their short wavelength, rail corrugations will not significantly contribute to low frequency ground vibrations. The difference between jointed rail and CWR is not likely to be more than 5 dB.

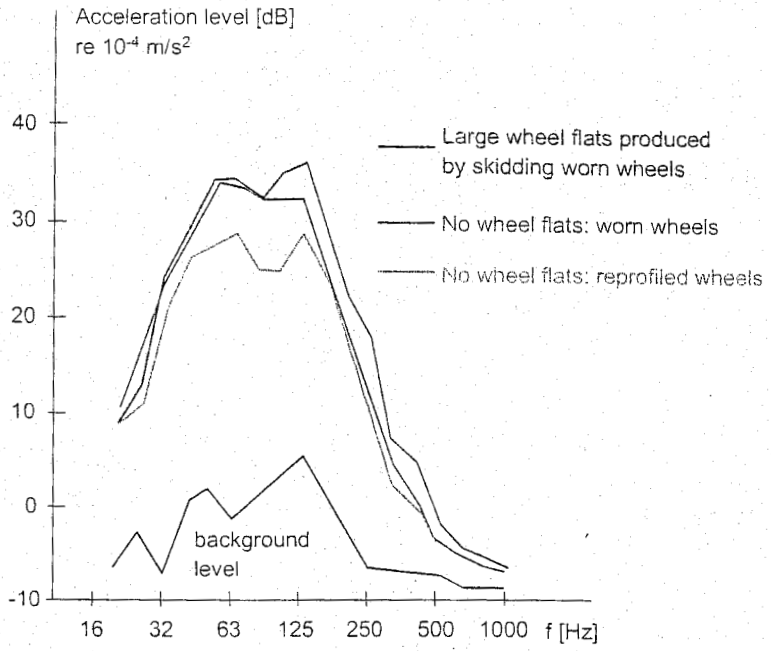


Figure 15.13: Measured vibration at a building near an underground railway

Measurements on the DB Hannover Würzberg high-speed line, presented in Figure 15.14, have shown that serious vibrations may occur at frequencies which correspond to the axle spacing of the bogies [215]. At 250 km/h rms-values per 1/3 octave band have been found of 78 dB with $v_0 = 10^{-8}$ m/s. This emphasises the importance of understanding the vibrations which will be caused by very high-speed trains.

Another example of railway-induced vibrations is the use of the Dynamic Track Stabiliser (DTS) described in Chapter 12. The high energy transmitted locally to the track results in strong soil vibrations, primarily radiated as Rayleigh waves in the surrounding field. NS carried out several acceleration measurements at different distances from the track [74]. The velocity equivalents were tested against the previously mentioned standard of 8 mm/s.

From the measuring results presented in Figure 15.15 it becomes clear that, according to DIN 4150 the DTS should not approach buildings closer than about 15 m. At about 20 m distance from the track the DTS causes the same vibration level as an ore train.

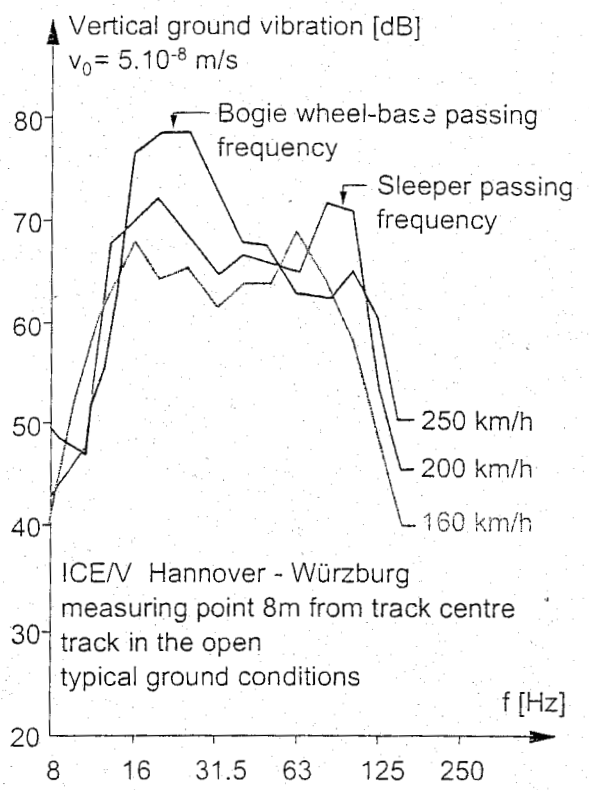


Figure 15.14: The effect of very high speeds on ground vibration measured by DB

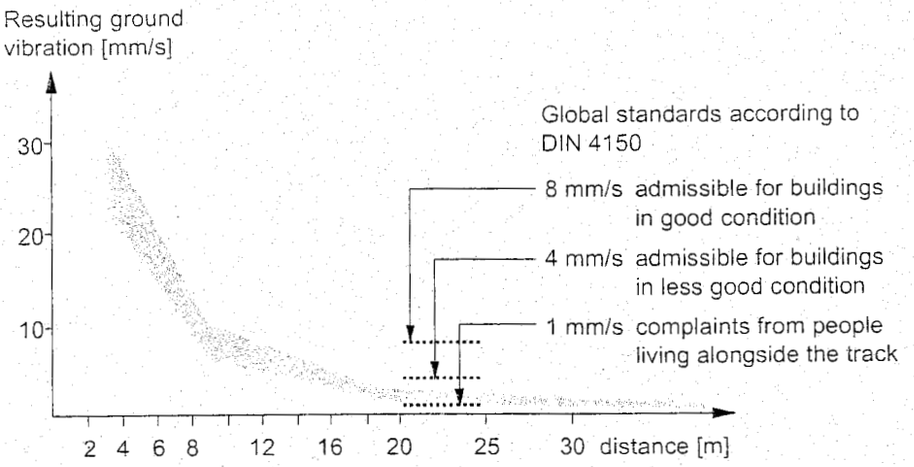


Figure 15.15: Measured ground vibration due to DTS

15.3.5 Vibration reduction

Vibration protection measures are largely confined to modifications to the track structure in tunnels, and to additional measures in the ground under or alongside tracks on open terrain.

The common principle in all the systems used in tunnels is to support the track mass by a resilient spring. The mass-spring systems act as a barrier to vibrations with a frequency greater than $\sqrt{2}$ times the natural frequency.

For an undamped 1-mass-spring system the dynamic amplification V, being the ratio between dynamic and static force or dynamic and static displacement, equals:

$$V = \frac{1}{\left|1 - \left(\frac{f}{f_n}\right)^2\right|} \qquad (15.12)$$

in which f_n is the natural frequency and f is the frequency of the excitation. The vibration reduction W can be expressed as:

$$W = 1 - V \qquad (15.13)$$

Figure 15.16 how both the amplification and the vibration reduction versus frequency for a natural frequency of 10 Hz. The vibration reduction in dB reads:

$$W_{dB} = 20\log\frac{1}{V} \qquad (15.14)$$

$$\cong 40\log\frac{f}{f_n} \quad for \quad \frac{f}{f_n} \geq 3 \qquad (15.15)$$

The most effective way of reducing vibrations is found in measures taken at the source. Of course the same is also valid for noise reduction.

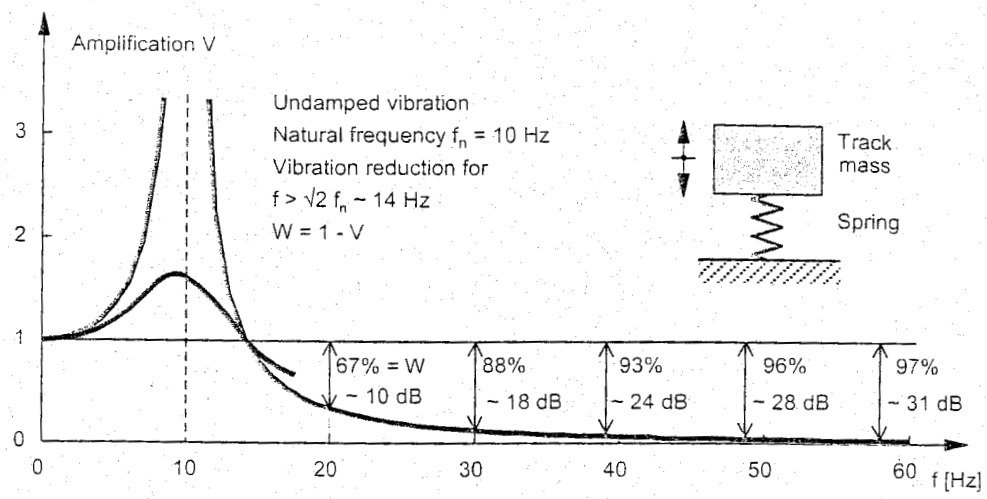


Figure 15.16: Principle of vibration reduction

15.3.6 Measures for ballasted tracks

With ballasted tracks, vibrations may be reduced by:

- increasing the ballast depth. DB tests have shown a reduction of 6 dB at frequencies below 10 Hz by increasing the ballast depth from 30 to 75 cm. This is not an attractive solution because of the costs, the weight and the extra height;
- installing resilient mats between the bottom of the ballast and the tunnel invert;
- installing sleeper soffit pads with the pads between sleeper and ballast;
- installing super-elastic fastenings systems, for instance the "Cologne Egg".

Figure 15.17 summarizes experimentally determined sample transfer functions which describe vertical vibration reduction versus frequency. The behaviour is quite similar and their effectiveness is larger at frequencies above 30 Hz.

For constructional details of these measures see Chapter 8.

15.3.7 Measures for slab tracks

Using ballastless track, particularly in tunnels of urban mass transit systems, is finding increasing support because of the opportunities to save on maintenance costs. However, unless certain precautions are taken, the substitution of a rigid concrete base by relatively loosely bound ballast can, compared to ballasted tracks, increase the degree in which the system transmits vibration.

Ballastless track systems may be designed to offer improved vibration attenuation by the interposition of elastomeric layers within the rigid track structures. These systems then approximate mass-spring systems as demonstrated in Figure 15.16.

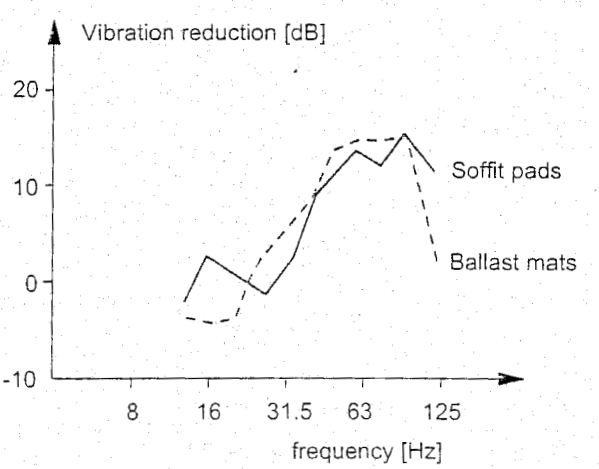


Figure 15.17: Vibration reduction achieved by ballast mats and soffit pads

Most slab track systems, which utilise baseplates in their structures, have the baseplates resiliently mounted relatively to the slabs. The vertical elasticity in either one or two stages is, however, generally designed to be similar to that of a typical ballasted track. Therefore, their vibration attenuation is not significant compared to ballasted track, and at low frequencies may be even worse.

However, some special types of baseplate system, such as the "Cologne Egg", mentioned above, have also been applied on slab track.

Slab track systems, which employ the principle of resiliently mounted blocks or sleepers, can show improved vibration isolation compared with direct-fastening baseplated systems by virtue of the increased mass, which is resiliently mounted when considered as a simple mass-spring-damper.

Probably the most well known example of this type of track system is the STEDEF-VSB system, which is also shown in Chapter 9.

Extending further the principle of increasing the track mass which is resiliently supported, thereby reducing the natural frequency and hence increasing the effectiveness of the system as a vibration isolation measure, slab track systems have been designed in which large precast or in-situ concrete slabs are resiliently supported from the tunnel invert.

All these constructions become beneficial at frequencies above 30 Hz. This means that these concepts can be effective for reducing vibrations radiated as P and S waves by underground railways. It is difficult to realise the mass-spring concept at an economic cost at lower frequencies of 10 to 30 Hz, where in practice the problems often occur.

15.3.8 Measures for tracks in the open

Tracks in the open mainly produce surface waves with vibrations at low frequencies. Especially for those frequencies below 10 Hz active isolation, by means of a mass-spring concept, will not be feasible. In such cases passive isolation may be considered, for instance by using open trenches in the vicinity of the structure exposed to vibrations, As the surface wave energy is mainly transmitted through a layer with a depth of one wavelength, intersection of this layer will therefore reduce the transmitted wave energy.

In [74] a number of tests based on this principle is described. The isolation effect appears to be more than 70% if the depth of an open trench exceeds about half the length of the surface waves. There are, however, a number of practical constraints, one of them being the trench depth. In fact, this solution requires the Rayleigh wavelength to be shorter than 20 m. Besides, it is often difficult to install open trenches permanently and to a sufficient depth. Franki has developed a system which allows a permanent gas-filled screen, consisting of pressurised plastic cushions, to be created in the soil.

Much work on vibrations has been initiated in the Dutch CUR/COB L400 project on designing so-called modules describing vibrations in civil engineering objects. Among them is a source module for the assessment of rail traffic-borne dynamic forces and vibrations resulting from irregularities of the track and/or the wheels.

15.4 Railway noise

Bridges

Moving trains radiate noise mainly as a result of wheel and rail interaction. In addition to structure-borne noise which was discussed in the previous section, this section will deal with air-borne rolling noise which is mainly radiated in the frequency band 500 - 2500 Hz.

Numerous measurements have been carried out by railway administrations to determine radiated noise levels. Various ORE Committees have also investigated noise problems. The noise level measured at a certain distance from the track is strongly dependent on the track structure type, i.e. wooden or concrete sleepers, track on steel bridges or concrete bridges, with or without ballast bed, and the condition of the rails, especially as regards corrugations and bad welds.

The ORE Committee C137 measured noise level near a great number of bridges and compared these values with levels in the open field alongside the track in the immediate vicinity of the bridge. These results were published in [204].

Figure 15.18 shows a typical example of the spectral composition in dB per 1/3 octave at 25 m distance from the track centre on the bridge and in the open field. These graphs clearly reveal that the vibrations of the bridge during train passage lead to an increase in low frequency noise energy levels and to a considerable rise of rolling noise. The precise figures strongly depend on the bridge type.

Above 1000 Hz bridge and plain track produce equal noise levels. When calculating the A-weighted equivalent noise energy level (L_{eq}) the open field gives 86 dB(A) and the bridge only 2 dB(A) higher whereas on a spectral level the differences are in the order of 10 dB. The discrepancy is caused by the A-weighting filter which removes almost all contributions below 200 Hz.

The L_{eq} values measured by ORE at 25 m from bridges vary between 80 and 101 dB(A). In the open field these values lie between 75 and 93 dB(A). The increase in noise level attributable to the bridge varies between 0 and 11 dB(A).

The noisiest bridges are always the all-steel types with directly laid track.

Side plate girders are a very effective screen against rolling noise, as they give a reduction in the order of 4.5 dB(A).

Owing to its mass, ballast is also an element which damps vibrations, and owing to its porosity absorbs air-borne noise. Compared to directly laid track ballast causes a sound level reduction of the order of 10 dB(A) on steel bridges.

Corrugated rails can cause an increase in noise level from 5 to 15 dB(A) [2]. Similar effects are to be expected from poor welds and joints, and also from wheel flats.

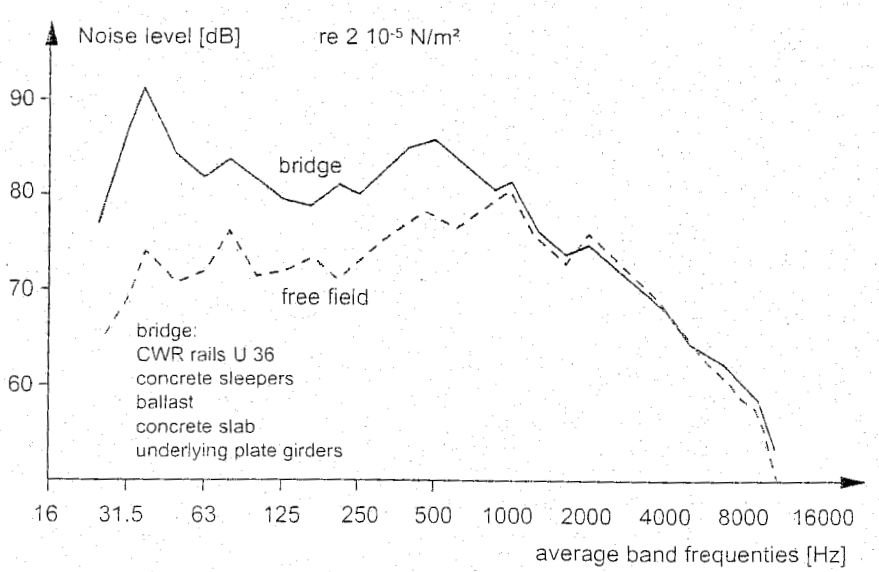


Figure 15.18: Comparison of noise spectra for bridges and plain track measured at a distance of 25 m from the track centre

Silent steel bridges

The noise radiation on bridges has lead to various developments to reduce the nuisence. The main sources are the wheels, the rails and the steel bridge girders. In The Netherlands the track specific solutions were primarily found in the application of embedded rail. Here, in the first place, the rail is to a large extend cast into Corkelast and therefore the radiation surface is relatively small. Figure 15.19 shows an example of the more traditional embedded rail concept on steel bridges.

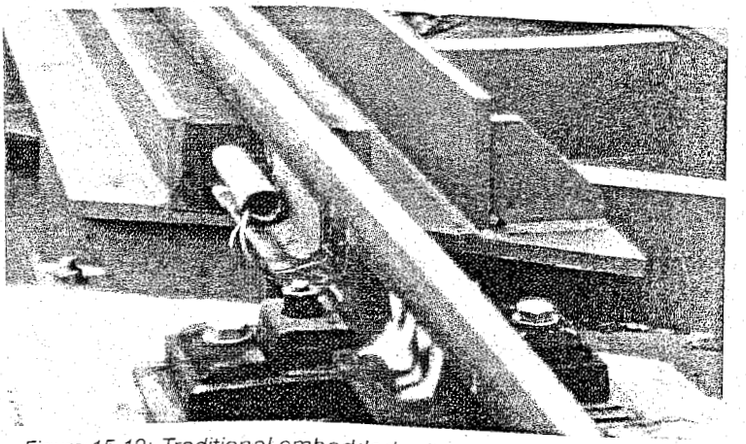
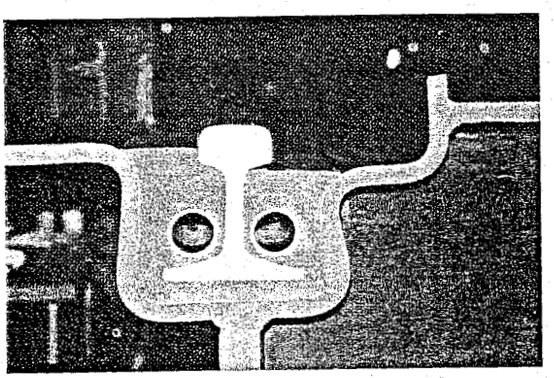


Figure 15.19: Traditional embedded rail concept on steel bridge

A more sophisticated solution is the concept depicted in Figure 15.20, representing an integrated design. The principle is based on supporting the rail structure by a very stiff spring, comprised of the main girder of the bridge, with an intermediate flexible spring consisting of corkelast between girder and rail.



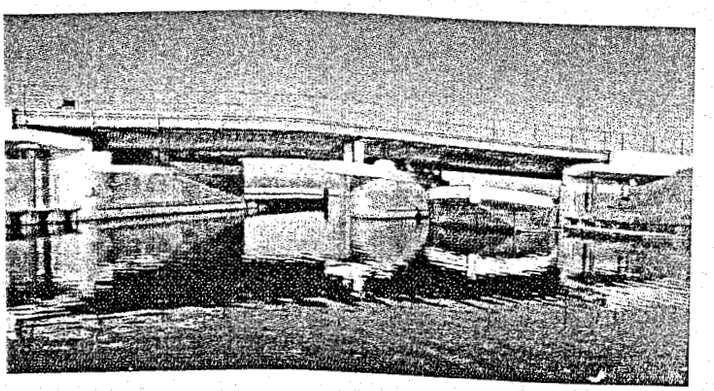


Figure 15.20: Silent bridge cross section and view of a bridge

As the rail trough forms a monolithic part of the main girder a very stiff support is created with an extremely low vibration level due to the dynamic decoupling between rail and bridge. Figure 15.21 shows a measuring example of a silent bridge and simultaniously presents the noise level spectrum of a conventional steel bridge. The difference is in the order of 10 dB.

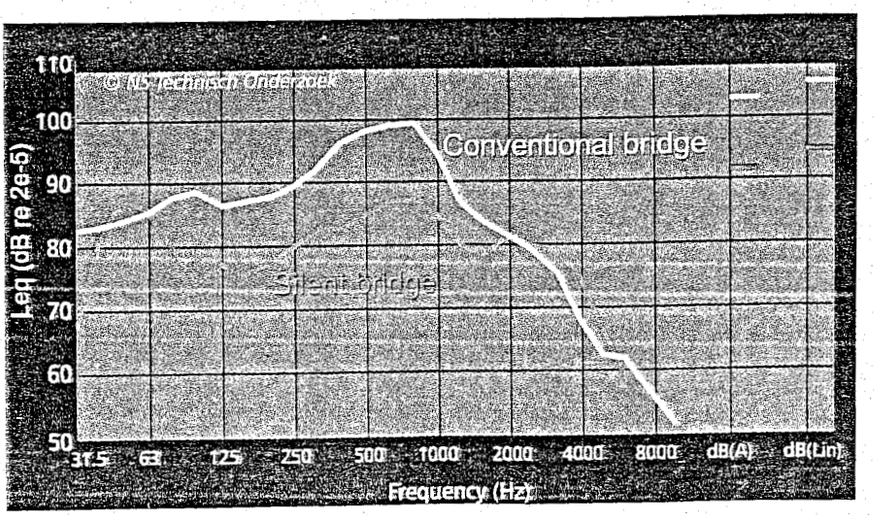


Figure 15.21: Measured noise energy levels at silent bridge and at conventional bridge

NS experience

NS has carried out a lot of noise measurements alongside tracks to investigate the noise level in relation to track type and rolling stock type. The SEL values obtained at a distance of 25 m from the track centre vary between 91 - 99 dB(A), which corresponds approximately to Leq values between 85 - 93 dB(A). The scatter due to the influence of the rolling stock amounts to 5 dB(A). A remarkable NS experience is the fact that on tracks with wooden sleepers the noise energy levels are systematically higher, of the order of 1 -3 dB(A), than on tracks with concrete sleepers.

Silent track

Noise problems associated with railway operations are nowadays attacked in a more structural way. In the Netherlands a major research project called 'Silent Train Traffic' was completed in 1998, where NS, in collaboration with industrial partners and research institutes, aimed at a noise reduction of 10 dB(A) for the whole system. For both subsystems, rolling stock and rail infrastructure, optimization were carried out regarding structure, components, and materials.

In the sub-project 'Silent Track' a new structure was developed with a target of reducing the noise production with at least 5 dB(A) compared to ballasted track and consequently 7 dB(A) compared to conventional slab track. The solution was found in an embedded rail structure. As most of the vibrating energy is in the rail and the slab is only vibrating at least an order of magnitude less, the solution went towards transferring more energy from the rail to the slab. This was achieve by using a relatively flexible block rail, cast into a relatively stiff corkelast (polyurethane mixed with cork). The final design, shown in Figure 15.22, was comprised of a block rail SA 42, with a weight of 42 kg/m, provided with a UIC 54 railhead.

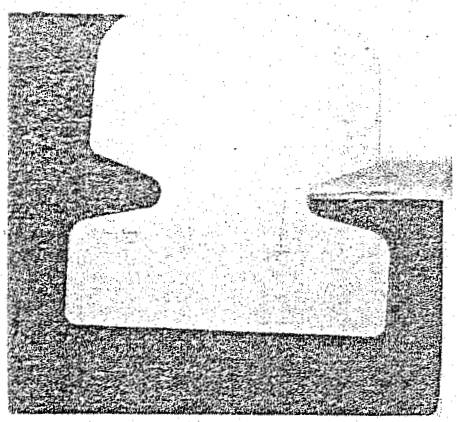


Figure 15.22: Rail SA 42

Acoustic calculations and tests at TNO [137] had shown that the noise reduction target was met. At TU Delft calculations and fatigue tests had been carried out [168], which proved the fitness of this assembly for 225 kN axle loads.

Finally this structure was tested in situ at the slab track near Best, described in section 9.8.3, where the claimed noise reduction levels could be confirmed.

16 INSPECTION AND DETECTION SYSTEMS

16.1 Railway Infrastructure Monitoring

Railway Infrastructure Monitoring (RIM) represents one of the most important parts of an Asset Management System (AMS). For more details on this subject please refer to Chapter 19. The overall managing capabilities of the AMS will greatly depend on the quality of the available monitoring systems. As the focus of this book is the railway track, track monitoring and condition assessment techniques will be given special consideration. However, since it was stressed that successful track condition analysis and consequent management could only be performed in combination with other railway infrastructure, monitoring of other infrastructure objects and means for their management will also be briefly elaborated on. Special consideration will be given to substructure monitoring and the monitoring of switches and crossings, as their influence on the track condition is significant.

The reason for monitoring is usually twofold. The first, immediate reason is obviously to detect irregularities that could endanger the safety and reliability of railway traffic. However, if a monitoring technique is continuous and fast enough to allow consecutive monitoring runs to be performed at regular time intervals, an extremely important temporal aspect is obtained which is of essential importance to a successful condition-based management. This means, that such a monitoring technique could provide insight into the infrastructure element's behaviour over time. And this could allow condition forecasting and consequent maintenance planning. This concept usually represents the ultimate goal of any condition monitoring.

16.2 Tunnel monitoring

Continuous tunnel monitoring represents one of the latest developments within the field of railway infrastructure management. Only recently, techniques for laser, thermal, and video scanning of the tunnel's inner surface became available and more widely used. Many railway tunnels worldwide are more than a 100 years old, and due to the extensive development above or adjacent to existing services there has been an increased emphasis on monitoring the integrity of tunnels.

Figure 16.1 shows the tunnel profile laser/visual/thermal scan processing system "ScanView". ScanView is a viewing utility which uses sophisticated data manipulation to display the information gathered by the TS 360 BP Scanner in a straightforward and intuitive manner. Amongst its features are:

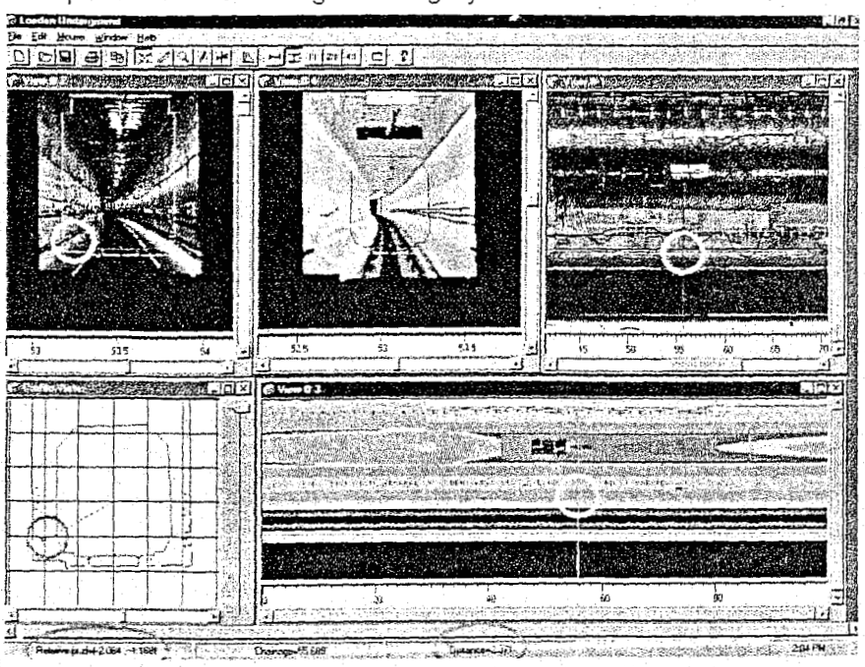


Figure 16.1: Tunnel profile laser/visual/thermal scan processing system - Scan View

- Plan viewing of visual, thermal, or profile data;
- Cross-section display for any chainage of the scan;
- Accurate real life distance measurements between features on the scan;
- Reconstruction of the three dimensional image viewed from any chainage;

- Annotation feature allowing areas of the scan to be highlighted and have other documents and applications hotlinked to them;
- Variable scale zooming to any part of the scan;
- Multiple Noise Reduction techniques to process the profile data;
- Automatic flight path calculation from track data;
- MDI Interface allowing multiple views to be compared.

Figure 16.2 shows another tunnel profile.

16.3 Bridge monitoring and management

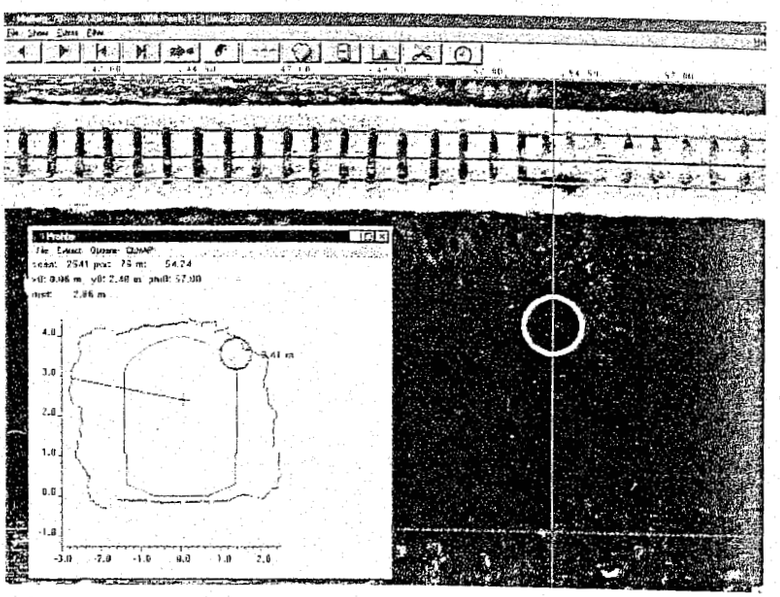


Figure 16.2: Another tunnel profile laser/visual/thermal scan processing system - TuView

Bridges belong to the most critical elements within the railway infrastructure, especially concerning safety. All over Europe there are tens-of-thousands of bridges, steel or concrete, which span up to hundreds of meters. Since national governments developed policies to stimulate rail transport and the need for mobility steadily rises, it is expected that rail traffic will considerably increase during the next decades. Increasing use, train frequencies, speeds, and axle loads lead to much heavier loads which will affect the design life in a negative way and might even jeopardise safety. This will increase the need for M&R work, which will affect availability and cause train delays and other traffic disruptions. It will also increase the chance of accidents involving maintenance staff working in the field.

Bridge management systems (BMS) provide both the owners and users with the knowledge and tools to analyse, ensure, and maintain bridges in the most LCC-effective way. These systems are based on the actual bridge condition, including deterioration and Failure-Mode Analysis, and are directed towards providing maximal track availability and reliability.

One of such BMS systems is a Pontis System, developed in the United States by the FHWA in conjunction with six state DOTs and the consultant joint venture of Optima, Inc. and Cambridge Systematics. It is a comprehensive bridge management system developed as a tool to assist in the challenging task of bridge management. It stores bridge inventory and inspection data, formulates network-wide preservation and improvement policies for use in evaluating the needs of each bridge in a network, and makes recommendations for what projects to include in an agency's capital plan for deriving the maximum benefit from limited funds.

Inspection records are stored for each bridge element. Further, for each bridge element Pontis develops models for how the element deteriorates over time (which may vary by operating environment), what preservation actions may be taken in response to deterioration, how much actions cost, and how effective actions are at preserving the element.

Figure 16.3 shows the framework Pontis uses in recommending projects for an agency's capital or maintenance programs. In making project recommendations, Pontis distinguishes between preservation and improvement projects. Preservation projects consist of bridge maintenance, repair, and rehabilitation (MR&R) actions performed on individual bridge elements. Pontis models how MR&R actions improve element condition, as well as how bridge elements deteriorate over time in the absence of MR&R actions. The overall objective of preservation projects is to maintain bridges at minimum long-term cost, without altering the bridges' functional aspects. The result is a recommended MR&R policy for minimizing long-term preservation costs for each bridge element. The general structure of the BMS software is presented in Figure 16.4.

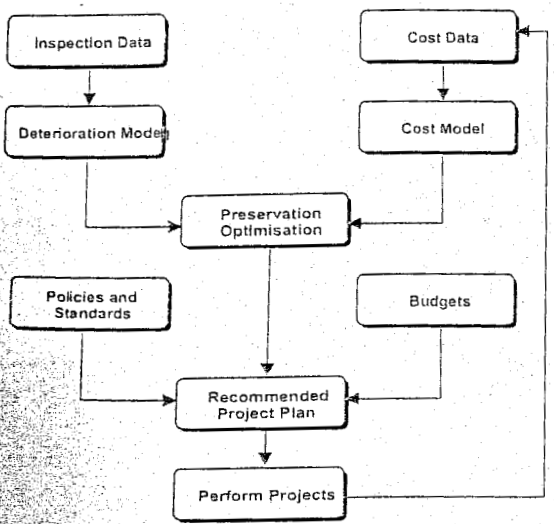


Figure 16.3: General structure of BMS software

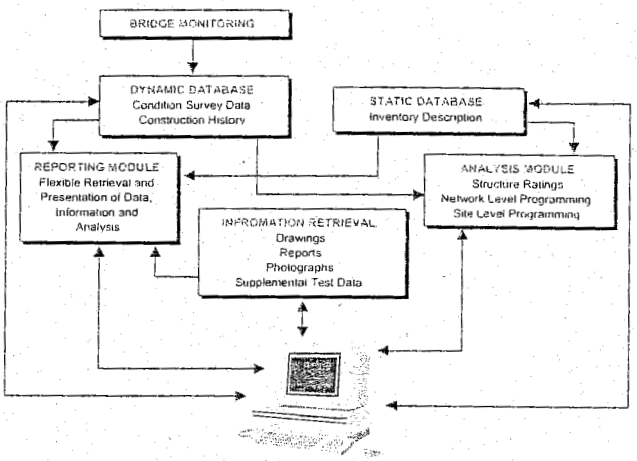


Figure 16.4: BMS framework system of Pontis

16.4 Substructure Monitoring

Very often problems identified at the superstructure level, e.g. track geometry problems and especially those severe and recurring ones, originate at the substructure layers (ballast, subballast, and subgrade). However, due to the lack of consistent knowledge about substructure behaviour and the inaccessibility of substructure for the purposes of continuous monitoring, usually the symptoms are treated rather than the actual cause.

What usually happens is that instead of the necessary maintenance activities, standard geometry maintenance is carried out by continual lifting, lining, and tamping. Unfortunately, problems recur very soon, and eventually, the geometry maintenance becomes too frequent and/or the rate of deterioration becomes so high that it is no longer acceptable. When that happens major corrective maintenance activity is needed, but in the meanwhile a lot of time is lost and the track has deteriorated to a much higher level than would be the case if the proper maintenance had been performed on time.

This is why it is obvious that the substructure significantly affects the track maintenance costs, both directly and indirectly. It also explains why an integrated approach is needed that would encompass both superstructure and substructure management. To make things worse, these corrective maintenance activities usually require a significant amount of time necessitating track closures. This is again in sharp contrast with the constant demand for railways to increase traffic for which a high availability of the track is needed. There are also indications that the stiffness of the subgrade could affect the life of the track components such as rails, sleepers, and ballast.

The two key objectives of such an integrated approach include the following:

1 To establish a reliable relationship between substructure conditions and track maintenance requirements;

2 To determine feasible means for altering the negative effects of substructure on the overall track structure in order to reduce maintenance costs.

This integrated approach is based on the following steps:

- Definition of substructure condition parameters to be monitored, which would serve as the basis for further analyses;
- Establishment of non-destructive methods and means for continuous substructure monitoring (easy and cost-effective to perform in periodical and continuous manner, possibly without traffic disruptions);
- Enabling simultaneous and joint analysis of all relevant aspects (inventory data, superstructure and substructure condition data, M&R history data, traffic characteristics, load and age information, site inspections and field and laboratory tests data, etc.);
- Definition of "rules" to govern the decision-making process dealing with optimal M&R works to be performed under certain circumstances, based on the previously mentioned joint analysis and economic analysis of the alternatives;
- Monitoring the performance of the proposed M&R works as feedback for further adjustments of the "rules".

16.4.1 Substructure condition parameters

Overview

There are various parameters that indicate the condition of the substructure. However, inaccessibility of the substructure makes the use of these parameters much more difficult. There is no doubt that the best condition data would be collected by means of excavation which provides direct evidence. However, this option is often either absolutely impossible or at least not preferred, as it would cause inevitable traffic disturbances. The parameters used so far and some of the prospective ones include:

– Track geometry data	– Moisture Content data
– Ground Penetrating Radar (GPR) data	– Infrared Thermographic inspection data
– Track Stiffness data	– Overall Drainage information
– Subgrade Strength data	– Petrographic examination data

As in case of superstructure, the ideal techniques for substructure monitoring would be the ones that could be performed in a continuous manner, at reasonably high speeds, without traffic disruptions, regardless of time of day and weather conditions, and could be repeatedly performed at regular time intervals so that the substructure behaviour could be captured.

However, as indicated earlier, it is extremely hard to monitor layers which are located deep down below the surface. But despite of that, to date a number of techniques have been identified which can assist in the substructure maintenance management. Some of them are mentioned below.

16.4.2 Ground Penetrating Radar

The principle of Ground Penetrating Radar (GPR) measurements is based on the radiation of very short electromagnetic impulses which are then reflected from the interfaces of materials of different dielectric properties and recorded. The detected signals offer potential for assessing foul ballast conditions and drainage problems hidden beneath the surface, as well as borders between different substructure layers, air voids, water inclusions, and other inhomogeneities. Reflections recorded by the receiving antenna by moving both transmitter and receiver along the selected track surface can be visualised as 2D images or radar grams, where the intensity of the reflected impulses is displayed in the form of a grey scale. These images can subsequently be analysed, both manually or automatically using computer-aided pattern recognition techniques, and could then provide further information on, for example, the thickness of the substructure layers. The principle is explained in Figure 16.5.

Furthermore, if GPR measurements were repeatedly performed at certain time intervals, they could provide information on locations where changes take place in the thickness of substructure layers, indicating possible problem places like, for example, ballast pockets. As in the case of the geometry car, radar data can be obtained on a continuous basis as the equipment is mounted on a suitable vehicle and moved along the track. This provides the opportunity to survey long distances and provide a 100 % coverage of the track with reasonable track possession.

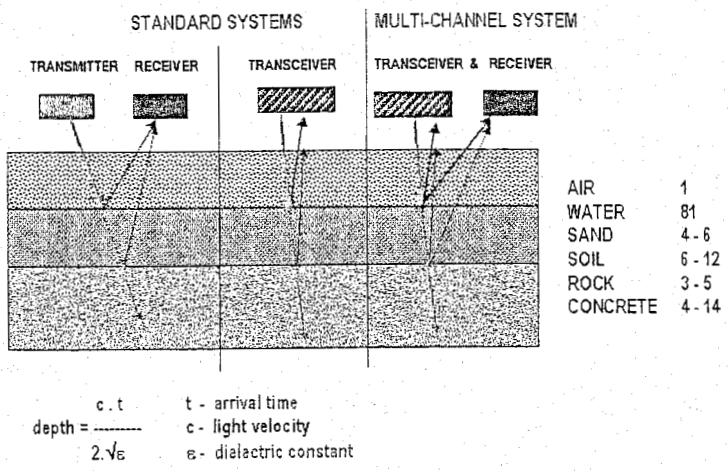


Figure 16.5: Basic principle of the multi-channel antenna GPR

During the last two decades many attempts were made worldwide to improve the use of GPR within the Track Substructure Monitoring with varying success.

The reliability of the interpretation of the radar scan readings remained the main problem and the key obstacle to the wider use of this method. Only lately, some new approaches have made the method seem more promising. One of these approaches took place in the Czech Republic, where a joint effort of the Czech Railways (CD) and a company G-Impulse Ltd. based in Prague, led to the design of a multi-channel GPR system and a connection between GPR and a track recording car. The basic principle of the multi-channel antenna GPR is presented in Figure 16.6.

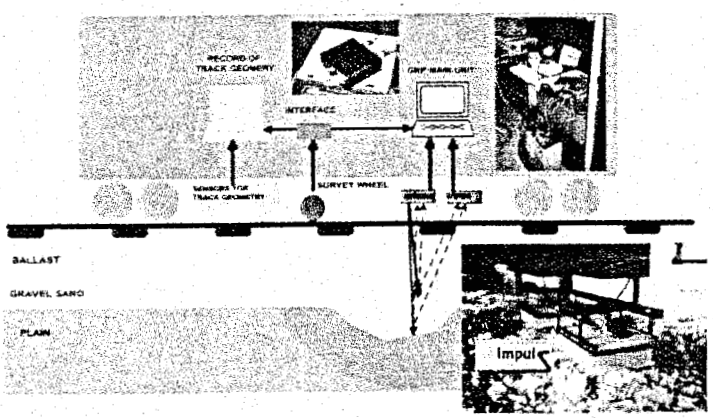


Figure 16.6: Incorporation of the GPR into a Track Recording Car

Some of the most important characteristics of the system are:

- Operation frequency 16 - 1500 MHz (for railway use 100 - 500 MHz);
- Sample frequency 100 scans/s;
- With running speed of 180 km/h, every 0.5 m can be measured;
- Penetration depth 10 m (at 100 MHz) or 3 m (at 500 MHz);

Current use for the purposes of Geotechnical Investigation:

- Determining the disturbed parts;
- Splitting up the track into quasi-homogenous blocks;
- Optimizing the number of necessary probes and their location;
- Determining technological layer thickness.

Prospects for the future use of GPR as a non-destructive continuous track investigation method:

- Forming a database of a railway network;

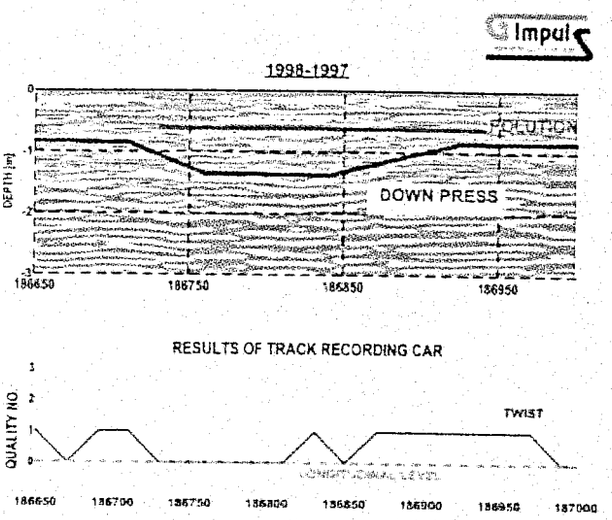
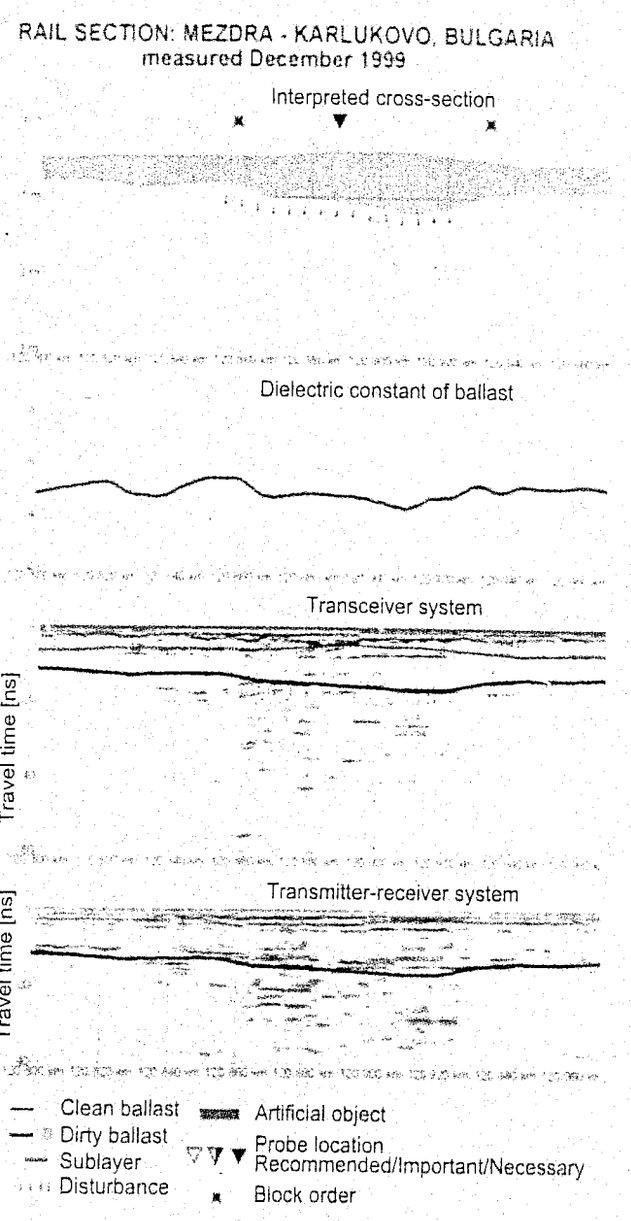


Figure 16.7: Joint interpretation of the GPR readings and Track recording car measurements

Figure 16.7 and Figure 16.8 show the basic capabilities of the system and the interpretation and use of GPR readings for the purposes of determining the number and locations of necessary probes.



16.8: GPR field records and interpreted cross-sections

16.4.3 Track Stiffness Measurement

Track deflection or stiffness

Another promising field measurement of the track's structural condition is track deflection or stiffness under vertical wheel loading, which has been found to be very useful for the purpose of determining the cause of certain substructure problems [56].

Unfortunately, in most of the cases railways do not possess the right equipment for this type of measurement and, thus, do not utilise the insight these measurements could have provided them with. The basic principle of this technique is that the vertical load is applied to both rails simultaneously using hydraulic cylinders, and a load cell is used to measure the vertical force.

At the same time, an optical level placed next to the track outside the zone of influence of the load, is used to measure the corresponding top of rail deflections by sighting targets attached to the rails. If collected over time, data about track stiffness and its changes over time, combined with the GPR measurements data on the thickness of layers and their changes, could provide the track manager with much better insight into the actual condition and behaviour of the track substructure.

High Speed Deflectograph

The High Speed Deflectograph (HSD) is a laser-based measuring device - currently being jointly developed by TU Delft and Greenwood Engineering, Denmark (the first prototype is expected to be available by the end of 2002) - which is designed to provide a fast, reliable tool for the continuous evaluation of the structural condition of railway track at network level. By means of non-contact laser Doppler sensors, it measures the velocity of deflections as a result of loading rather than the absolute deflections. As the rate of deflection tends to increase with increasing travelling speed, the High Speed Deflectograph is well suited for measuring at high speeds (up to approx. 130 km/h).

Measuring and analysis method

A number of laser Doppler sensors mounted on a beam positioned in front of one of the wheels of a heavy vehicle (Figure 16.9), emit laser rays which scan the rail surface. Vehicle loading will cause the rail surface to deflect, the velocity of which is then registered by the sensors.

The minimum number of sensors required is two. One sensor (A in Figure 16.9) is placed at a distance from the load where deflection data is to be obtained. The other (B in Figure 16.9), a reference sensor, is placed further away in a position where no deflection occurs. Sudden changes in deflection velocity reveal discontinuities in the structure. The locations where these discontinuities occur can subsequently be examined more closely, e.g. by visual inspection or by means of the Hammer Excitation Test.

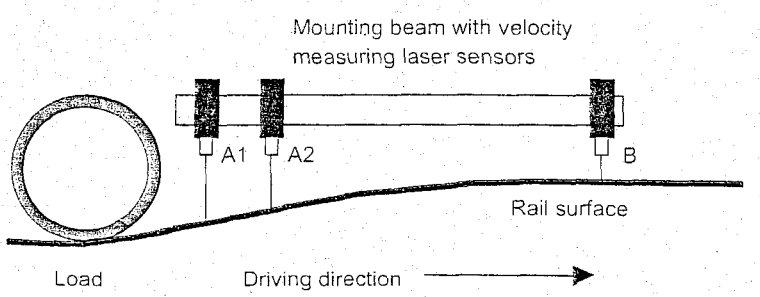


Figure 16.9: Schematic layout of the measuring principle of the High Speed Deflectograph, as an example, shown with three velocity measuring laser sensors

Since the velocity is measured in the direction of the laser rays, the incident of the laser rays on the rail surface should ideally be exactly perpendicular. If the incident is not perpendicular, a component of the travelling speed will be registered as part of the measurement. Calculation of the magnitude of the component of the travelling speed requires the angle of the incident to be measured accurately along with the travelling speed. The movements of the laser sensors and the angle of incident are measured by means of an inertial unit comprising three accelerometers and three fibre gyroscopes. The driving speed is measured using an odometer.

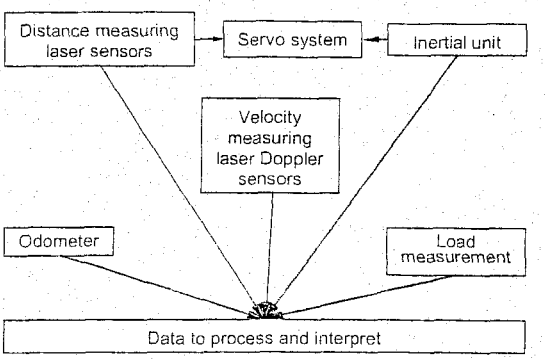


Figure 16.10: Schematic overview of data recording system of the High Speed Deflectograph

Movement of the laser sensors is limited and controlled by a servo system which is also mounted on the rigid steel beam and, thus, ensures that the laser sensors are focused at all times. Input data for the servo system is provided by the inertial unit and a distance measuring laser located at each end of the beam. Figure 16.10 gives an overview of the data recording system.

Doppler principle

As noted earlier, the measuring principle of the High Speed Deflectograph is based on the use of laser Doppler sensors. Here follows a short explanation of the Doppler shift effect on which the velocity measuring sensors are based. The shift of the observed frequency relative to the emitted source frequency is the result of a difference in the velocity of the sensor and the deflection velocity of the surface. Doppler already defined this principle as follows in 1842:

$$f_{observer} = f_{source} \pm f_{Doppler} \tag{16.1}$$

$$f_{Doppler} = f_{source} \cdot \frac{v}{c} \tag{16.2}$$

where:

$f_{observer}$: Observed frequency;
f_{source} : Emitted source frequency;
$f_{Doppler}$: Doppler shift in frequency;
v : Relative velocity between source and observer;
c : Speed at which the wave propagates (via an intermediate medium, e.g. air).

If the source moves slower than the wave propagates, the wave fronts ahead of the source emerge closer together than in the static case, and the wave fronts behind the source are further apart. Consequently, transmitted waves (e.g. light) from moving sources will appear to have different wavelengths depending on the relative motion of the source and the observer.

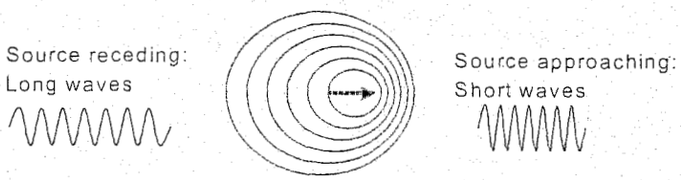


Figure 16.11: Doppler shift principle

The Doppler shift is illustrated in (Figure 16.11) and may be explained as follows. When one, for example, looks at a source that is receding, waves are observed which have longer wavelengths than when they were initially emitted (a red shift). However, when one looks at an approaching source, waves are observed which are shifted to a shorter wavelength (a blue shift).

Interpretation of measured data

Using suitable models for the interpretation of the measured data, real-time information about the structural condition of the track can be obtained which makes it possible to detect problems at an early stage. The data obtained by means of the High Speed Deflectograph can be used to supplement data obtained by means of other measuring systems, thus enhancing planning and cost-effectiveness of maintenance works.

A number of possibilities exists in order to interpret the measured data, for the choice of method is dependent on the number of laser Doppler sensors used and their respective positions. The data can be interpreted for the purpose of revealing sudden changes in deflection velocity. Data can also be processed using a moving average routine to provide more accurate data which describe the displacements of the structure.

Furthermore, obtaining the frequency response function (FRF) would allow for a more detailed picture of the track condition. If the High Speed Deflectograph were configured with a larger number of laser Doppler sensors, it may be possible to study wave propagation (the latter is subject of further testing of the equipment and of investigations into feasible models for data interpretation).

Example of data application

The following example reveals the applicability of the measured deflection velocity data. The example initially uses data from deflection bowls obtained by simulations with the computer program RAIL, developed at TU Delft (see Chapter 6). A classic track and a single load moving at 20 m/s was used as input to the program.

The shapes of the deflection bowls were determined by polynomial fitting. Figure 16.12 shows a sixth-order polynomial that fits the data obtained by means of the RAIL program.

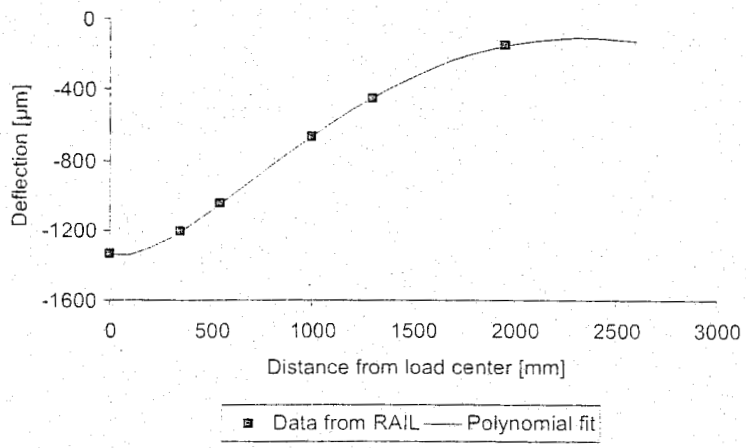


Figure 16.12: Sixth order polynomial fit applied to deflection data obtained by means of the RAIL program

The deflection bowl based on the polynomial fit and its derivative with respect to time is presented at a travelling speed of 20 m/s in Figure 16.13. This shows how it is possible to use differentiation and integration to transform deflection velocity data into a deflection bowl, and vice versa.

The above example requires the device to be configured with one sensor for each point in the deflection bowl where the deflection velocity must be known.

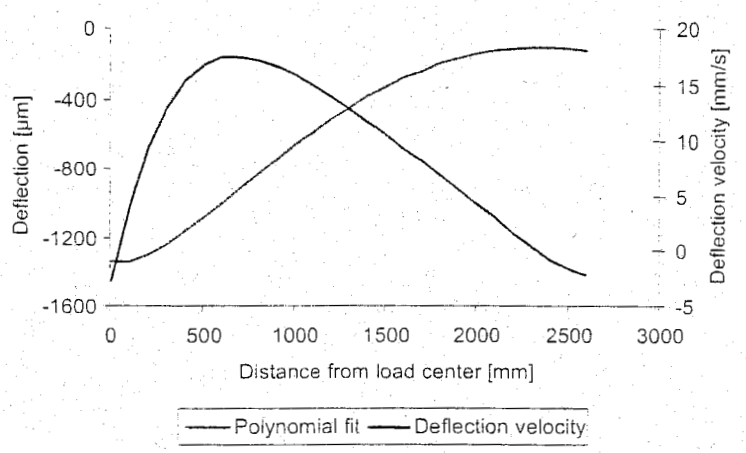


Figure 16.13: Polynomial fit and deflection velocity at a travelling speed of 20 m/s

Field of application

As opposed to stationary measuring systems, the High Speed Deflectograph provides continuous measurements which make it feasible to survey a complete network at regular intervals. It can therefore be used as a screening tool to detect discontinuities in bearing capacity, but also in order to detect the presence of loose rails or sleepers which could pose a potential safety risk.

The HSD prototype was developed for the application on roads and has proven to be very successful. The prospects for the success of the HSD prototype development for railways are also very good, because it is generally easier to measure on rail surfaces than on road surfaces. The rough surface of a road introduces more noise in the measurements compared to rail surfaces. Also the interpretation of the data is expected to be simplified compared to the interpretation of the road data.

16.4.4 Infrared thermographic inspection data

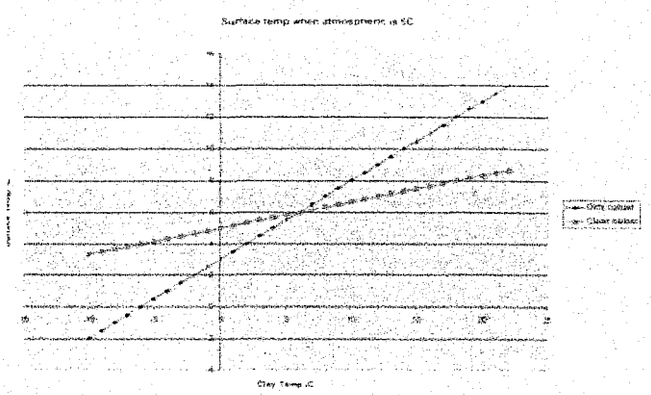


Figure 16.14: Surface temperature against soil temperature when atmospheric temperature is 6 °C

Infrared thermography [34] represents a possible means for determining the condition of the track bed. This technique is very cost-effective and large surveys can be undertaken quite rapidly. It is also non-intrusive and causes no environmental impact.

The basic idea behind infrared thermography is that in the dirty ballast the fines and small particles will fill up all the air voids, and, thus, allow the dirty ballast to be considered as homogeneous. In contrast, the clean ballast is assumed to contain voids. Using this assumption, it is possible to estimate the heat flux through different types of ballast.

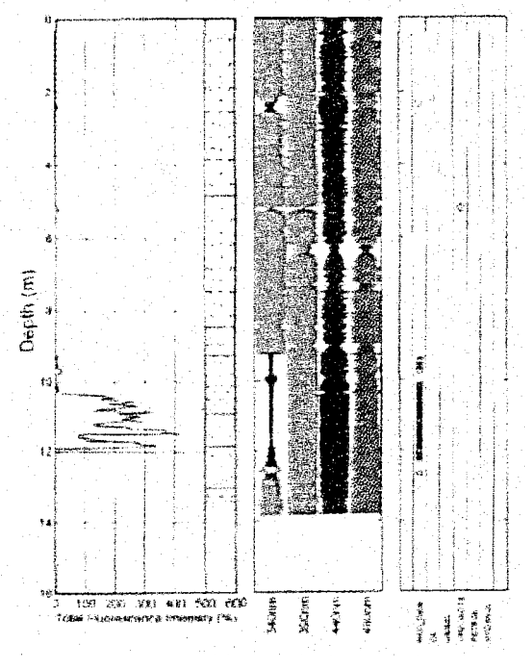


Figure 16.15: LIF results used to define contamination

However, the heat flux equations do not involve the surface temperature of ballast, which is the one aspect that can be recorded by the infrared camera. The heat flux through the whole plane is assumed to be the same as that through any individual part. This means that heat flux through the entire ballast is the same as that from the surface to the air. It is expected that it will be possible to theoretically predict the surface temperature of the dirty and clean ballast based on knowledge of the temperature of the ground and the atmosphere. Both of these temperatures can also be measured.

This means that the ideal condition to conduct an infrared survey is when the temperature difference between the clean and dirty ballast is greatest. Figure 16.14 shows the two surface temperatures of clean and dirty ballast in relationship to the soil temperature at an atmospheric temperature of 6 °C. Since the relationship is linear, the slope of the graphs does not change when the atmospheric temperature changes; only the cross-over point does. It can be seen that the temperature difference between the atmospheric temperature and the soil will have to be greater than 5 °C to achieve a 2 °C difference between the two different types of ballast.

Figure 16.14: Surface temperature in contrast with infrared soil temperature when atmospheric temperature is 6 °C. The results found so far using infrared camera analysis indicate that this technique could be very promising in order to identify different areas of ballast deterioration.

16.4.5 Laser Induced Fluorescence (LIF) Cone Penetrometer measurement

Refuelling operations in railway sidings can give rise to diesel spillage. The advantages of the LIF cone [124] are that the two channel cone penetrometer (tip resistance and sleeve friction) provides a continuous profile of the ground condition, while the LIF element mounted behind the cone provides real time data on location and relative concentrations of diesel oil. Subsequent analysis of the data in the office enables the extent of the contamination to be defined with the aid of contour plots of relative concentrations of diesel oil, together with combined plots of ground conditions and associated contamination (Figure 16.15).

16.4.6 Non-invasive moisture monitoring

So far we were able to witness some experimental monitoring systems using electrical resistivity as a proxy for imaging changing moisture content in engineered earthworks. The moisture content of soil controls both its density and strength, making it the key property when considering the stability of embankments containing soils.

Earthworks are used for embankments carrying both road and rail links. Non-invasive geophysical methods [135] have potential to monitor moisture content. Electrical resistivity, for example, is sensitive to the moisture within a soil. In the field, resistivity tomography utilises arrays of simple metal stakes as electrodes in schemes that enable the measurements of electrical resistance to be inverted into electrical resistivity of the subsurface in two and three dimensions.

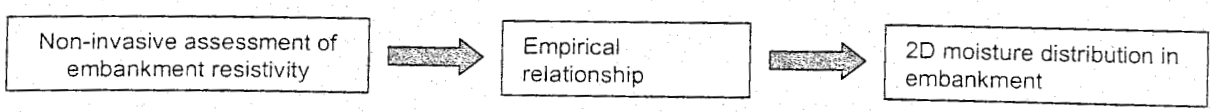


Figure 16.16: General approach to assessing ground properties

The electrical resistivity of soils and porous rocks is controlled by:

- Density / Porosity;
- Saturation;
- Pore fluid resistivity;
- Matrix resistivity;
- Pore space morphology.

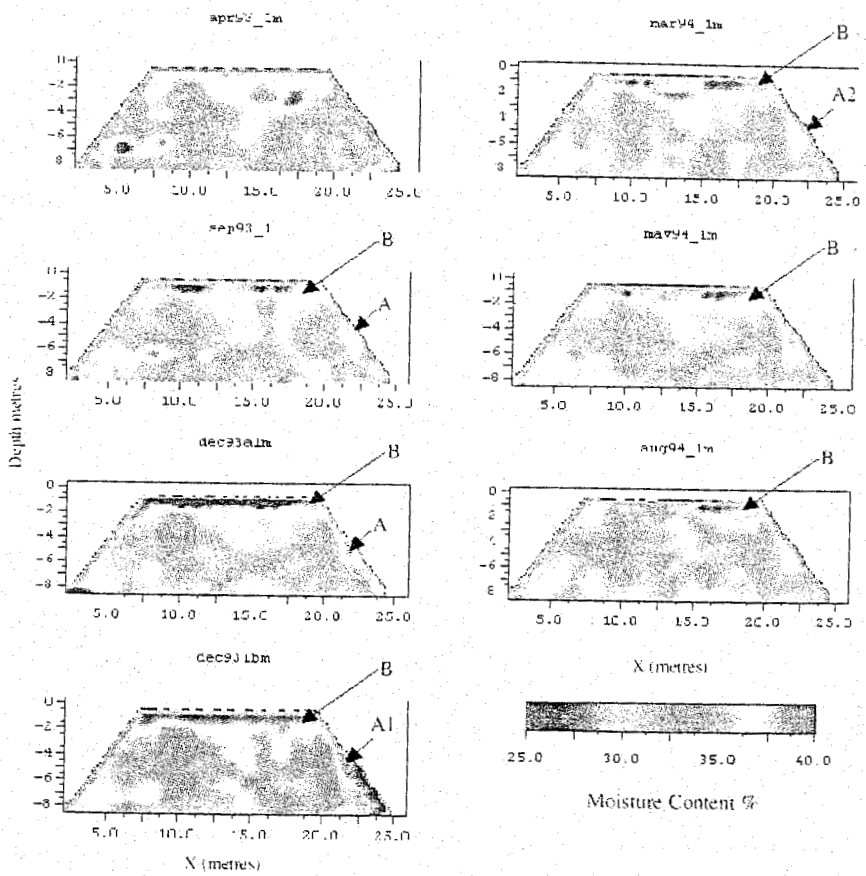


Figure 16.17: Moisture content

In order to use electrical resistivity as a proxy for moisture content, it is necessary to establish a relationship that applies to the embankment concerned. The use of electrical resistivity to assess the water content of fully and partially saturated reservoir rocks has been established more than half a century ago in the oil industry. If matrix conduction is small, such as is the case in saturated sands, the original results are applicable and provide an initial model based on which the effects of changing moisture content can be predicted. Matrix conduction is a significant process for clay-rich soils, but no generalised model exists. In such cases empirical laboratory relationships may be necessary. Figure 16.16 shows the general approach to assessing ground properties.

Moisture content inferred from the resistivity Monitoring performed at Red Soil Road Embankment in Molo Olenguruone, Kenya, in the period April 1993 - August 1994 is presented in the following Figure 16.17.

16.5 Monitoring and management of switches and crossings

16.5.1 Introduction

Switches and crossings (S&C) are often found to present a very large problem with sometimes as much as 25 % of the total M&R budget being spent on them each year. Practically all the railway companies, regardless of their organisational structure and technological level, permanently face this problem of cost.

There have been attempts to resolve this problem by applying more or less sophisticated S&C Maintenance Management Systems, some of which where computer-aided and some paper based. However, although these systems often managed to alleviate the problem to a certain degree, they never seemed to solve the problem completely. The reason for this could be that the general methodology for S&C Management has never been fully developed and defined. On the other hand, this could well be the consequence of the fact that the diversity of the technical solutions of S&C is far larger than in the case of plain track. This diversity of solutions has to a large extent hampered the development of a consistent methodology and clear guidelines for S&C management.

Following the concept of condition-based maintenance management, the key steps that would have to be taken in order to achieve better S&C management would be:

- Define the critical elements of S&C that account for most of the S&C failures and maintenance expenditures and represent the biggest danger to the safety of traffic;
- Define key parameters which define the condition of S&C or their elements in an as universal as possible form for various S&C technical solutions;
- Investigate best possible methods for (possibly continuous) monitoring and data-collection of the previously defined parameters;
- Investigate and evaluate existing maintenance activities applied on S&C and the possibilities for their improvement in order to make them more condition-based rather than suited to the predefined fixed time intervals;
- Investigate possibilities of defining "rules" (similar to those used in ECOTRACK) and thresholds that would be based on the collected data and captured deterioration pattern and would govern the subsequent decision-making process directed towards improving safety, reliability, and cost-effectiveness of S&C.

Future will most probably witness the entire S&C Management being done using exclusively automatic S&C condition monitoring. Nowadays however, although the capabilities and reliability of the automatic S&C condition monitoring are rapidly increasing, it seems that it still can not be entrusted with the total S&C management. Rather, a fair amount of standard (but computerised) geometry recording and visual inspections still has to be regularly performed and coupled with automatic S&C condition monitoring in order to arrive at the optimal S&C management system.

16.5.2 Switches and crossings monitoring by EURAILSCOUT

S&C represent highly important object of railway infrastructure. On NS network for example there are some 7000 S&C in place, which must be regularly inspected in order to be properly maintained.

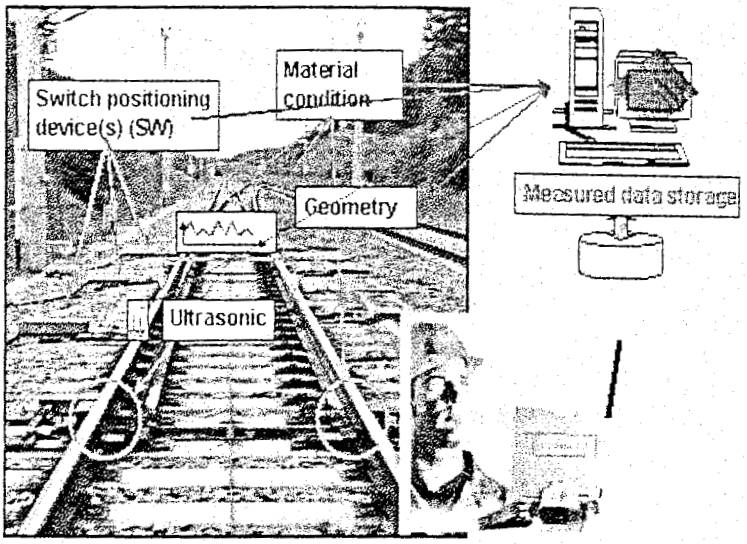


Figure 16.18: Different types of inspection performed on S&C

Therefore, a systematic approach to the S&C maintenance management had to be taken. All the monitored places on a switch have been presented on the Figure 16.18.

Beside the previously described sophisticated UFM 120 measurement and inspection vehicle, Eurailscout also works with hand held equipment to conduct important measurements on rail geometry and ultrasonic rail inspection of S&C. The mobile Rail and switch equipment SGMT5 measures, records and presents the rail geometry and switch parameters like gauge, superelevation, twist, the switch and the distance covered. MT95 ultrasonic hand equipment is used to inspect tracks for hidden faults. The MT95 equipment is primarily used to inspect critical spots such as welds, switch points and rail joints. Hidden faults in the rail, reported by other, high speed, measurement systems, can also be evaluated precisely with the MT95 in order to take all necessary correction measures.

After all the data from the various monitoring devices have been collected, they are processed and displayed by the SwitchView software program, developed by Erdmann Softwaregesellschaft (Figure 16.19 and Figure 16.20).

The goal of SwitchView is to bring together all the different data collected about the S&C into one system and to present them to the infrastructure manager in an obvious and user-friendly manner to facilitate decision-making about the S&C maintenance management.

Figure 16.20 shows the measured data for two switches performed in two consecutive time sections, showing the change in quality that occurred between them.

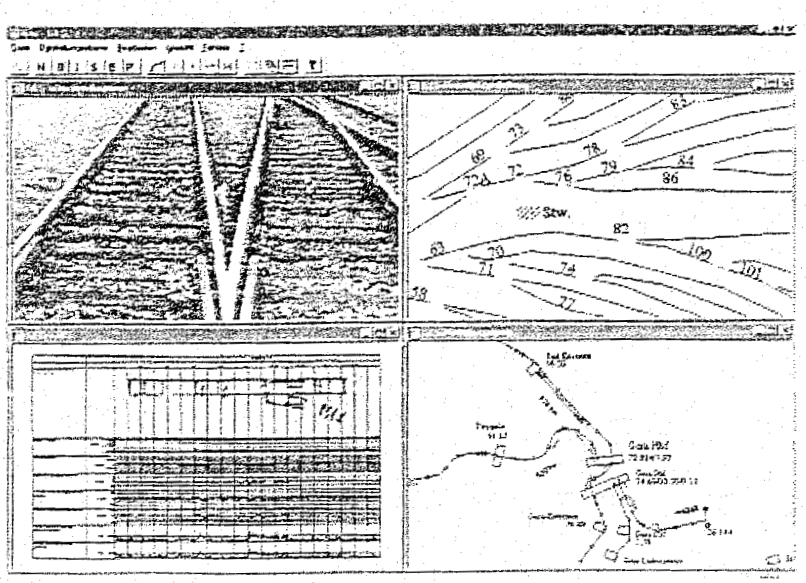


Figure 16.19: SwitchView example

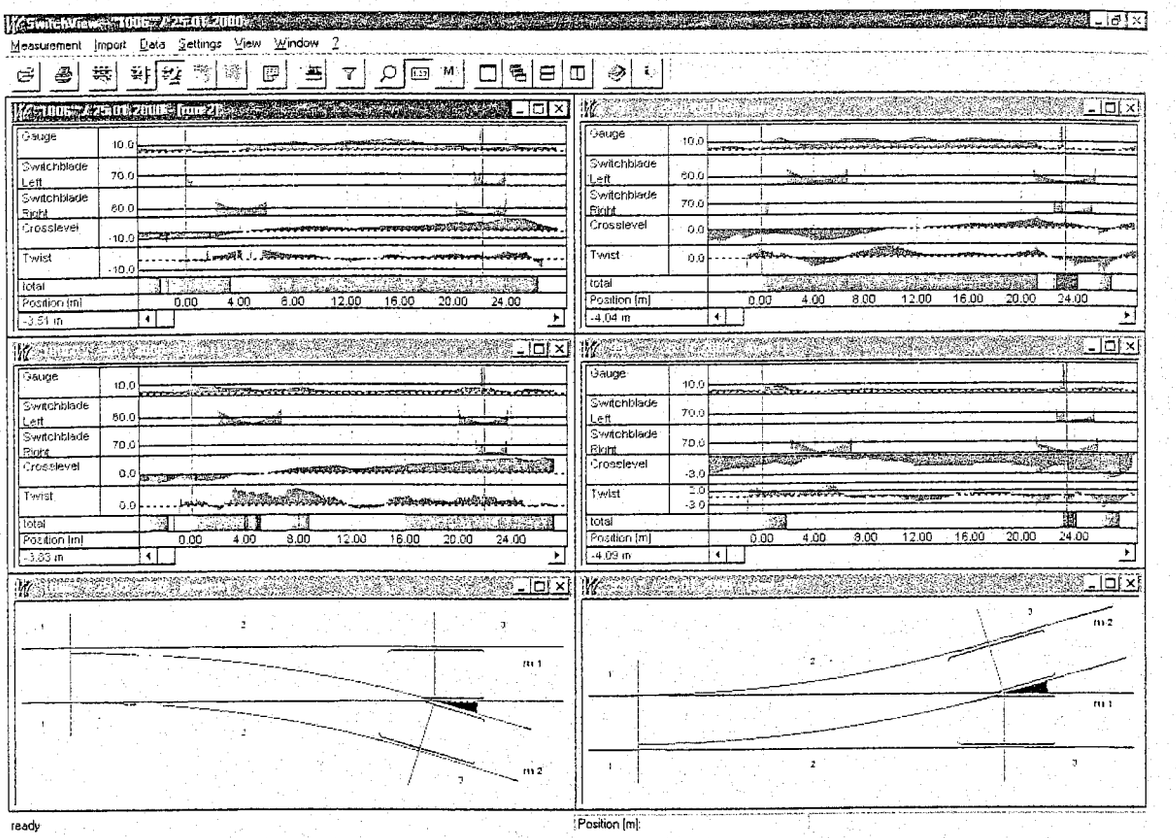


Figure 16.20: Measured data for two switches performed in two consecutive time sections

16.5.3 SwitchView

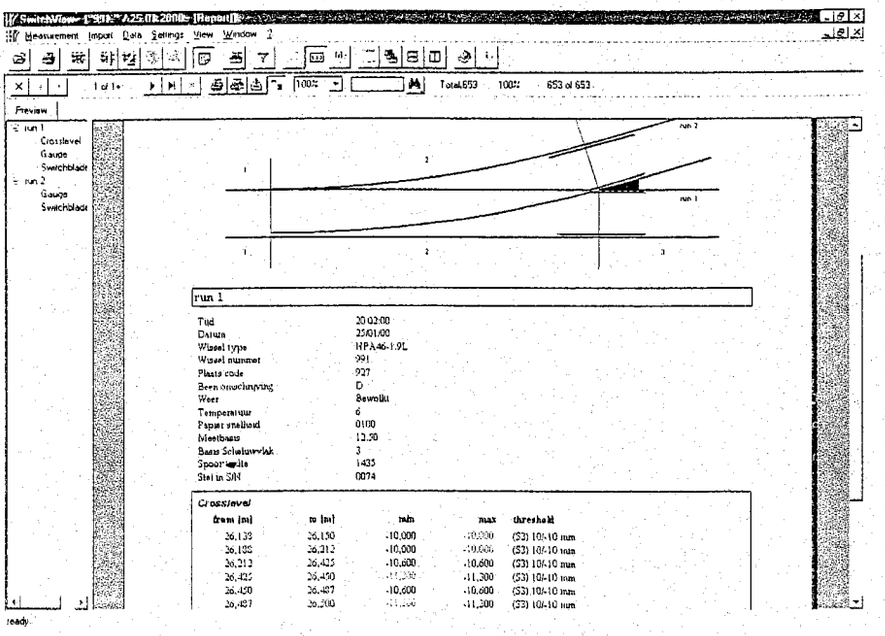


Figure 16.21: SwitchView alphanumerical print-out

SwitchView represents a software program for viewing switch inspection data in a clear and understandable fashion for the purposes of further analysis (Figure 16.21).

Switches are very important assets within the rail infrastructure. On Netherlands Railways, for example, there are some 7000 switches, which at regularly base need to be inspected on the basis of maintenance regulations and contracts.

An unambiguous inspection method, accompanied with an optimal data processing contributes to an efficient conduct of maintenance.

The main purpose of SwitchView is to gather various kinds of switch inspection data in one system, in order to present them in a well-organised manner, as a support for the conducting of maintenance and renewal (Figure 16.22).

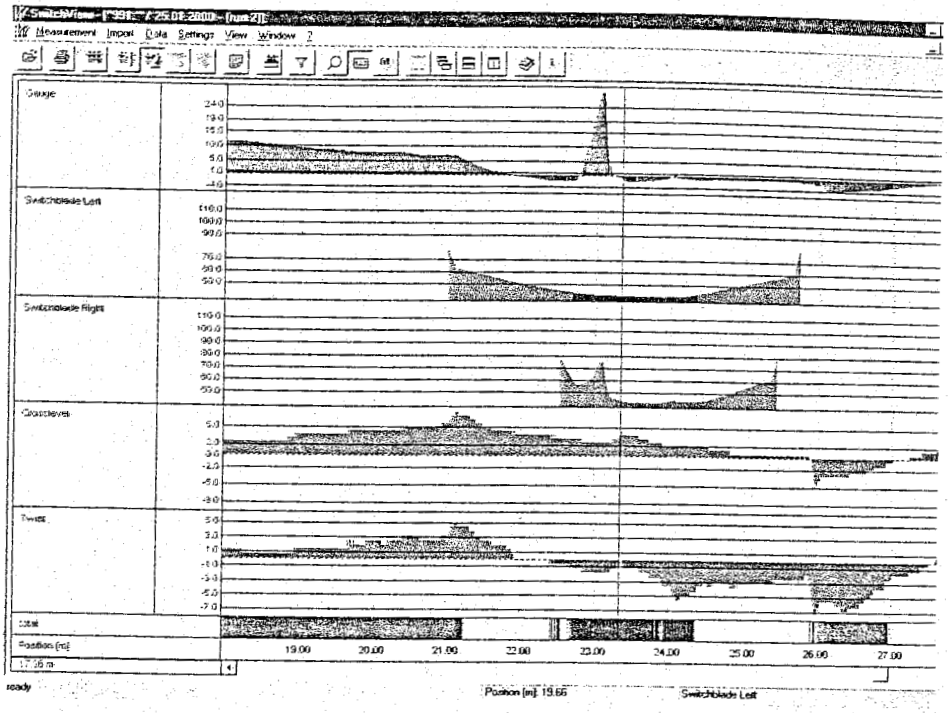


Figure 16.22: SwitchView switch quality parameters display

After reading in the data, SwitchView displays in various colours, the location and the gravity of the Exceedence at the corresponding switch symbol. It can also display the measured values of the track gauge, left and right groove width, superelevation and twist per each and every measurement pass. Several ways of presentation are possible, all in order to enhance the data analysis. Of course, it also possible to print the data in a form of a report(s), in different shapes.

Finally, the user must be able to navigate from a rough to a fine-meshed levels within the available information, for instance from date of construction to the current state of the switch.

16.5.4 Condition monitoring and maintenance management of switches

The most common general needs of railway companies meant to achieve improved returns on their investments in the RAMS (Reliability, Availability, Maintenance, and Safety) area can be recognised as follows:

- Cost reduction of maintenance by introducing condition monitoring where technically and economically feasible in order to decrease the number of necessary preventive maintenance (PM) actions;
- Cost reduction of track maintenance by applying an optimization procedure which takes the Life Cycle Costs of all the equipment used and of all the maintenance activities applied into account, and helps to select the most appropriate maintenance activity and determine the optimal PM interval;
- Necessity of keeping the safety level unchanged.

As these tasks are very complex and difficult, they have prompted the European Commission to support a research project called REMAIN within the Fourth RTD Framework Program (1994-1998) [234]. The aim of this project was to create practical and realistic methods and tools to evaluate RAM (Reliability, Availability, and Maintenance) in railway systems.

Subtasks of this project were the definition of requirements for those railway components which are crucial for the safety and efficiency of rail transport, the development of a system that would use up-to-date information in combination with historical data in order to predict future statuses, and the strategic planning of maintenance operations.

The main vessels used to achieve the preset goals were:

- Systematic analysis of maintenance needs and organisation, e.g. using Reliability Centred Maintenance (RCM);
- Introduction of condition monitoring methods;
- Life cycle cost (LCC) analysis in order to confirm the anticipated effect on costs and delays when utilising these methods. Safety analysis methods were used to confirm that there had been no reduction in safety level.

One of the first steps in fulfilling the objectives was to analyse the distribution of failures of various infrastructure components provided by REMAIN partners, both on conventional railway lines and high- speed railway lines, and detect which components accounted for most of the problems. The statistical data shown on Figure 16.23 and Figure 16.24 include all kinds of failures caused by the equipment itself. Failures due to external causes (i.e. damages produced by maintenance machinery) are excluded. As can be derived from Figure 16.23 and Figure 16.24, most of the failures (ca. 55 %) are caused by signalling equipment and turnouts. ("Signalling equipment" here includes signals, track circuits, interlockings, ATP or LZB, and the traffic control centre.)

Since switches were found to be important components in the condition monitoring approach of any railway, and since they obviously were a frequent cause of failure in both conventional and high speed lines, switches were selected as the key elements for further investigation within the REMAIN project.

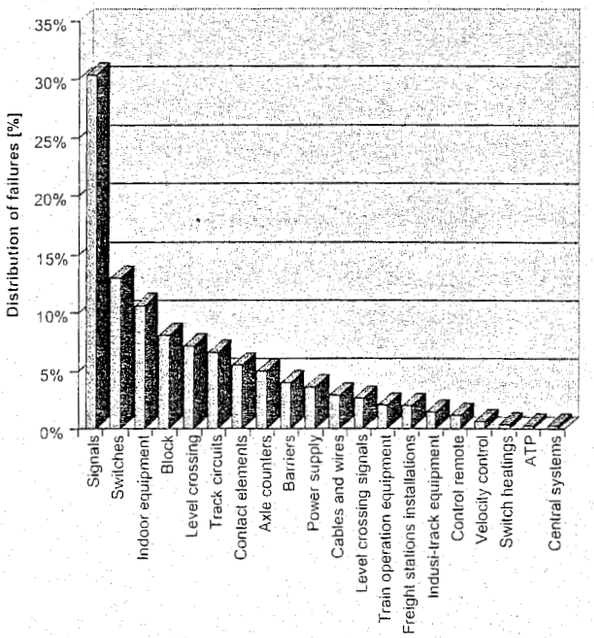


Figure 16.23: Distribution of failures of track infrastructure components (conventional line)

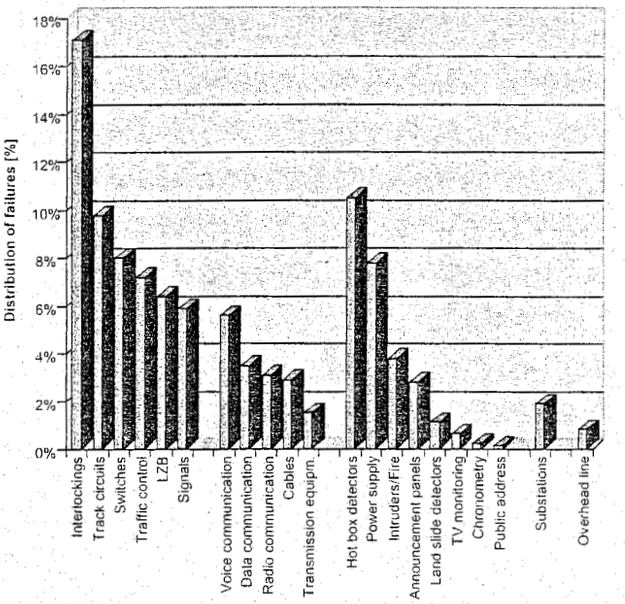


Figure 16.24: Distribution of failures of track infrastructure components (high-speed line)

Condition monitoring and maintenance management system concept

The proposed system comprises three major functional components:

- Monitoring based on:
  - Acquisition of data, their storage, and retrieval of information, and
  - Communication of data and information needed for RAMS management;

- State Prediction, using:
  - Signal analysis and pattern recognition for components;
- Maintenance Management using:
  - Strategic planning of maintenance,
  - Introduction of condition monitoring,
  - Observation of safety and cost (LCC) restraints.

Condition Monitoring of Track Components

Usually, inspections and M&R actions are scheduled at fixed time intervals. In order to not endanger the availability and safety, these intervals are often made too short. On the other hand, since inspections usually require a certain amount of human interaction, they are often quite expensive. Therefore, a lot of money could be saved in the long run if the time between inspections could be extended without effecting the availability and safety. At present there is often no up-to-date information available on the true state of components between inspections, which is another reason for scheduling inspections at fixed intervals. However, if information about actual wear would always be available, at least for the critical components, the time between complete inspections could be extended and adapted to the actual need.

Concepts for monitoring and diagnosing the related components are discussed in the following section based on a list of about 30 switch and track failures compiled by VAE. Some of the failures can already be detected by the Roadmaster 2000, while the automatic early detection of others is still a problem that has to be investigated further.

The values of measurements at a component and the allowable intervals between inspections while maintaining the necessary level of safety and availability, can sometimes be derived from a mathematical model, but it is often necessary to base them on the experience of a human expert. Data supplied by experts should not replace operational data, but should represent a valuable supplement to data measured during operation as well as give essential information in cases when operational data are not available.

The list of about 30 discovered switch failures can be subdivided into 3 groups:

- Failures that cause looseness resulting in wrong geometry;
- Malfunctioning of moving parts, e.g. the failure of a switch motor;
- Failures due to wear of rail or frog point.

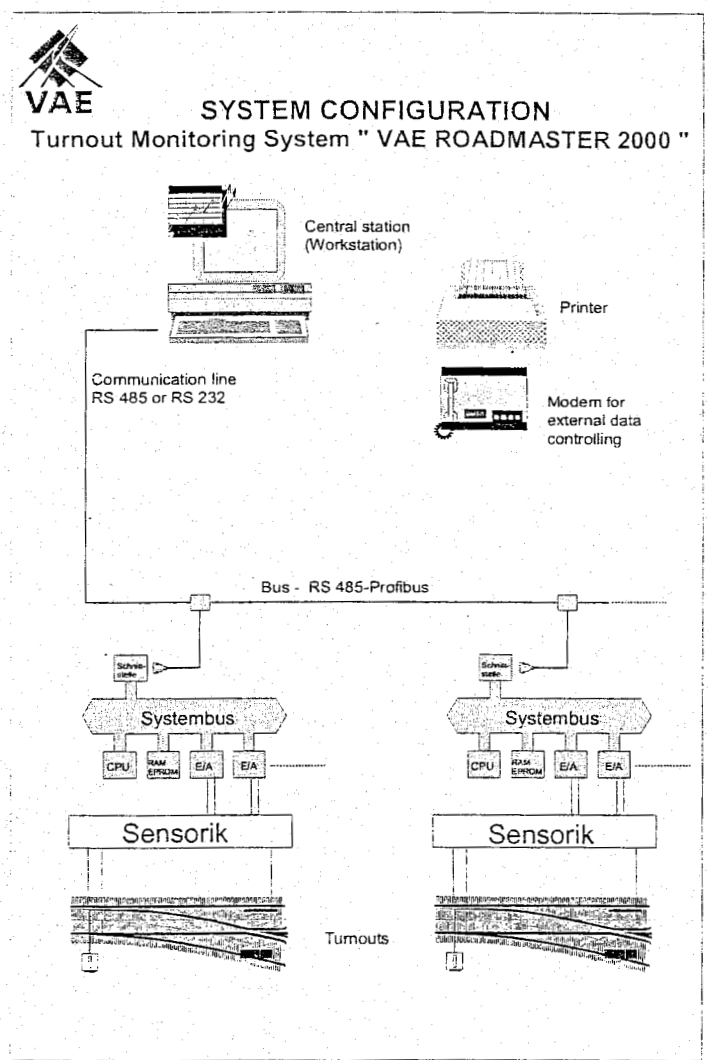


Figure 16.25: An overview of the VAE Roadmaster

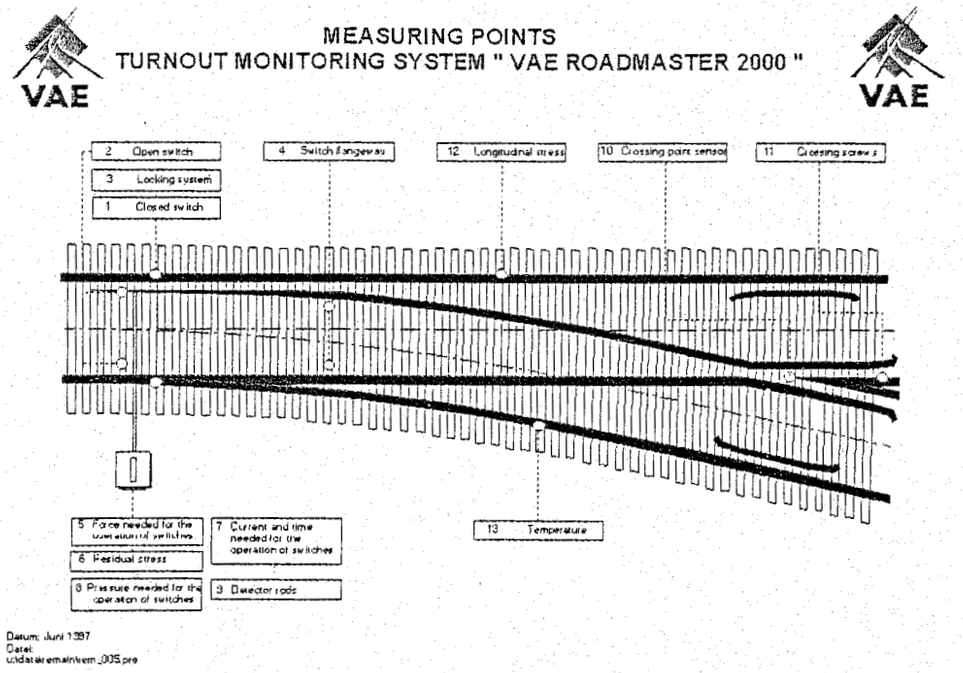


Figure 16.26: Location of the measuring stations in the turnouts (switches)

Measurement Principle of Remote Supervision

The overall function of remote supervision is split up into several levels and is achieved by means of the combined use of hardware and software. An overview of VAE Roadmaster 2000 is given in Figure 16.25.

On the acquisition level, the measured data are collected by appropriate sensors and analogue systems in a digital electronic collection system with bus modules. The data are then preprocessed and prepared for transmission. On the user level, the measured data are visualised in the form of tables and graphs. The result is that the output of such messages is suitable to recognise trends.

Measuring Points in the Turnout

The locations of measuring points are illustrated in Figure 16.26. The measuring points are:

- Monitoring of switch-stock rail contact area;
- Monitoring of open switch;
- Monitoring of switch operating rods and/or switch locking system;
- Monitoring of minimum distance between open switch and stock rail (switch flangeway);
- Direct measurement of the force needed for each single operation of switches;
- Residual stress in the switch and/or the rodding (retaining force);
- Current and time needed for the operation of switches;
- Monitoring of the pressure needed in the switch machine for the operation of switches;

- Monitoring of the position of the detector rods;
- Monitoring of the strikes at the crossing point, indicating wear of the check rail and/or wing rail;
- Monitoring of screw preloading forces in bolted compound crossings;
- Longitudinal forces in the rail;
- Rail temperature and ambient temperature.

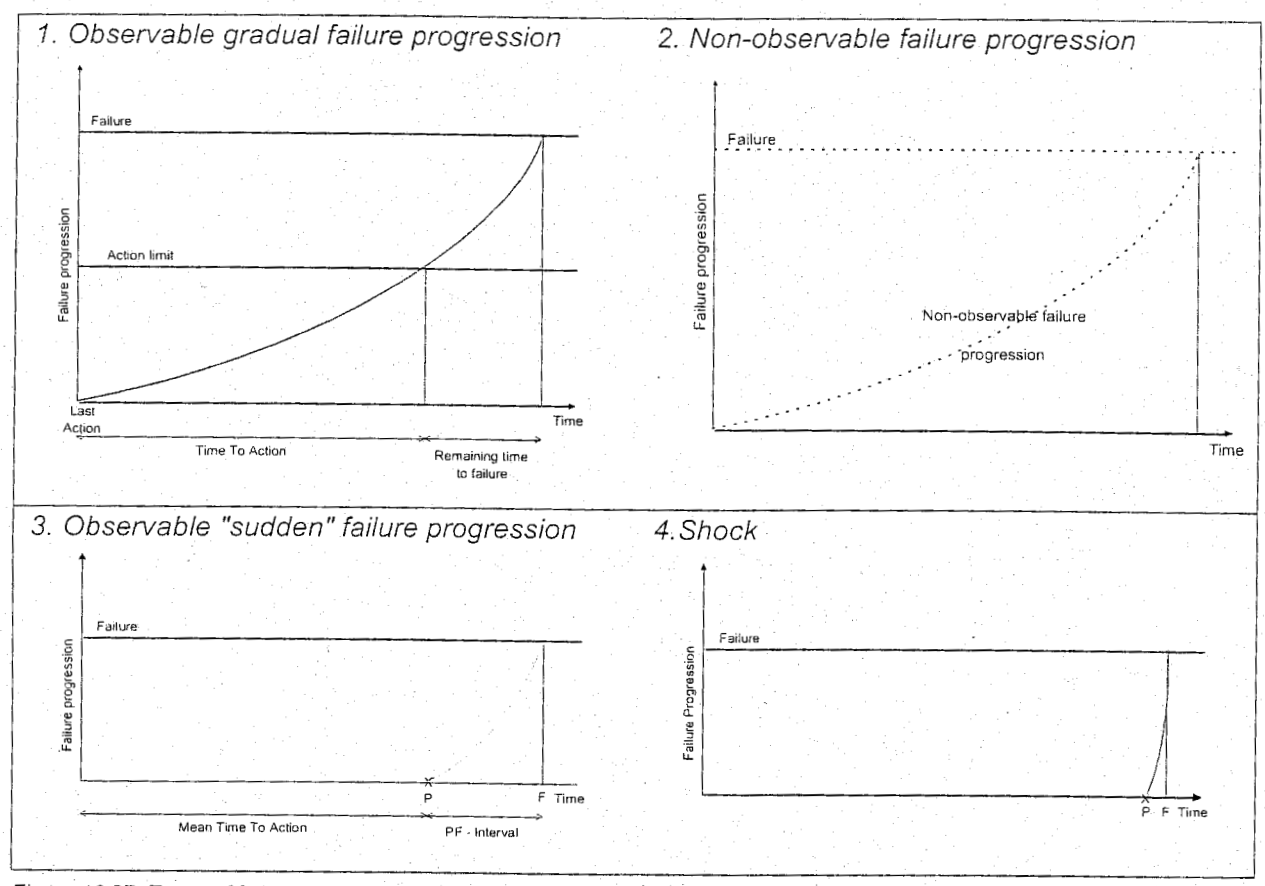


Figure 16.27: Types of failure progression

Results of example analyses

The type of failures are summarized in Figure 16.27. During the analyses it was found that some of the sensors in the Roadmaster 2000 system were superfluous, i.e. they did not prove to be cost-efficient. Figure 16.28 summarises the annual contribution to the LCC for various options with respect to including Roadmaster 2000 sensors. These data were obtained for a conventional line. In Figure 16.27 the "Present" bar represents the annual LCC prior to the installation of Roadmaster 2000. This is the "baseline" system, representing the period prior to the introduction of Roadmaster 2000 (in 1997).

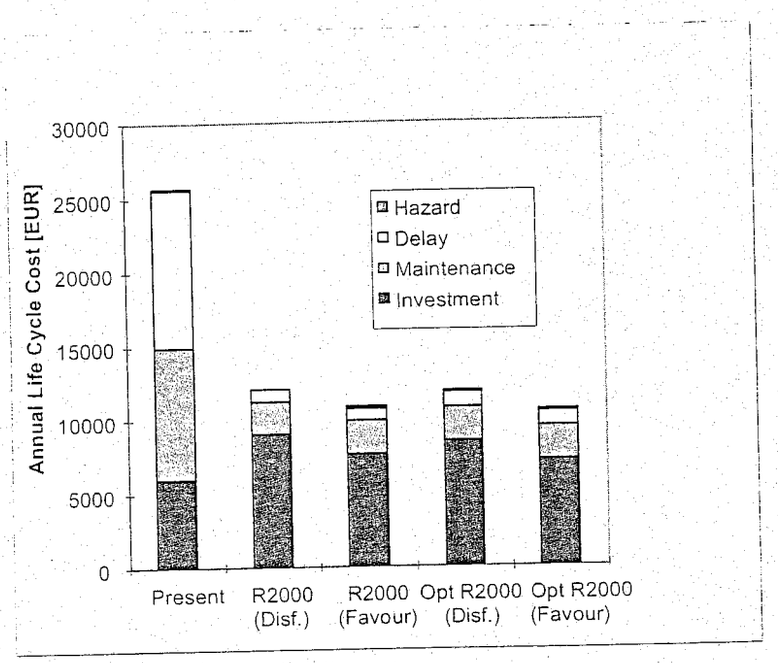


Figure 16.28: Summary of LCC results (conventional line)

The "R2000" bars represent the costs if Roadmaster 2000 were to be installed today. The "Opt. R2000" bars represent an optimized version of Roadmaster 2000 in which some of the sensors are omitted from the installation in order to minimise annual LCC. For each of these two options two different LCC calculations (representing a sensitivity analysis) have been performed. The calculation labelled "Disf" is based on assumptions (about service lives, etc.) that disfavour Roadmaster 2000, whereas the calculations labelled "Favour" are based on assumptions more favourably inclined towards using Roadmaster 2000.

In the illustration of Figure 16.28 there are four columns that refer to the introduction of condition monitoring, and one column that illustrates the costs prior to the introduction of R2000. Thus, the possible further reduction in LCC by using RCM analysis is not illustrated here.

The results of the analysed conventional line showed that introducing condition monitoring with Roadmaster 2000 could bring about possible changes that could lead to a reduction of LCC costs by a factor of 2.

16.5.5 CEDIAS - Railway Lines Diagnostic System

A CEntralised DIAgnostie System (CEDIAS) has been developed for the railway lines in Italy [233]. The system is installed in the Terontola-Foligno line of the Ferrovie dello Stato spa (FS) in the Centre of Italy for the centralisation of monitoring of the several parameters of the line. The goal was to obtain real time information about the efficiency of the line signalling equipment, track and civil works.

CEDIAS is based on diagnostic modules called ADAM (Automatic Diagnostic Acquisition Module) which are installed in each station of the line. ADAM is composed of a fast data acquisition unit, which collects data from the station hardware (switches, signals, fixed and coded track circuits, UPS, ancillary circuits, meteorological data, etc.) in real time through non-invasive sensors. A computer, with a use of diagnostic software, evaluates collected data. Further, ADAM processes data from sensors along the line (rail temperature, lateral and longitudinal track displacement, level crossing protection systems, track lane inclination and landslide, bridges and tunnels). Data are evaluated by a computer using specially developed diagnostic software.

ADAM currently consists of two computers (Figure 16.29), and in future it will be possible to use only one computer. The first unit is used for the acquisition of station and line sensors' signals and their processing using specific algorithms. The second unit is used for the data transmission with a periodic time interval to a Centralised Unit, which controls the whole railway line. The system is fully operational and has been commissioned at the end of 1998. Four years of R&D activity served as the base for the design.

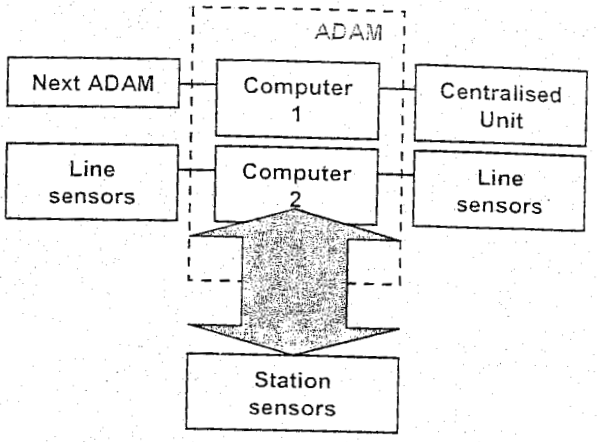


Figure 16.29: Organisational structure of ADAM

The development of the predictive software is the most important and critical point. Detection of the failures is fairly easy, while the prediction of the same failures is much more difficult. The prediction of preventive maintenance in track circuits with fixed and coded currents can be made with sensors for the measurement of the time variations of currents and voltages. Should the currents and voltages exceed some given values, a warning message is issued, asking for a specific maintenance of the circuits.

For other railway equipment, the prediction is more difficult. In the case of the switches for example, it is necessary to keep under observation the temporal behaviour of current and power consumption for each operation, normal or reversed, of the switch in order to determine the appropriate maintenance eventually required.

The centralisation of the diagnostic system is intended for the purposes of a remote control of the railway equipment. The centralisation is performed at different levels: at a line tract, at a full line, at a regional and at a national level.

16.6 Ultrasonic rail inspection

16.6.1 Introduction

Ultrasonic rail inspection is the common method to check rails for internal defects. This type of inspection was first introduced for safety reasons, with the aim of detecting internal failures which could not be seen from the outside. Another increasingly important aspect concerns detection of defects at an early stage so that repairs can be scheduled in the regular process, without interrupting train operations. For this reason, emphasis is laid on the monitoring function and on the supply of data in order to plan activities.

Figure 16.30: Ultrasonic hand equipment

In principle there are two ways of carrying out ultrasonic inspection: with hand equipment and using a special ultrasonic train. To inspect large networks efficiently, operation of an ultrasonic train is of course inevitable. The measuring speed depends on the type of system and the track quality, but is in general limited to 75 km/h.

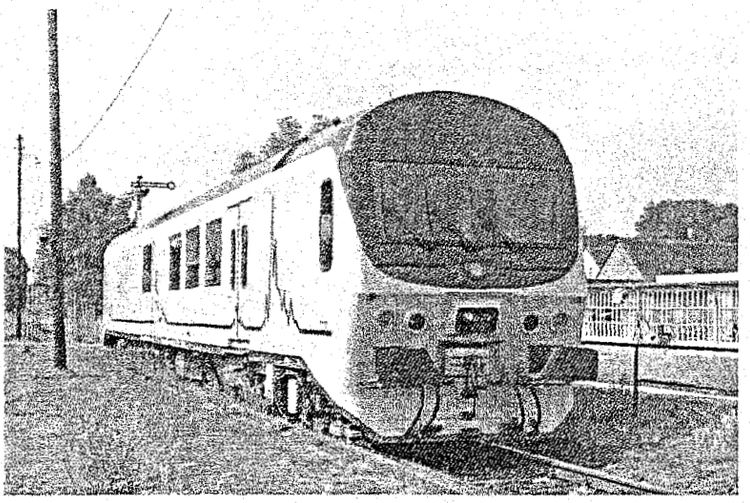


Figure 16.31: EURAILSCOUT Ultrasonic measuring car UST96

Special structures, such as switches and transitions for moveable bridges, not to mention junctions and sidings, are normally inspected with ultrasonic hand equipment as shown, for example, in Figure 16.30. Moreover, it is common practice to use hand equipment to check indications given by an ultrasonic train before the final decision is made on removing a defect.

A number of inspection trains are in use for ultrasonic rail flaw detection. In Europe Railtrack, DB, SNCF and NS have their own cars, while Matix offers the possibility of renting their car, as do DB and NS. Outside Europe Sperry is a well-known firm which hires out ultrasonic trains. Figure 16.31 shows the ultrasonic train UST96 used on the NS network and operated by the firm EURAILSCOUT Inspection & Analysis.

16.6.2 The EURAILSCOUT ultrasonic train

Description of the ultrasonic train

Since 1976 Netherlands Railways have been using a special train with ultrasonic equipment for the detection of rail defects, now denoted to as the "UST76". After having been in service for over 10 years the electronic part became obsolete and had to be replaced. At this stage the whole measuring concept was critically re-examined with respect to quality and reliability, primarily with the aim of reducing the total cost of ultrasonic inspection. This investigation had resulted in a new concept consisting of a computer-controlled measuring system, to which an off-line, but on-board data analysis system had been added. The main parts of this system were completed in spring 1989.

In 1997 a third generation of the Ultrasonic Rail Inspection System, the UST96, depicted in Figure 16.31, was commissioned. The UST96 is an inspection vehicle for ultrasonic inspection of tracks with great precision at high speed and is specifically used to detect rail defects. The UST 96 is equipped with its own diesel hydraulic power and can be used on all standard gauge lines. It has been put into action on European railroad tracks (Belgium, Switzerland, Germany, Denmark, Italy, Norway, the Netherlands and Spain). The specifications are represented in Table 16.1.

- Year of construction: 1996;
- Measurement speed: 100 km/h (maximum);
- Transport speed: 140 km/h (maximum);
- Traction: independent, diesel hydraulic;
- Length buffer to buffer: 24.5 metres;
- Track gauge: 1435 mm;
- Wheel base: 2.5 metres;
- Distance between bogie centres: 17.2 metres;
- Weight: 57 tonnes:
- Water capacity: 6000 litres;
- Diesel fuel: 1000 litres;
- Measurement and analysis system: 16 channels, computer-controlled, eight transducers per rail, 2 x 0° SE + 70°, -70°, +70° satellite, -70° satellite, +35°, and -35°;
- Data storage: CD-ROM.

Table 16.1: Technical information of the EURAILSCOUT UST96 train in a nutshell

With its own diesel engine it can run with a maximum speed of 140 km/h and measurements are carried out at speeds up to 100 km/h. The UST 96 has been equipped with the latest in measuring and computer technology referred to as the URS (Ultrasonic Rail inspection System), developed by the TNO Institute of Applied Physics (TNO-TPD). The rail is inspected at different angles with the help of 16 special ultrasonic detection heads.

The control electronics ensure permanent optimisation of the signal levels, the rail profile type is automatically taken into account in the optimisation. It is also completed with a new measuring system for detecting so called "head check" flaws. Accurate determination of the position on the rail track is done with a wheel encoder (odometer), together with magnetic sensors, D-GPS satellite navigation and manual input of location. Measurements are carried out every 2 or 3 mm. The accuracy of the global position of reported defects is about 1 meter. Depth of a defect is reported with an accuracy of about 1 mm. The defect reporting performance is very high: more than 95 % of reported defects are confirmed to be defects. Well controlled calibration runs demonstrated the outstanding performance of the system. Data are stored digitally and can be imported or exported to the data bases. Digital and/or paper reporting is provided to the customer within 24 hours of the inspection run. Critical anomalies may be reported within a few hours, or even minutes, after inspecting a section. Several hundreds of kilometers of track are inspected and reported each day, depending on the quality of the track.

16.6.3 Architecture of the URS

The UST96 train houses the URS inspection system. This system is composed of two subsystems: the Incident Detector (IDS) and the Report Generation System (RGS), which are both being housed onboard the UST96. More precisely, the route from the rails to the reporting encompasses the chain presented in Figure 16.32. Besides these parts, there is also the CAN subsystem, which consists of two dedicated keyboards and an LCD display. This is used for efficiently entering special codes, markers, location numbers and controlling the measurement parameters while measuring. The on-line real-time parts SEL+IPR+WST+CAN together constitute the IDS (Incident Detection System), as opposed to the RGS, which represents an off-line PC system connected to the IDS by a network link onboard the train.

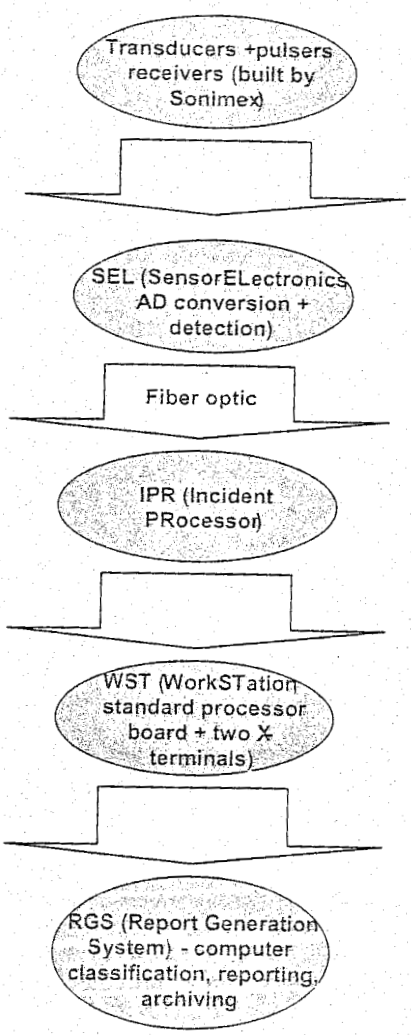


Figure 16.32: Process flow within the URS

Within the IDS, the SEL, IPR and CAN controllers consist of dedicated electronics designed and constructed by TNO TPD. In fact, there are two SEL's, one for the left and one for the right hand rail. The SEL's have 8 processor boards each with analogue gain components, AD converters and dedicated processors. On the motherboard of each SEL there is a DSP to set parameters, get data off the processor boards, and handle the data and control transfers over the optical serial links.

The IPR is a VME-bus card are sitting in a VME crate together with the processor board. The IPR is an embedded DSP-based computer itself and combines all SEL channels of both SEL's, puts data in hardware fifo's, reads real-time markers, and communicates over the VME bus to the WST.

The Workstation (WST) refers to the VME-bus crate with the processor board. It is a standard DEC-Alpha AXP single-board computer. It runs the Digital Unix operating system and deals with all the pre-processing, data display, graphical user interfaces, on-line data storage, and serial communications. It has neither display nor keyboard of its own. Via the onboard network two X-terminals are connected to this computer displaying the two user interfaces: 1. data visualization (rail plot) and 2. control. The two SELs and the CAN are controlled via separate (optical) serial links. Also, a 3rd party DGPS system is connected via a serial link.

The data connection between sensor electronics and SEL's is analogue, whereas the data connection between the SEL's and the IPR is digital using optical fibre.

In addition, analogue signals from the transducers to the operators desk are provided via an optical link. A digital storage oscilloscope is connected to these outlets.

The software consists of real time, soft real time and batch processing modules. For the IDS there is software running on the SEL, on the IPR and there are dedicated processes running on the WST communicating over unix sockets with multi-processor scalability. On the RGS, software is running for visualization, automatic classification, interactive classification and reporting. Expert systems like rule-based classification processing and case-based classification processing have been implemented [138]. Dedicated multi-lingual reporting modules are present, developed according to the wishes of the different international end-customers. The software on the RGS is running under Windows.

16.6.4 Probe system

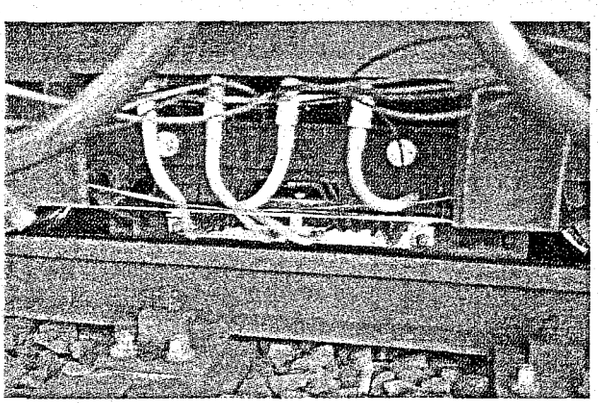


Figure 16.33: Probe system

The UST96 train has two probe blocks per rail containing two 0°, two 70°, two 35° transducers and two extra 70° transducers for inspection of the inner side of the rail head. To achieve the optimum signal-to-noise ratio, the electronics for the transducers are mounted in a housing on the bogie adjacent to the transducers. Figure 16.33 shows the probes mounted in the carrier, with the black BNC cables running to the pulser/receiver boxes. The 4 thick tubes supply the water.

The probe positioning and guidance system has been developed in collaboration with the NS Rolling Stock Division. The system ensures optimum contact with the rail and the correct positioning of the transducers on the rail at high speeds. The carrier is positioned relative to the centre of the rail by means of a pneumatic system. During measuring, the guide wheel pushes against the running edge of the rail head and in this way fixes the position of the probes. During transport the probe carrier and the guide wheel are free from the rail. The probe system is mounted in a small trolley fixed to one of the bogies. This trolley always remains on the track.

Water is used as contact fluid. For this purpose the train carries a water capacity of 6000 liters. The water is transported under pressure to the chamber in the probe carrier, which is sealed by a rubber ring. In this way a water film is formed between probe and rail to create the acoustic connection.

The 0°-probe is used to scan part of the head, the whole web and the central part of the foot perpendicular to the rail as shown in Figure 16.34. Due to the scattering of the transmission pulse in the water film and the properties of probes currently used the defect echo's in the upper part of the rail head cannot be detected directly. This zone, formed by the transmission pulse area and the water interface area, is the so-called dead zone and can be reduced to about 5 to 8 mm by applying the SE technique.

The 70°-probes have an intermediate layer in the perspex probe block. The test area is formed by the central part of the rail head and part of the web. The maximum depth in the perpendicular direction amounts to about 90 mm.

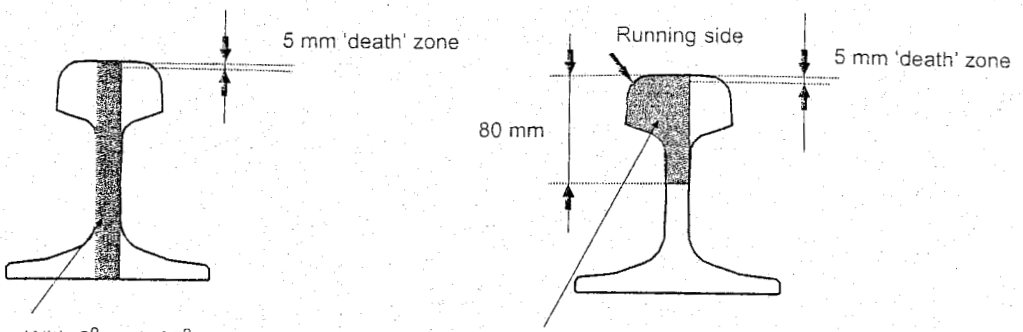


Figure 16.34: Ultrasonic measuring areas

The signal-to-noise ratio is continuously monitored so that the equipment retains its optimal settings. The measurement interval is automatically set according to the type of transducer. Depending on the measurement speed a measurement interval of 2 or 3 mm is chosen. A defection area can be set for each transducer type. The IDS automatically changes measurement interval and detection area parameters when going from one rail profile to the other.

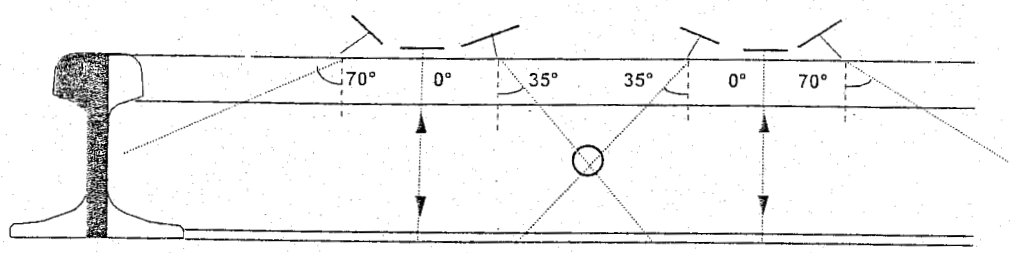


Figure 16.35: Transducer arrangement of the rail testing vehicle UST 96

The transducer arrangement of the UST96 is presented in Figure 16.35. The echo pattern, as presented on an oscilloscope, is shown in Figure 16.36.

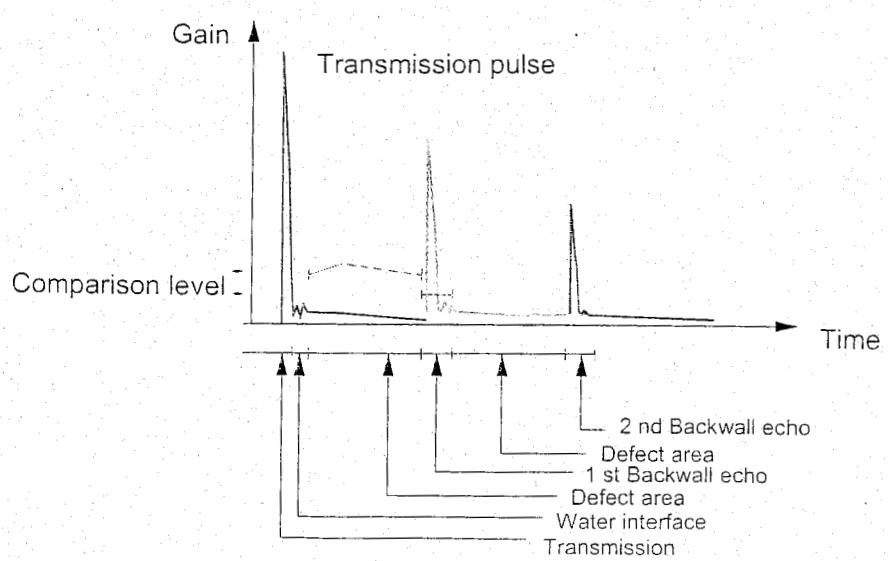


Figure 16.36: Ultrasonic echo pattern

The following four areas can be distinguished:

- transmission pulse area, in which the transmission pulse is emitted and the receiver amplifier is still overloaded;
- water interface area, in which the sound strikes backwards and forwards between probe and rail surface. This effect leads to pulse elongation for both the transmission pulse and the echo signals, which limits detection of defects at the top of the rail head;
- defect echo area, in which defects can be detected directly. The 0°-probe scans these cracks in the central part of the head, the web and the foot. The 70°-probes scan transverse cracks, whether perpendicular or slightly inclined, in the head and the upper part of the web;
- backwall echo area. This concerns exclusively the 0°-probes. The occurrence of multiple echo's means there are more backwall echo areas. However, only the first defect area is used for detection.

16.6.5 Sensor electronics

The sensor electronics consist of the pulser/receivers and the dedicated SEL hardware. The following parts can be distinguished:

- Pulser/receivers. These are the probe electronics, which cover all the transmission and receiving circuits for the probes. For an optimum signal-to-noise ratio these are mounted in a special housing on the guiding trolley, near the probes, in the grey boxes as shown in Figure 16.37. In this environment the electronics are subjected to large vibrations and great fluctuations in temperature and humidity. The transducers and pulser/receivers have been designed and constructed by the Dutch company Sonimex.

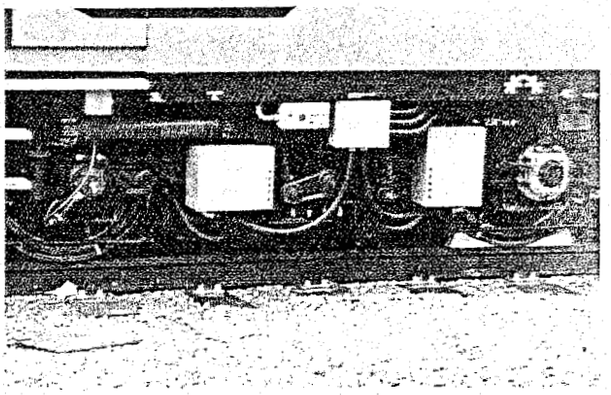


Figure 16.37: Pulser/receiver boxes (gray). Probes not mounted

- Amplification. On the SEL processor boards (Figure 16.38) there are both fixed gain and time variable gain (TVG) amplifiers present. Per channel, these amplifiers normalize the echo gain which is depth-dependent due to acoustic losses. The differences in electric and acoustic sensitivity of the individual transducers can thus be compensated for via software.
- Detection. The TVG-amplified signals are digitized and are compared to detection levels on the SEL processor boards. The primary function of detection is to reduce the high flow of echo data to exclusively observations of incidents. The amplification settings as well as the comparison levels and echo zone definitions determine signal detection. All parameters may be set on-line via software.

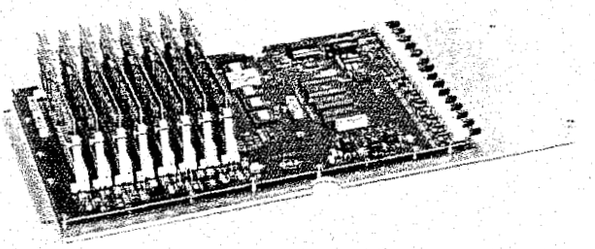


Figure 16.38: Single sensor electronics (SEL) motherboard with the 8 channel processors (left) and optical links (right) clearly visible. In the URS there are two SEL's present

- Incident generation. A channel incident is a 32 bit value representing the data of a detected event. It contains the incident type (here: "channel"), channel number, amplitude and time-of-flight of the event. The SEL central processor collects the incidents from each of the processor boards.

- Test signal generation. The SEL can re-generate pre-stored ultrasonic signals and treat them as coming from the pulser/receivers. Running the SEL in a special test mode, the data acquisition and processing chain can be tested and calibrated without the presence of the pulser/receivers and probes.

- Communication to IPR. The SEL sends the incidents to the IPR for further processing.

16.6.6 Incident Processor

The Incident Processor (IPR) is a dedicated VME card residing in the VME crate in a 19" rack. The IPR has a dedicated processor (DSP). The IPR provides all clock signals for the two SEL's, digitises the tacho signal, combines incidents from both SEL's and concentrates all the information into "incident reports". These reports are communicated to the main processing board via a FIFO data buffer and the VME bus. An incident is not restricted to just echo patterns of the defects in rails and constructions, but may also refer to marking signals generated by the system, the driver or the operator during the inspection run. These "real-time markers" are also inserted into the data stream by the IPR.

The transducers are pulsed synchronously using the tachogenerator, which is connected to a wheel of the measuring trolley and delivers one pulse per mm covered. Normally the transducers are driven every 2 or 3 mm covered. This situation always allows for inspection at the highest possible speed. The IPR computes the appropriate clock signals from the tachogenerator with which the SEL drives the pulser/receivers.

To test the system, the IPR has an additional FIFO buffer. In normal mode the data FIFO is used to transport data from the SEL to the main processor. In test mode the second FIFO is filled with previously recorded data by the main processor and is subsequently emptied by the IPR as if the data was originating from the SEL. In this way, the functioning of the data acquisition and processing chain can be tested from the IPR upwards, without the presence of the SEL, pulser/receivers and probes.

16.6.7 On-line control and data interpretation

An inspection run starts by filling in an administration page on the configuration and control console and printing a so-called configuration page containing all administrative and parametric data of the measurement run. Furthermore, on the configuration and control console, measurement parameters can be optimized during the measurements by using the special keyboard (Figure 16.39).

Figure 16.39: : Configuration and control console

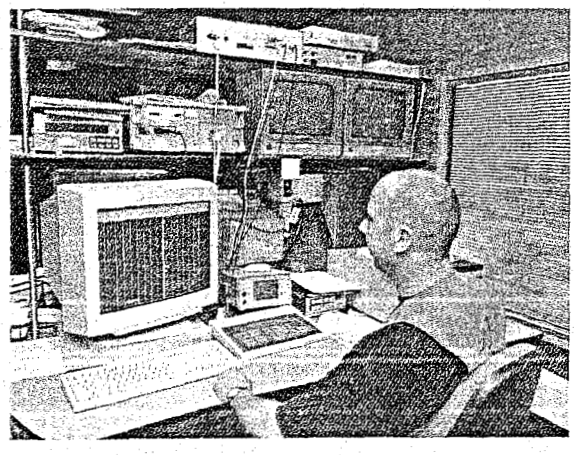


Figure 16.40: Observer console

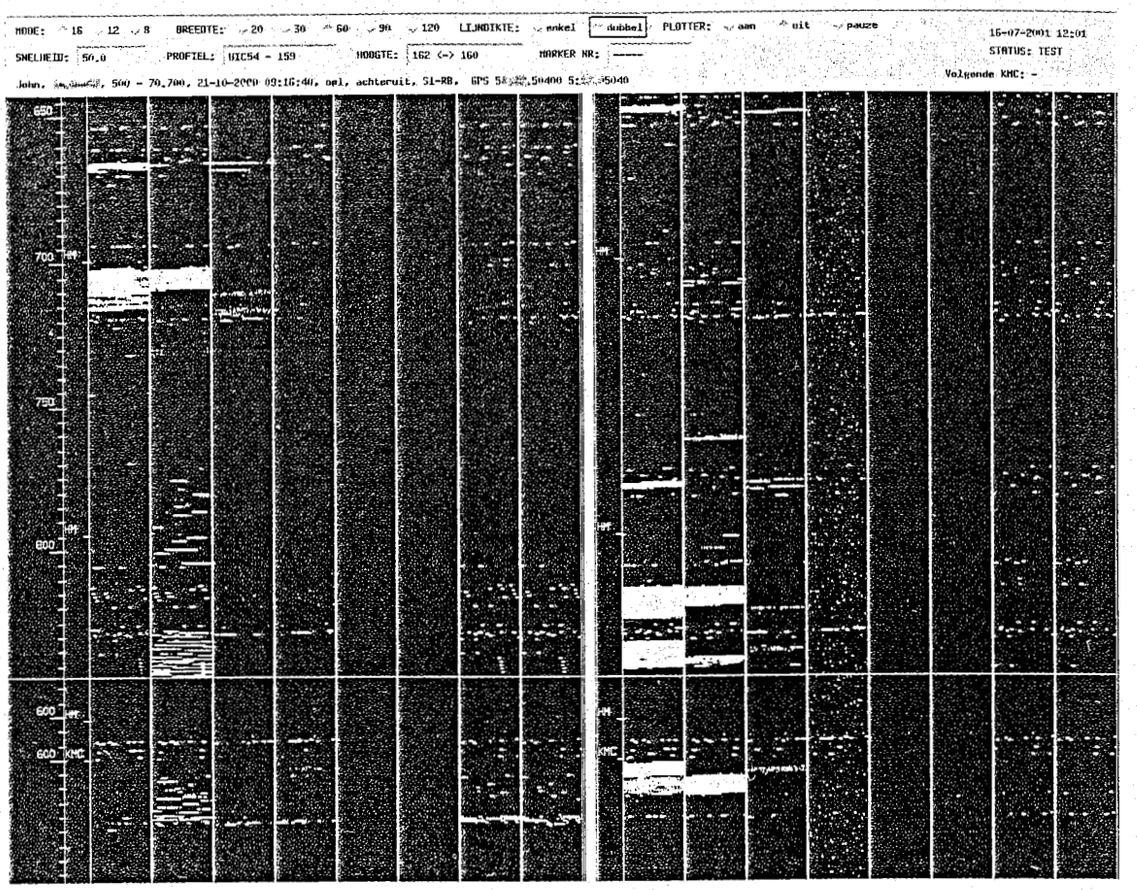


Figure 16.41: Example of on-line observer screen, as used for on-line data quality monitoring. White arrows are drawn over the screen dump to point at indications of joints with bolt-holes. In this screen, a small defect is difficult to spot. A trained and experienced expert, however, can identify all kinds of instructions like welds, bolt-holes, switches, certain types of surface wear, severe defects, etc. Note that for demonstration purposes a relatively old track was chosen: there are lots of echo's present. On modern rails like those of high-speed lines there are only few echo's visible, because there are no bolt-holes and few construction elements. In that case, anomalies are much easier to spot.

During inspection an on-line rail plot is monitored for the observer on a video screen (Figure 16.40). A large colour monitor displays the measurement data over several hundreds of meters of track on-line (Figure 16.41). Per transducer, the results of every 60 cm of rail are represented by a single horizontal line. There are sixteen channels depicted, eight for the hand rail and eight for the right hand rail. Per rail, the left most channel indicates the disappearance of the backwall echo from the rail bottom. The other channels indicate when an echo, appearing in the detection zone for that channel, exceeds the preset comparison level for the 0°, +70°, -70°, +30° and -30° transducers respectively. The operator can thus permanently monitor the quality of the measurements. Larger defects can be noticed immediately.

In the rail plot picture, the km-posts are reproduced at the extreme left. Some marks are also indicated, per rail. The marks are abbreviations for structures like LL for glued insulated joint, TL for thermit weld, HM for hectometer post and KM for change of position. Marks can be input manually or by the system. The data as drawn on the rail plot screen is simultaneously printed on paper for reference purposes.

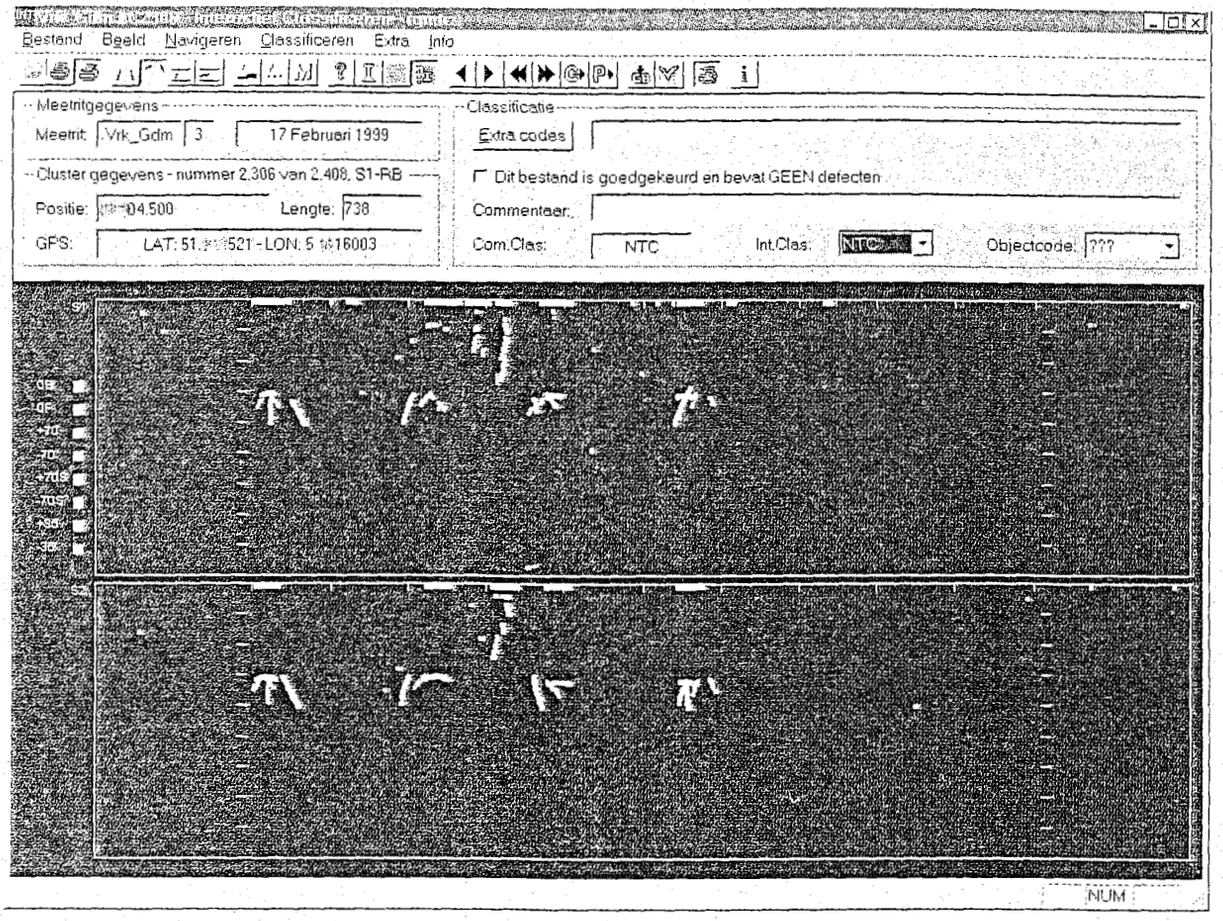


Figure 16.42: Typical off- line data analysis screen showing two fish plated joints, one for the left leg (upper part) and one for the right leg (lower part). This visualization is a "side view". The vertical axis shows depth (0-180 mm), the horizontal axis distance. One meter of data is shown, horizontal tick marks are present every 100 mm. Each dot is a single echo. Where objects are present, the echo dots cluster into lines and curves. The white horizontal indications on the top of the rails are regions where the bottom echo is not detected. The yellow curves depict the top of the 30mm wide bolt holes.

16.6.8 Off-line data analysis and report generation

For the analysis of measured data the URS system is complemented with the Report Generating System (RGS), which runs on a separate PC on the US-train. Since 1996 this PC has already been upgraded to a dual-CPU server machine. All data coming from the IDS are first processed by the RGS system using an advanced rule-based expert system. This computer classification procedure offloads the operator from having to sift through non-defect objects, commonly found in the rails, by hand. After the computer classification step, objects not identified by the expert system and all possible defect objects are shown to the operator. The operator always makes the final judgement. Finally, the defects are reported and printed on paper. Lists of found defect objects can be printed out. Also, the detailed echo's of an object can be printed in a colour coded format with 1 meter of echo's on one sheet of paper, similar to the screen dump shown in Figure 16.42. Reports can be produced in various languages and can also be supplied in electronic form (floppy, CD).

TNO TPD has developed a highly automatic classification procedure, which can automatically spot the difference between cracks and structural perturbations and flag up potential flaws. The algorithms are based on years of experience with the first generation train and on research carried out in the field of expert systems. However, the final classification of potential flaws is always performed by the human expert. Overall, this enables a detection rate to be achieved of over 95%.

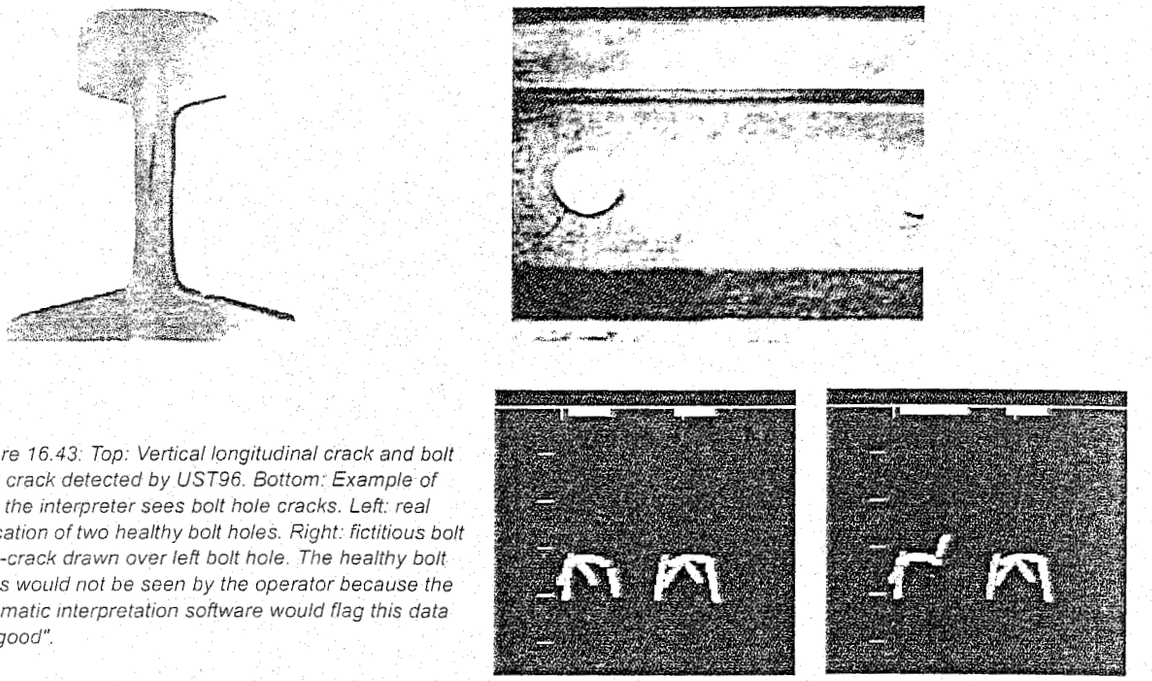


Figure 16.43: Top: Vertical longitudinal crack and bolt hole crack detected by UST96. Bottom: Example of how the interpreter sees bolt hole cracks. Left: real indication of two healthy bolt holes. Right: fictitious bolt hole-crack drawn over left bolt hole. The healthy bolt holes would not be seen by the operator because the automatic interpretation software would flag this data as "good".

For most of the time the results of a measurement run is reported to the customer within 24 hours. The system not only gives the location of the defects found, but also offers detailed information on their nature. Figure 16.43 shows an example of two cracks detected by the UST 96.

16.6.9 NS Ultrasonic inspection program

The UST96 measures NS main lines with daily tonnage above 40.000 tonnes twice a year. The remaining tracks are measured annually. Data are stored digitally and can be imported or exported to the data bases. Parts which are not measured by the US-train are measured using ultrasonic hand equipment. Also, Eurailscout operates hand held equipment for special tasks like switch inspection or acceptance checks of rail welds. Data are handed over to compatible software modules like "Switch-View". The incidents reported by the US-train are stored in a computerized planning system. This system controls the overall planning and production. It also provides the information needed by the manual inspection team.

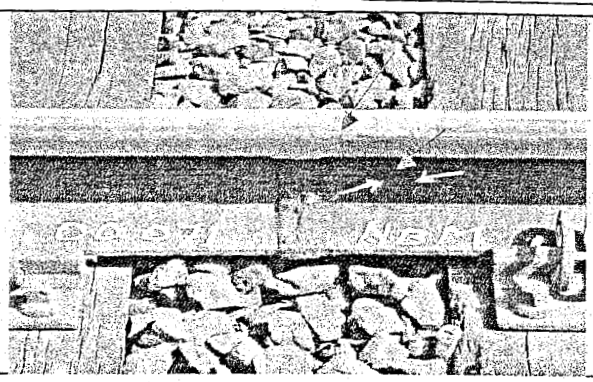
On NS the number of rail defects replaced annually per 100 km of main track amounts to about 7. Globally speaking, 25 % of these defects originate in plain rail and the remainder in welds and joints. About 80 % of the rail defects are found by ultrasonic inspection, which comes down to approximately 4 defects per 100 km of track inspected. The frequency of Ultrasonic Inspection on the NS is given in Table 16.2.

Type of tracks	Times per year
All tracks	1
Tracks > 40.000 tons per day	2
Switches and crossings in open main track	2
Transitions of movable bridges dependent on age	2 - 4

Table 16.2: Interval of Ultrasonic Inspection on the NS

One of the contracting companies in The Netherlands, BAM NBM Rail, providing manual ultrasonic rail inspection services, use the rail failure report form presented on the Figure 16.44.

Date: 18/04/00



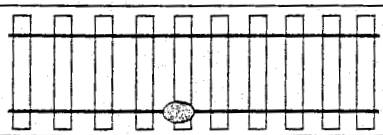
Contracted area	RTD	
Work area	DT	
Date of inspection	18/04/00	
Adviced date of repair	18/04/00	
NBM US-number	00.071	
Reported by	A.v.d.Linden	
Geometric code	112	
Scoop number	34900	
Certificate number	2652	
Probe type	70 dgr.	

Location	Main Track	Yes	Max km/h	140	
	Train yard	No	Tracknumber	B	
	Km:	74.200		☐ L.Rail	
				☑ R Rail	
Track	From	DELFT	To	SCHIEDAM	

Location of incident

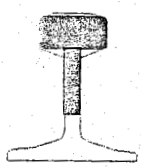
☐ Left

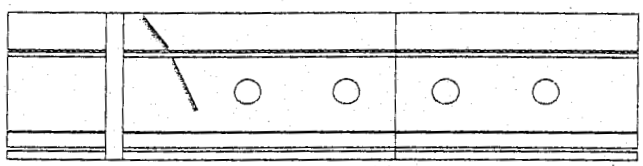
☑ Right



Distance to fish-plate joint: 16 meter Thermit joint — 30 meter Butt joint

Draft of construction





Rolingmark			Right:	
Profile			Profile	UIC 54
Manuf.mark			Manuf.mark	VILRU
Month			Month	10
Year			Year	79
Hardness			Hardness	O

Railconstruction	☑ Wooden sleeper	☐ Concrete sleeper	☐ Other	
Type of track	☑ Longwelded	☐ Fish-plated joints		
Curves	☑ No curve	☐ Low rail	☐ High rail	

Welded constructions: ☑ none ☐ Re-surfacing ☐ Cable joint

Location of discontinuity:

☑ Rail	☐ Rail joint support		Emergency fish-plate	☑ heel
☐ thermit welded joint	☐ Expansion joint			☐ heeled
☐ Butt joint	☐ Insulated joint			☐ N/A
☐ welded joint	☐ Rail Joint			
☐ thermit welded (skv)	☐ Fishplated joint	☐ SB ☐ WB	☑ When Emerg. Fish-plates are mounted, repair within 3 mnths	
☐ Expansion joint				

Adviced term of repair:

☑ 1 (direct)	☑ Vissible		BBS code	
☐ 2 (three weeks)	☑ Not vissible			
☐ 3 (three months	☐ Non regular			
☐ 4 (Contr.engineer)	☑ Vertical incident	0-130	mm	
☐ 5 (Contr.inspector)	☐ Horizontal incident		mm	

Signature: Inspector A/L — Operational staff C.Hagens

Notes: Speed limitation 40 Km/h
Advice: Make-up Rail (4 meter)

Figure 16.44: An Ultrasonic Rail Failure Report Form used by the Dutch Contracting Company BAM NBM Rail

16.7 Track Recording Cars

16.7.1 Introduction

Track recording cars are manufactured on an industrial basis by a few suppliers. The most important manufactures according to the number of delivered cars are Plasser & Theurer and Matisa. Recording cars exist in the following modes:

- Self propelled or hauled railway vehicles;
- Non-contact measuring or measuring axle principle;
- Inertial (absolute) or chord measuring system;
- Track recording only or combination of track-, rail- and overhead wire recording;
- Recording speed between 30 and 250 km/h.

16.7.2 Track recording systems

Optical measuring system

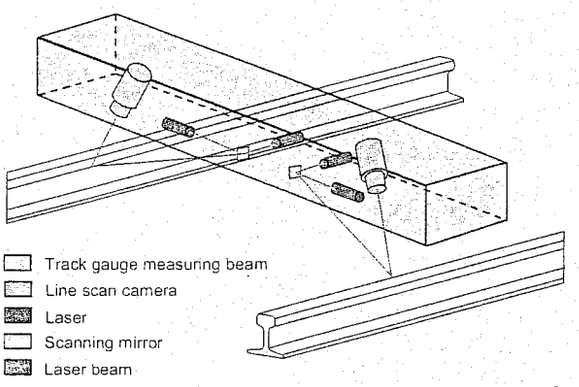


Figure 16.45: Optical measuring system for gauge and track alignment

These are non-contact optical systems for inertia-free scanning of the rail running edges and the rail head profile. One such system as used by Plasser & Theurer, is shown in Figure 16.45. The optical measuring systems are positioned in a measuring beam which is installed in a bogie between or in front of the running axles. Two optical measuring systems are housed in this measuring beam.

Each of the two systems consists of a laser, a deflecting mirror which aligns the laser beam in relation to the actual position of the measuring beam, and a line scan camera which determines the horizontal distance between the light spot formed by the laser beam on the rail running edge and the measuring beam.

Measurement by telescopic measuring axles

The classical and still most applied systems for track measurement use telescopic measuring axles for gauge, horizontal, and vertical versines. Located in front of and between the main axles, the measuring axles are linked to the vehicle frame and the running axles, respectively. When preparing for the measuring run, the measuring axles are lowered individually onto the track.

On bogie-mounted vehicles the telescopic measuring axles are located between the two axles of each bogie. During the measuring run pneumatic cylinders press the wheels horizontally and vertically against both running edges of the rails. This ensures accurate scanning of the track geometry over the entire range of permissible speeds up to 120 km/h.

Guiding skids prevent the measuring wheels from derailing due to the lateral pressure when travelling over a switch or crossing (Figure 16.46). A safety mechanism will automatically lift the measuring axle during a measuring run if a measuring wheel should climb up the rail.

Linear transducers are used as signal transmitters to measure the displacements of each half of the telescopic measuring axle towards the other and towards the vehicle frame.

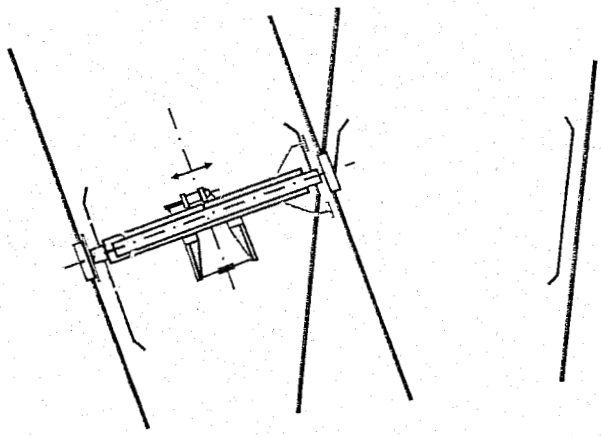


Figure 16.46: Principle of the telescopic measuring axle

Measurement of cross level - gyroscopes

When measuring the cross level gyroscopes are used which are mounted either on the machine frame or directly on the axles or bogie frames. If mounted on the machine frame, the relative movement between wheels and frame must be measured and compensated for. For low measuring speeds of 30 - 40 km/hr level meters can also be used.

Inertial measurement

Inertial measurement uses the known measuring technique by which the spatial position of a measuring sensor is determined either by double integration of acceleration measurements or by using a strap-down measuring platform.

The special advantages of inertial measurement are:

- Almost undistorted representation of the track geometry and its faults;
- Measurement of long-wave faults of more than 100 m which is of great importance, particularly on high-speed lines;
- Direct use of measuring data in order to calculate track correction values and input in track maintenance machines.

The inertial measuring principle on the basis of acceleration measurement is described later in this chapter. An advanced system is the strap down inertial platform (Applanix system).

Contrary to acceleration measurement, even at very low measuring speeds, the strap-down inertial measuring platform supplies accurate measuring results. The measuring transducers of this system can be installed directly in the bogie. The system consists of a three-dimensional laser gyroscope and acceleration measuring system which enables the spatial position of the axles of the bogie frame to be recorded directly in the three-dimensional space (Figure 16.47).

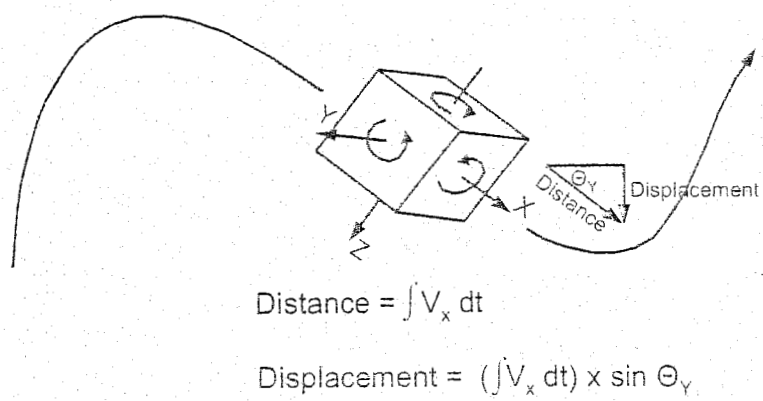


Figure 16.47: Applanix track measuring system

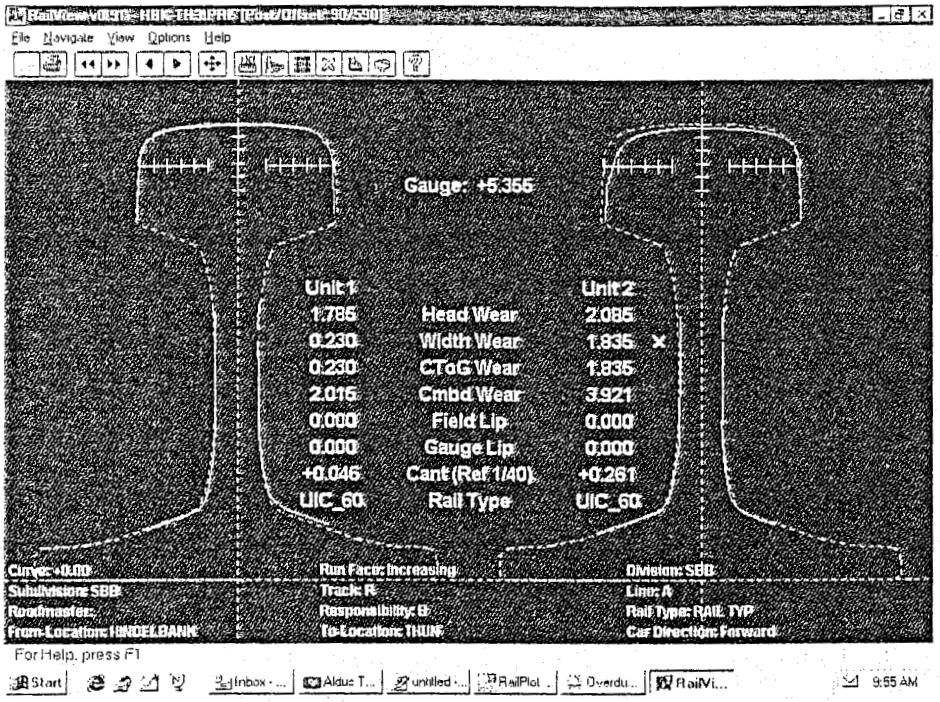


Figure 16.49: Real-time diagram of the rail cross-section

Rail surface

The exact data for planning rail grinding is acquired using special corrugation measuring cars of the company Speno. Track recording cars can be equipped with axle box acceleration measuring devices which give a good indication of the location of the errors. In some cases they are also equipped with precise corrugation measuring devices.

16.7.4 Overhead wire recording

To evaluate the geometry of the contact wire it is necessary to measure the height of the contact wire above the rails and the horizontal position of the contact wire and to know the relative position between track and overhead wire. It is, therefore, practical to have the catenary measuring units on the track recording car.

The non-contact laser measuring device, shown in Figure 16.50, is a very compact and precise device which is mounted on the roof of the measuring vehicle and can easily be integrated into a track recording car. In addition, a measuring pantograph for measuring dynamic reaction forces should be available.

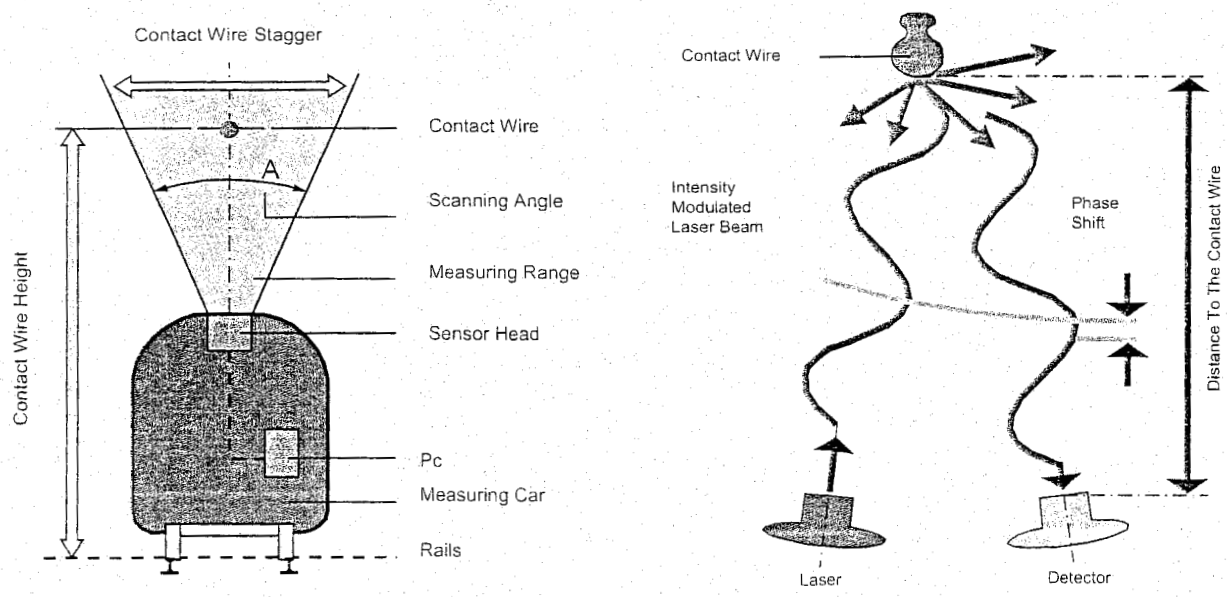


Figure 16.50: Overhead wire measuring system of Plasser & Theurer

16.7.5 Video inspection

Video inspection systems enable visual supervision of the section of track both during the run as well as later. With new, high-resolution systems automatic fault recognition is also possible.

Inspection using a video system can be performed instead of the cost intensive, time consuming, and also dangerous line inspections on foot.

Video inspection systems also enable a visual track database to be established which makes it much easier for the engineer in the office to link track recording information to the local environment.

16.7.6 Processing and recording the measured data

The basic goal of processing and recording measured data consists in converting the individual sensor data into track geometry data. To achieve this it is necessary to set the data in relation to the distance measurement and km marking.

The distance is measured by a rotational pulse generator which is either integrated in the telescopic measuring axle or mounted on a running axle. The voltages produced by the sensors, like the linear transducers, are transmitted directly to the analogue digital converter interface which serves as input interface to the computer. Adaptation, conditioning and calibration of the sensor signals is performed in the computer digitally and in real time. The event marking of stations, crossings, bridges or tunnels is performed either manually by an operator using an event keyboard or directly from a route data file which is loaded into the computer before the measuring run.

The track geometry data elaborated in this way form the basis for the evaluation of the track sections measured. The overhead line measurements are conditioned, recorded and processed for the evaluation in the same way.

On modern track recording cars, the complete processing, recording, evaluation of measured values and storage of the data is performed on an onboard PC computer network in real time controlled by complex software. The menu-controlled software can be operated and supervised by railway technicians operating the vehicle.

Due to the modular structure of such software, additional measuring systems can easily be incorporated for the catenary and rail profile measurement as well as for video inspection. Furthermore, in the light of further developments it is easy to make adaptations within the PC operating system.

Analysis systems

The recorded data is analysed on board and a track quality report for each section of 200 to 500 meters of track is issued. The most common used TQI's (Track quality indices) are based on standard deviations, but there are also indices in use like the Plasser & Theurer ADA 2 index, which is based on virtual deviations of acceleration and, therefore, considers also the permissible line speed.

Besides the analysis immediately during the measuring run, further evaluations can of course be undertaken at a later date. Here too, the elaboration of interfaces to the offboard analysis systems provided by the customer is an important factor.

16.7.7 Track recording cars

High speed track recording cars

Track recording cars for speeds of 200 km/hr and more are usually standard railway coaches equipped with non-contact recording systems. One of the most advanced cars is the EM 250 of Austrian Railways ÖBB (Figure 16.51 and Figure 16.52). The basic vehicle is a four axle RIC passenger coach with Minden-Deutz 524 bogies which is pressure tight and air conditioned.

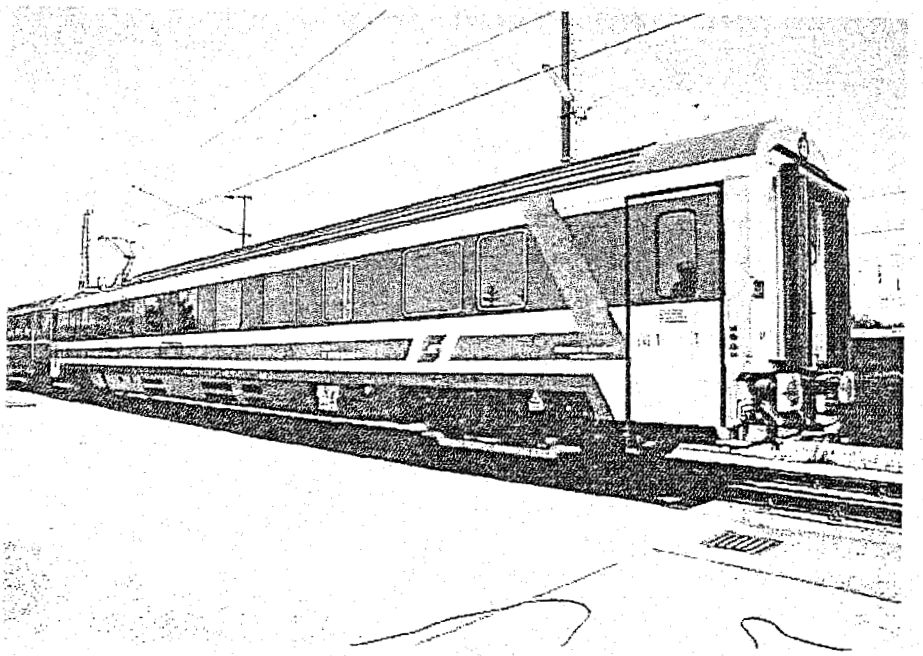


Figure 16.51: High speed track recording coach EM 250 of ÖBB

For level and alignment, the basic measuring unit is the Applanix strap down inertial measuring platform backed up by a laser gauge measuring unit. The geographical position of the coach and the assignment of the measuring data to the line position is organised with a GPS unit.

Rail profiles are checked with an Orian system. In addition, ÖBB has developed a program to evaluate the equivalent conicity of defined vehicles in relation to the recorded track.

That car has, furthermore, force measuring bearings for vertical and lateral wheel forces and video cameras on both ends.

The track data is analysed on board with the ÖBB/ ADA 2 analysing system and all measuring data is submitted to the ÖBB data base by means of wireless LAN stations.

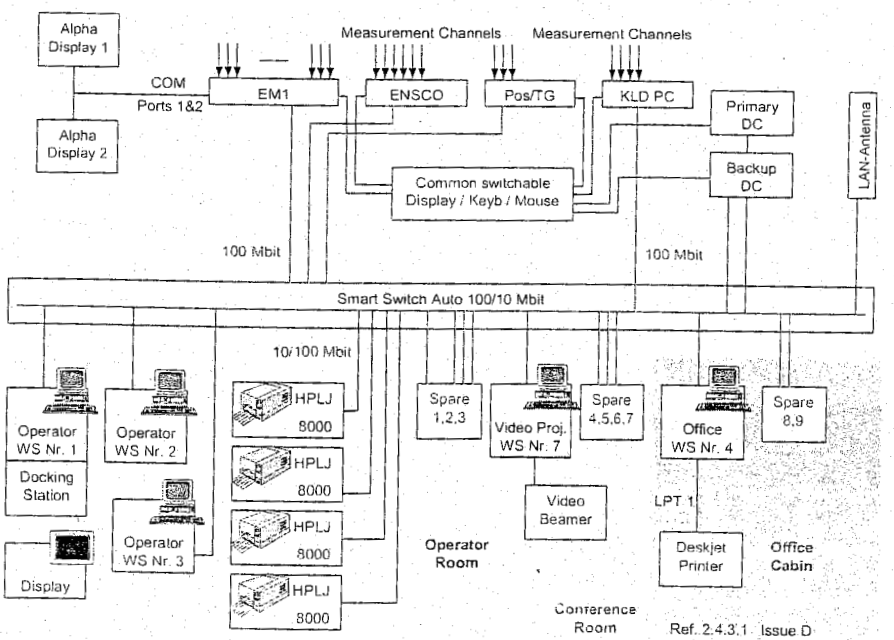


Figure 16.52: Data flow of EM 250

Recording cars for standard speeds

For standard networks recording cars with measuring speeds varying between 80 to 160 km/hr are used. Some of them are towed cars, but the majority are self propelled units. The measuring systems which are used are non-contact as well as telescopic measuring axles.

– *UFM 120 on Netherlands railways*

Since 1999, tracks in the Netherlands are recorded with the UFM 120 of the company Eurailscout (Figure 16.53). This self-propelled non-contact measuring car for 120 km/hr recording speed is equipped with the same measuring equipment as the EM 250 of ÖBB for track and rail measurements. In addition, it has a video rail defect scanning system, a non-contact laser measuring unit for the overhead wires, and a video catenary inspection system.

– *EM 130 on Belgium railways*

Also since 1999, a self-propelled four axle measuring car with 130 km/h measuring speed is used to measure the TGV tracks and standard tracks in Belgium (Figure 16.54). The track measuring system is a chord measuring system based on telescopic measuring axles and measuring bogies. This is an example of the coexistence of both measuring systems in the field of modern track measurement.

For rail profile and overhead wires non-contact laser measuring systems are used.

Figure 16.53: Track recording car UFM 120

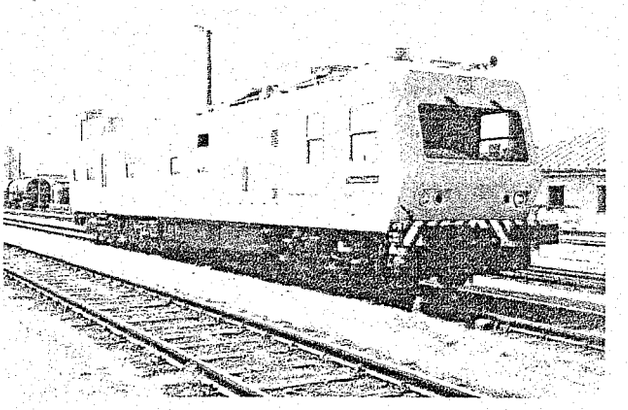


Figure 16.54: Track recording car EM 130 of SNCB

Figure 16.55: Track recording car EM 130 of SNCB

Recording cars for smaller networks

The majority of measuring cars for smaller networks, like urban transport systems and local networks, are self-propelled two axle vehicles with telescopic measuring axles (Figure 16.55). Measuring speeds vary between 30 to 80 km/hr, sometimes up to 120 km/hr. The measuring parameters can be the same as described for the other measuring cars.

16.8 Track recording systems

16.8.1 Introduction

For efficient and adequate control of track maintenance processes, fast and accurate measuring systems are indispensable. This chapter primarily describes those systems in use on Netherlands Railways, such as the recording systems for short-wave rail geometry and long-wave track geometry.

Nowadays almost all administrations make use of automatic measuring systems either by renting the equipment or owning it themselves. As many railways have developed their own track recording cars there is a large variety of makes, but specialised firms, such as Plasser & Theurer, also sell recording cars. Basically there are two principles, i.e. with mechanical transducers and contact-free by means of inertia systems.

For short-wave rail geometry measurements the variety is much narrower. Speno in Switzerland rents out a recording car which measures displacements directly. An attractive alternative, which can be applied well at high recording speeds, consists of axle box acceleration measurements. This is, for instance, applied by NS and Railtrack.

16.8.2 Some aspects of geometry recording

When looking at track irregularities different wavebands can be considered. Faults in the 0 - 3 m band are mainly due to the rail shape and welds, whereas the longer waves originate from ballast and subsoil. It is, therefore not possible to represent track quality properly by one simple figure. Figure 16.56 represents the different wavebands in the vertical geometry which are significant for the maintenance of rail and track quality. This information is nowadays commonly measured by recording cars.

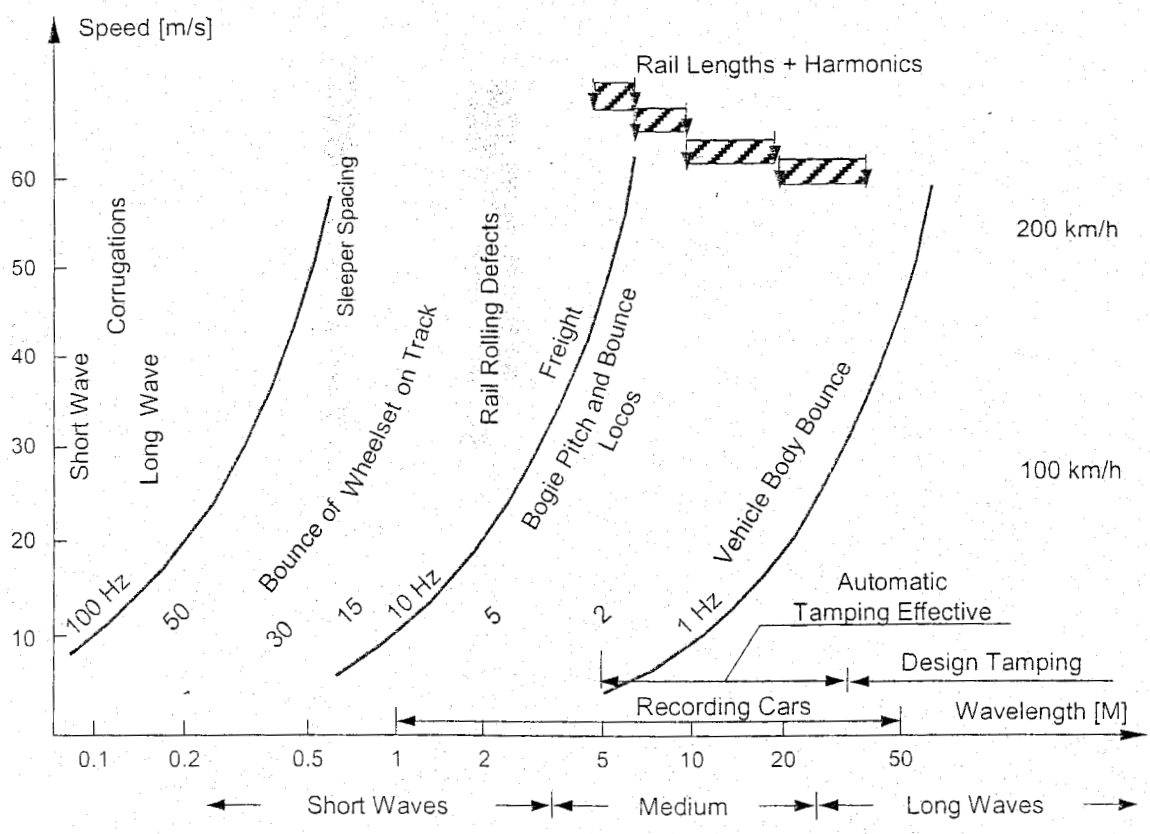


Figure 16.56: Wavelengths in vertical geometry relevant to vehicle interaction

Table 16.3, for example, surveys the different wavebands which were used in the NS track recording system BMS. The irregularities are split up according to the maintenance processes available for correcting rail and track geometry.

Waveband l	Phenomenon	Detection	Remedy
1 – 30 cm	Corrugation	Axle box acceleration	Grinding
1 – 100 cm	Poor weld geometry	Axle box acceleration	Bending
	Poor wheel threads	Rail accelerations	Turning off
2 – 3 m	Rolling defects	Axle box accelerations/ deflections	Grinding up to 2 m
3 – 25 m	Moderately long waves	Conventional	Automatic tamping
25 – v/f m	Long waves	Stabilized platform	Design tamping
70 - ∞ m	Geometric design	Quasi-static	Design tamping

Table 16.3: Different wavebands with corresponding phenomena, detection and remedy possibilities

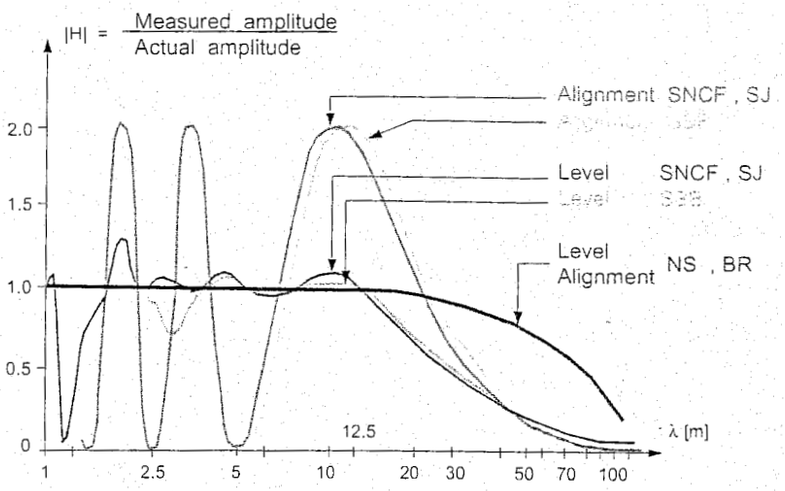


Figure 16.57: Transfer functions of various recording cars

The measuring bandwidth of track recording cars is generally restricted to wavelengths ranging from a few metres to 20 or 30 metres. This information is represented by the transfer function. Figure 16.57 shows some examples of transfer functions for recording systems used in Europe. These functions represent the ratio of measured and actual amplitude. Obviously, the systems based on a mechanical transducing principle cause relatively large distortions. Inertia systems, like those of NS and BR, do not possess these drawbacks.

It is important to note that the effectiveness of modern tamping machines is restricted to the same waveband.

Complementary measuring methods were developed to collect short- wave information on rail geometry:

- Trolleys or straightedges with a short reference base for measuring rail rolling defects or corrugations and weld geometry;
- Vertical axle box accelerations, often measured, as for example described in [84] and [158], on geometry inspection cars, and filtered in such a way that the signals produced are as far as possible independent of speed.

In order to quantify track quality with reasonable accuracy, the measuring systems are required to have an overall measuring accuracy [88] depending on the spectral content according Table 16.4.

Waveband	0-5m	5- 10m	10-20m	20- 40m
Typical spectral content [mm²]	0.05-0.2	0.3-1.2	0.4-1.6	0.3-1.0
Required measuring accuracy [mm]	±0.05	±0.1	±0.2	±0.3

Table 16.4: Required measuring accuracy for track geometry

16.8.3 Assessment of track quality for maintenance decisions

To make decisions on track maintenance, it is necessary to have representative data available along the track. Geometrical quality can be expressed in various ways. The simplest way is to count the defects exceeding a given amplitude level. Using this method the quality index especially reflects the maximum defects for each parameter. For instance, the P-value used on the Shinkansen network is the ratio of the track defects exceeding 3 mm and the total number of defects [287]. A total index can also be obtained by combining the various values from several geometrical parameters as in the DB-method described in [290]. The advantage lies in making large local defects more noticeable.

The most common way of representing quality is to compute standard deviations. This method is further applied in this book. In this case, the quality index directly represents the total energy of the measured geometrical parameter. Alternatively, the mean absolute value, which is proportional to the standard deviation, can be easily obtained by analogue filtering as applied by SNCF [136]. These indices do not give information about the shape of the defects, but are very well-suited to continuous maintenance decision-making.

A third kind of assessment consists of looking at vehicle reactions. For Japan's high-speed lines, for instance, the geometry is weighted according to the vehicle response characteristics described in reference [243]. Real time vehicle calculations are carried out on NS and this process is described further on in this chapter.

16.9 Universal measuring coach EURAILSCOUT

16.9.1 Introduction

As mentioned previously, since 1999 tracks in Netherlands are recorded with the EURAILSCOUT UFM120 measuring coach (Figure 16.53). Using an integrated system that measures, processes, analyses stores and interprets track quality data.

Eurailscout operates several high-tech measuring cars and devices for measuring track geometry, overhead wire, rail surface, ultrasonic checks and switch inspection.

The measuring systems of Eurailscout now have interfaces to several visualisation and software packages like the Plasser "Off-board System", to Erdmann's "GeoView" and to the "IIS" (track quality database of DB AG). Measured data are exported to these systems directly.

The UFM 120 represents a universal measuring car designed for the track inspection of European normal gauge lines (1435 mm). It stands for a new type of fast running high-tech inspection cars. With a single run it can measure all the relevant quality data of the track geometry, the overhead wire and rails as well. The UFM 120 is equipped with the latest measuring devices connected to a local area network (LAN).

All of the devices operate in a contactless manner, so there is no interaction with railway signal devices nor safety problems at high speed. Measurement results are compared at a common point of reference. Precise allocation of a test result to a rail is possible via accurate DGPS location reference. The car is self-propelled by a diesel engine. It operates in both directions with a maximum speed of 120 km/h. Basic technical information on the UFM120 are given in the Table 16.5.

Date of construction	1998
Railway gauge	1,435mm
Vehicle length over the buffer	23,000 mm
Centre casting distance	17,500 mm
Axle base	2,500 mm
Minimum turning radius in movement	90 m
Vehicle weight approximately	70 tonnes
Maximum speed Own drive Train	 120 km/h 120 km/h
Maximum measurement speed	120 km/h
Sifa	Present
Vehicle control (PZB)	Present
Railway warning system (Mesa 2002, TeleRail)	Present

Table 16.5: Technical data UMF120

16.9.2 Track geometry measurement

For investigating the condition of the track quality UFM120 is equipped with a contactless geometry measuring system. The heart of this system is an inertial-measuring system (IMS), actually consisting of 3 high precision gyro systems and accelerometers for all 3 spatial directions x, y, z. Through the combination with an optical track measuring system, using laser distance sensors and a GPS location system, the required long-time accuracies are established. This geometry measuring system is in the front side of the car body, positioned between the axles, and can emit the following parameters:

- Gauge;
- Vertical profile;
- Horizontal profile (alignment);
- Superelevation (cross level);
- Twist (ORE-twist or twist against any base length).

These track parameters are measured at every 25 cm, stored in an onboard computer system processed and printed immediately during the ride. All the measured data can also be plotted as a graphical representation, as well as be exported to the external data medium.

It is obvious that the precise allocation of test results represents a basic need. It is necessary for triggering the subsystems, to combine all information at the local area network and to calculate several parameters (e.g. twist). It is especially important when data are processed later at the office. There are several position detecting systems on the UFM 120 acting at different levels:

- Automatic detection by 2 separate encoders;
- Automatic detection of known positions with a magnetic sensor (e.g. for magnets at the NS lines);
- Automatic detection of known positions with an optical sensor (e.g. for "Indusi" safety brakes at DB AG lines);
- Automatic synchronisation by means of satellite navigation (D-GPS);
- Manual triggering.

16.9.3 Overhead wire measurement

To evaluate the geometry of the contact wire it is necessary to measure the height of the contact wire above the rails and the horizontal position of the contact wire and to know the relative position between track and overhead wire. It is therefore practical to have the catenary measuring units on the track recording car.

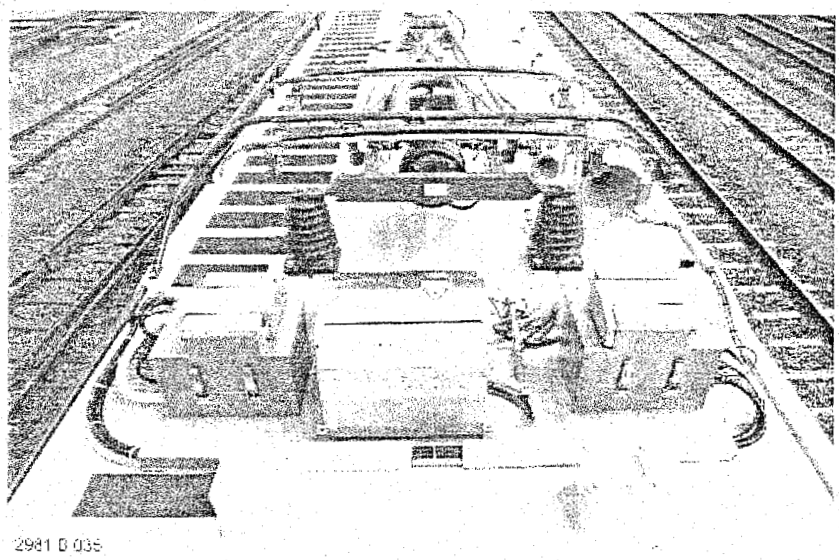


Figure 16.58: The UFM120 non-contact laser measuring device

The non-contact laser measuring device, which is part of the UFM120, is shown in Figure 16.58. The system was developed by TNO-TPD and is called ATON ("Automatic Thickness measurement of Overhead wires Netherlands Railways").

Benefits of the new system go beyond recording of images of worn wires. It can also detect individual flaws, such as abnormal patches of abrasion. Processing software used to identify and detect such patches has also been developed and is continuously refined, resulting in a large reference data set. By correlating the detected pattern with the known characteristics of a flaw, more accurate pronouncements on the nature of the abraded surface can be made. This in turn enables more effective prognoses regarding overhead wire wear.

System overview

The overhead wire inspection system consists of two parts; a recording system in the train which determines all the basic data, and the processing software which generates information from this for decision-making relating to maintenance. The recording system in the train (Figure 16.59) consists of the optical wire imaging system plus the sensors for superstructure movement, location and pantograph acceleration.

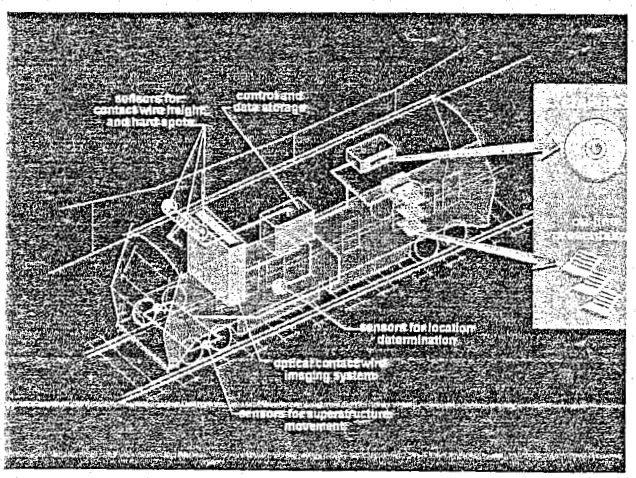


Figure 16.59: ATON system overview

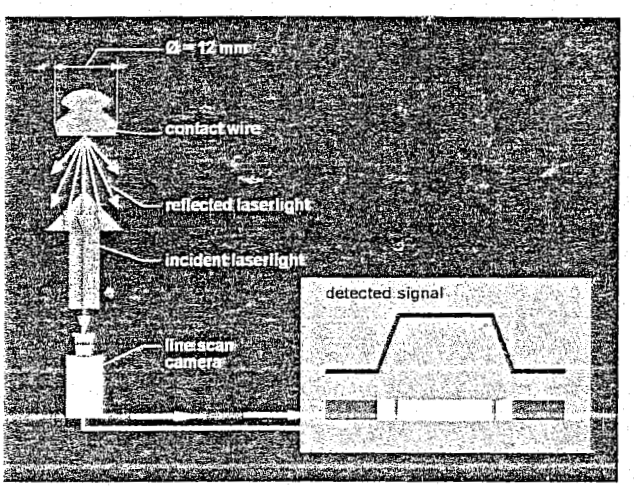


Figure 16.60: ATON system overview

The type of contact wire used by many railway companies has a round cross-section. The width of the worn underside represents an indication of the degree of the wear of the contact wire. The overhead line inspection system exploits this fact and records the worn surface of the contact wires. Thickness estimates are produced from the worn surface recordings using algorithms developed for this purpose.

The recording system in the train

At the heart of the overhead line inspection system is an optical system with which the contact wires above the coach are imaged at a high resolution (Figure 16.60). Up to a running speed of 180 km/hour, the resolution in the direction of travel is 2 cm, and 0.15 mm perpendicular to the direction of travel. The system employs lasers for illumination and CCD linescan cameras for the imaging of the contact wires. The images are stored along with many other parameters on CD-ROMs. These other parameters include contact wire height, hard spots, lateral deviation of contact wires and superstructure orientation and position along the track. The accuracy and reliability of the determination of wire position is extremely high. The reason for this is that the overhead line inspection system produces a complete image of the contact wire.

Processing of the measurement data

The processing procedure of the measurement data which is stored on CD-ROM onboard the train is shown schematically in Figure 16.61. The CD-ROMs first go through a pre-processing step in which a number of characteristics from the stored images are calculated. Together with the other measurement data, these are graphically presented in reports called '1.5 km plots'. The four most important types on these plots are:

- Height: The distance of the contact wires in relation to the rails (4.45 - 6.30 m);
- Position: The location of the contact wires in relation to the centre of the track;
- Minimum thickness: This is the lowest thickness of each 0.5 metres of the route covered;
- Hard spots: This is an indication of the resistance that the pantograph experiences, in fact an acceleration sensor.

The 1.5 km plots are first interpreted manually, and points for attention that are found are verified with the assistance of the raw data viewer. The knowledge from this interpretation is fed into an expert system software so that these points for attention can be found automatically. An important aspect in this is the integration of contact wire reflection with other relevant information such as hard spots. The result of this process is a number of attention point forms. Verification of these points for attention is partly undertaken by an operator and partly automated. Each year more information is supplied to the software, so that the operator's effort reduces annually.

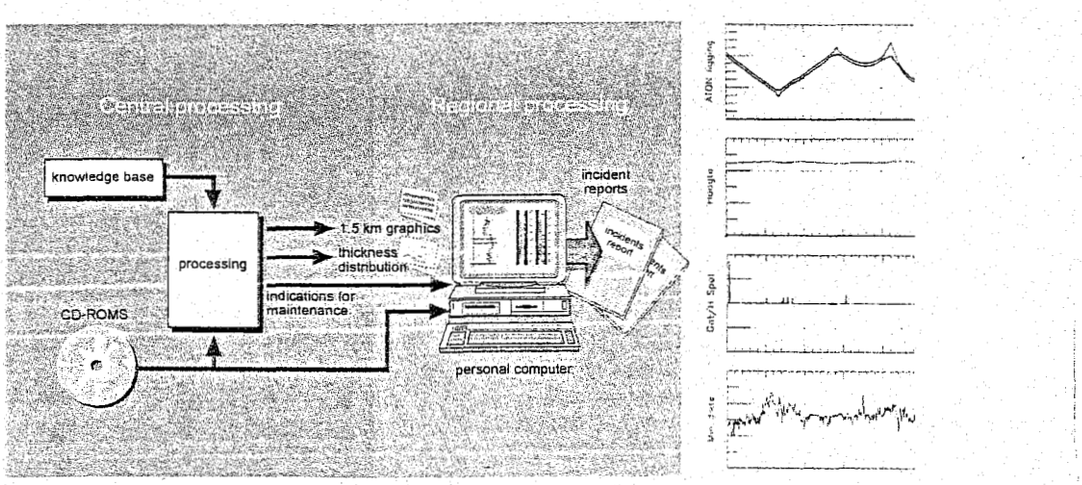


Figure 16.61: ATON system overview

Another output from the pre-processing is 'thickness distributions', which represent estimates of the thickness of the wire divided into classes. The percentage length of the wire that falls into a particular class is calculated for each contact wire section.

Reporting of indications for maintenance

The attention point form gives the maintenance staff the precise location of the indicated problem and a visual impression of its nature. The points for attention are presented as shown in Figure 16.62. The position of the contact wires over a length of 600 metres is shown. The area between the two horizontal lines in the centre of the positioning is 15 metres long, and the point for attention is located in the centre of this area. The distance of the point with respect to the nearest poles is calculated by the software and is stated at the top. A reflection image of this 15 metres area is shown at the right, in colour, displaying the reflecting contact surface of the wires. The reflection strengths are colour coded. A wear surface width is determined from the reflection patterns, which can in turn provide an estimated contact wire thickness. All reflection patterns are lined-up on their left side. This form of presentation allows a thickness scale to be introduced on the right-hand side. On the attention point form these lines indicate the 11 mm, 10 mm, 9 mm and 7.5 mm wire thickness limits.

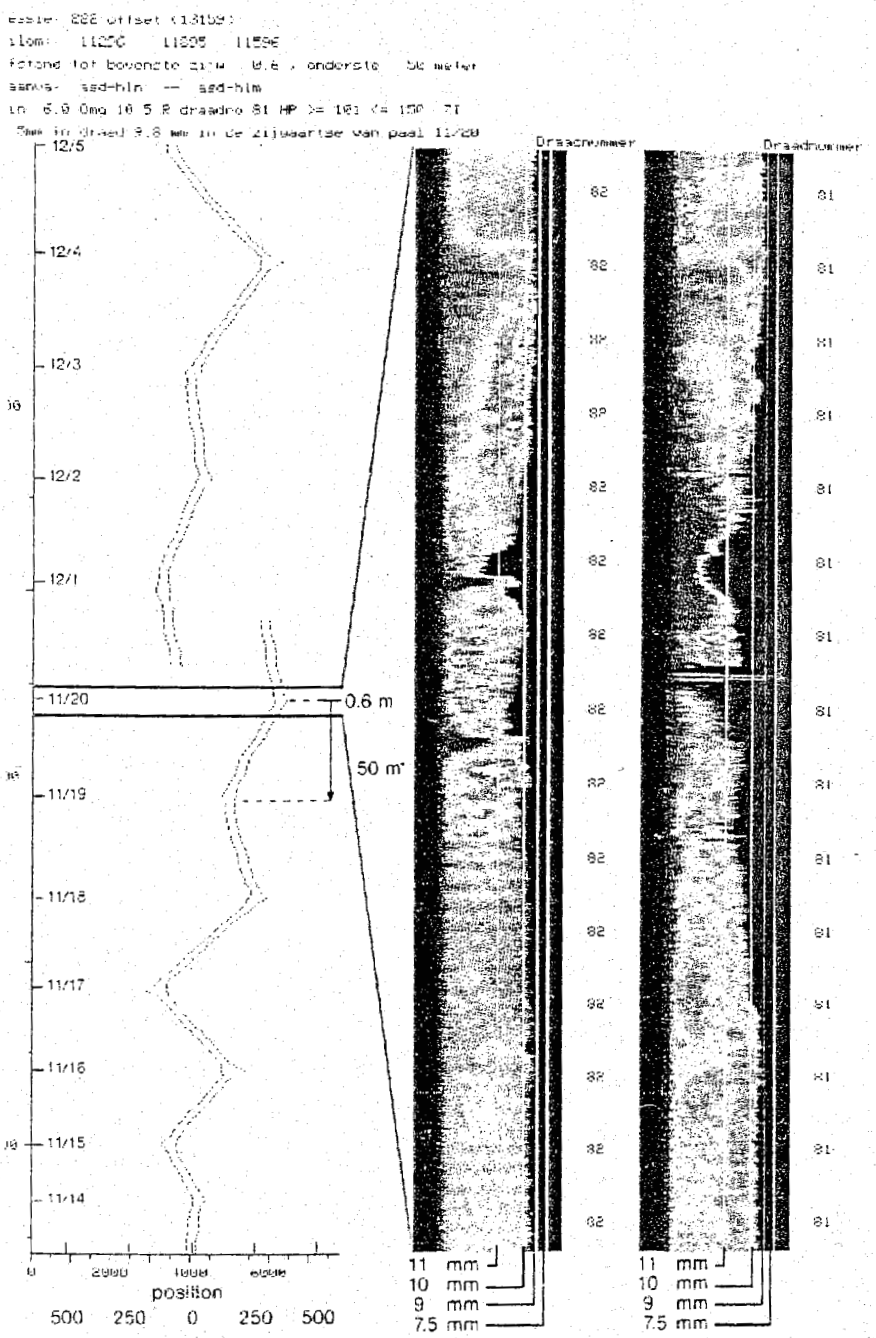


Figure 16.62: ATON system overview

Planning of contact wire replacement

The aim is to be able to spread the replacement of contact wire sections over the average lifetime (25 years in The Netherlands) of such sections. To achieve this, each year a certain number of contact wire sections must be replaced. The worst sections can readily be selected with the assistance of 'thickness histograms'.

16.9.4 Rail Profile measurement

The rail profile measuring system provides a graphical display of the rail profile contour as well as all rail parameters calculated from it. The measuring system consists of four laser systems and four CCD-cameras, which are mounted under the vehicle at the front of the car body. With the use of these cameras in extremely high sequence cross sectional profiles of both rails are taken and displayed in a real time after processing (Figure 16.63). For each track side the following parameters are investigated (Figure 16.64):

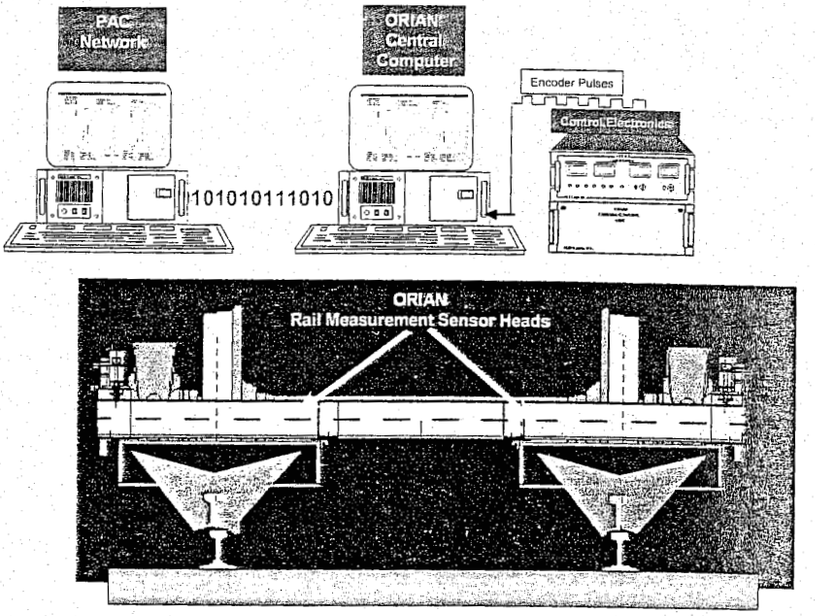


Figure 16.63: EURAILSCOUT UFM120 Rail Profile measuring system

- Head loss (in mm² and percent);
- Rail type;
- Track width (gauge);
- Gauge-face width;
- Rail head width and wear width (side wear);
- Rail height and wear size (wear height / vertical wear);
- Lip and cant (inclination) of the rail;
- Wear profile of the rail;

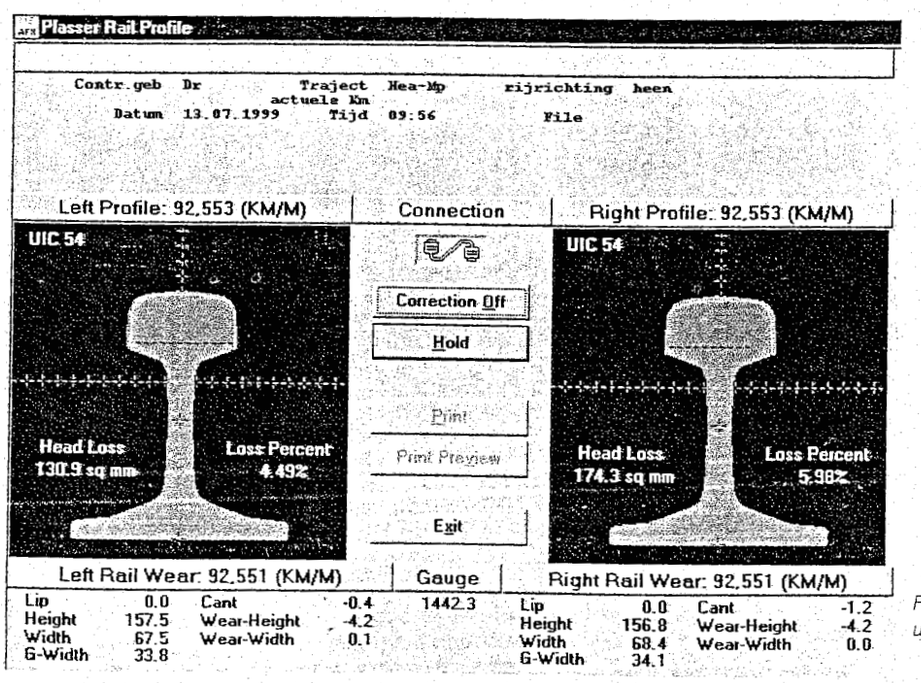


Figure 16.64: Rail Profile measurement display

Exceedences of the preset tolerances are elaborated and displayed. Moreover, the special wear characteristics are being measured and established.

Both rail profiles and the calculated values could be directly displayed on a screen in real time. Subsequent processing and reporting of various statistical calculations is also available.

16.9.5 Rail Check System

Rail Check System performs a contactless inspection of the rail surface and fastenings. It consists of two CCD cameras with special lighting and accompanying image processing. During the ride UFM120 performs using Rail Check System a discontinuous observation of the rail top surface. Corresponding images are directly captured by cameras with resolution of 2mm and transferred to the appropriate data processing method.

The following (recognition) steps are processed for every single captured picture, so that typical rail failures are recognised and marked:

- Rail failure (rupture and impact);
- Short wave patterns;
- Worn-out welds;
- Round or oval imprints;
- Missing (steel) components e.g. fasteners;
- Short wave corrugations.

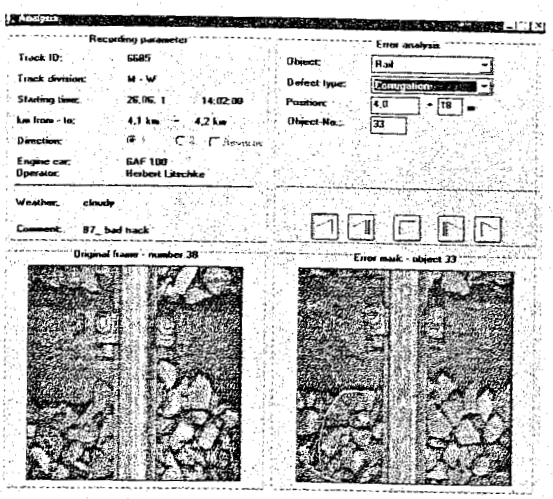


Figure 16.65: Detecting corrugation

The pictures, in which clear faults have been observed, are marked and stored for documentation and archiving (Figure 16.66 and Figure 16.65). After measuring runs, the marked and stored pictures can be transferred to external data storage equipment. But the UFM120 can also be loaded with the measuring data. In such a way, all failure images can be re-established during maintenance at the exact measuring position (kilometre point x). Moreover a general failure assessment can be performed over the full measuring run data.

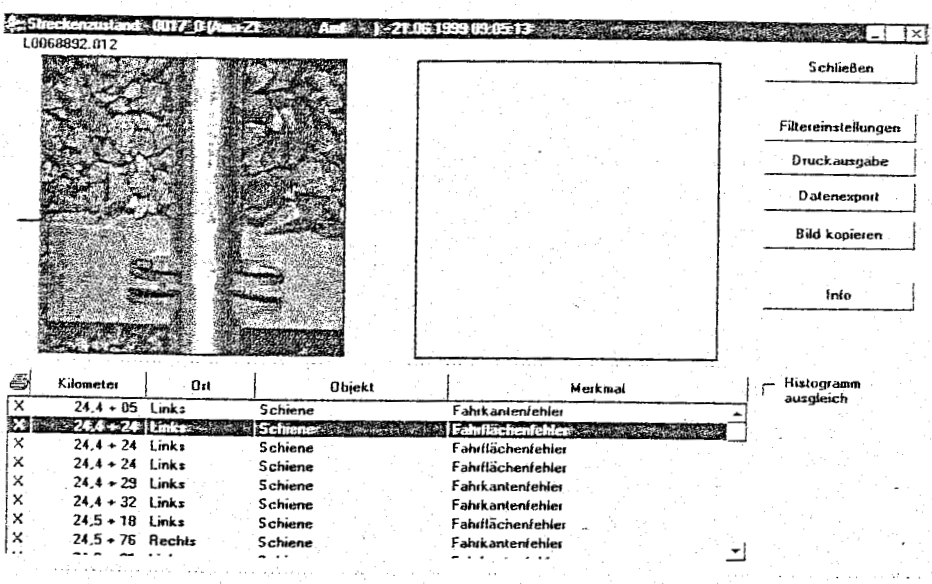


Figure 16.66: Detecting rail surface defects

16.9.6 Video inspections systems

UFM120 hosts 3 video cameras. These cameras are used for visual inspection of track's direct surroundings, but they also serve the train driver. Two of these are positioned in the front in order to control the objects ahead of the train. The third camera is located on the roof for permanently recording the wires (Figure 16.67).

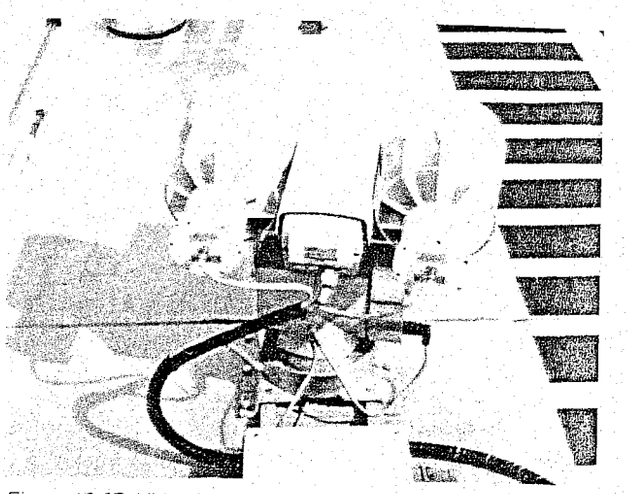


Figure 16.67: Video inspection system

The cameras are mounted on platforms that can rotate 360° around the vertical and the lateral axis. The focus orientation is controlled inside the train. Together with automatic zooming, the cameras can capture all relevant and desired objects along the track.

Every camera platform is equipped with high performance lights in order to improve the contrast and the visibility at night or in tunnels.

The pictures captured by the cameras are displayed on two monitors during the run. The recordings take place on standard video-tapes. The UFM120 train computer adds track information (e.g. kilometre point x, track direction and position) to the recorded pictures. Moreover external data can be added to these video pictures as well as marks, so that interesting and critical points can be recovered very easily when performing post analysis.

Figure 16.68 and Figure 16.69 show some video inspection shots.

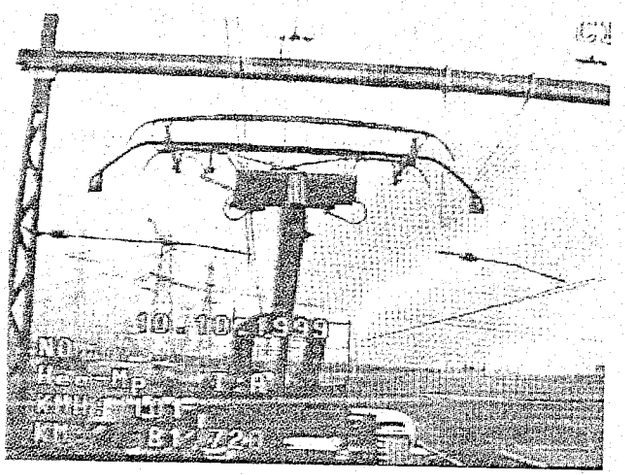


Figure 16.68: Video inspection of the catenary structures, overhead wires and suspension devices

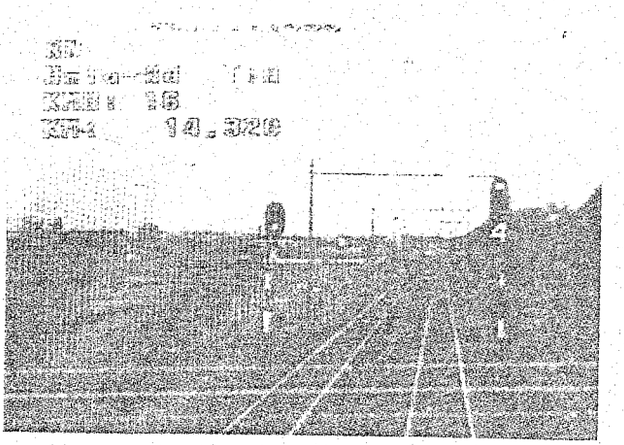


Figure 16.69: Video inspection of the track, vegetation, surroundings, ballast bed

16.9.7 Data processing and storing

All values measured during inspection are displayed at once on the monitor as track charts (Figure 16.70) and directly documented on the printer.

The set-up of the display, ordering of the signals, scale factors, calculations, thresholds, links and analyses can be configured freely and modified to the special wishes of each client as well as colour, display text, language, reports and so on.

Such modifications for the customers can be done without extra programming and within minutes.

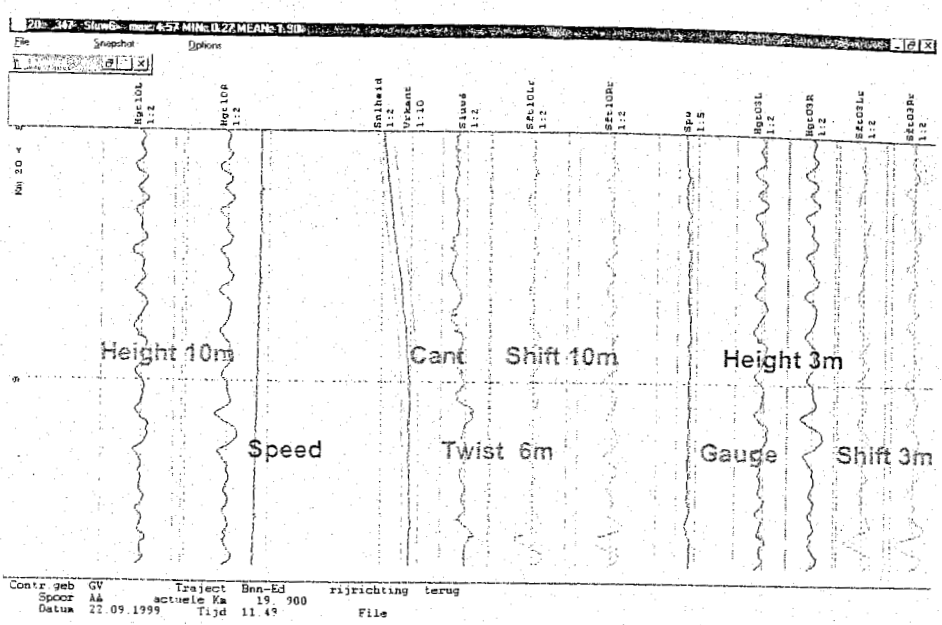


Figure 16.70: Track geometry measurement system. (typical result)

Figure 16.71 and Figure 16.72 show screen shots from the GeoView software.

GeoView combines different information of a rail track, displays all measured data and calculates quality data, while the exceptions and quality indexes of several thresholds are represented by different colours.

One can zoom into the map of the track, stretch the display, calculate statistics. By a click to a track location data values are read back and pictures can be displayed. Subsequent measuring can be compared to estimate trends and conclude decision suggestions for maintenance.

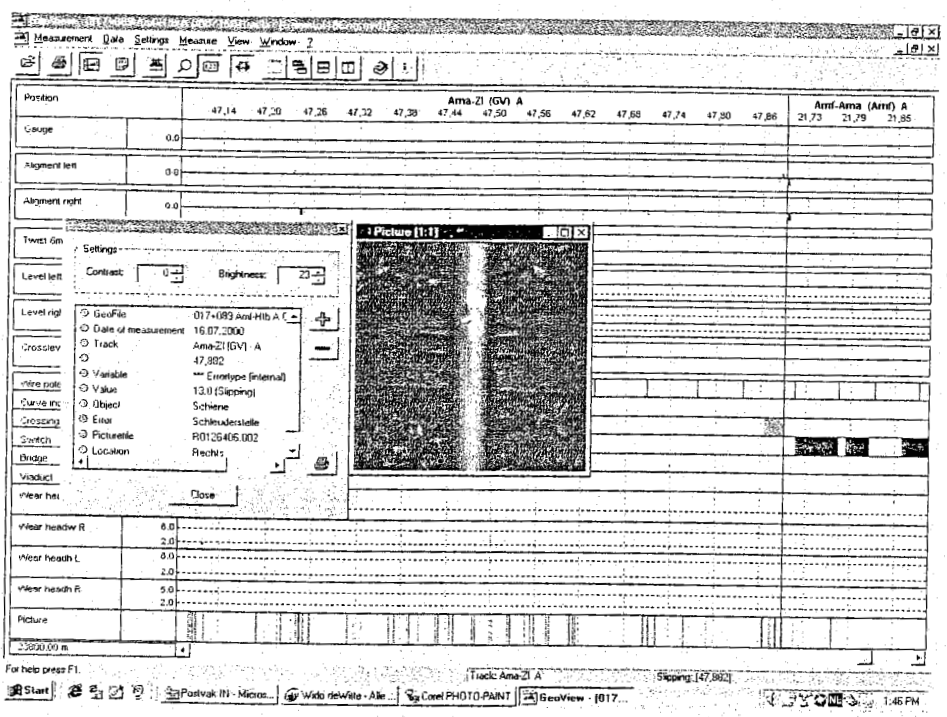


Figure 16.71: A screen shot from the GeoView Software

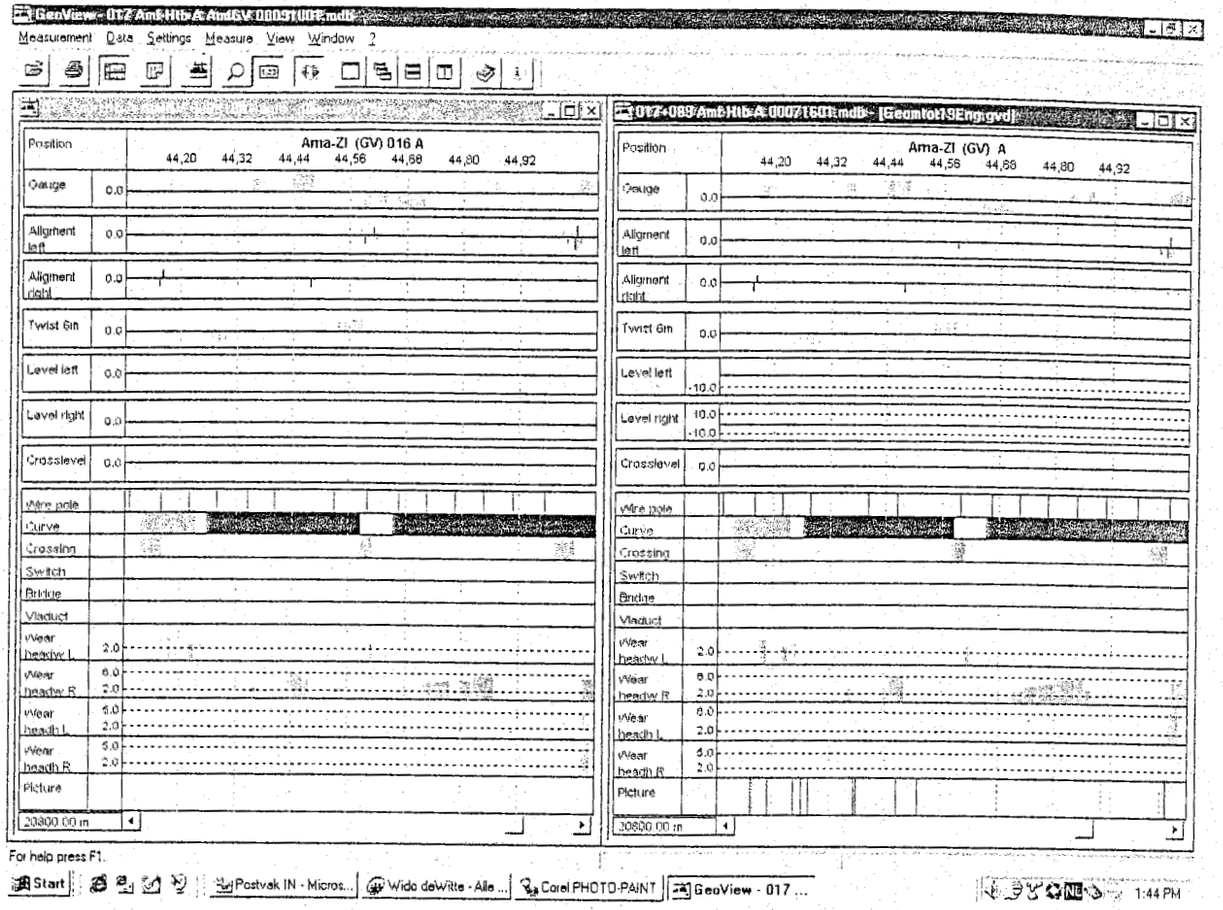
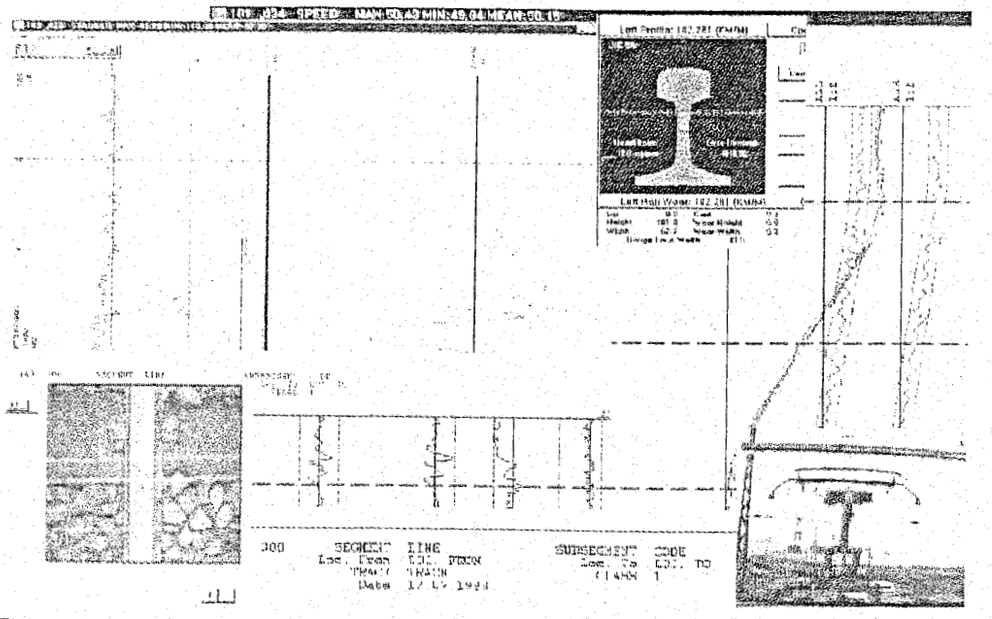


Figure 16.72: A screen shot from the GeoView Software

Additional values can be calculated and displayed at once, for example a floating mean over a base length (e.g. 20 m). Cord measurement curves, as they usually come from measuring axle systems (e.g. 4:6 or 5:5 m) are calculated from space curve points.

All values can be compared to different sets of thresholds. Exceeding of a threshold can cause a change of colour in the charts and generate an exception report. Exception reports can be stored and sent to the screen or printer as well Track quality indicators and summary reports are calculated.



All measured data can be tied-up together correspondingly. As a consequence of that, the inter-dependencies between different parameters could also be evaluated, Figure 16.73.

Figure 16.73: Overlaying various corresponding measurements

Since all the data are stored digitally, they can be reviewed again in a playback after an inspection run and analyzed with other settings.The post-processing equipment is shown in Figure 16.74, whereas the process control is explained in Figure 16.75). Examples of alternatives to be analyzed are:

- Calculation of the twist according to the different base lengths;
- Evaluation according to different thresholds;
- Calculation of the quality indexes for track segments.

The UFM 120's measurement technology and software allows for the first time ever the extreme accurate allocation of rail parameters in different ranges - geometry, contact wire, rail profile and surface.

Moreover, for the first time accuracy of geometry values is high enough to simulate other measuring vehicles (e.g. cord systems with measuring axles) without extra expense.

In such a way, UFM 120 measurements can be compared with those from older vehicles and thus precise data allowing smooth transition from an old to the new measurement technology can be obtained.

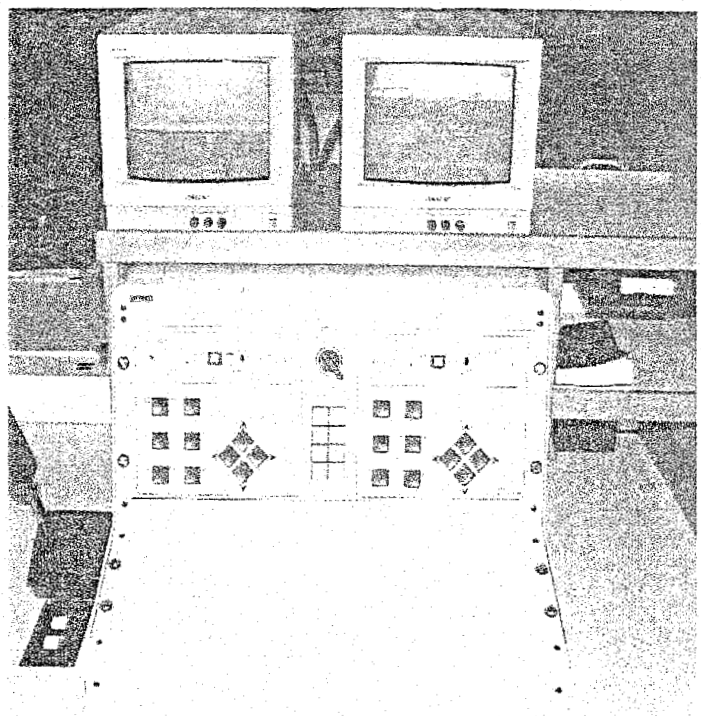


Figure 16.74: Video inspection image processing

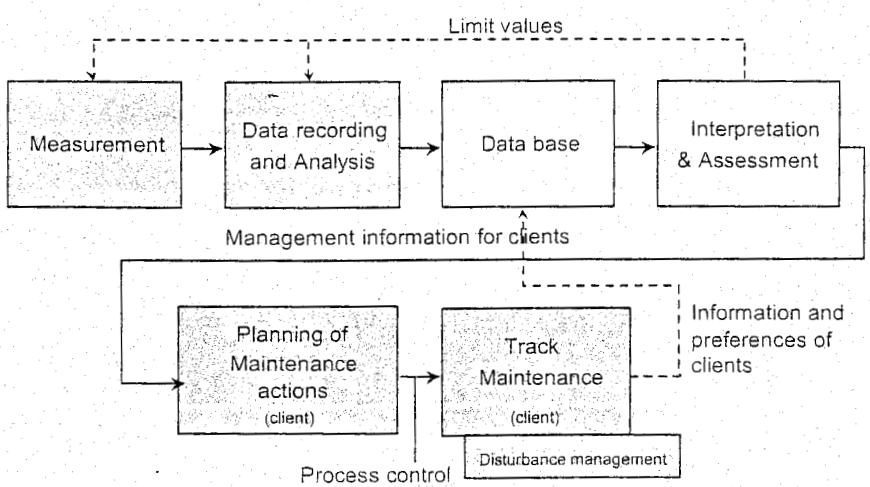


Figure 16.75: Analysis of data and subsequent maintenance actions

16.10 The NS track recording system BMS

16.10.1 Short-wave recording via axle box accelerations

For recording short-wave rail irregularities basically two principles exist: direct recording of displacements via a laser system and measurement of axle box accelerations. In this section the last principle is described.

Because of the strong speed dependence, the axle box accelerations are transformed on line to a fixed reference speed of 90 km/h. In Chapter 6 the relationship between rail geometry and axle box acceleration is derived for a theoretical model. The transfer function, according to (6.71), is composed of contributions from wheel, rail, and Hertzian contact spring. As a first approximation, this relationship may be considered as linear.

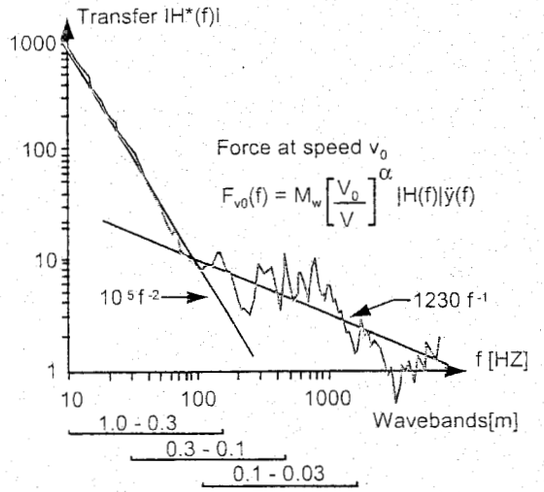


Figure 16.76: Estimated transfer function between axle box acceleration and rail geometry

As the different parameter values are unknown, an algorithm was developed to estimate the transfer function based on a number of reproducibility runs at different running speeds. The result is displayed in Figure 16.76.

The relationship in the frequency domain between vertical axle box acceleration and vertical rail irregularity can be assumed according to:

$$G(f) = H^*(f)\ddot{Y}(f) = cf^{-\alpha}H(f)\ddot{Y}(f) \qquad (16.3)$$

in which:
G(f) = vertical rail geometry
$\ddot{Y}(f)$ = vertical axle box acceleration
H*(f) = total transfer function
H(f) = transfer function cleared of trend
f=v/λ= (time) frequency
v = speed [m/s]
c = constant

16.10.2 Inertial measuring principle

Modern recording systems, based on an inertial principle, can be operated in a speed range from approximately 40 to 300 km/h. The lower limit is determined by the reduction in acceleration signals and, thus, in measuring accuracy due to the reduction in speed.

According to the inertial measuring principle, the position of the measuring car is determined by means of double integration of the acceleration measured. The track position is then measured in relation to the resulting coach position by using relative displacement transducers. Figure 16.77 shows this principle in a very simplified version. At a certain velocity the change in level amounts to:

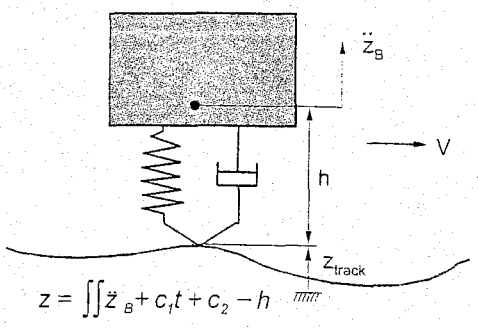
$$z = \iint \ddot{z}_B + c_1 t + c_2 - h \qquad (16.4)$$

If the static value does not need to be measured it can be removed by high-pass filtering, which at the same time eliminates the constants c_1 and c_2. This creates a reference line which runs abreast of the track and against which the track position is measured.

During integration, small off-sets of the accelerometers will also be integrated and this will soon lead to inadmissibly large errors. Non-linearities produce the same effects. These problems are also avoided by high-pass filtering.

In principle, it might be possible to place an acceleration transducer vertically above the wheel, rigidly connected to the axle box, and in this way obtain the vertical rail position. In practice, this solution is not possible as the high accelerations require a wide transducer measuring range. Small amplitudes at long wavelengths, however, only produce very low accelerations which, again on account of non-linearities, cannot be measured in this manner.

The accelerations are therefore measured at a point on the coach. As they are considerably lower, a more sensitive transducer can be used. At the same time the relative vertical displacement of the acceleration transducer in relation to the axle box is measured using displacement transducers. The track position can be derived from these parameters.



High Pass filtering to remove :

- constants
- offsets
- non-linearities

and to create selected wavebands:

Integration ⟶ Time

Filtering ⟶ Length (λ)

Figure 16.77: Principle of an inertial measuring system

Thus, here the coach and its suspension perform the function of a mechanical low-pass filter. The high frequencies, which cause the high accelerations, are filtered before they reach the acceleration transducer.

It will be clear that those wavelengths of the track which, at a certain running speed, produce frequencies lower than the natural frequency of the coach, are principally determined by means of the acceleration transducer, whereas the shorter wavelengths, which produce frequencies higher than the natural frequency of the coach, are principally determined by the relative displacement measurement. Of course, no phase differences are allowed between signals which are to be combined.

In The Netherlands the following track geometry signals are produced: cant, level, alignment, and gauge in the 0.5 - 25 m and 0.5 - 70 m wavebands, while the quasi-static component in the 70-∞ m waveband is also determined for cant, curvature, and gauge. Moreover, the track twist is calculated from the difference in cant on the two bases of 6 m and 2.75 m.

16.10.3 Dynamic signals

Vertical

How the dynamic measurements are actually carried out is shown diagrammatically in Figure 16.78 and will be explained using the example of vertical level. The measured vehicle body acceleration is doubly integrated with respect to time in order to obtain the absolute spatial car body displacement. For reasons explained previously, the long waves are cut off electronically at 70 m (attenuated by 3 dB) using a fourth-order Butterworth filter. The purpose-built electronic control system continuously adjusts the time-defined parameters of the Butterworth filter to the speed, so that the characteristic, based either on spatial frequency or on wavelength, is not altered. As the system integrates three times at the most, high-pass filtering should be performed with a filter at least one order higher to eliminate the effect of drifting due to the integration of offsets.

As the accelerometer is mounted rigidly on the floor of the car body, the acceleration is always recorded perpendicular to the floor. In actual fact, though, it is the vertical acceleration which is involved. This is derived from the measured signal by making an electronic correction as a function of the car body rotation due to cant.

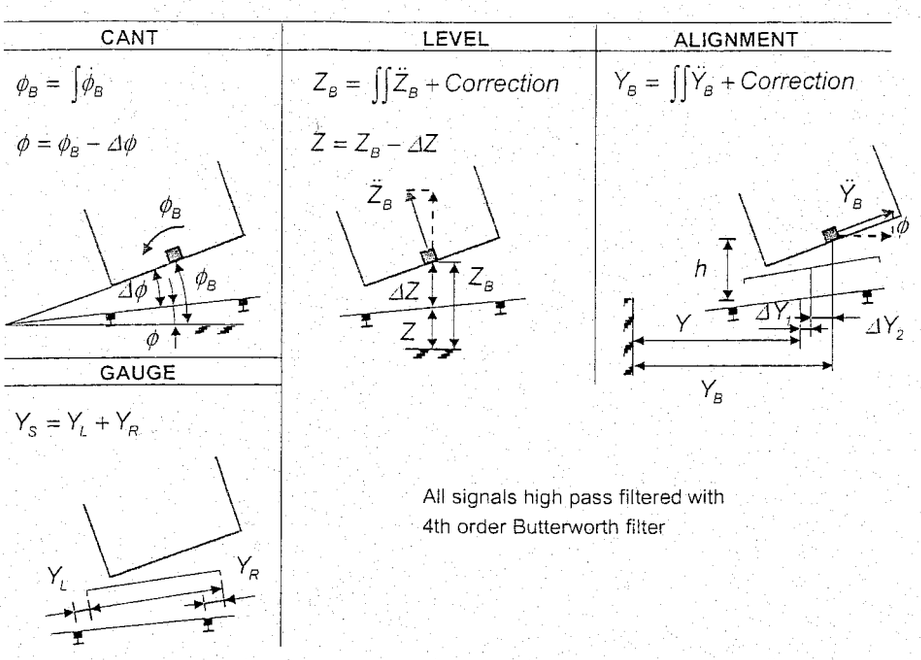


Figure 16.78: Recording principle of BMS-1 dynamic signals

The relative body displacement between track and vehicle body is subtracted from the thus determined absolute body displacement in order to obtain the vertical level of the track. This difference in displacement is actually measured as the change in length of a string joining the axle box and the vehicle roof in the manner shown in Figure 16.79. The advantage of this solution is that the corrections resulting from lateral displacement and rotations of the vehicle body remain small.

LVDTs with a non-linearity of 0.15 % for approximately half range, were used to measure the various relative displacements.

The acceleration transducers are the most important transducers and they must meet very high requirements. If at 70 m wavelengths an amplitude of 1 mm has to be determined at a running speed of 12 m/s, the acceleration measured will amount to 10^{-3} m/s^2, while under normal operating conditions the transducer must be able to process accelerations with an effective value of the order of 0.3 g, distributed over a large frequency range.

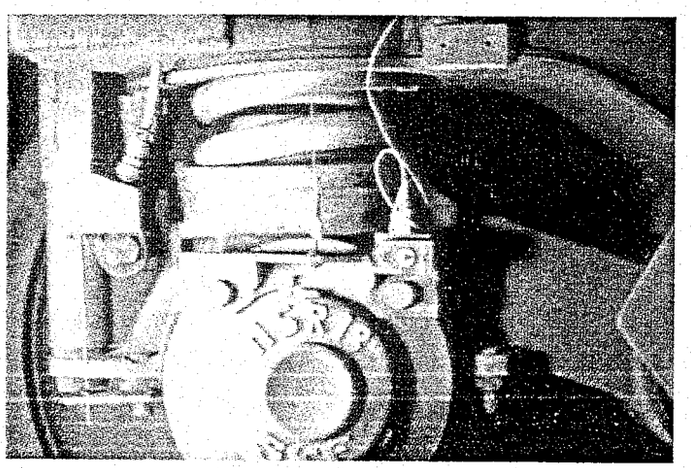


Figure 16.79: Transducer for displacement between axle box and car body

Non-linearities can also be the cause of large errors. For example, a non-linearity of 0.1 % means an additional output of the order of 0.5 mm. The transducers must therefore meet very high linearity requirements, with a non-linearity of less than 0.01 %.

The bandwidth is 700 Hz to prevent phase problems occurring in the most important low-frequency area.

Cant

To measure cant the change in car body rotation per unit of time can be measured by means of a so-called rate gyro, which is the most expensive transducer in the whole system. The rate gyro has a measuring range of +/- 20 °/s, 24,000 rpm, a natural frequency of 20 Hz, and a damping ratio of 0.65.

The energy of the rate gyro is mainly around 0.5 Hz and is practically negligible above 2 Hz. As the phase error in this range is inadmissibly large a phase correction is carried out with a second order filter. This restores the phase up to about 7 Hz.

Lateral

In the lateral direction the differential displacements are measured in two stages, i.e. between both rails and bogie frame by means of a laser system and between bogie frame and car body by means of linear displacement transducers. As a result of body rotation there is a rotation in the accelerometer. In order to prevent gravity causing a virtual lateral acceleration, the measuring system electronically eliminates the gravity component.

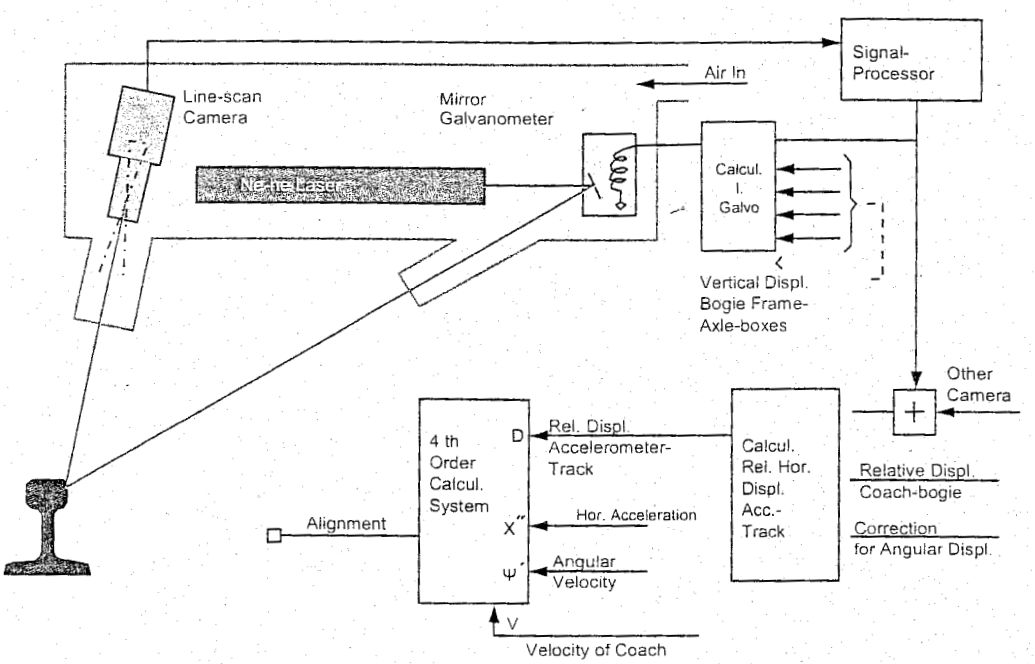


Figure 16.80: Principle of laser system

The horizontal displacement between bogie frame and track is measured by means of line scan cameras of the Reticon LC 100 type which contain an array of 256 photodiodes. The point on the rail is marked by a light-spot with a 1.6 mm diameter originating from a 5 mW NeHe Spectra Physics model 105 laser. The laser beam is projected onto the rail, 14 mm below the running surface, via the mirror of a General Scanning G 108 galvanometer. The lay-out of this assembly is annotated in Figure 16.80. A control system adjusts the angular rotation of the mirror in such a way that the desired position is maintained regardless of the movement of the bogie frame.

Line-scan camera, NeHe laser, and mirror galvanometer are incorporated in one unit on a 15 mm thick aluminium frame. Since the laser beam and the array of the line-scan camera must be in one plane, whereas the rotation axis of the galvo must be exactly perpendicular to this plane, at least two of the three components must be adjustable. All three components have in actual fact been fixed in such a way that they can be adjusted by screws. Adjusting is done in the workshop, after which the unit is mounted in the bogie in one piece as shown in Figure 16.81.

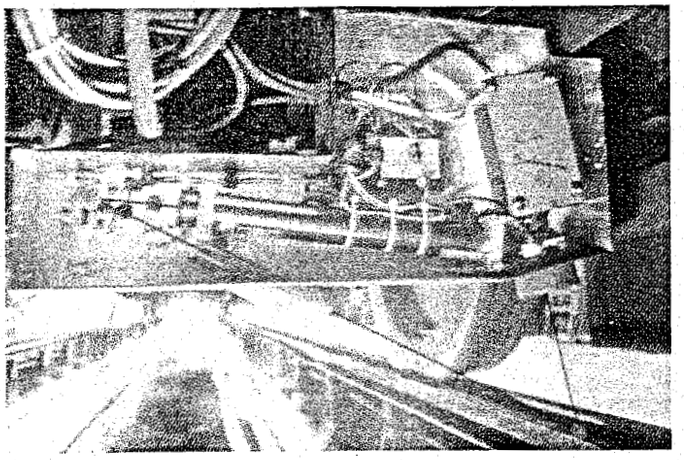


Figure 16.81: Laser system mounted in measuring bogie

The mounting frame is provided with an aluminium cover with rubber gaskets for sealing and has two openings which are protected by nozzles. The camera "looks out" through one of the openings and the laser beam exits via the other. A third opening admits filtered air under pressure at a rate of approximately 100 m^3/h. As this air exits through the other openings at considerable speed, it prevents dirt and dust from entering through these openings.

The cameras are on the bogie frame and are directed at an angle of 1: 5 to the gauge face of the rail head at the point where the light-spot is located. The cameras are equipped with an array of 256 photodiodes, providing a measuring range of 50 mm. The light falling on a diode is integrated for about 1 ms.

Signal processing partly takes place in the hardware and partly by a microprocessor. The latter determines the sliding average of the last 8 scans.

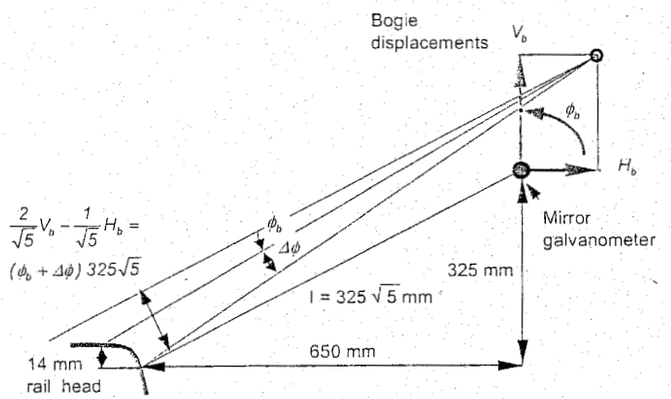


Figure 16.82: Principle of galvo control to position laser spot on rail head

The laser spots are kept at the correct level of 14 mm below the rail head by controlling the current through the mirror galvanometer. The galvanometers have a natural frequency of about 1500 Hz and are critically damped. To control the galvo angle, the bogie displacements relative to the track are used as depicted in Figure 16.82. For a vertical displacement V_b, a horizontal displacement H_b, and a rotation ϕ_b the galvo angle should be corrected by:

$$\Delta\phi = -\phi_b - \frac{H_b}{\ell\sqrt{5}} + \frac{V_b}{\frac{1}{2}\ell\sqrt{5}} \qquad (16.5)$$

in which $\ell = 325\sqrt{5}$ mm.

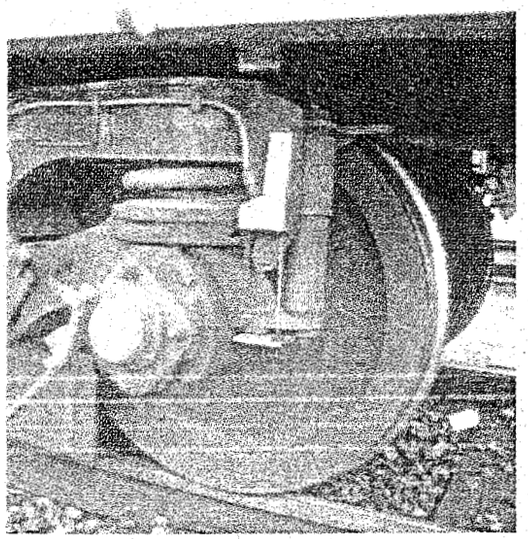
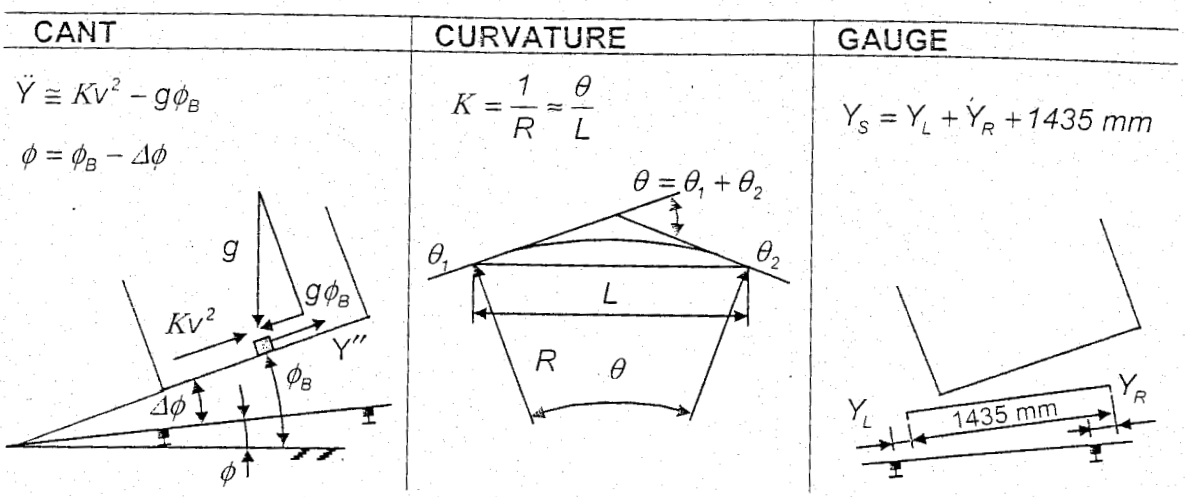


Figure 16.83: Transducer for vertical displacement between axle and bogie frame

The lateral bogie displacement is obtained from the camera output. Vertical bogie displacement and bogie rotation are determined by means of 4 linear displacement transducers, as shown in Figure 16.83, which have been mounted between bogie frame and axle.

16.10.4 Quasi-static signals

In addition to the dynamic signals, quasi-static signals are produced for cant, curvature, and gauge. The measuring principles are presented schematically in Figure 16.84. Problems due to drift and lack of initial conditions mean that the quasi-static cant is not determined by integration of the rate gyro signal, but by making use of the lateral car body acceleration, curvature, and recording speed as indicated in Figure 16.84.



All signals low pass filtered with 3rd order Bessel filter

Figure 16.84: Recording principle for BMS-1 quasi-static signals

Two other transducers are, in addition to the two linear displacement transducers between car body and bogie frame in the measuring bogie, installed in the second bogie to record curvature according to the principle sketched in Figure 16.84. The quasi-static gauge is directly derived from the line-scan camera signals.

All three quasi-static signals are low-pass filtered with a 3rd-order Bessel filter. The phase relationship is linear and causes a distance delay of 27.5 m. As the curvature recorded in fact corresponds to the car centre, the delay of the signal in relation to the measuring bogie amounts to 35 m or 20 m for forwards and backwards running respectively. With the introduction of VRA this delay can be corrected.

16.10.5 Signal combination for determining track parameters

The track parameters produced by the former BMS system of NS are derived from a combination of signals provided by 16 transducers, i.e. 1 rate gyro, 2 accelerometers, 10 LVDTs, 2 line-scan cameras, and 1 tachometer. The location of the various transducers is indicated in Figure 16.85. The track parameters are derived from the following expressions:

level:

$$Z_r = \iint \ddot{z} - \frac{z_1 + z_2}{2} + \Delta z \qquad (16.6)$$

cant:

$$\Phi_r = \int \frac{\dot{\phi}}{0.7} + z_1 - z_2 \qquad (16.7)$$

alignment:

$$Y_r = \iint \ddot{y} - \iiint 9.81\dot{\phi} - \frac{y_1 + y_2}{2} + \ell_{10}\phi_r + \Delta y \qquad (16.8)$$

curvature:

$$K_r = \frac{y_3 + y_5 - y_4 - y_6}{\ell_2 \ell_2} \qquad (16.9)$$

static cant:

$$C_r = \frac{K_r v^2 - \ddot{y}}{0.7 \cdot 9.81} + z_1 - z_2 \qquad (16.10)$$

corrections:

$$\Delta z = \left[\frac{y_3 + y_4}{2} + \frac{\ell_1}{2\ell_2}(y_3 - y_4)\right]^2 \frac{1}{2\ell_3} \qquad (16.11)$$

$$\Delta y = -\frac{y_1 - y_2}{\ell_9}(\ell_5 - \ell_8) - \frac{z_1 - z_2}{2}\left(\ell_8 + \frac{z_1 + z_2}{2}\right) \qquad (16.12)$$

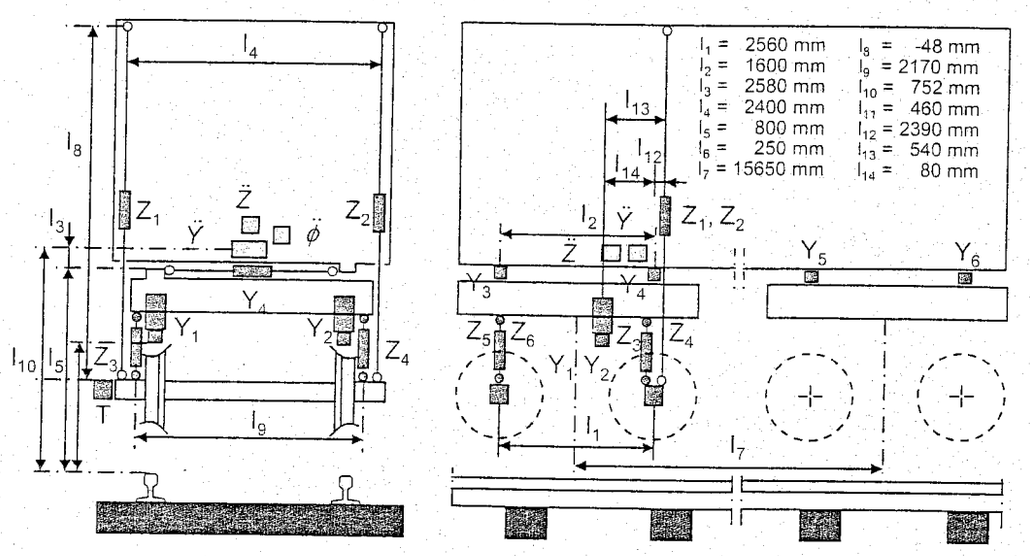


Figure 16.85: Location of various BMS-1 transducers

The Δz correction originates from the requirement that car body movements relative to the measuring axle may not produce any output in level. The Δy term is related to the removal of the gravity component in the alignment signal. The twist signals are created directly in the PDP computer as the difference in cant and the cant signal delayed over 2.75 m and 6 m respectively.

The dynamic signals are high-pass filtered by a 4th-order Butterworth filter with a transfer function governed by the following expression:

$$H(s) = \frac{\lambda_c^4 s^4}{\lambda_c^4 s^4 + c_1 \lambda_c^3 s^3 + c_2 \lambda_c^2 s^2 + c_3 \lambda_c s + 1} \qquad (16.13)$$

in which:

s = complex spatial frequency = i/λ;
λ_c = cut-off wavelength at - 3dB = 70 m;
c_1 = c_3 = 2.613;
c_2 = 3.414.

The alignment Y_r governed by (16.15) is obtained by the network presented in Figure 16.88. This network is composed of integrators and multipliers which take care of integration and filtering. In the frequency domain the network provides the following alignment output:

$$Y_r(p) = \left[-9.81\frac{\Phi(p)}{p^3} + \frac{\dot{Y}(p)}{p^2} + \Delta Y_c(p) \right] H(p) \qquad (16.16)$$

in which:

$$H(p) = \frac{1}{1 + \frac{a_1}{p} + \frac{a_1 a_2}{p^2} + \frac{a_1 a_2 a_3}{p^3} + \frac{a_1 a_2 a_3 a_4}{p^4}} \qquad (16.17)$$

p = complex radial time frequency [rad/s].

The coefficients of the 4th-order high-pass Butterworth filter described by (16.13) are produced by means of the multiplication factors of the network according to:

$$\begin{aligned} a_1 &= c_1\beta \\ a_2 &= c_2/c_1\beta \\ a_3 &= c_1/c_2\beta \\ a_4 &= 1/c_1\beta \\ \beta &= v2\pi/\lambda_c \\ v &= \textit{running speed of the recording car [m/s]} \end{aligned} \qquad (16.18)$$

Thus, by continuously adapting the multiplication factors of the network to the speed, a spatial wavelength cut-off at 70 m is maintained.

16.10.6 Signal analysis

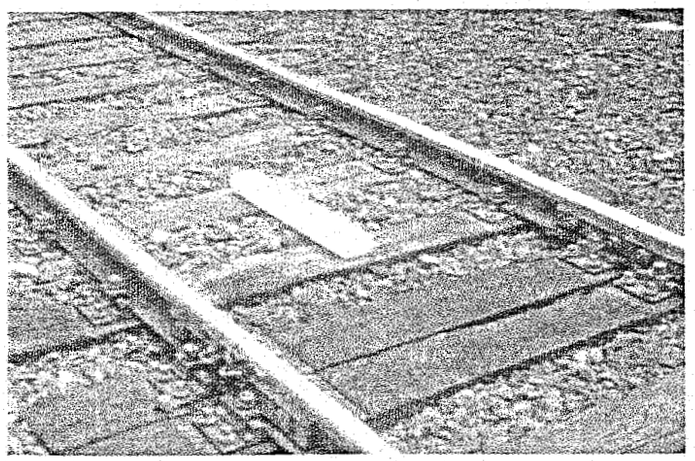


Figure 16.89: Reference magnet

A reliable system for determining exact location forms an important aspect of the analysis of recorded data. The NS Permanent Way Department has for many years used the magnets shown in Figure 16.89 for this purpose. There are about 400 of these magnets in NS main tracks, which are also used as reference points by the ultrasonic inspection train.

When the measured signals are analysed, it must be remembered that the information should be separated according to the type of track maintenance. This will enable the particular geometry component in question to be corrected.

For mechanical maintenance the analysis is based on standard deviations determined over a length of 200 m, whereas, for manual maintenance, locations where a predetermined threshold value is exceeded are used as basis.

The analysis is carried out for all signals in the 0.5 - 25 m waveband, whilst the 25 - 70 m waveband is analysed for alignment only. The long waves are especially important regarding vehicle dynamics, as they generate low frequency steady-state car body vibrations in straight, or almost straight, tracks. This long wave information serves to determine whether it is necessary to carry out track lining with the aid of lasers and the DRIVER system [77].

Standard deviations based on the short waveband are decisive for determining use of tamping machines in the automatic mode. It is also this short waveband from which exceedences are deduced, as they are necessary for manual maintenance. Above all, though, this short waveband is used to check whether a track may be considered safe.

The analogue measuring signals are also reproduced graphically in duplicate.

Quality indices

Track recording cars normally generate standard deviations, often per 200 m section in various wavebands. Local defects are disregarded here. As standard deviations are hard to interpret for non-experts, preference could be given to a QI deduced from the standard deviation. The starting point with regard to QIs can be summarized as follows:

a. significance for each signal is identical;
b. significance for each quality class (QCL) is identical;
c. choice designated such that the maximum value is 10 and the minimum value is 0;
d. the break-even point for maintenance lies in the vicinity of QI= 6.

To satisfy the above requirements as well as possible, NS took the national geometry distributions for spring 1983 as a reference. Figure 16.90 shows the distributions of standard deviations per 200 m related to the wavebands produced by BMS track recording car and the 3 QCLs. They indicate what percentage of the length has a value less than the value on the x-axis.

In order to achieve the best possible match, the curves are transformed by changing from standard deviation to QI, via an exponential function according to:

$$QI = 10\left[\frac{QI_{80}}{10}\right]^{\sigma/\sigma_{80}} \qquad (16.19)$$

This function already satisfies the requirement that a QI of 10 applies when $\sigma = 0$.

Via the transformation just one more parameter can be chosen. Quality assessment does not require an accurate knowledge of either the very poor or the very good sections. In fact only those close to the maintenance intervention level are of interest. On average, 20 % of the NS network per annum is recommended for tamping. Although arbitrary, it is an obvious choice to assign the same QI, referred to as QI_{80}, to the standard deviations at 80 %, known as σ_{80}, for each QCL.

As the lowest value of each of the 3 QIs in the waveband 0 - 25 m is always chosen as overall QI per 200 m section, the 80 % point of this distribution for each QCL is, by definition, set at $QI_{80} = 6$. With this choice the 80 % values for the individual signals in the 3 QCLs appear to lie at a QI_{80} of respectively 6.75, 6.85 and 6.65; this of course only applied to the calibration campaign of spring 1983.

The σ_{80} values have been determined for the three QCLs and four geometry components from the distributions shown in Figure 16.90. They are grouped together in the table in Figure 16.91. With the relationships between QI and standard deviation the national distributions as a function of QI are obtained, as also shown in Figure 16.90. They all pass, by definition, through the point QI_{80} and consequently have the same significance. This intermatching should, of course, only be done once.

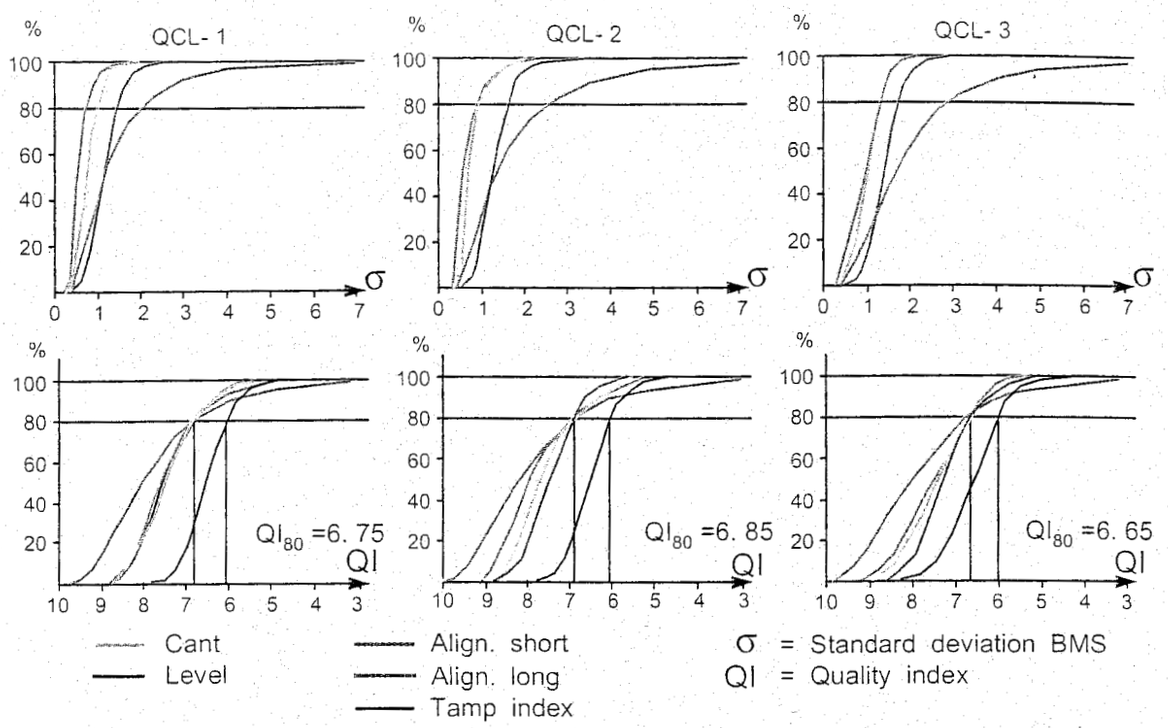


Figure 16.90: Cumulative distribution functions for NS network spring 1983

During the recording runs a quality index report is printed for cant, level, and alignment. As the indices of the previous run are available on floppy disk, the difference between the previous and the actual results is also printed out. If desired, it is also possible to print out the standard deviations, for instance, during a test run.

Furthermore, mean value and standard deviation are printed for gauge, but, in addition, maximum value in curves is also printed. The latter is used to derive lateral rail wear data. If any exceedences are found these are reported too, but the details are printed on a separate list. To facilitate interpretation the maintenance section numbers are printed as well.

Normalized standard deviations and peak values

The normalizing procedure employed on NS had been chosen because of the school marking system used in Holland. Such systems are, however, different from country to country. An interesting alternative is therefore to apply normalized standard deviations and peak values. In this approach standard deviation s and peak value P should be divided by the corresponding standard. If these standards are referred to as σ_{norm} and P_{norm} respectively then the normalized values become:

$$\bar{\sigma} = \sigma / \sigma_{norm} \qquad (16.20)$$

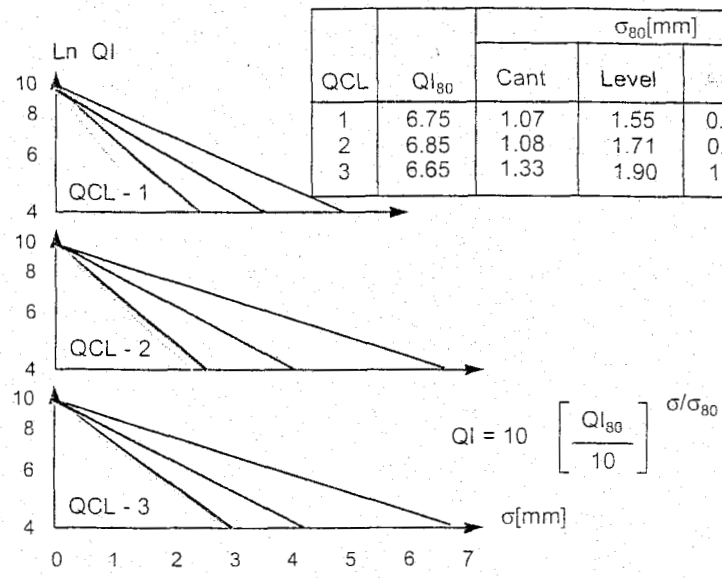
$$\bar{P} = P / P_{norm} \qquad (16.21)$$

If this is consistent, all the normalized values have, like the quality indices, the same significance. Normalized values less then 1 comply with the standards, whereas normalized values beyond 1 reflect that maintenance action is required. In fact the σ_{80} values discussed previously can be regarded as the norm in this approach. The advantage of normalized standard deviations is furthermore that a better resolution is obtained, especially in poor situations. In chapter 17 dealing with high-speed tracks further attention is given to the precise interpretation of normalized values.

Quality classes

In the past, maintenance advices were established manually from quality indices per 200 m section. For automated planning and control it is indispensable to start from a number of quality classes (QCLs) to differentiate between jointed and long-welded (CWR) tracks on the one hand and between high and low speeds on the other.
The NS network has been split up into 3 QCLs according to Figure 16.91. The QIs for each class all have the same significance with regard to the quality level.

QCL-1 pertains to main lines with a permissible speed V > 130 km/h. Main lines allowing 100 km/h < V < 130 km/h, with mainly long-welded (CWR) tracks are classified in QCL-2, whilst all jointed tracks and long-welded tracks with V < 100 km/h, forming part of the main line network, belong to QCL-3.



QCL	QI_{80}	σ_{80}[mm] Cant	Level		AI - I
1	6.75	1.07	1.55	0.78	2.14
2	6.85	1.08	1.71	0.95	2.78
3	6.65	1.33	1.90	1.31	3.05

Figure 16.91: Relationship between QI and σ

Tamping

From the point of view of ride comfort and track forces, the quality of the various geometry components for each MAINS should not only be as good as possible, but also as uniform as possible. Since, in principle, all QIs have the same significance, the lowest index of the three indices pertaining to cant, level and alignment is retained as the maintenance index or tamping index for each 200 m section. If normalized standard deviations are used the highest value per 200 m should be taken.

The next step necessary to determine from the 200 m sections a quality index for each MAINS is as depicted in Figure 16.92. The tamping index for each 200 m section will not be constant throughout the length of the MAINS. In order to allow for this variation, not the mean value but the value at 80 % is adopted so that 80 % is better and 20 % is worse. The QI per MAINS in the waveband 0.5 - 25 m thus established, or alternatively the σ, furnishes the final assessment as far as automatic tamping/lining is concerned.

QI per 200 m section			
0 - 25 m			25 - 70 m
Cant	Level	Alignment	Alignment
6.3	5.7	6.7	6.5

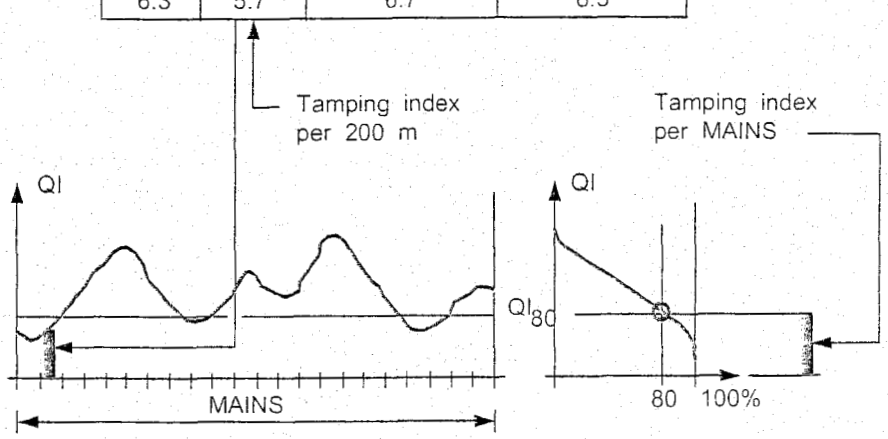


Figure 16.92: Determination of tamping index per MAINS

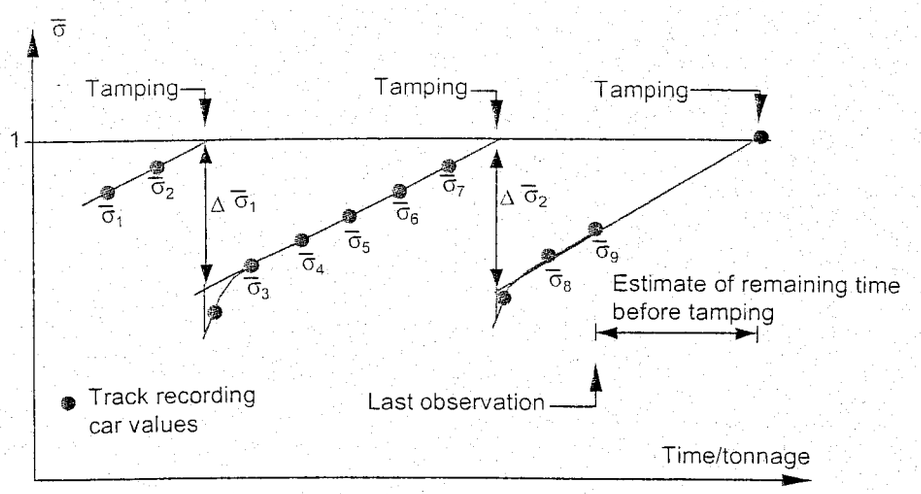


Figure 16.93: Examination of the time to tamping using normalized standard deviations

Figure 16.93 shows a typical plot of the σ_{80} value for a MAINS as a function of time or tonnage borne. The tamping intervention level is unity by definition. From the average rate of deterioration for a particular MAINS, determined from the preceding values, the intersection with the maintenance intervention level can be estimated and consequently the date that tamping will be necessary.

As indicated in Figure 16.93 average quality σ_A, average improvement due to tamping $\Delta\bar{\sigma}_{tamp}$ and average rate of deterioration $\Delta\bar{\sigma}/\Delta t$ may be estimated as follows:

$$\bar{\sigma}_A = \frac{1}{N}[\bar{\sigma}_1 + \bar{\sigma}_2 + \ldots + \bar{\sigma}_N] \tag{16.22}$$

$$\Delta\bar{\sigma}_{tamp} = \frac{1}{N}[\Delta\bar{\sigma}_1 + \Delta\bar{\sigma}_2 + \ldots + \Delta\bar{\sigma}_N] \tag{16.23}$$

$$\Delta\bar{\sigma}/\Delta t = \frac{1}{N}\left[\frac{\bar{\sigma}_2 - \bar{\sigma}_1}{T_2 - T_1} + \frac{\bar{\sigma}_3 - \bar{\sigma}_2}{T_3 - T_2} + \ldots\right] \tag{16.24}$$

With the determination of the average rate of deterioration only those points which are not influenced by rapid deterioration following tamping, i.e. at least 0.7 MGT away from the tamping operation, should be considered.

For overall analysis of the quality of all or part of a network, for instance for planning the disposition of maintenance resources, present and predicted average quality figures over entire routes are very suitable. For this purpose priority lists may be created with the MAINS in the sequence of the tamping index: the MAINS with the lowest quality is found at the top of the list and the MAINS with the highest quality at the bottom.

Apart from the tamping index, which dictates the sequence, the 80 % values for cant, level, and alignment in the 0.5 - 25 m waveband and alignment in the 25 - 70 m waveband are calculated in an analogous way.

Figure 16.94 shows the cumulative distribution functions of the mean deterioration of 4 successive recording car campaigns, starting in spring 1984. The distribution of the tamping index reveals that 80 % of the MAINS deteriorate less than 0.23 units every six months and 50 % less than 0.11 units. How this works out for the individual track geometry components can be seen from the distribution functions in the right-hand part. The deterioration in mm standard deviation per year for 80 % of the network is less than 0.14 for cant, 0.20 for level and 0.16 for alignment. The 50 % values are as follows: cant 0.08, level 0.14, alignment 0.10.

As a further example of useful management information consider Figure 16.95 which shows how the quality of geometry has altered over the years 1984 - 1988 on NS. Here the 80, 50 and 20 % quality levels are shown for cant, level and alignment. This type of information is useful in monitoring the effects of new and improved methods of maintenance. Furthermore it is of great interest to note that the curves hardly display any variation.

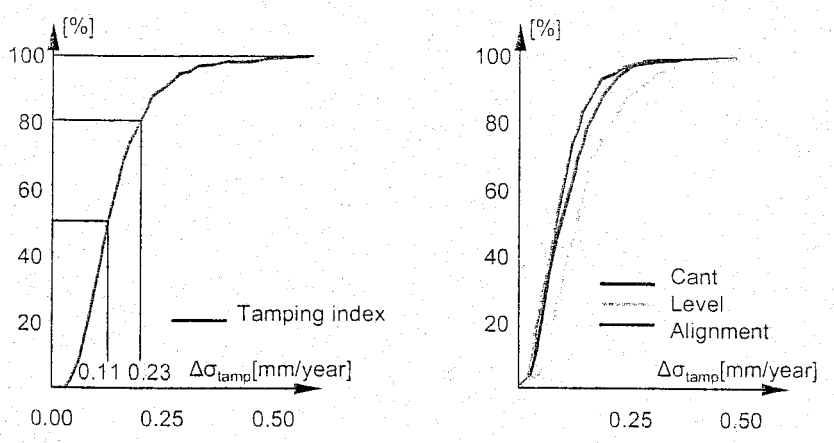


Figure 16.94: Cumulative distribution functions of track deterioration on NS

With the empirical QIs only a relative tamping sequence or priority can be given. In fact the whole maintenance process is focused on retaining the arbitrarily chosen reference situation of spring 1983. Of course this is far from ideal and so there are many advantages in using vehicle reactions for the assessment of track quality.

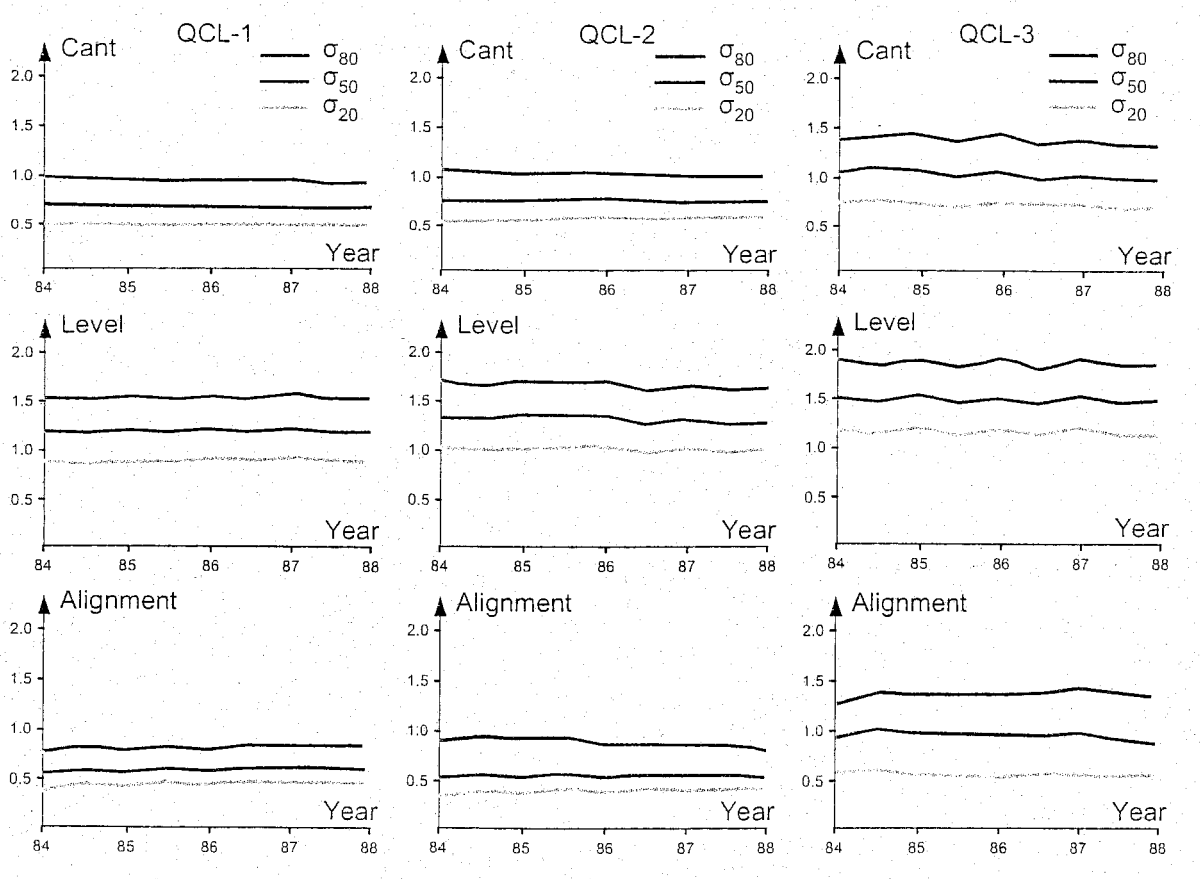


Figure 16.95: Trends in track geometry on NS

Curves

In curved tracks vehicle dynamics are more of a transient nature and long-wave resonance phenomena rarely occur. Thus, in sharp curves the short waves in the 0.5 - 25 m waveband and the quasi-static signal will normally be the main concern.

In transition curves a fault-free alignment still produces an output in the measured alignment signal. Although the aim is to measure the deviation from the theoretical transition, the recording system presents the sum of the deviation and theoretical value. Tamping machines should only correct the first component and not the design values, and therefore there is a need to separate both contributions. In the 0 - 25 m waveband the contributions from the design line are negligibly small, but in the 25 - 70 m waveband they may prevail. With the introduction of the array processor for vehicle response calculations (VRA), part of the theoretical alignment can be estimated in a relatively simple way. For each interval of 10 m the difference in curvature:

$$\frac{\Delta K}{\Delta x} = y'''(x) \tag{16.25}$$

is determined from the quasi-static curvature, which corresponds to the third derivative of the alignment signal. The values are collected if $\Delta K/\Delta x > 0.5\ 10^{-5}\ m^{-2}$. These values are approximated by step functions according to Figure 16.96.

The BMS-1 system records the design alignment contribution in the frequency domain as follows:

$$Y_{BMS}(s) = H_{70}(s)\, Y(s) \tag{16.26}$$

$$= \frac{1}{(2\pi s)^3} H_{70}(s)\, Y'''(s) \tag{16.27}$$

$$= H_{\kappa}(s)\, Y'''(s) \tag{16.28}$$

In combination with (16.13), the transfer function to be applied to the first derivative of the curvature is found to be:

$$H_{\kappa} = \frac{1}{(2\pi s)^3} H_{70}(s) \tag{16.29}$$

$$= \frac{\lambda_c^4 s/(2\pi)^3}{\lambda_c^4 s^4 + c_1 \lambda_c^3 s^3 + c_2 \lambda_c^2 s^2 + c_3 \lambda_c s + 1} \tag{16.30}$$

in which

s = complex spatial frequency = i/λ
λ_c = cut-off wavelength at - 3dB = 70 m
c_1 = c_3 = 2.613
c_2 = 3.414

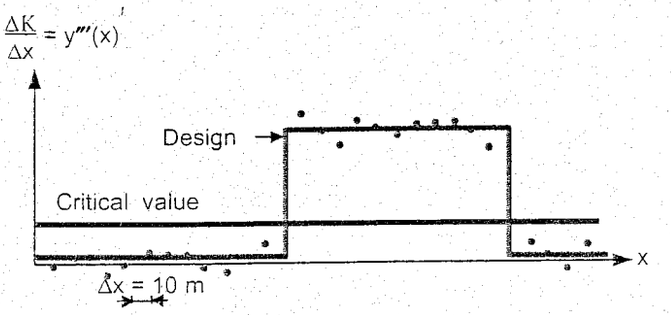


Figure 16.96. Principle of determining the design curvature from the quasi-static curvature

The solution adopted in BMS consists of transforming the y'''(x) signal, simultaneously with the other signals for VRA, to the frequency domain. This complex function is then multiplied by the triple-integrated 4th-order Butterworth filter (16.30) and is back-transformed. The variance of this signal is determined for the 200 m section concerned and subtracted from the variance of the total alignment in the 0 - 70 m waveband. After also subtracting the variance in the 0 - 25 m waveband and

taking the square root, the corrected standard deviation in the 25 - 70 m waveband is obtained including a value reflecting the amount of irregularities to be removed by track maintenance.

With the aid of the quasi-static signals, BMS-1 checks whether the non-compensated lateral acceleration complies with a specified standard according to:

$$\ddot{y}_{nc} = \frac{V_m^2 |K_r|}{12.96} - \frac{|C_r|}{146.28} + \frac{3\sigma_{cant}}{146.28} < norm \qquad (16.31)$$

in which:
K_r = quasi-static curvature [1/m];
C_r = quasi-static cant [mm];
σ_{cant} = standard deviation of dynamic cant [mm];
V_m = admissible line speed [km/h];
norm = admissible non-compensated lateral acceleration [m/s^2];
= 0.4 - 0.8 m/s^2 depending on ballast and track conditions.

The admissible non-compensated lateral accelerations correspond to a cant deficiency of 60 - 120 mm. It is, furthermore, verified that the cant value C_r remains smaller than 150 mm. Moreover, the derivatives of cant and curvature are tested against permissible values which follow from the variation in non-compensated lateral acceleration according to:

$$\frac{\Delta \ddot{y}_{nc}}{\Delta t} = \left[\frac{V_m}{3.6}\right]^3 \frac{\Delta K_r}{\Delta x} - \frac{g}{1500} \frac{V_m}{3.6} \frac{\Delta C_r}{\Delta x} \qquad (16.32)$$

In this expression both terms are confined to the value a_{lim}. For the variation in cant and curvature this leads to respectively:

$$\frac{\Delta C_r}{\Delta x} < \frac{550}{V_m} \dot{a}_{lim} \qquad [mm/m] \qquad (16.33)$$

$$\frac{\Delta K_r}{\Delta x} < \frac{46.66}{V_m^3} \dot{a}_{lim} \qquad [1/m^2] \qquad (16.34)$$

These expressions have been implemented in BMS for the NS standards, with $\dot{a}_{lim}$ = 0.18 m/s^3, according to:

$$\frac{\Delta C_r}{\Delta x} < \alpha\ 10^3 \qquad [mm/m] \qquad (16.35)$$

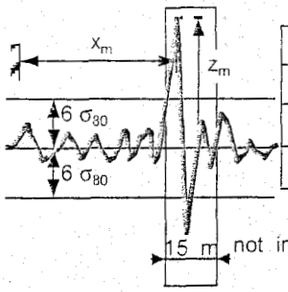
$$\frac{\Delta K_r}{\Delta x} < \frac{\alpha\ 10^3}{11.8 V_m^2} \qquad [1/m^2] \qquad (16.36)$$

in which:

$$\alpha = \frac{1}{10\ V_m} \qquad \text{if } V_m \geq 80 \text{ km/h} \qquad (16.37)$$

$$\alpha = \frac{1}{600} \qquad \text{if } V_m < 80 \text{ km/h} \qquad (16.38)$$

If within a 200 m section one of the parameters exceeds the specified standard, then the exceedence report is printed.



	Level 1 Maintenance	Level 2 Safety	Level 2 Recommended
Twist	10 mm	20 mm	14 mm
Cant	6 σ_{80}	20 mm	12 mm
Level	6 σ_{80}	20 mm	18 mm
Alignment	6 σ_{80}	20 mm	18 mm
	Minimum 8 mm		

Figure 16.97: Principle of determining local defects

Local defects

The examination of isolated defects is limited to cant, level, alignment, and twist in the 0.5 - 25 m waveband. A value of $6\sigma_{80}$, with a minimum of 8 mm, is used as exceeding level for detecting local defects pertaining to cant, level, and alignment within a maintenance section. A fixed value of 10 mm is applied for twist. Outside the maintenance sections a fixed threshold of 12 mm is applied, except for twist which is again set at a value of 10 mm.

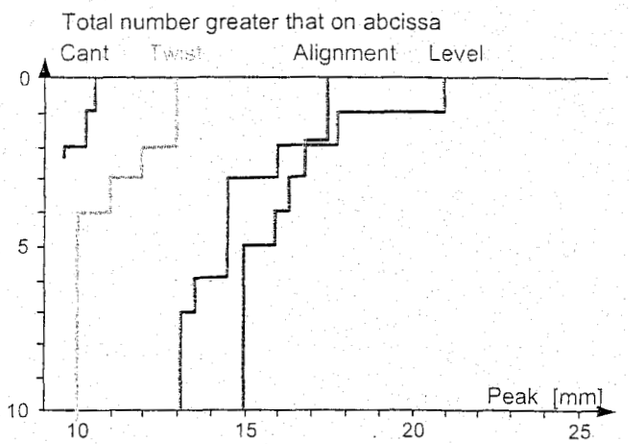


Figure 16.98: Total number of exceedences beyond 10 mm on NS tracks as of spring 1985

When one of the signals exceeds the corresponding level-1 value mentioned in Figure 16.97, the system searches for the maximum of all signals over the following 15 m. The information on this 15 m is not included in the calculation of the standard deviation for the relevant 200 m section, on the principle that a contribution should only be included in the calculations for either mechanical or manual maintenance.

The local defects beyond a peak value of 10 mm, found during the BMS campaign of Spring 1985, can be read from Figure 16.98. The maximum values obtained on NS main line tracks for cant, level, alignment, and twist were 10.5, 21, 17.5, and 13 mm respectively. From these findings it can be deduced that the level-2 limits should be altered according to the values specified in Figure 16.98.

Gauge tolerances

From the point of view of stable vehicle movement gauge variations should be limited. This applies especially to the lower limit. For new tracks NS applies the following gauge tolerances:

Mean gauge per section of 200 m 1435^{+3}_{-1} mm

As within a 200 m section the variation of the mean value is normally not greater than +/- 1 mm, gauge can always lie between 1433 m and 1439 mm. In terms of BMS-1, these limits correspond to the mean value minus three times the standard deviation and the mean value plus three times the standard deviation respectively.

The target for gauge remains, however, 1435^{+2}_{-0} mm

16.11 Vehicle response analysis according to VRA

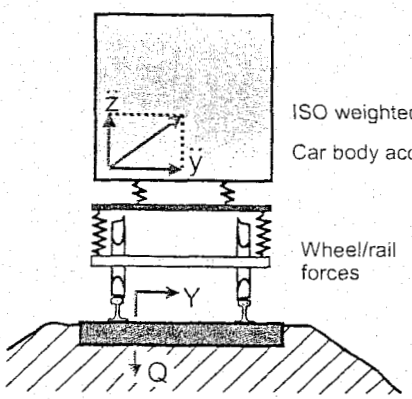
16.11.1 Introduction

The great problem in the analysis of track geometry lies in the combination of the various components and the assessment of these values. In the final analysis, vehicle reactions are clearly the decisive factors. Passenger comfort criteria are all based on weighted car body accelerations. Tracks experience loads in vertical and lateral directions which lead to stresses in the track components and ballast bed causing deterioration of geometry and fatigue of components. Safety criteria are also specified in terms of Y and Q forces. If all these aspects are to be considered then the governing vehicle reactions have to be determined. This was the reasoning behind the development of VRA which has been in use as part of the BMS system since 1988.

The VRA system is described in detail in Chapter 6 together with the procedures for estimating the transfer functions for the vehicles considered, and the frequency-dependent weighting applied to the calculated car body accelerations.

16.11.2 Principle of calculation

As indicated in Figure 16.99, VRA calculates 4 vehicle reactions, i.e. vertical and horizontal car body acceleration weighted according to ISO-2631, but also the vertical and horizontal track load for 3 types of rolling stock at 5 running speeds. The inputs are digitised by BMS and consist of cant, level alignment, and gauge. In the frequency domain the Fourier transforms are multiplied by the corresponding transfer functions to create the Fourier transforms for the 3 vehicles and 5 speeds. Subsequently, the 60 response records are back-transformed.



Calculation of vehicle reactions in real time for:
- 3 vehicles: locomotive
  passenger coach
  freight car
- 5 speeds: 40-160 m/h

Calculations via array processor:
- Fourier transform of 4 geometry signals
- Computation of vehicle reactions via 240 transfer functions
- Determination of largest variance per speed range: these decisive 12 values are printed
- Back transformation of selected 12 responses
- Vectorial summation of car body accelerations
- Recording resulting in 9 response signals
- Determination of exceedences

Figure 16.99: Vehicle reactions calculated by VRA

Subsequently, a number of calculations per 200 m such as determining standard deviations are performed in the spatial domain. To limit the amount of data, the output is restricted to one speed at which maximum variance is attained, i.e. one speed for Q, one for Y, and one for the acceleration vector. Every 200 m, or 4 seconds at 180 km/h, the AP 500 thus has to carry out 60 Fourier transforms of 1024 data points and a great number of multiplications and additions.

As far as exceedences are concerned, a level-1 and a level-2 exceedence threshold can be specified per vehicle, as well as the mass per axle. Together with the non-compensated lateral acceleration obtained from the quasi-static signals according to (16.31), the quasi-static Q and Y forces are computed and added to the dynamic components provided by VRA. These total forces are used to test the derailment criterion according to:

$$Y / Q_{2m} < L_d \qquad (16.39)$$

If L_d = 0.6 is exceeded over a length greater than 2 m, this is reported as level 1. If the ratio exceeds L_d = 1.2, this is indicated as level 2. The Prud'homme criterion is tested against:

$$\frac{\sum Y_{2m}}{0.85\left[10+\frac{2Q}{3}\right]} < L_p \qquad (16.40)$$

If over more than 2 m this ratio exceeds a value of L_p = 0.5, this is indicated as level 1, and for L_p = 1.0 as level 2.

16.12 Results from BMS campaigns

16.12.1 NS distribution functions

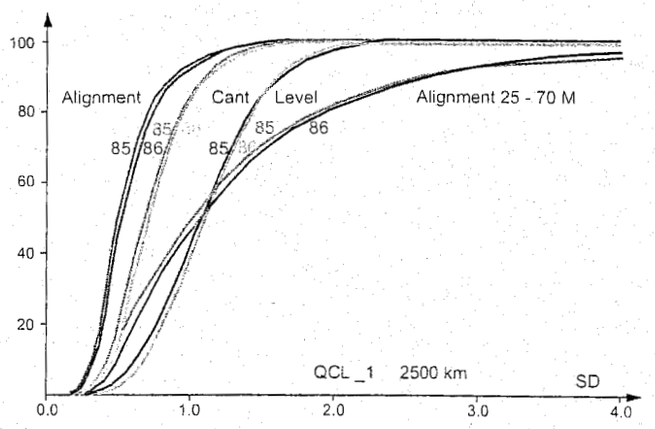


Figure 16.100: National distributions of standard deviations BMS per 200 m on NS network in autumn 1985 and autumn 1986

In this section a number of examples will be given of data produced by the BMS-1 track recording system belonging to NS. After completion of a recording car campaign the standard deviations for every 200 m were collected in distribution functions representing the overall quality of the network. Figure 16.100 shows the distribution functions for the NS tracks recorded in autumn 1985 and autumn 1986. The differences between the results of both recording campaigns are clearly quite small.

16.12.2 Results from the ORE D 161 Europe Tour

In Autumn 1983, ORE Committees D161 and B12 conducted a measuring tour through several European countries to test the influence of increasing axle loads from 20 to 22.5 tonnes. The NS track recording car formed part of the test train for the purpose of measuring track geometry.

Table 16.6 contains a survey of the track geometry measuring results collected in the different countries expressed in terms of the percentage points of the cumulative distribution functions of the standard deviations per 200 m.

D161.1	Cant [mm]	Level [mm]	Alignment [mm]	Cant [%]	Level [%]	Alignment [%]
DB	0.81	1.42	0.97	79	94	135
CSD	1.73	2.62	1.78	168	174	247
SNCB	1.32	2.14	1.70	128	142	236
SNCF	1.13	1.53	1.23	110	101	171
SBB	1.14	1.92	1.32	111	127	183
NS Q1	1.03	1.51	0.72	100	100	100
NS Q2	1.04	1.67	0.89	101	110	124
NS Q3	1.29	1.85	1.23	125	122	171

Table 16.6: 80%-points for various national distributions

Equivalent standard deviation values with respect to NS' BMS system, produced by the recording cars of various administrations, were published in [88] and are summarized in Table 16.7

Network	NS	DB	CFF	SNCF/ SJ	CSD	BR	FS	CFR	PKP
Level	1	1.24	0.91	0.91	1.52	1.14	1.33	1.40	0.73
Alignment	1	1.41	1.44	1.47	1.77	1.20	1.72	1.95	–

Table 16.7: Conversion table for track recording car output

16.12.3 Track geometry spectra

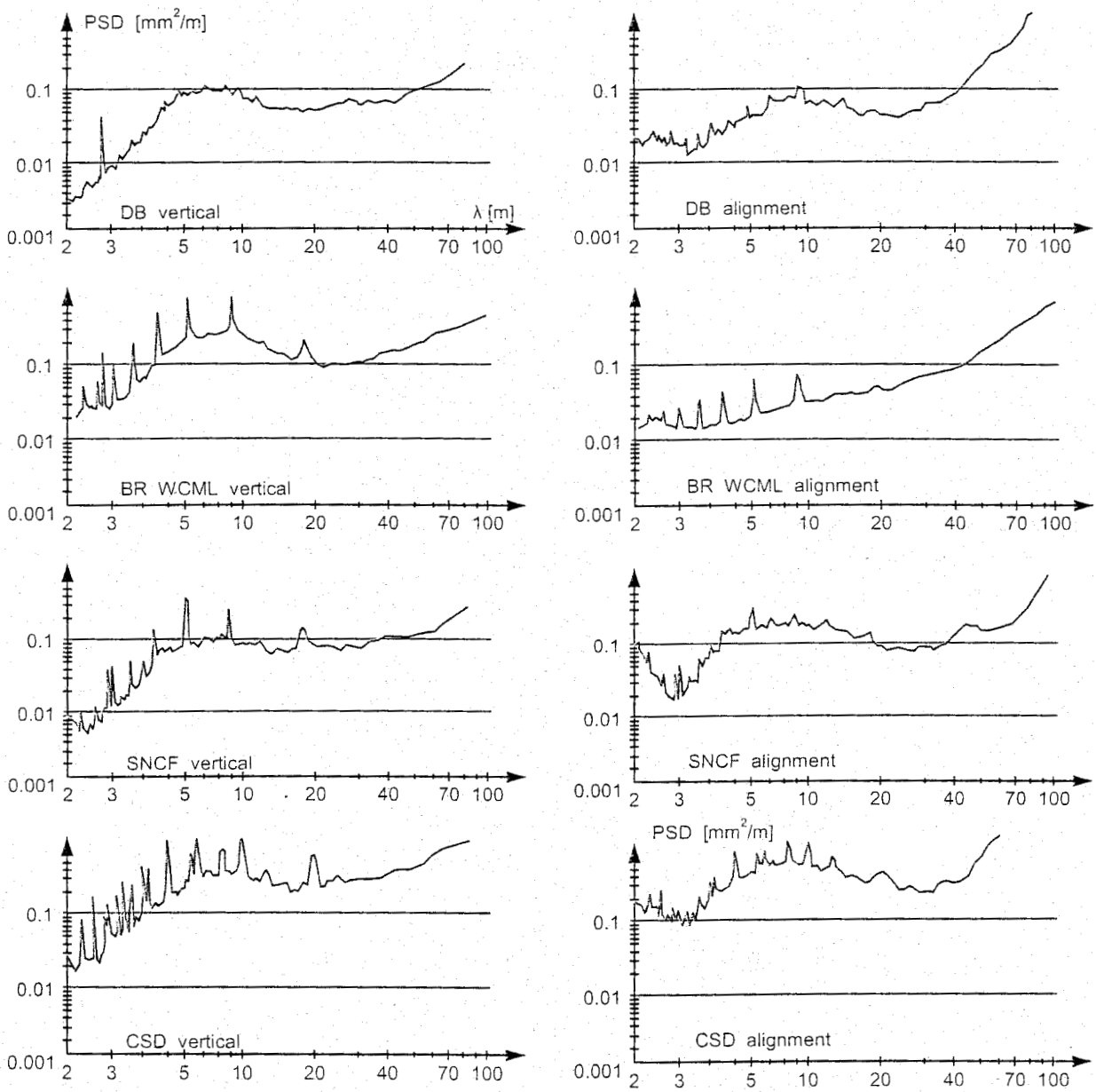


Figure 16.101: Psd-functions for vertical track geometry based on BMS measurements carried out for ORE D 161

A more detailed way of looking at track quality is by means of power spectral density functions, which show how the variance or energy of a signal is distributed over the wavelengths. The theory of how these functions can be determined is treated in Chapter 6.

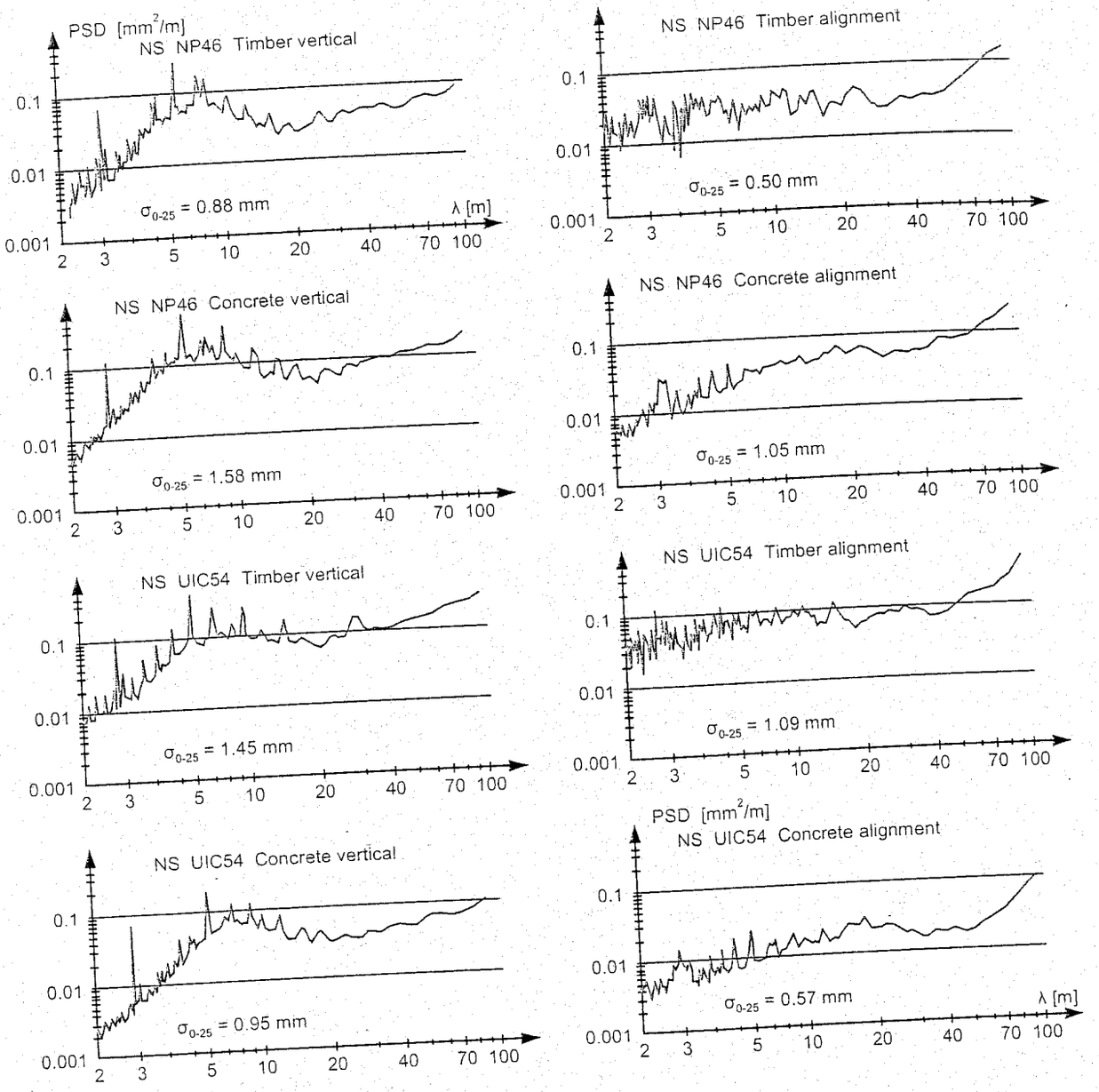


Figure 16.102: Average psd-functions for NS tracks

During the above-mentioned European Tour conducted by ORE D 161, a lot of information was collected on various railway administrations. The power spectral density functions for vertical and horizontal track geometry were determined for a number of tracks as collected in Figure 16.101. Analysis of vertical track spectra, as described in [252], indicates that at high speeds the main areas of influence of the track on the vehicle, when ISO-weighted, are centred around wavelengths of 10 and 40 m. In general terms, the vertical track spectra for different track qualities, presented in Figure 16.101, are parallel, that is, the power for all wavelengths increases proportionally as the track quality decreases. Nevertheless, although this is reasonably true for the 0 - 25 m waveband, the longer waves around 40 m show more variation.

Finally, some power spectral density functions for CWR tracks on NS are summarised in Figure 16.102. The vertical spectra are all very similar regardless of rail and sleeper type and clearly reveal the characteristic rolling wavelength introduced during rail manufacture. A pronounced peak at a 6 m wavelength is also observed, which coincides with the characteristic wavelength for vertical bending. The spectra increase linearly up to 8 m and then decrease slightly until 20 m, followed by an ascending trend. The alignment spectra increase linearly up to 20 m, remain constant up to 50 m and

then increase again. The increasing trend for long waves for both vertical and lateral spectra is explained by the recolouring operation which was carried out to compensate for the 4th-order Butterworth filter used by BMS.

16.13 T-16: FRA's High Speed Research Car

16.13.1 Introduction

In November 2000 the Federal Railroad Administration (FRA) began operating the high-speed research car T-16 (Figure 16.103) to investigate methods for providing a safer and smoother ride for passenger and freight trains travelling at higher speeds. FRA's Office of Research and Development uses T-16 to study the dynamics of wheel-rail behaviour, to investigate methods for improving track inspection, and to assess potential high speed rail corridors. The T-16 also assists Amtrak's local and regional commuter authorities, and individual railways in assessing their routes.

The T-16 is a former Amtrak Metroliner passenger car which was refurbished and instrumented with advanced technology, with capabilities for measurement and data collection at speeds up to 160 m.p.h.. Measurement capabilities include track geometry, rail head profile, ride quality, and wheel-rail forces.

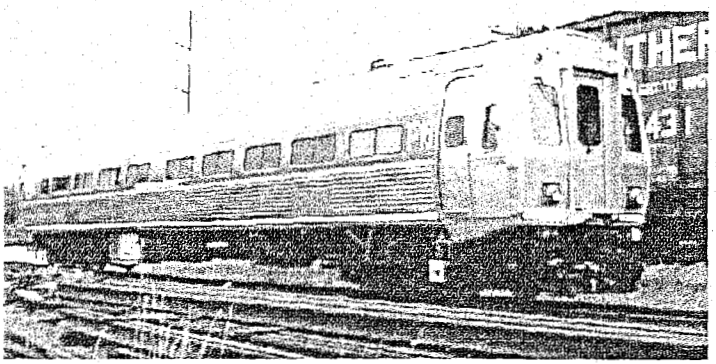


Figure 16.103: High speed research car T-16

To better assess the speed potential for planned high speed corridors, to learn from Amtrak's faster operations in the Northeast Corridor, and to improve understanding of the requirements for operating at speeds above 110 m.p.h., FRA needed a research car with high speed and advanced technological capabilities. Through arrangements with Amtrak, the FRA obtained a surplus passenger car which was rebuilt and instrumented to serve as its mobile high speed research laboratory: car T-16.

16.13.2 Instrumentation and measurement capabilities

The T-16 is equipped to measure track geometry, rail head profile, ride quality, and wheel-rail forces at speeds up to 160 m.p.h. Its measurement systems and data analysis techniques incorporate the results from FRA- sponsored research as well as work conducted by the Volpe National Transportation Systems Center in Cambridge, MA. Capabilities and instrumentation include:

- Inertial systems for measuring track gage, profile, alignment, and crosslevel over a range of defect wavelengths.
- Gage-face optical rail profile system (supplied by KLD Laboratories, Inc.) for measuring rail wear and wear patterns. At 150 m.p.h., rail profiles are measured every 15 feet, with smaller increments at lower speeds.
- Automated Track Data Alignment System (ATDAS) to highlight changes in track condition between surveys. This system provides a real time comparison between measurements made as the car is running and those from selected past test runs made over the same track.
- 32-channel general purpose data acquisition system for time or position-based measurements, which allows the car to be used for a variety of special tests.

- Integrated computer network system to provide a high speed real-time interface for data viewing, storage, and analysis.

- Figure 16.104: instrumented wheels for measuring wheel to rail contact forces

- Full truck of instrumented wheels (two wheelsets) for measuring wheel to rail contact forces (Figure 16.104). These are installed when needed for special testing.

- Neural network to determine track geometry limits based on vehicle design parameters.

- Video system to observe and record testing and track conditions.

- Differential Global Positioning System (DGPS) to tag all data with track position information; inertial dead reckoning in tunnels; automated mapping of vehicle path.

The neural network installed on the T-16 is a typical artificial intelligence (AI) recurrent network. It is trained to predict wheel-rail interaction forces generated from the T-16 as it travels over track with known geometry. Vehicle characteristics and track geometry data were used to train the network to predict the vertical and lateral forces of the T-16 on other known sections of track.

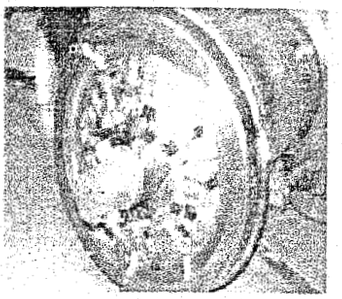


Figure 16.104: Instrumented wheels for measuring wheel to rail contact

16.14 Rail Profile Management

Rail Profile Management has traditionally often been considered separately from the rest of the track elements due to its extreme influence on the safety of traffic, but also due to pronounced economical reasons. A number of integrated maintenance management systems as well as decision support systems were developed specifically for rails in the past. Since then a great deal of research and consultant work has been pointed in this direction. Only lately rail management has gradually begun to be considered more as an integral part of the track maintenance management systems like ECOTRACK and, consequently, as a constitutive part of asset management systems.

Rail management, be it dedicated or integrated into TMMS, is closely related to two important groups of problems:

- **Safety problems**, involving in service failures and prevention of derailment and related accidents and

- **Maintenance problems**, including rail wear and costs related to its management, e.g. grinding, lubrication, and consequent rail renewal planning.

These problems are resolved by adequate monitoring techniques and proper wheel/rail interface management, which involve important issues like track geometry, friction, wheel and rail profile measurements, train dynamics, wheel/rail metallurgy, and, finally, operation factors like speed, loads, braking forces, etc.

Rail profile monitoring represents a vital component of rail management. The reason for this is that properly controlled profiles for rail and wheels reduce the risk of derailment and minimise the dynamic interaction between vehicle and track. In addition to this, monitoring also improves passenger comfort while reducing material stresses and wear of track and vehicle components. Regular recording of both wheel and rail profiles makes it possible to monitor changes and control maintenance in an optimum way, in order to reduce maintenance costs and improve safety.

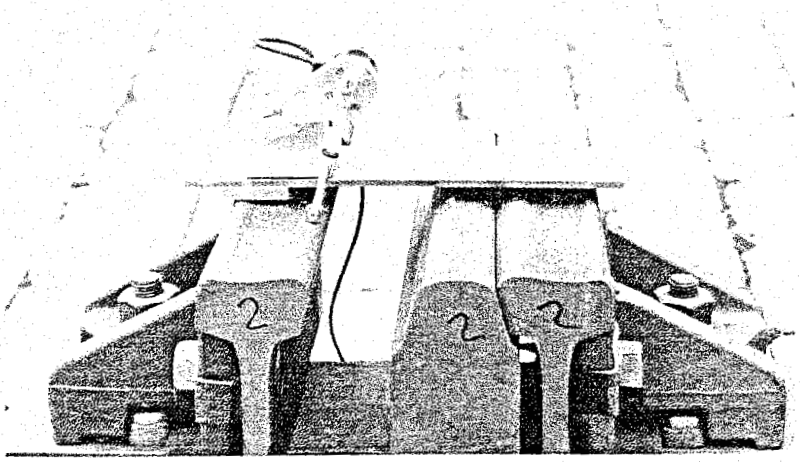


Figure 16.105: MINIPROF switch device

Various laser and mechanical systems are currently used to monitor the rail and wheel profiles. One of such systems, which is widely used worldwide, is the MINIPROF system. In addition to the standard system the so-called TWINHEAD MINIPROF was designed to measure both wheel and rail profile simultaneously. A special version, referred to as the MINIPROF SWITCH (Figure 16.105), was developed for measuring multiple profiles in turnouts. This version facilitates the recording of switch blades, stock rails, and frogs, as well as multiple profiles in crossings. This makes it also extremely useful for S&C management. For further details please refer to "MINIPROF" on page 554.

Integral wheel/rail and geometry monitoring is also used for optimizing and planning grinding activities. With the constant increase of wheel loads, rail grinding becomes essential in order to control the excessive fatigue and consequent deterioration of the rail running surface. Gauge corner rail shelling raises the possibility of rail breaks. This situation can be monitored and kept under control with a well-planned grinding program.

Such a grinding program is especially important since too much grinding could lead to an early consumption of the rail steel, while too little or inappropriate grinding, on the other hand, could result in a bad wheel/rail interface and cause accelerated wear rates. Accurate wheel/rail profile data is again used here to optimize the wheel/rail interface. From this information, the optimum shape can be inferred of either the wheel or rail for any particular situation. Furthermore, measured wheel/rail data can be used to compare the performance of different rail, bogie, or wheel types in specific circumstances. This allows one to choose the most appropriate and cost-effective ones.

As for the safety considerations, derailment potential can also be assessed from the wheel/rail profile data. With the wheel flange angle and the resultant gauge corner angle known, the potential for wheel climb can be determined using wheel/rail data.

16.15 Rail Defect Management

16.15.1 Introduction

Similar to Rail Profile management, Rail Defect Management can be considered a separate branch of track management or, which is a better alternative, within systems of wider scope like ECOTRACK. Whichever way one chooses to consider it, Rail Defect Management also constitutes a part of an AMS. The problem of Rail Defect management can additionally be complicated by the introduction of issues like the probabilistic estimate of possible catastrophic rail failures, inferred train derailments, and accident costs as well as consequent train delays and remedial costs. It is obvious that all of these aspects have a very close connection to the AMS.

Rail fatigue failure development goes through three phases. The first phase is the phase of crack initiation. The second one is the crack growth phase and in the third one the rail finally breaks. During the first two phases, which last a certain amount of time, loading cycles accumulate before the rail breaks in the third phase. The whole idea of rail defect management is based on the ability of track engineers to spot and control the problems during the growth period and keep them confined within the allowable limits. In order to do that, regular inspections are needed as well as the implementation of prescribed actions when defects, including fatigue cracks, are found.

Worldwide no major trend changes are reported in number and types of rail fatigue defects. In most cases the total number of rail failures is fairly steady. Many defects caused by manufacturing faults and shell defects originating from subsurface inclusions, are expected to drop due to advancements in steel and rail manufacturing technology. Aluminothermic welds still remain a major problem for most railways. However, it is expected that improved processes will provide adequate solutions in the near future.

Furthermore, a continued increase in the use of mobile flash welding instead of aluminothermic welds should additionally alleviate the problem. However, an increase in surface initiated rolling contact fatigue defects, like squat head checks, has been observed, especially in Europe and Japan. These failures indicate the inability of steel to sustain the imposed operating conditions. This inability makes the problem quite complicated, for the solutions can only be found in the domain of improved wheel/rail interface and the improved properties of the used material.

Some of the conclusions drawn from the NS network were that the aluminothermic welds represent the most frequent cause of rail fatigue failures, followed by flash butt welds, and, finally, the rails themselves. Another frequent source of failure found on the NS network were the insulated joints, but this situation has recently stabilised. The final observation was that the rail defects were more frequent on timber sleepers than on concrete sleepers. However, this observation could be affected by factors like the type of traffic, fastening, rail age, maintenance policy, etc. [66].

Rail defect monitoring includes various regular or specialist visual inspections, ultrasonic testing, and other methods like magnetic particle and eddy current or electromagnetic testing, or even radiography testing in cases of aluminothermic welds. One of the methods currently used in the USA is the so-called real time and verify inspection, which means that defects indicated by an ultrasonic test vehicle are verified visually and/or by hand held equipment. If necessary, repair or renewal of the defective rail takes place within one day of detection or a temporary speed restriction is applied within one to two days after detection.

The response time, of course, depends on the severity of the defect. Another method called continuous inspection is applied mostly on European and Australian networks. The basic idea behind this method is that continual inspection is followed by defect reporting and consequential programmed action, depending primarily on the defect type and size. On NS network, for example, the ultrasonic test vehicle produces a number of reports which include real-time exception reports with defects requiring immediate attention, detailed exception reports with defect prioritisation, off-line summaries, and statistical reports.

The onboard computer provides extended defect classification by a rule-based expert system. However, the final classification is still performed by a human expert. Except for larger defects, which are dealt with immediately, hand equipment is used within two weeks to verify the existence of defects, their classification, and the type of action to be taken.

Within ERRI D229 [66] several rail failure models have been investigated, including database statistical models and mechanistic models with probabilistic simulation. However, the conclusion was that comprehensive models of rail failure still do not exist, but that many of the building blocks, at least for detailed fracture and squat type defects, have been explored.

It was also concluded that the comprehensive models which provide a means of improving and supporting rail defect management would have to include:

- Materials
- Track design
- Vehicle parameters
- Crack initiation
- Crack propagation
- Fracture
- Inspection criteria
- Action criteria
- Pro-active measure
- Rail fracture/derailment risks
- Risks and costs

16.16 Ballast monitoring and management

Ballast represents one of the track elements which behaviour is least known and hardest to monitor. Some of the most important parameters, both from the monitoring and behavioural point of view, are the ballast soiling level, petrographical content, percentage of organic components, weedy ballast, surface pollution, grain geometry, grain distribution, hardness and stiffness, abrasion, age, and accumulated tonnage. However, most of these parameters are very hard to monitor, especially in a continuous manner and on a large scale. This to a large extent obstructs successful ballast management.

From the managing point of view, the old ballast is the most interesting. The new ballast is relatively stiff; the elastic deformation and, therefore, the loads on the superstructure are, within the acceptable limits, used as a basis for dimensioning. Under these circumstances, the ballast bed exhibits markedly elastic behaviour. However, as it ages, ballast absorbs an increasing amount of energy which it uses for its internal changes. Ballast grains become rounded and fine material is formed.

As a result of these changes, the stiffness of the material falls, leading to a gradual increase in elastic deformations. The result is that the rail track's bedding becomes increasingly poor. The loads on the rails, the rail fastenings, and the sleepers clearly increase. Due to decreasing stiffness of the grain structure, the track quality obtained by tamping is not achieved, i.e. above a certain level of ballast soiling it is no longer possible to achieve any lasting improvement in the track bed by purely tamping. The old, defective track bed keeps coming back. The ballast bed's ability to spring back is reduced by the amount of energy consumed by the material. Since the amount of energy consumed gradually rises as the level of pollution increases, this process can be compared to the weakening of a spring.

According to the laboratory tests performed within the ERRI D182 studies [63], the properties of the ballast material become markedly poorer at soiling levels of approximately 50 - 70 %, depending on the weighting of the parameters. This could be taken as one of the criteria for evaluating the useful life of railway ballast.

However, the same study indicates that a certain distance should still be maintained from the critical soiling level of 50 % to 70 % determined in the laboratory. The reason given was that the ballast bed soiling is very uneven. Thus, the most heavily soiled zones are the most interesting ones from the point of view of soil mechanics, since they are the ones that determine the load bearing and deformation performance of the ballast. However, since specimens taken from the ballast bed usually contain both material from heavily and less heavily soiled zones, this understandably produces lower degrees of soiling than the values determined during the research.

The study also indicated that it was not possible to study the full range of stones and rocks present in reality, as well as simulate all other variables which occur in reality (different axle loads and speeds, weather conditions, etc.). However, the final suggestion still is that it seems justifiable to start with ballast cleaning activities at the soiling levels of 30 % to 40 %. This makes ballast soiling a desirable parameter to monitor.

Techniques like Ground Penetrating Radar can also be applied on ballast. The reliability of this technique has still not been fully proven. However, a lot of scientific research has pointed in this direction lately, the reason being an ever present lack of a quick and continuous means of ballast monitoring. The prospects for the success of this technique are particularly promising if coupled with trial pits and cores which combined should contribute to a much better and reliable ballast assessment. So far, at the sites where the ballast was relatively clean, such as on mainline tracks, the base of ballast has been clearly imaged by GPR. In particular, changes in the quality of the image correlated well with the marked changes observed in trial pits and track geometry records [131]. So far, according to the reports coming from the companies dealing with the ballast GPR scanning, it has been successfully proven that the GPR ballast scan can be used to:

- Profile a continuous profile image of the ballast base;
- Identify sections containing attributes which can indicate that the ballast is in good condition;
- Identify sections of ballast which require specific maintenance or further investigation;
- Target the ballast sampling or other monitoring techniques more efficiently based on identifying sections which have similar image characteristics;
- Provide quality control after ballast cleaning.

16.17 Hand-held inspection equipment

16.17.1 Ultrasonic Hand Equipment MT 95

MT 95 ultrasonic hand equipment is used by EURAILSCOUT Inspection & Analysis to inspect tracks for hidden faults. The equipment is primarily used to inspect critical spots such as welds, switch points and rail joints. Hidden faults in the rail, reported by other, high speed, measurement systems, can be evaluated precisely with the MT 95 in order to take all necessary correction measures. Several EURAILSCOUT mobile rail testers operates this light-weight equipment.

The hand inspection equipment is also used for quality inspection of all rail welds. The MT 95 has a digital memory oscilloscope, which allows to extensively measure, evaluate and store data and make hard copy of the images of the defects found on site using MP-2000 mobile dotmatrix printer. Corresponding error documentation is produced on the HPC-2000 handheld PC also if the data are loaded for external processing and storage.

16.17.2 Hand-held Georadar

Hand-held Georadar allows swift, inexpensive and precise inspection of a rail's foundations. Contrary to expensive point drilling, a complete image of the ground structure up to five meters deep is obtained, with depths up to 20 meters possible. Measurements, accurate to a centimetre, of the strength of the ballast bed, PSS, FSS, pollution, ballast bags, silt, and water in the foundation layer are possible. Georadar is the ideal means for quality inspection after construction, soil sanitation and so on. The flexible, compact Georadar hand measurement system is used for special inspection on short route sections, switches and stations. The portable, on-rail platform is pulled manually and can be lifted off the rail in the shortest possible time. Inspections on longer route sections is undertaken at up to 40 km/h on the EM-SAT inspection vehicle or SKL installed system.

16.17.3 AUTOGRAPH

Rail AutoGraph is the latest in a line of lightweight track geometry surveying systems, which measure the actual geometry of the track. It differs from its predecessors in that it has been enhanced to measure track alignment and gauge.

AutoGraph (Figure 16.106) is a Track Geometry Survey Trolley, which measures actual track geometry. Weighing only 75kg with the heaviest single component of 19kg and being only 140 cm long, can be easily lifted clear of the track by 2 men and assembled/dissembled in less than 5 minutes and carried in back of a car when dissembled. It can be pulled at the speed of 5 km/h in two ways: manually (power assisted if required), or by a host machine.

Autograph measures the following parameters:

- Distance along track;
- Alignment - left and right;
- Profile - left and right;
- Versine - vertical and horizontal;
- Gauge;
- Cross level;
- Twist over any chord length.

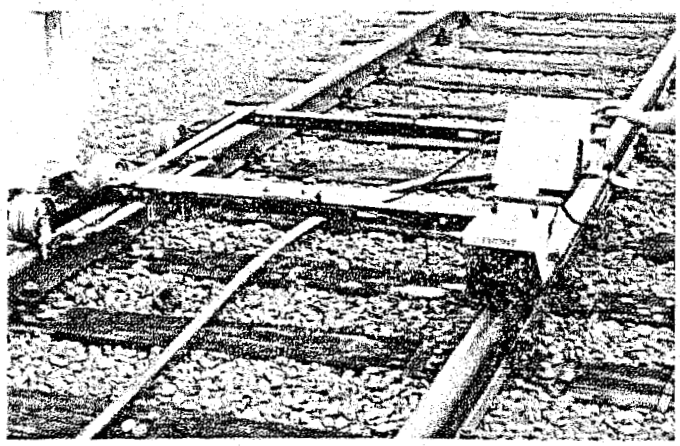


Figure 16.106: AutoGraph Trolley

AutoGraph (Figure 16.107 and Figure 16.108) comprises of two track geometry measuring components:

- Measurement of Vertical Geometry and Cross level;
- Measurement of horizontal alignment and gauge.

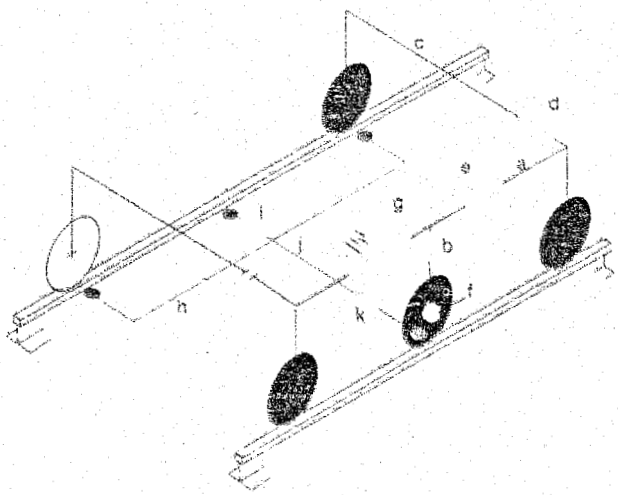


Figure 16.107: AutoGraph measuring principle

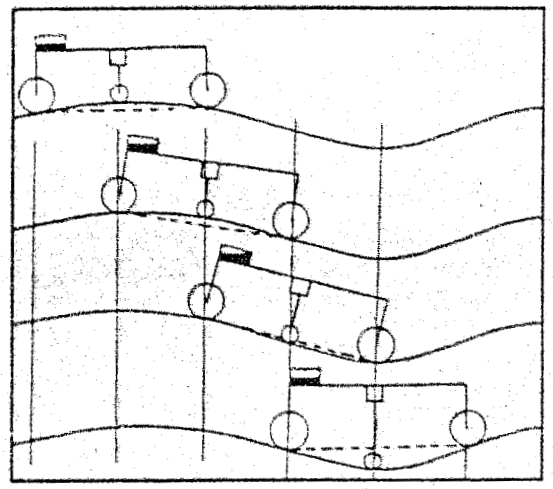


Figure 16.108: Principle of measuring vertical geometry via overlapping versines and inclinometers

Typical applications

Typical applications of AutoGraph would include:

- Assessment of the quality of a tamping operation;
- Assessment of work with respect to compliance with contractual requirements;
- Follow up suspect track conditions highlighted by 'rough ride' or track recording car reports.

16.17.4 MINIPROF

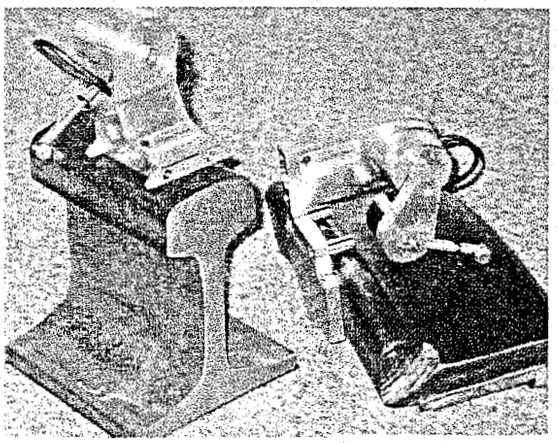


Figure 16.109: Photograph of MINIPROF equipment

Profile measuring systems for wheel and rail are already in use for quite a long time. Originally these were mechanical pen recorders. The next stage comprised of storing the information in a memory for subsequent downloading. The accuracy and repeatability have always been the critical points.

For a high accuracy the transducers should follow the profile in a natural way. In this respect systems with a straight guiding bar are less favourable. The user is not only interested in downloading data, but would also like to see immediately what is it that he is measuring and how does it compare to reference profiles.

The MINIPROF system, produced by Greenwood Engineering in Denmark, consists of a notebook computer connected to either a wheel measuring unit or a rail measuring unit. Figure 16.109 shows a picture of the system.

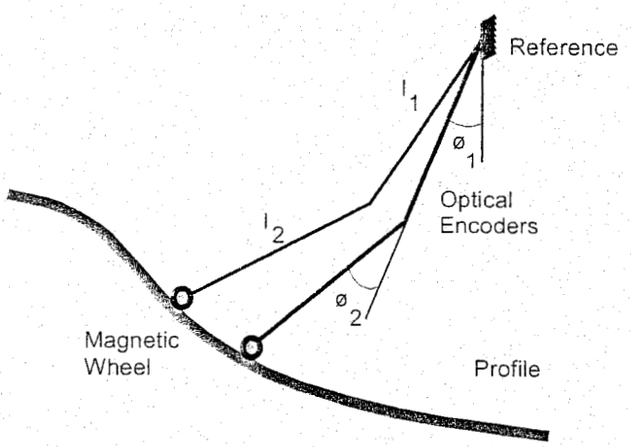


Figure 16.110: MINIPROF measuring principle

The sensing element, which is identical for both the wheel and rail unit, consists of a small magnetic wheel, with a diameter of about 12mm, attached to the extremity of two joint extensions. This magnetic wheel guarantees a safe contact during measurement. By moving the magnetic wheel manually the extensions rotate as demonstrated in Figure 16.110. In fact this is a polar coordinate system with 2 degrees of freedom. The two angles are measured with the aid of optical encoders, having an extremely high accuracy in the order of some microns.

The computer samples the transducer data, which are in polar coordinates, and calculates the profile in Cartesian coordinates. The resolution is in thousands of millimetres. By averaging closely spaced values accuracy is further improved.

In fact the lotus of the magnetic wheel centre is recorded (Figure 16.111). The software automatically calculates the real profile by subtracting the magnetic wheel radius perpendicular to the recorded path.

From the recorded wheel profile the following parameters are calculated, stored and monitored on the screen:

- flange height f_h;
- flange thickness f_d;
- flange gradient represented by qR.

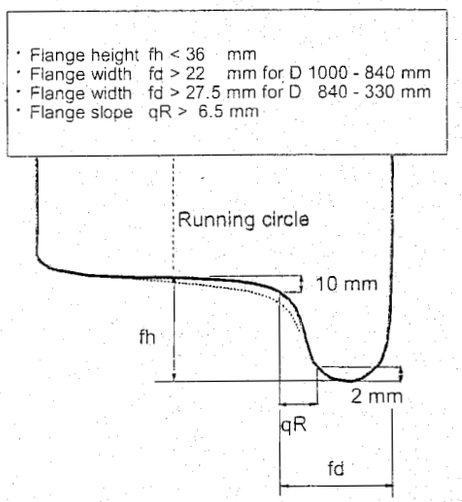


Figure 16.114: Admissible wheel parameters according to UIC leaflet 510-2

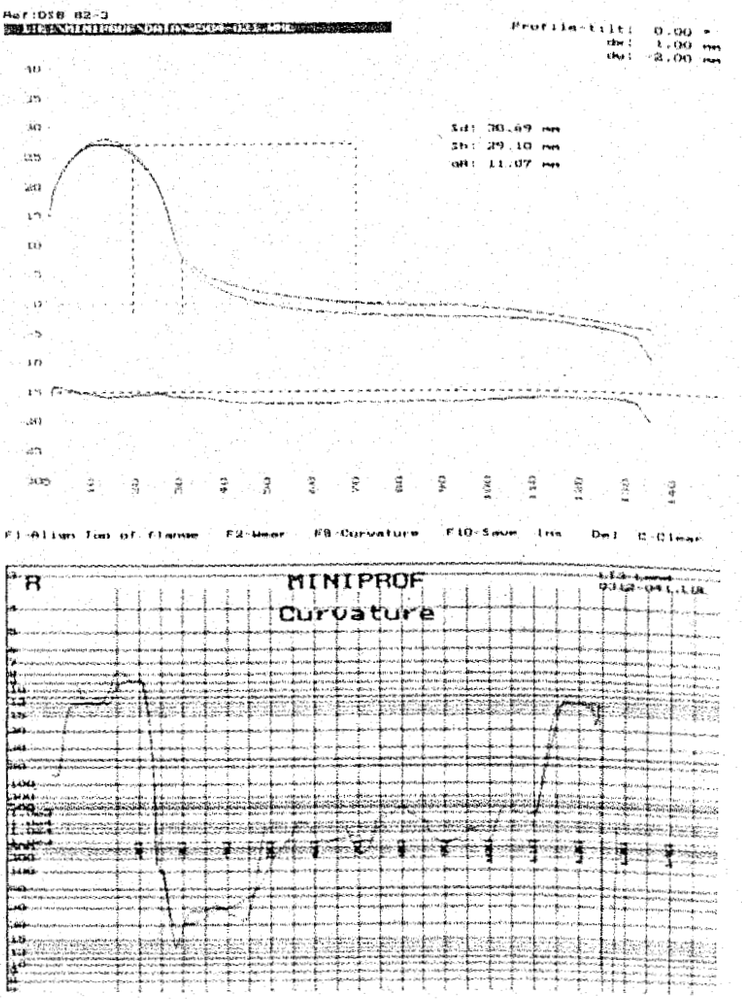


Figure 16.115: Example of measured wheel profile

The definition of these parameters together with the admissible values according to UIC and RIV are given in Figure 16.114.

The recording of one wheel profile takes about half a minute. Figure 16.115 shows an example of a wheel profile measured on London Underground Limited (LUL), together with the LT3 reference profile. In the same picture the curvature of both profiles is displayed.

A special version of the recording unit was developed for measuring profile and structure of brake disks for wear studies at Danish State Railways.

Rail profile measurement

The rail measurement principle is basically the same as for wheel measurement. The rail unit uses the opposite rail as a reference by using a telescopic fixture. The weight of the rail unit is 1.1 kg.

Figure 16.116 shows details of the measuring head. A special version with supports tailored to the railhead was developed for use of profile checking at rolling mills (Figure 16.117).

The principle of railhead measurement is schematically depicted in Figure 16.118. First the gauge face is recorded. After releasing the magnetic wheel, the arms are reversed followed by measuring an overlapping part of the running surface plus the field face. The software automatically recognised the second part of the profile and takes care of assembling the complete profile

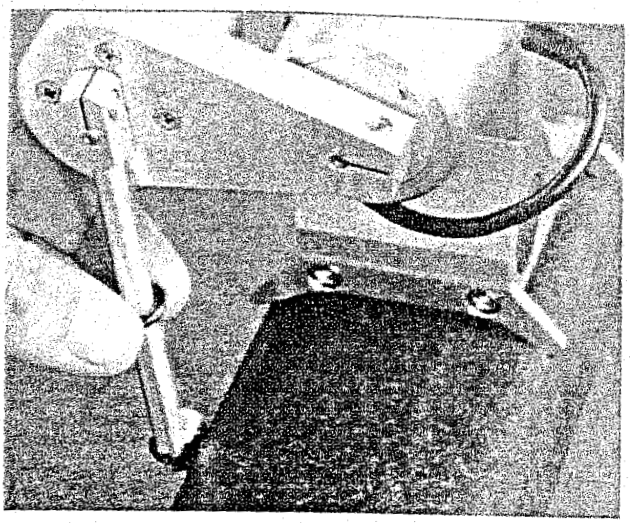


Figure 16.116: Rail measuring head

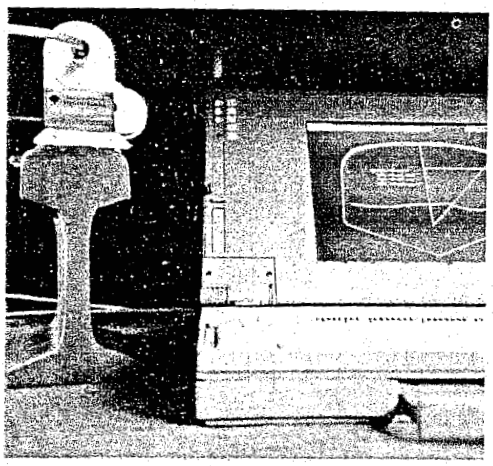


Figure 16.117: Rail measuring unit plus computer

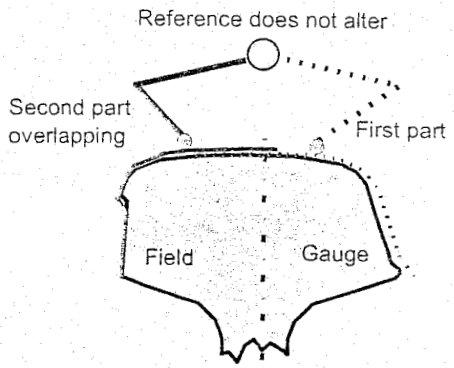


Figure 16.118: Principle of rail head measurement

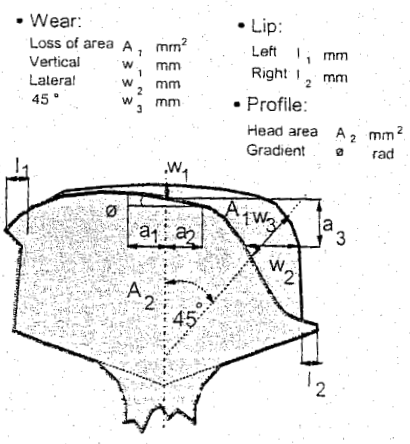


Figure 16.119: Rail parameters calculated in real time

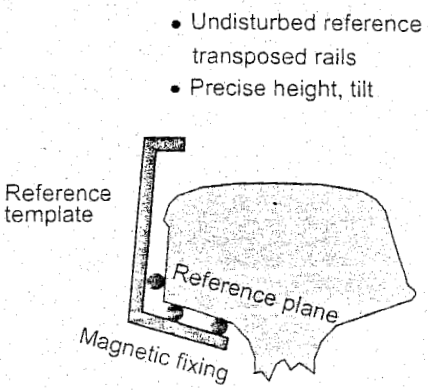


Figure 16.120: Rail reference template

Combined diameter and profile measurement with MINIPROF

In addition to the cross sectional profile MINIPROF can also provide the wheel diameter, by assuming that the train wheel is perfectly round. When the diameter is known at one point of the cross-sectional profile, the diameter will be known at all points of the profile, i.e., it is possible to give the diameter at the top of the flange and at the running line in one measurement (Figure 16.123 and Figure 16.124).

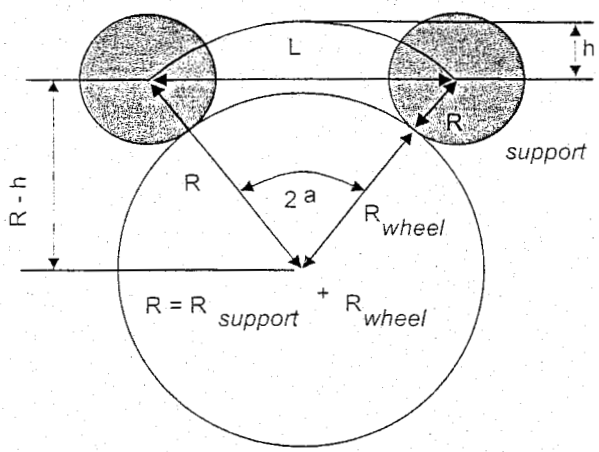


Figure 16.123: Principle of wheel diameter measurement

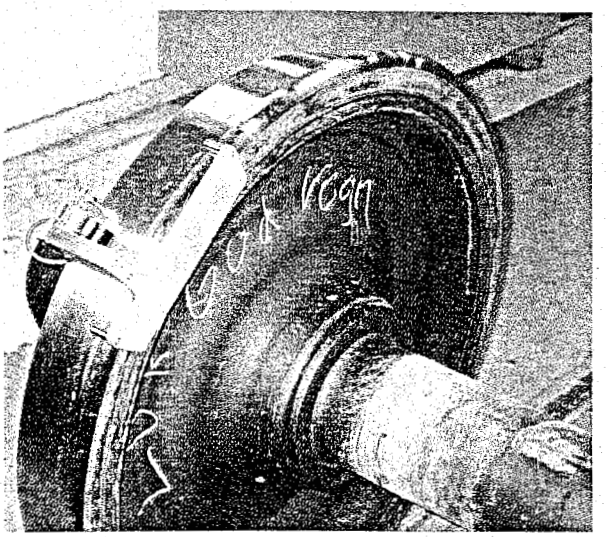


Figure 16.124: Wheel diameter measurement

The wheel diameter can be determined from the measurement of the versine. To simplify the calculations, the versine is considered $R_{support}$ "above" the wheel:

h the measured versine.
L the distance between the centre of the supports.
R_{wheelt} the radius of the wheel.
$R_{supportt}$ the radius of the supports.
R equals $R_{wheel} + R_{support}$.

The radius of a circle through the centres of the two supports (Pythagoras, see Figure 16.123):

$$R^2 = (0.5L)^2 + (R-h)^2 \Leftrightarrow R = \frac{(0.5L)^2 + h^2}{2h} \tag{16.41}$$

$$R_{wheel} = \frac{(0.5L)^2 + h^2}{2h} - R_{support} \tag{16.42}$$

The versine *h* can be expressed as a function of *Diameter D*:

$$h = R - \sqrt{R^2 - (0.5L)^2} = \frac{D}{2} - \frac{\sqrt{D^2 - L^2}}{2} \tag{16.43}$$

The accuracy of the diameter can be calculated from the first derivative of the versine:

$$Diameter\ accuracy = \frac{Versine\ accuracy}{\frac{\partial h}{\partial D}} = \frac{Versine\ accuracy}{\frac{1}{2} - \frac{D}{2\sqrt{D^2 - L^2}}} \tag{16.44}$$

The accuracy of the diameter depends on *L*, the distance between the centres of the supports, and the precision of the measured versine (Table 16.8).

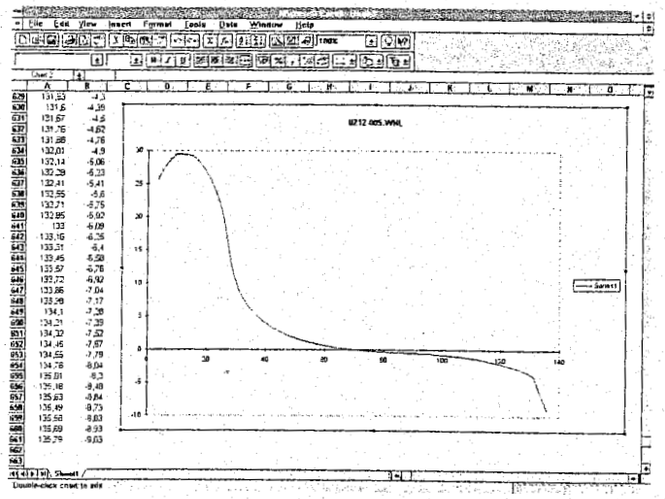


Figure 16.125: Example of a spread sheet application

(Versine precision = 0.005 mm)			
	D = 800	D = 900	D = 1000
L = 250	0.19	0.24	0.31
L = 150	0.55	0.72	0.89
L = 50	5.23	6.61	8.15

Table 16.8: Accuracy of diameter in mm

As all the measurement data are produced in ASCII format it is easy to import these in a spreadsheet like for instance EXCEL. Figure 16.125, for example, is showing a wheel profile imported in this way.

Rail grinding

The reference profiles can either be new profiles or grinding templates. The program calculates residuals, i.e. the difference between the actual profile and reference profile in curvilinear coordinates.

This means the offset is given perpendicular to the reference profile and the distance is measured along the reference profile.

The program can automatically position a rail template according to the CEN standards in such a way that the template touches at the top of the rail and coincides with the point 14 mm below the rail crown.

Figure 16.126 shows the template relative to the measured rail, whereas Figure 16.127 displays the offsets measured perpendicular to the template in a curvilinear coordinate system.

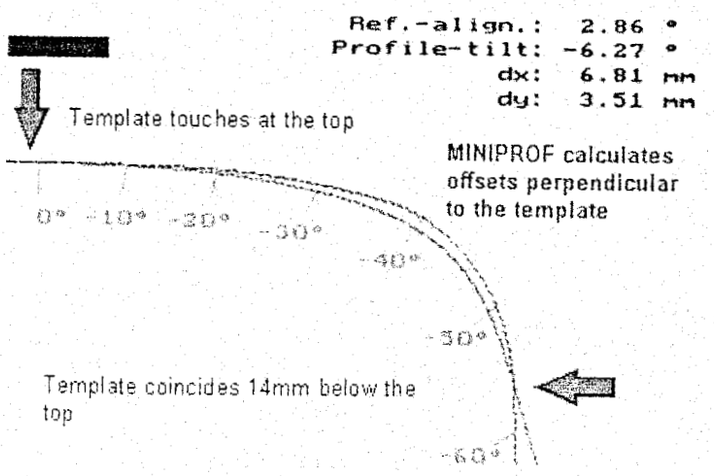


Figure 16.126: Grinding template according to CEN

Typical applications of the MINIPROF device would include:

- Wheel maintenance;
- Rail maintenance;
- Checking rail profiles before and after grinding for:
- Studies for optimizing wheel rail contact from the point of:
  - Contact mechanics;
  - Wear;
  - Steering/curve negotiation.

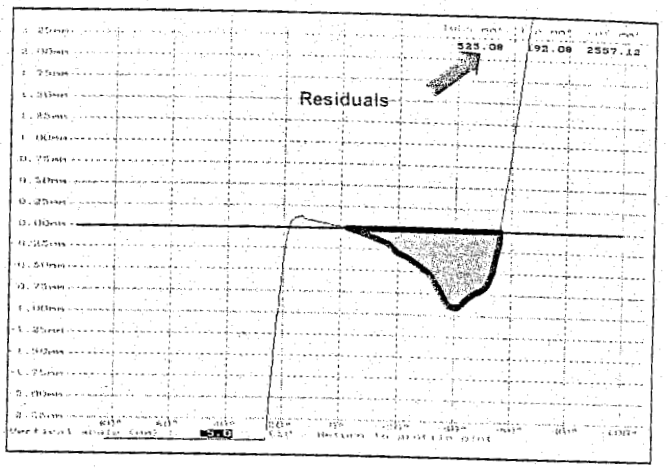


Figure 16.127: Offset of profiles in Figure 16.126

16.17.5 RAILPROF

Rail and weld geometry are dominant factors for track service life. A poor geometry leads to high dynamic forces and consequently to a rapid deterioration of track components. A substantial increase of forces already develops at relatively small geometrical deviations.

When speaking of short wave irregularities, better known as corrugations, one should typically think of peak - peak values in the 0 - 20 cm waveband in the order of 0.05 mm. Similar limit values hold for steps at welds.

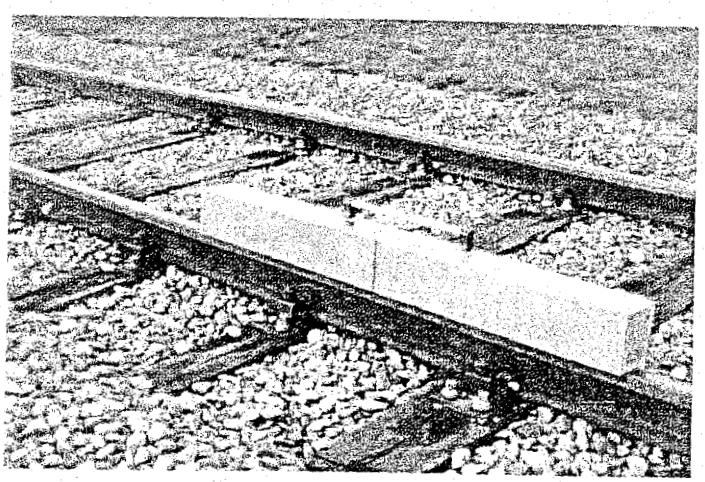


Figure 16.128: Photograph of RAILPROF

Although corrugations with amplitudes of some hundredths of a millimetre can be recognized by the human eye, it is extremely difficult to measure them sufficiently accurate, in the order of 0.01 mm. However, it is of great importance to recognize the development at an early stage so as to be able to start with remedial grinding actions as soon as possible. At small amplitudes the amount of metal removal is small, but if maintenance is postponed, or the phenomenon is not observed early enough, then heavy grinding at large costs is required.

Long term rail defect statistics of Netherlands Railways (NS) have shown that in particular weld defects are often the dominating factor in decisions for a track renewal. On average 75 % of the rail defects on the NS network were found at welds and joints.

In the early 1980s NS started an integral approach to the improvement of rail geometry control, specially dedicated to welds and joints. To make such a control effective accurate and versatile measuring methods are inevitable. For this reason it was decided to develop a straight edge measuring technique for weld geometry monitoring. The extreme conditions as regarding accuracy and environment were the reason that earlier developments failed. The RAILPROF system, however, complies with the tight specifications of NS. Figure 16.128 shows a photograph of the equipment.

Measuring principle

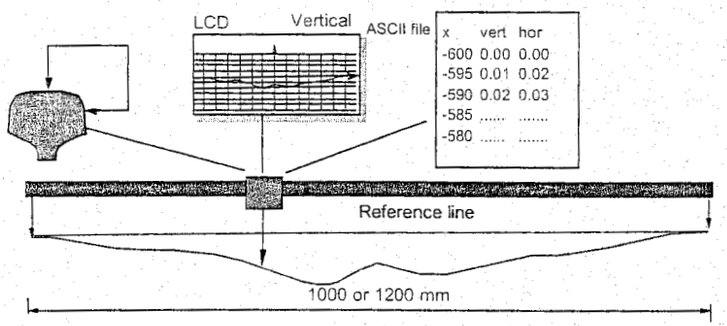


Figure 16.129: Measuring principle

RAILPROF is a tool for the accurate measurement of the longitudinal rail geometry. Low weight and robustness could be realised through the use of high-tech material developed in space technology.

The measuring sensors are based on a non-contacting principle. Micro roughness is smoothed out by averaging over an area of about 5 mm by 5 mm. In vertical direction the top of the rail is measured, in lateral direction the gauge face at 14 mm below the top of the rail.

The two sensors are located inside the housing and driven by a step motor. The latter is responsible for the results not being influenced by the operator. Besides, the one button push to carry out a measurement has great advantages for safe operation. The measuring principle is explained in Figure 16.129. The measuring base can be either 1000 mm (RAILPROF 1000), or 1200 mm (RAILPROF 1200). In the 1200 version the measuring base can be reduced to 1000 mm via software.

RAILPROF contains a built-in PC compatible processor. For practical reasons RAILPROF is provided with a limited keyboard and a small screen. They allow the operator to retrieve data like measured profile, trend, short waves, versine, grinding measure, etc. Immediately after a measurement conclusions can be drawn without having to wait on downloading the data into a central computer.

The data is sampled at an interval of 1 mm, and then condensed till data points every 5 mm. After each measurement the result is displayed on the LCD screen for direct assessment, both graphically and digitally.

Definition of parameters

First the longitudinal profile is determined, consisting of the deviation from a straight line connecting the first and the last measuring point. This is displayed without further data processing. For the other parameters the profile is split into long waves (trend) and short waves.

The trend in every point is the mean value over a length of 10 cm at either side. The versine of the weld is calculated as the maximum value of the trend, so independent of the short waves. A measure for the amount of material to be ground is derived from the weighted difference between the recorded profile and the calculated post grinding profile.

The short waves are calculated as the difference between profile and trend and play an important role for corrugation. The average wavelength is determined by dividing the measuring length by two times the number of zero crossings. The mean amplitude is the sum of all maximum values plus the sum of all minimum values divided by the total number.

There are basically two options for processing the measured data to be selected by 'measuring welds' and 'measuring corrugations'. The measured data points are processed as follows:

- Determine the high pass component of the signal by applying a rectangular filter with the edges at plus and minus 10 cm;
- The high frequent part is used when selecting corrugations;

- The residual part, i.e. the low pass geometry, is used in combination with welds to determine the largest inclination (first derivative) and the largest versine in absolute sense.

Corrugation

The high pass part of the geometry is displayed on the LCD screen together with the following condensed information:

- Maximum positive versine of trend [mm];
- Minimum negative versine of trend [mm];
- Maximum positive inclination of trend [mm/m];
- Minimum negative inclination of trend [mm/m];
- Grinding index [mm];
- Average amplitude [mm].

The grinding index is calculated as the average value of the material to be ground off. For the precise definition reference is made to the next section on welds.

Welds

The low pass part of the signal is used to characterise the weld geometry. On NS the following limit values are applied:

- Versine < 0.3 mm;
- Inclination < 1/1000.

The weld geometry can either be displayed by the total geometry, or the low pass filtered geometry. Figure 16.130 shows an example of total weld geometry. This information is directly presented on RAILPROF's LCD display and can also be reproduced after downloading the measured data in a PC.

The condensed information consists of the same parameters as in case of corrugation.

The grinding index is determined as follows:

- For the profile after grinding the highest parabola, just touching the measured profile, is taken, with a maximum height of 0.3 mm in the middle, and a horizontal line as a minimum;
- The area between the sine wave and the weld geometry, referred to as A [mm²], is calculated;
- The grinding index, being the average amplitude between both profiles, is then calculated as A/1000, or A/1200 [mm], depending on the base length.

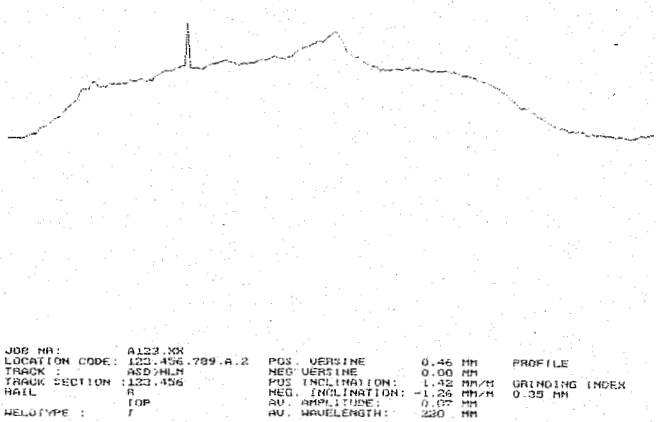


Figure 16.130: Screen dump of data downloaded in PC

The choice of using a PC board makes it very easy to modify, or extend, all software functions discussed above to the specific needs of the user.

Measuring reports

A report for each measurement is automatically stored for downloading in a computer. RAILPROF can store 430 measuring reports. The ASCII format is such that the report can be imported in the generally known database and spreadsheet packages. The measuring report per direction contains:

- The RAILPROF serial number, specifying the instrument measured with;
- A location code;
- Date and time of the measurement;
- Optionally an operator code;
- All measured data, i.e. 200, or 240 numbers for RAILPROF 1000 and RAILPROF 1200 respectively;
- Versines, average inclinations and grinding index for welds;
- Mean wavelength and mean amplitude.

Repeatability

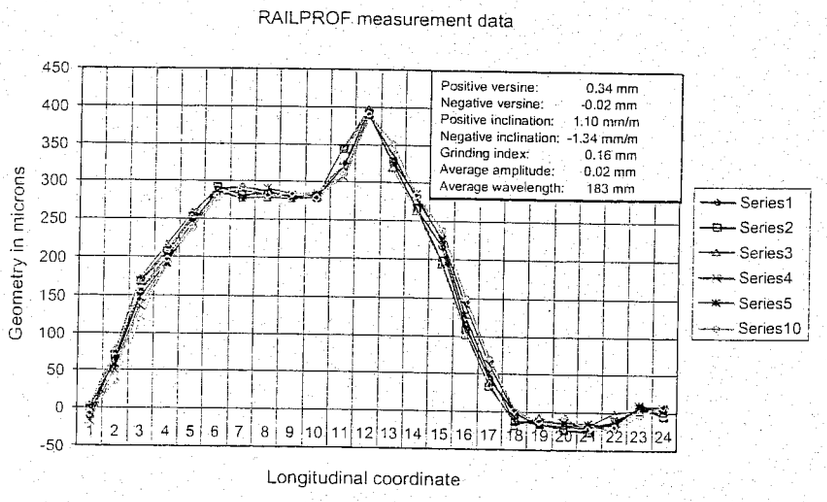


Figure 16.131: Results of repeatability tests

Apart from the absolute accuracy it is especially important for the user to have consistency in the measuring based decision. This is primarily reflected by the repeatability. Figure 16.131 shows six recordings of the same weld. The maximum deviations are less than 50 microns. From these results it can be concluded that the repeatability is extremely high so that consistency in the measuring based decisions is guaranteed.

Applications

Typical applications of RAILPROF would include:

- Checking and acceptance of flash butt, thermit welds, and glued insulated joints;
- Positioning of rails before welding;
- Production of glued insulated joints;
- Assessment of corrugation;
- Monitoring progress of grinding process;
- Acceptance of grinding work;
- Acceptance of rails for longitudinal geometry, including straightness of rail.

16.18 Pandrol Jackson SYS-10 Rail Flaw Detector

The SYS-10 represents a 24-channel portable ultrasonic rail flaw detection system displayed in Figure 16.132

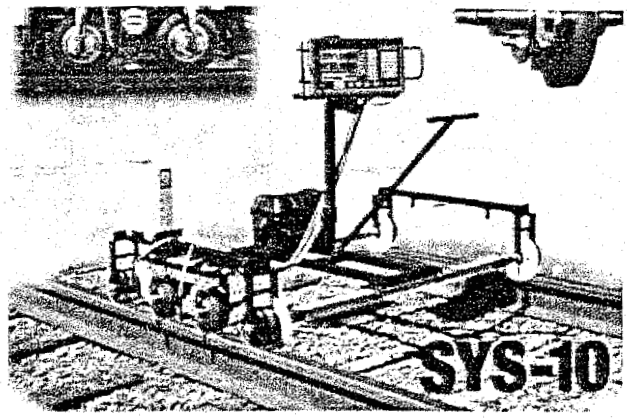


Figure 16.132: Pandrol Jackson SYS-10 Rail Flaw Detector

This automatic, intelligent instrument can test one rail or two rails simultaneously at up to 5 m.p.h., and significantly reduces human error. The lightweight unit can easily be assembled and operated by one person, making it suitable for rail inspection in areas where railbound or hi-rail vehicles cannot operate efficiently.

Durable wheel probes conform to the shape of the rail, improving sensitivity and reducing use of couplant. Electro-luminescent display gives near analog scope performance of a digitized analog A-scan waveform with an update rate greater than 60Hz for flicker free viewing.

Extended temperature ranges (-20 °C to 55 °C) and environmentally sealed case provides reliability in harsh environments. Ability to detect a loss of echo from the base of the rail provides a continuous ultrasonic signal check. Timing the signal gates off of the interface echo gives reliable defect detection by compensating automatically for rail surface variations. Audio and visual alarms warn the operator when rail anomalies are detected. 'Hand Test Mode' eliminates the need for a second test instrument.

17 HIGH-SPEED TRACKS

17.1 Introduction

17.1.1 Vehicle reactions

Despite vehicle running stability, wheel/rail forces and car body accelerations at high speeds should be confined to acceptable limits. As far as the track is concerned, these vehicle reactions can be influenced by the track geometry. In addition to the quasi-static components, which occur in curves, the response components comprise a dynamic part. The dynamic components can be further split up into low frequency steady-state contributions and high frequency impact loads occurring locally at welds and generated by wheel flats.

On account of the quasi-static and low frequency loads, the track may not displace permanently in the lateral direction, i.e. the Prud'homme criterion should be met. To guarantee safety against derailment, the Y/Q ratio should be less than a specific value, normally 1.2. For the sum of the quasi-static and low frequency Q-force, i.e. the 97.5% value, DB apply a standard of 170 kN. According to [252], BR calculations use 340 kN for the sum of the quasi-static, low frequency, and high frequency Q-force. In the case of the TGV, a maximum Q-force of 137 kN was attained for the quasi-static Q-force supplemented by twice the standard deviation of the low frequency dynamic component [231]. This 67% value of the Q-force is practically the same as the value applied to the German ICE.

As far as passenger comfort is concerned, the quasi-static and low frequency dynamic car body accelerations are dominant. In extreme cases a non-compensated lateral acceleration of 1.5 m/s^2 is allowed. For both the TGV and the ICE the absolute maximum for the total peak value of the car body acceleration is set at 2.5 m/s^2. Under normal conditions the standard deviation of the car body accelerations will be limited to 0.2 m/s^2.

In the case of the various high-speed projects, extensive series of measurements have been performed to check that the limits discussed earlier are not exceeded. A summary of the DB measurements on wheel/rail forces, published in [147], was discussed in Chapter 4. The 97.5% value of the Q-force attributable to locomotives appears to increase up to 150 kN at 250 km/h. Freight wagons with 22.5 t axle loads exert the same Q-force on straight track at 120 km/h. According to [231], during measurements on TGV trains a 67% value of the Q-force of 134 kN was found.

In [27] measurements on the German ICE are described. Figure 17.1 shows the ΣY_{2m} forces as a function of speed, measured in a curve with R = 3400 m and in a curve with $R \geq 495$ m and a cant deficiency of 140 mm. In all cases the Prud'homme criterion was met. No acceleration at car body level was found that exceeded 2.5 m/s^2.

In order to test the viability of the system periodic acceleration measurements should be carried out. The safety limits according to the SNCF are set as follows:

Transverse bogie acceleration	6 m/s^2	
Transverse body acceleration	2.5 m/s^2	v < 350 km/h
Vertical body acceleration	3 m/s^2	

Table 17.1: Safety limits for high-speed operation

The above values are absolute safety criteria. Under normal conditions the values of Table 17.2 should not be exceeded. With the opening of new lines these values are used during the so-called homologation runs in which the speed is step wise increased until the maximum line speed plus 10 % is achieved. These measurements are periodically repeated. If the values of Table 17.1 are exceeded the SNCF should report this to the Ministry of Transport.

Transverse bogie acceleration	4 m/s^2	
Transverse body acceleration	1.5 m/s^2	if cant deficiency < 120 mm
Vertical body acceleration	1.75 m/s^2	if cant deficiency > 120 mm

Table 17.2: Maximum accelerations to be adhered to in high-speed operation

For a stable running performance of high-speed trains the equivalent conicity is a prime factor. According to the European Train Standards for Interoperability (TSI) the following values should be adhered to:

Speed range	Maximum at starting operation	Maximum value in service
230 < V < 250 km/h	0.25	0.30
250 < V < 280 km/h	0.20	0.25
280 km/h < V	0.10	0.15

Table 17.3: Maximum values for the equivalent conicity in high-speed operation

The SNCF's policy is to start with a very low conicity of 0.025, achieved by conical wheels with an inclination of 1/40. During service the conicity increases till approximately 0.10, with exceptional values of 0.13, the maximum value the SNCF allows. The DB is starting with a much higher equivalent conicity, in the order of 0.1, associated with the philosophy of worn wheel profiles. DB's maximum value is 0.15 for high-speed operation.

17.1.2 Track geometry

As far as the track is concerned, tight tolerances should be imposed to restrict the dynamic vehicle reactions. If the transfer functions between geometry and reactions are known, the vehicle reactions can be calculated using the recorded geometry. The VRA system discussed in section 16.11 covers this aspect.

In principle, standards can only be defined for vehicle reactions and not for track geometry as each track geometry component contributes to a specific reaction. It is in particular this combination of geometrical contributions which may be decisive, although each geometrical deviation as such should not necessarily lead to vehicle reaction standards being exceeded.

In the absence of a VRA system, track standards have been developed based on practical experience. For instance, for 200 km/h lines BR applies a maximum value for alignment of 1.8 mm σ_{BMS} and for level a value which is 1.5 times greater, i.e. 2.7 mm. This is based on a vertical car body acceleration of σ = 0.2 m/s^2.

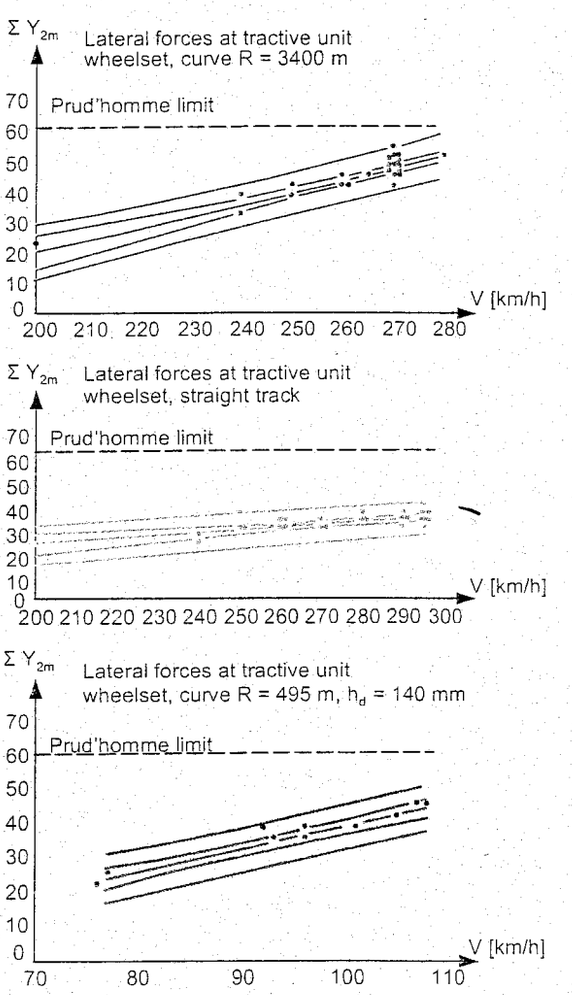


Figure 17.1: Measured lateral track force exerted by the German ICE

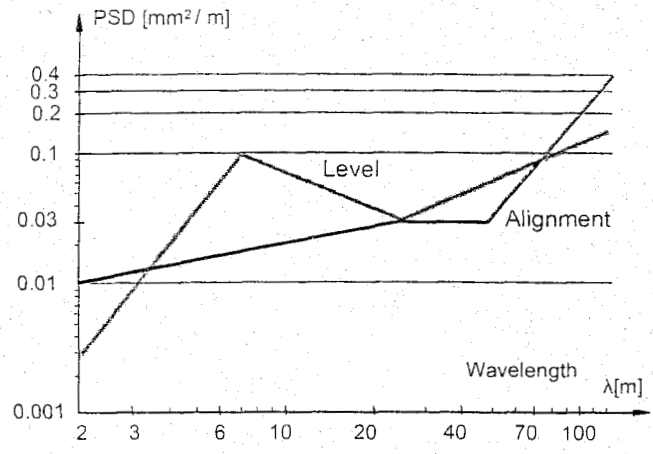


Figure 17.2: Target psd-function for level and alignment for tracks operated at 300 km/h

According to [186], SNCF applies a mean absolute value of 0.6 -0.8 mm, while local peaks in cant and level should be kept to 10 - 12 mm. These Mauzin car values apply to an extended measuring base of 33 m obtained by means of recolouring.

In fact such an approach is too global as these standard deviations refer to a waveband of 0 - 25 m. This waveband covers the Q and Y forces quite well, but car body accelerations originate from much longer waves. For natural frequencies of the order of 0.7 to 0.9 Hz the dominating wavelengths at 300 km/h are 119 - 93 m. For this reason the total measuring range should be extended to a wavelength of about 120 m. This waveband is too long to determine one standard deviation and, therefore, should be split up, for instance, as follows:

- 3 - 25 m: short-wave geometry associated with Q and Y forces, automatic, tamping actions, and exceedences due to local irregularities;
- 25 - 70 m: associated with car body accelerations at medium speeds;
- 70 - 120 m: associated with car body accelerations at high speed.

In view of the lack of precise knowledge of the transfer functions of the rolling stock and the lack of experience in calculating vehicle response on-line as described in section 6.7, power spectral density functions of the track geometry have been examined and from these a waveband composition with corresponding standards for variance of standard deviation has been derived.

In Figure 16.102 psd functions for different types of track are given. These functions have been estimated over a track length of about 100 km each. Regardless of the track type, the shapes are very similar. There is also a distinct correlation between short-wave and long-wave irregularities. The psd functions for NS UIC 54 concrete track are extremely good and may be considered as the best possible in practice. Figure 16.101 shows another set of psd functions. Except for the BR ones, they were recorded on conventional speed lines. The BR spectra refer to tracks for V = 200 km/h. The target spectra for 300 km/h lines, presented in Figure 17.2, have been derived from this information.

Standard deviations which are to be used as target value and as limit value according to Table 17.4 have been derived from these spectra.

Waveband	Level		Alignment	
	σ_{target} [mm]	σ_{limit} [mm]	σ_{target} [mm]	σ_{limit} [mm]
3 – 25 m	1.0	1.5	0.7	1.0
25 – 70 m	2.0	3.0	1.3	2.0
70 – 120 m	2.7	4.0	3.4	5.0

Table 17.4: Target and limit standard deviations for high-speed tracks operated at 300 km/h

17.1.3 Rail geometry and weld geometry

For reasons of dynamic impact, deviations in the short waves should be kept to a minimum. The NS tolerances, specified in Chapter 16, are well suited to high-speed tracks. According to [186] the permissible deviations in rail and weld geometry for the TGV track amount to 0.3 mm on a basis of 1.6 m.

The forces in the 0 - 150 Hz frequency band, which are associated with unsprung mass, are very detrimental to the track. Axle box accelerations represent these forces very well. On ground tracks axle box accelerations of the order of 10 m/s^2 have been found on the TGV, and on non-ground tracks these values reach a maximum of 25 m/s^2. In order to sufficiently limit the transfer of these dynamic forces to sleeper and ballast, SNCF employ 9 mm thick high-resilience pads.

17.1.4 Track quality standards for 300 km/h

This section specifies the standards which should be applied in order to maintain the required quality level on high-speed tracks. The standards are basically formulated in terms of vehicle reactions, i.e. wheel/rail forces and car body accelerations. If the transfer functions between track geometry and vehicle reactions are known, the vehicle reactions can be calculated directly making use of the measured track geometry. This requires an ultra-fast on-line computing system as described in Chapter 16.

Again, if the transfer functions of the rolling stock are known it is, in principle, possible to make some estimates of the admissible geometry to prevent vehicle reactions exceeding their limit values. In this process one should realise that several geometrical components will contribute to a specific reaction and there is, therefore, no unique relationship between maximum vehicle response values and admissible track geometry.

The steady-state vehicle reactions have a random character and are expressed in terms of standard deviations. Isolated or local irregularities possess a deterministic character and are expressed in terms of peak values.

Rail geometry

The recommended standards for rail geometry are specified in Table 17.5.

If welds are produced or inspected manually, it is recommended that a 1.2 m straight edge provided with two LVDTs be applied 0.2 m apart as described in Chapter 10. The recommended admissible deviations are:

- versine: ± 0.2 mm;
- step: 0.1 mm.

If these standards are exceeded the welds should be straightened and ground, for instance using Plasser and Theurer STRAIT/GWM220 machines.

Waveband	Signal	Standard	
		Peak P_{norm}	σ_{norm} 200 m section 90 %-value MAINS
0.03 – 0.3 m	Vertical axle box acceleration If displacement is measured		12 m/s² 0.05 mm
0.3 – 1.0 m	Peak accelerations due to welds	72 m/s²	
1.0 – 3.0 m	Vertical rail irregularities due to the rolling process		0.14 mm

Table 17.5: Rail geometry standards for 300 km/h

Vehicle reactions

The recommended standards for vehicle reactions are specified in Table 17.6. The ISO filter characteristics are presented in Figure 6.26 and Figure 6.27. Non-compensated lateral acceleration is determined according to (16.31). The derivative of the lateral acceleration is used in expressions (16.33) and (16.34). The Prud'homme ratio follows from (16.40).

Waveband	Signal	Standard	
		Peak P_{norm}	σ_{norm} 200 m section 90 %-value MAINS
3 – 120 m	Q force	10 kN	12 m/s²
	ΣY force	6 kN	0.05 mm
	Vectorial sum of ISO-weighted vertical and lateral car body acceleration	1.0 m/s²	0.2 m/s²
70 – ∞ m quasi-static	Non-compensated lateral acceleration a_{NC}	0.8 m/s²	
	Derivative of lateral acceleration $\dot{a}$	0.2 m/s³	
2 – ∞ m quasi-static + dynamic	Prud'homme ratio	1.0	
	Total acceleration on carbody level, to be considered as an absolute maximum	2.5 m/s²	

Table 17.6: Vehicle reaction standards for 300 km/h

It should be realised that the above-mentioned standards can only be applied directly if a vehicle reaction analysis system, as discussed in section, is available. If the track standards specified in the next paragraph are applied, it is likely that the above-mentioned vehicle reaction standards will be met, although no guarantee whatsoever can be given.

Track geometry

Based, for instance, on the power spectral density functions in Figure 17.2, the recommended standards for track geometry are summarized in Table 17.7. The various parameters are defined in Chapter 16.

Interpretation of standards

The standards to be applied are specified in Table 17.5, Table 17.6, and Table 17.7 and pertain to standard deviations of dynamic signals and peaks of dynamic, quasi-static, and total values. The standards for quasi-static and total values should be considered as absolute limit values which may not be exceeded in any circumstances for safety reasons.

The standards for the dynamic signals can be interpreted as maintenance intervention levels and may, therefore, be exceeded. It is recommended that if the standard is exceeded this amounts to less than 50 %. The interpretation of the standards is summarized in Table 17.8. The normalized values are obtained by dividing the actual value by the norm according to:

$$\bar{\sigma} = \sigma / \sigma_{norm} \qquad (17.1)$$

$$\bar{P} = P / P_{norm} \qquad (17.2)$$

Waveband	Signal	Standard	
		Peak P_{norm}	σ_{norm} 200 m section 90%-value MAINS
0 – 25 m	Cant Level Alignment Gauge Twist	5 mm 8 mm 6 mm 5 mm 6 mm	0.7 mm 1.0 mm 0.7 mm 0.7 mm –
25 – 70 m	Cant Level Alignment	– – –	1.3 mm 2.0 mm 1.3 mm
70 – 120 m	Cant Level Alignment	– – –	2.0 mm 2.7 mm 3.4 mm
70 – ∞ m	Lateral rail wear in curves Gauge Curvature κ_r a_{NC} Cant C_r a_{NC} Cant variation $\dot{a}$ Curvature variation $\dot{a}$	 10 mm 1435^{+2}_{-0} 0.8 m/s² 0.8 m/s² 0.2 m/s³ 0.2 m/s³	

Table 17.7: Track geometry standards for 300 km/h

	Signals	Normalised value	Classification
MAINTENANCE	• Corrugations • Rolling defects • Track geometry • Vehicle reactions	$1.00 < \bar{\sigma} < 1.25$ $1.25 < \bar{\sigma} < 1.50$ $1.50 < \bar{\sigma}$	Moderate | Light Poor | Medium Very poor | Heavy
	• Welds • Track geometry • Vehicle reactions	$1.00 < \bar{P} < 1.25$ $1.25 < \bar{P} < 1.50$ $1.50 < \bar{P}$	Light Medium Severe
SAFETY	• a_{NC} • $\Delta\kappa/\Delta x$ • $\Delta C/\Delta x$ • a_{total} • Prud'homme	$0.50 < \bar{P} < 0.75$ $0.75 < \bar{P} < 1.00$ $1.00 < \bar{P}$	Moderate Severe Alarm

Table 17.8: Interpretation of track standards

Requirements for checking track geometry

Waveband	Signal	Accuracy
0.03 – 0.3 m	Vertical axle box acceleration per rail at a fixed reference speed of 90 km/h Vertical displacement of rail running surface derived by dividing the acceleration signal by the transfer function between acceleration and geometry	0.5 m/s² 0.005 mm
0.3 – 1.0 m	Peak value of vertical axle box acceleration per rail to describe poor welds	3 m/s²
1 – 3 m	Vertical displacement per rail to describe rolling defects	0.02 mm
3 – 25 m	Track geometry components • Cant • Level • Alignment • Gauge • Twist on a basis of 2.75 m Peak values	 0.1 mm 0.5 mm
25 – 70 m	Track geometry components • Cant • Level • Alignment	0.2 mm
70 – 120 m	Track geometry components • Cant • Level • Alignment	0.3 mm
70 – ∞ m	Track geometry components • Maximum gauge • Gauge • Cant • Curvature derived from cant and curvature • a_{NC} • $\Delta\kappa/\Delta x$ • $\Delta C/\Delta x$	 0.5 mm 0.5 mm 2.0 mm $10^{-4} \cdot m^{-1}$

Table 17.9: Accuracy specifications for recording of high-speed track

To be able to check the specifications discussed earlier for high-speed tracks, very accurate recording systems are required. This section gives the main specifications for a recording system which is able to monitor the quality of tracks operated at 300 km/h. In principle, such a recording system comprises three parts:

- recording system to measure irregularities in the surface of both rails. The measuring principle consists in measuring axle box accelerations;
- recording system to measure track geometry in a series of wave-bands. The longest wave is adapted to the lowest natural frequency of the rolling stock and the maximum operating speed, in this case 300 km/h;
- ultra-fast computing system which enables calculation in real time of all the relevant vehicle reactions to be carried out for a number of representative vehicles and a number of possible running

speeds up to 300 km/h. The measured track geometry serves as input based on which, with the aid of transfer functions, the response values are calculated in the frequency domain.

The relevant wavebands to be covered by the recording system have already been discussed in previous sections. They are summarized in Table 17.9 together with the measuring accuracy required if they are to be reproducible. This accuracy is expressed in terms of standard deviation for dynamic signals and amplitude for quasi-static signals and peak values.

17.2 The Korean High Speed Railway Project

17.2.1 Introduction

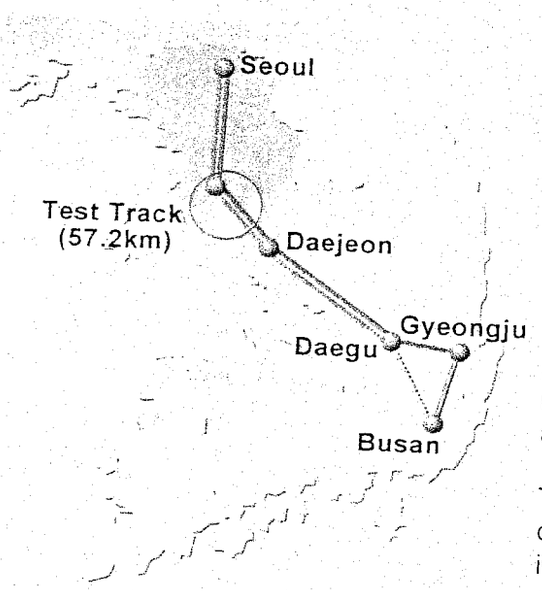


Figure 17.3: Seoul-Pusan High Speed Line

The Korean High Speed Line between Seoul and Pusan is linking the major economic and cultural centers of Korea. The line has intermediate stations at Chonan, Taejon, Taegu and Kyongju, and consists of 412 km double track, including 112 km (27%) of at-grade sections, 109 km (27%) of viaducts and 191 km (46%) of tunnels (Figure 17.3). The train sets are based on TGV technology. The commercial speed will be 300 km/h and the travelling time 1hr 56 minutes between the two terminal stations. The daily transport of passengers is estimated 520,000. The annual tonnage amounts to about 60 MGT.

To cut down the initial investment, the construction of the first phase of the high-speed line terminates in Taegu, while in the interim the upgrading and electrification of the existing conventional line between Taegu and Pusan will allow the revenue service starting in April 2004. Construction of the Taegu–Pusan section via Kyongju as well as underground stations in Taejon and Taegu comprise phase two, which is scheduled for completion in 2010.

In March 1992 construction was started of a 57.2 km long double track test section located between Chonan and Taejon on which test runs at 300 km/h commenced in December 1999.

17.2.2 Civil Works

One of the main features of the project is the large number of the civil structures. Standardized structures and construction methods have been adopted to optimize the construction processes. Typical viaduct structures comprise two or three continuous 25 m spans or two 40m continuous spans. The application of those typical spans limited the length between bridge deck expansion joints to less than 80 meters, thus avoiding the installation of rail expansion joints in most instances.

The Precast Span Method (PSM) was adopted for the bridge deck construction of several long viaducts. This method was state of the art construction technology, which involved the prefabrication, transport and installation of 25 m long Precast PC Box Girders. The 600-ton girders are cast in a temporary casting yard located near one end of the bridge, and transported on specially designed carriers to the point of installation. The concrete box girders are lifted from the carrier and moved along a launching beam. The launching beam is supported at the front end on the pier head, on which the girder is to be installed, and at the rear end on the previously installed girder.

As the construction progresses, the installed girders serve as the access route for the special carrier. This method, of course, is particularly advantageous for long bridge structures (See Figure 17.4).

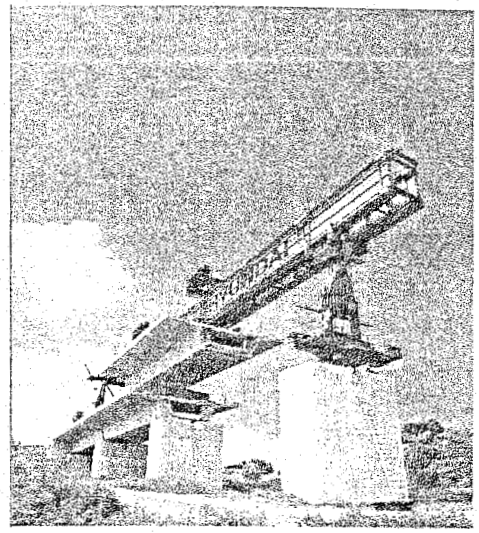


Figure 17.4: Installation of 25m long Precast PC Box Girder

17.2.3 Track Characteristics

The general track characteristics are as follows:

Ballast:	Crushed stone 31.5/50 mm Thickness min. 35 cm
Sleeper:	Prestressed concrete mono-block
	L = 2.6 m
	Weight = 3 kN
Sleeper spacing:	60 cm
Fastening:	Elastic Pandrol e-clip, 25 kN/rail Elastic Pandrol Fastclip (future), 20 kN/rail
Rail pad:	Studded rubber pad, 10 mm thickness, stiffness 65-95 kN/mm
Rail:	UIC 60, grade 880 N/mm2
Track gauge:	1435 mm
Speeds:	300 km/h (no freight)
Axle loads:	170 kN
Curve radii:	Min = 7,000 m
Cant:	Normal = 130 mm Max = 180 mm
Cant deficiency:	Normal = 65 mm Max = 85 mm
Grades:	Normal maximum = 25 per mille Exceptional maximum = 30 per mille
Vertical curve radius:	25 - 40,000 m
Transition curve length:	630 m
Track distance:	5.0 m

17.2.4 Track Laying

For the 57 km long test section concrete sleepers in ballast were selected. The prestressed mono-block sleepers were fitted with Pandrol e-clip fasteners. The rail profile was UIC60.
On all bridge decks ballast mats with a thickness of 25 mm were installed on top of the protective waterproofing layer over the bridge to reduce the track stiffness. All the track components have been tested, either in Korea or abroad, to ensure that they meet the requirements of the high-speed track performance specifications.

17.2.5 Track Installation

On the test section track installation started in February 1999 at a rate of 400 meters a day. The Osong Depot located 120 km south of Seoul was the operational center of the test section and served as the site office for the civil construction and the work base for the track and catenary installation. It includes 25,700 m of tracks, a maintenance workshop, storage areas and site offices. The welding facility that welds 25m long rails into 300m long rails is also located at the Osong site (See Figure 17.5).

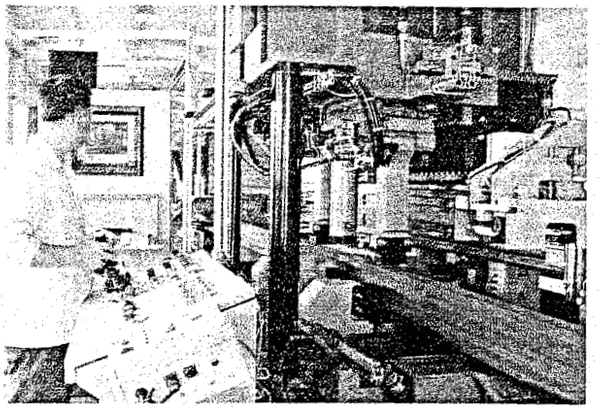


Figure 17.5: Grinding facility at the Osong Welding Depot

Track installation started by laying a temporary track in order to provide access to the work trains. When a sufficient length of temporary track was available, the special train transporting the 300-meter long welded rails unloaded the rails on both tracks (see Figure 17.6). The train was fitted with special rail support devices to transport up to thirty rails, and negotiates curves with a minimum radius of 150 m in the Osong depot.

Track installation proceeded with the laying of PC sleepers and the assembly of the track components. Two gantry cranes traveling on the long rails were used, first to remove the temporary track panels and second to place the concrete sleepers on the track. Then, the long rails were moved onto the sleepers and fastened with Pandrol rail clips. Welding of the rails using Thermit welding completed the operation.

In six successive steps, ballasting and tamping operations ensured that the track is lifted to its final level, providing a minimum ballast thickness of 350 mm. Dynamic stabilizing is carried out three times during the lifting process.

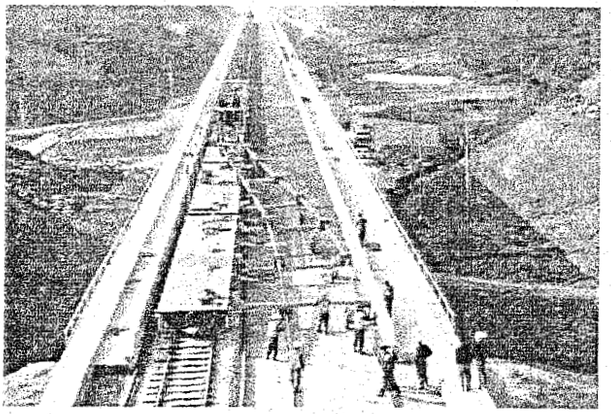


Figure 17.6: Transportation of Long Rail

The turnouts are installed after the fourth step of the lifting process. The turnouts were pre-assembled and inspected at the turnout factory. Each turnout was temporarily divided into 3 or 4 segments and loaded on special wagons to be transported to the installation site. Assembling and laying on concrete sleepers was performed using lifting units. Finally, turnouts were incorporated into the continuous welded rail, thus ensuring the continuity of the track.

Destressing of the rails and final tamping completed the track installation. The use of a tamping machine equipped with a system to record the track geometry allowed for an accurate control of the quality achieved. Finally, a track recording car measured the track parameters to confirm that the track quality complied with the high-speed rail operation standards.

17.2.6 Catenary and Systems

Erection of the catenary for the two 25 kV systems was completed for the test section during November 2000, together telecommunications, signalling and train control systems. The signalling system and Automatic Train Control system – the TVM 430 –allows trains running at 300 km/hr with a headway of three minutes between trains. Early 2000 high-speed testing and commissioning could commence. At the end of 2000 an overall construction progress of 50 % was achieved. Figure 17.7 shows a picture of the KTX running on an elevated track between Chonan and Taejon.

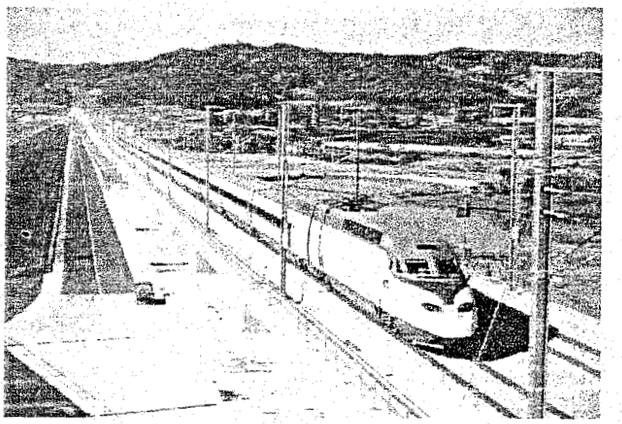


Figure 17.7: KTX running on elevated track

17.3 Dimensions of railway tunnels

17.3.1 Introduction

In the design of tunnels the dimension of the cross section is an important factor for the total costs of the tunnel. In tunnels the resistance is greater than in the open air. At the entrance and the exit there are sudden air pressure variations, which are uncomfortable for passengers because they can cause problems like ear pain and headache. These air pressure variations have an effect in the whole tunnel because they cause air-pressure waves in longitudinal direction. Especially the introduction of high-speed trains, with speeds greater than 300 km/hour and with sometimes many long tunnels made a new approach of this item necessary. In the next part a short explanation is given of the problem associated with running at high-speed trains through tunnels. At first there is a comparison between the situation in the open air and in a tunnel.

17.3.2 Air resistance in the open field situation

In the open air the train resistance is normally considered under the assumption that the air is incompressible. The air resistance normally consists of two components as shown in Figure 17.8:

- Difference in air pressure at the front and the rear of the train due to air pressure variations along side the train;
- Difference in friction alongside the train as a result of the surface resistance of the train.

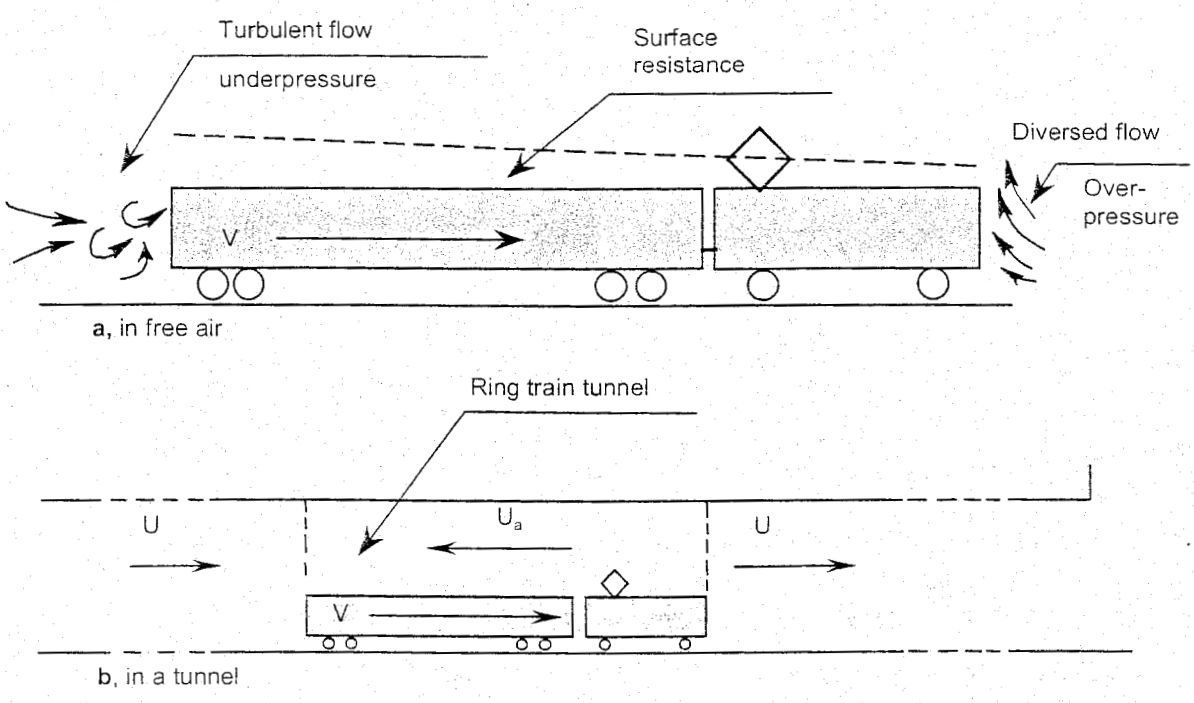
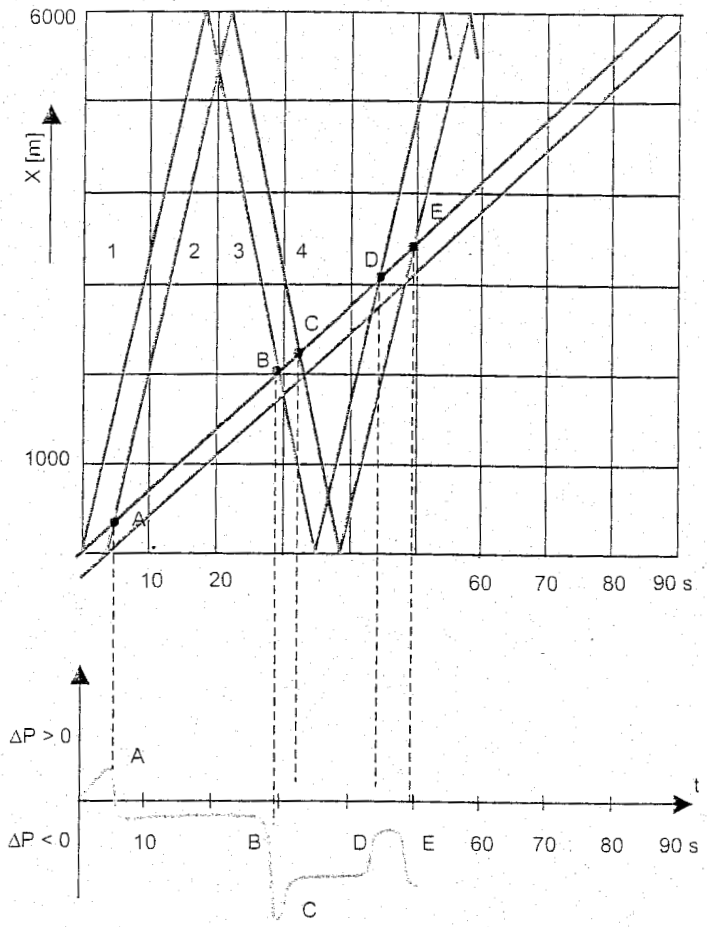


Figure 17.8: Two air pressure situations



Line 1: compression wave (blue line) caused by the front of the train when entering the tunnel

Line 2: depression wave caused when the rear of the train is entering the tunnel

Line 3: position of the front of the train in the tunnel

Line 4: position of the rear of the train in the tunnel (red line)

Figure 17.11: Wave forms in tunnel

Figure 17.11 shows the characteristics of the air pressure in tunnel (external pressure) and train (internal pressure) in case of a sealed train and a standard train. In a sealed train the internal pressure differences are much lower than in the standard train. The differential values are in the order of some kPa, being a few percent of the atmospheric pressure.

Line (1) gives the position of the compression wave (front) caused when the train entered the tunnel. This compression wave propagates with the sound speed (340 m/s) to the end of the tunnel and is then reflected as a depression wave (red line (3)).

Line (2) shows the position of the depression wave (front) caused when the rear of the train entered the tunnel. This depression wave also propagates with the sound speed (340 m/s) to the end of the tunnel and is then reflected as a compression wave (red line (2)).

These waves move between the tunnel ends while a train is passing through the tunnel. The points A to E show the position and time where the train front negotiates the compression and depression waves.

During the train passage in a tunnel the alternating air pressure waves may resonate, which may cause great changes in air pressure. Measurements have shown values of 7.5 kPa. The tunnel length and cross section, the train dimensions and speed determine the external air pressures on the train. The air tightness of the train determines how the external air-pressures are transformed into internal air-pressure variations causing passenger discomfort.

17.3.4 Basic design criteria for tunnels

There are four main criteria for tunnel design:

- Passenger comfort in the train at line speed;
- Passenger safety if the exterior of the train is open by any reason;
- Safety of maintenance personal in tunnels;
- Strength of the train exterior.

17.3.5 Calculations of external air pressures on the train

For the calculation of external air pressures the following train parameters should be considered:

- Train speed;
- Cross section of the train;
- Contour of the train;
- Train length;
- Train shape;
- Exteriour train friction.

For the calculation of air-pressure variations the following parameters are to be considered:

- Tunnel section;
- Tunnel contour;
- Tunnel length;
- Tunnel lining friction;
- Shafts (Figure 17.12);
- Tunnel wall perforations (Figure 17.12).

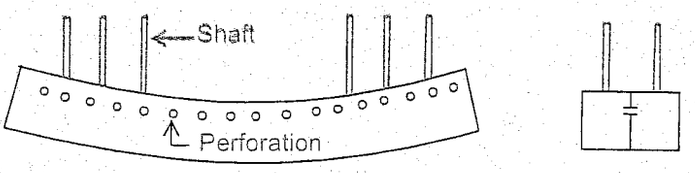


Figure 17.12: Shafts and tunnel wall perforations

17.3.6 Modeling of the tunnel

Prof. Vardy of the University of Dundee in Scotland developed a one-dimensional model THERMO-TUN/5, with the parameters mentioned above, for calculating the external air pressure variations of the train. This model, based on a pipe-system, was validated by measuring results from a French tunnel with a length of 1500 m and also from tests at the Dutch Air- and Space Laboratory with a model of a tunnel at a scale of 1:170.

External air pressure can be reduced by different design solutions. Vertical shafts and lateral wall perforations with an area of 0.01 m^2/m in tunnels comprised of two pipes give good results. The influence to the air-pressure is shown in figure 6a and figure 6b where each figure shows a comparison between the calculated value with Vardy's model and measurements of the Dutch Air- and Space Laboratory. Figure 17.13 shows the air-pressure in the model without shafts and perforations for a tunnel with a section of 50 m^2, while in Figure 17.14 the air-pressure in the same tunnel with shafts and perforations is shown.

Calculations reveal that tunnels with a cross section of 50 m^2 provided with shafts are equal to tunnels with a cross section of 70 m^2 without shafts. Perforations in the wall between the two pipes also give a greater effective section.

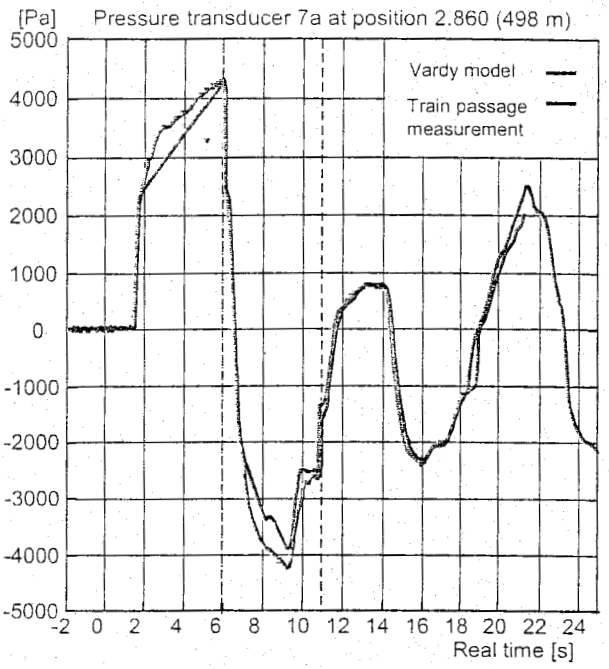


Figure 17.13: Influence of air pressure without shafts and perforations

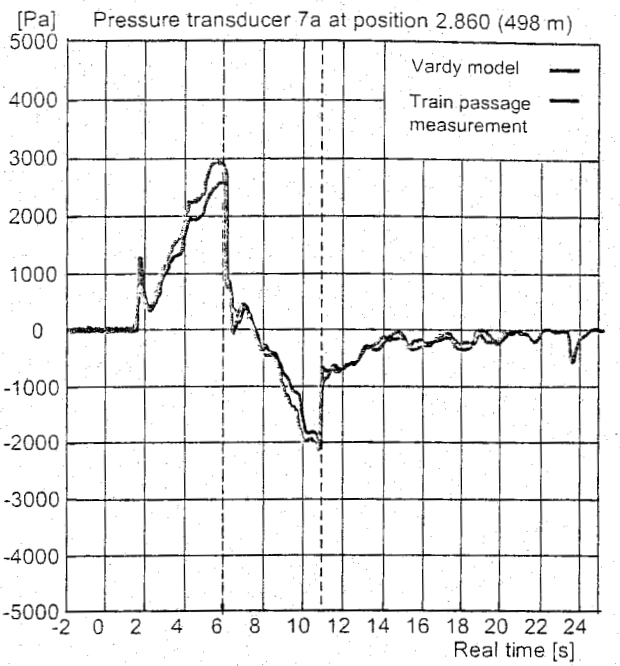


Figure 17.14: Influence of air pressure with shafts and perforations

17.3.7 Calculation of air-pressure variations in trains

External air-pressures cause internal air-pressures because a train is not air tight. The speed of reducing the difference in air-pressure in- and outside is depending on the form of the leak in the exterior of the train and on the magnitude of the air-pressure-difference itself.

The project organisation High Speed Line South in The Netherlands developed for the new railway line from Amsterdam to the Belgium border a model in which the train is a permeable and compressible box. Depending on the shape of the leak the stream will be viscose or not.

Not-viscose stream

A not-viscose stream is proportional to the square root of the difference of the in- and external pressure.

$$\frac{dp_i}{dt} = -C_1 sgn(p_i - p_e)\sqrt{|p_i - p_e|} \qquad (17.3)$$

p_i = internal air-pressure
p_e = external air-pressure
C_1 = constant factor, to be determined with a static test for the specific leak
Sgn: If $p_i - p_e > 0$ the value is 1 and if $p_i - p_e < 0$ the value is -1.

Viscose stream

If the flow is viscose the pressure variation is directly proportional to the difference in air pressure:

$$\frac{dp_i}{dt} = -C_2 sgn(p_i - p_e) \qquad (17.4)$$

C_2 is dependent on the leaks, the sound propagation speed and the volume of the train.

Air pressure variation by deformation

The shape of the train is changed by the external air-pressure. According to Klingel the inside volume variation of the train is linear with the air pressure variation if the train is air tight

$$\frac{dp_i}{dt} = k\left(\frac{dp_e}{dt} - \frac{dp_i}{dt}\right) \Rightarrow \frac{dp_i}{dt} = \frac{k}{k+1}\frac{dp_e}{dt} \tag{17.5}$$

k is a compressibility factor with a value of 0.1 according to Klingel. If k=0 the shape can't be changed. If k= ∞ the air pressure outside and inside the train is identical. The constant factor k can be determined by simultaneous measurements of the internal $p_i(t)$ and external air-pressures $p_e(t)$.

Combined model

The three models represented by the equations (17.3), (17.4) and (17.5) can be combined to:

$$\frac{dp_i}{dt} = -\frac{C_1}{k+1} sgn(p_i - p_e)\sqrt{|p_i - p_e|} - \frac{C_2}{k+1}(p_i - p_e) + \frac{k}{k+1}\frac{dp_e}{dt} \tag{17.6}$$

If $p_e(t)$ is constant, meaning that dp_e/dt is 0, than equation (17.6) will become:

$$\frac{d(\Delta p)}{dt} = -\frac{C_1}{k+1}\sqrt{\Delta p} - \frac{C_2}{k+1}\Delta p \tag{17.7}$$

With $\Delta p = p_i - p_e$

The solution of this equation is:

$$\sqrt{\Delta p} = C_0 e^{-C_2 t/(2(k+1))} - \frac{C_1}{C_2} \tag{17.8}$$

In a test the three points t_0, t_1, t_2 with Δp_1, Δp_2 and Δp_3 are determined as shown in Figure 17.15.

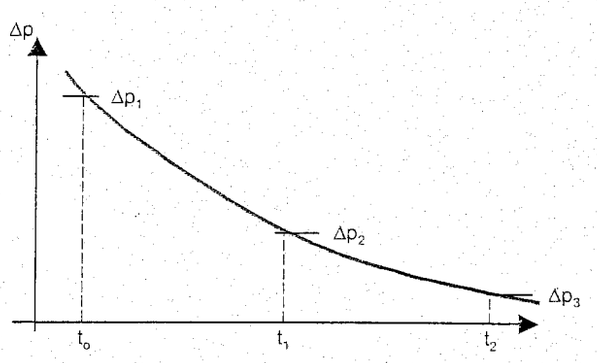


Figure 17.15: Characteristic pressure course in a leak test

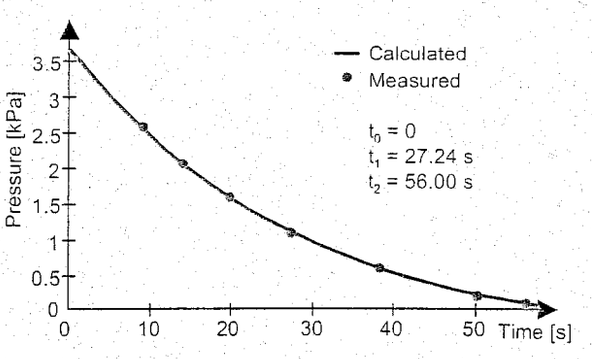


Figure 17.16: Static leak test: comparison of measuring/calculation

After this C_1 and C_2 can be computed. With C_1 and C_2 and the function of $p_e(t)$ provided for instance by the program THERMATUN/5, $p_i(t)$ can be calculated.

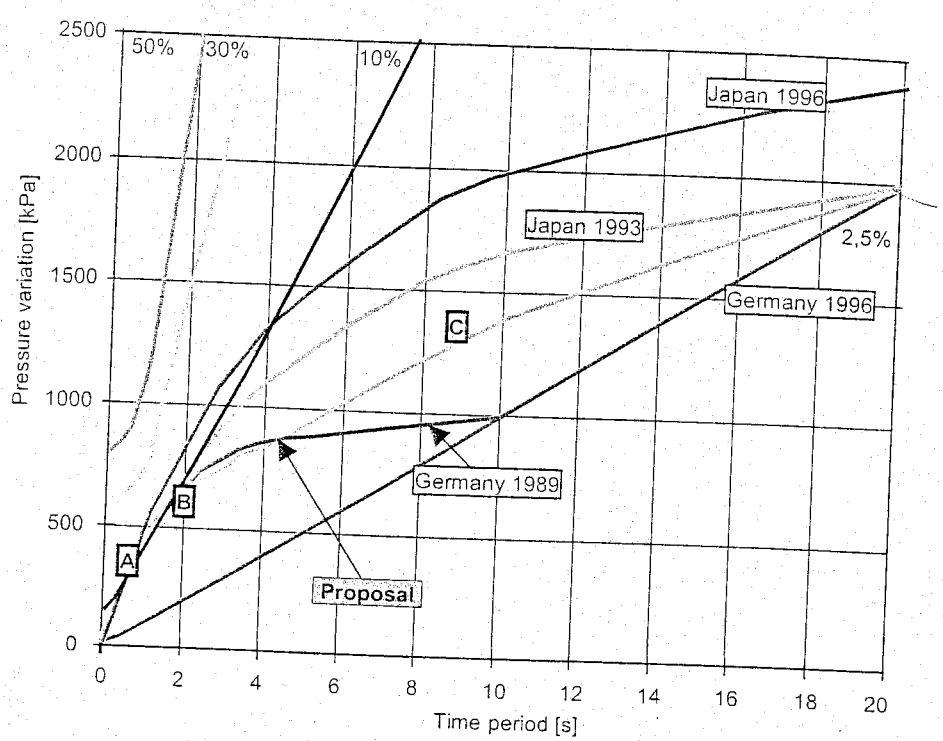


Figure 17.17: Proposed Dutch criterion for Δp in train by solo trains in tunnels

The time for reducing of the inside air-pressure from 3.0 kPa to 0.1 kPa is 51 s. The factor k is estimated 0.07 based on a French TGV-R. With these values the constant factors C_1 and C_2 are then 0.03408 and 0.02661 respectively. (Figure 17.16)

17.3.8 Criteria

In different countries comfort is experienced differently. Different criteria for the human perception of pressure variations annoyance are shown in Figure 17.17. In The Netherlands Table 17.10is proposed for comfort. These values correspond to line ABC in Figure 17.17.

Time after Δp is caused (p_e-p_i is maximal at t =0)	Interior pressure variation	Exteriour pressure variation
Time in seconds	Δp (p_t-$p_{t=0}$) in train (kPa)	Δp (p_t-$p_{t=0}$)in train (kPa)
1	0.50	0.85
4	0.85	1.35
10	1.40	2.10
20	2.00	3.00
30	2.40	3.60
40	2.80	4.20
50	3.20	4.80

Table 17.10: Comfort values proposed for The Netherlands

For safety reasons (for example if a window is broken) a maximum external pressure variation of Δp = 10 kPa is attained.

For trains a model exists (not shown) with a Wöhler curve which shows a relation between the number of air pressures on the train and the allowed intensity of these air pressures.

17.3.9 Results of calculations for tunnels in the HSL in The Netherlands

Table shows computed results for the required square meters of the cross section of two tunnels in the HSL South with a different length and a different combination of shafts and perforations.

Tunnel	Design speed TGV/SMT	Length	Square meters of cross section required		
	km/h	M	TS (1)	TS (2)	TS (3)
GH	300/220	7200	2*80	2*60	2*50
OM	300/220	1369	2*59	2*49	2*45

TS (1) Two single track tunnels without shafts and perforations
TS (2) Two single track tunnels with only shafts
TS (3) Two single track tunnels with shafts and perforations

GH = "Groene Hart"-tunnel (to be bored)
OM = tunnel under the river "Oude Maas"
TGV = High Speed Train (design speed = 300 km/h)
SMT = Standard Modern Train (also used on existing network, design speed = 220 km/h)

17.4 Maglev Applications

17.4.1 Introduction

There is a growing interest in the possible use of magnetic trains for very high-speed traffic. Test tracks exists in Germany and Japan and application in operational lines is considered. With magnetic trains the supporting action is achieved by electromagnetic levitation while the propulsion is also performed magnetically by means of a linear motor. In test tracks very high speeds were reached up to 500 km/h.

There is a basic difference between the German Transrapid system and the Japanese system. In the German system the levitation is achieved by attracting magnets in the vehicle and in the guideway, resulting in lifting the vehicle. In the Japanese system the levitation forces are generated as a result of the speed of the train, where the super conducting magnets in the vehicle interact with the coils in the sidewalls to generate the lifting forces. This means that first the train is running on wheels and is levitated after a certain speed has been exceeded.

17.4.2 The Japanese system

In the Japanese system the "8" figured levitation coils are installed on the sidewalls of the guideway. When the on-board superconducting magnets pass at a level of about several centimeters below the center of these coils, an electric current is induced within the coils, which then act as electromagnets temporarily. As a result, there are forces which push the superconducting magnet upwards and ones which pull them upwards simultaneously, thereby levitating the vehicle. The principle is illustrated in Figure 17.18. The distance between vehicle and guideway is depending on the speed and lies in the order of 10 cm.

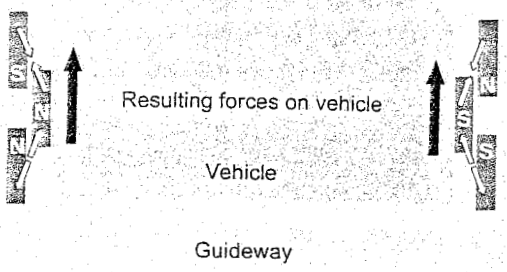


Figure 17.18: Levitation principle

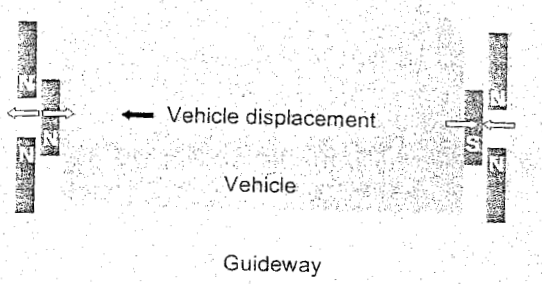


Figure 17.19: Lateral positioning principle

For the lateral guidance levitation coils facing each other are connected under the guideway, constituting a loop. When a running vehicle displaces laterally an electric current is induced in the loop, resulting in a repulsive force acting on the levitation coils of the side near the car and an attractive force acting on the levitation coils of the side further apart from the car. Thus, a running car is always pushed to the center of the guideway as shown in Figure 17.19.

The propulsion principle is explained in Figure 17.20. A repulsive force and an attractive force, induced between the magnets, are used to propel the vehicle via the superconducting magnet. The propulsion coils located on the sidewalls on both sides of the guideway are energized by a three-phase alternating current from a substation, creating a alternating magnetic field on the guideway. The on-board superconducting magnets are attracted and pushed by the shifting field, propelling the vehicle.

The guideway consists of a permanent way structure provided with ground coils corresponding to the conventional motor, as displayed in Figure 17.21. This is a vital element of maglev. For the Yamanashi Maglev Test Line three methods of installing the ground coils for propulsion, levitation, and guiding to the guideway were adopted.

In the beam method, the sidewall portion was constituted solely of concrete beams. The entire process from beam manufacturing to installation of the ground coils take place at the on-site factory. A finished beam is transported to the work site within the guideway, to be placed on two concrete beds set up in advance there.

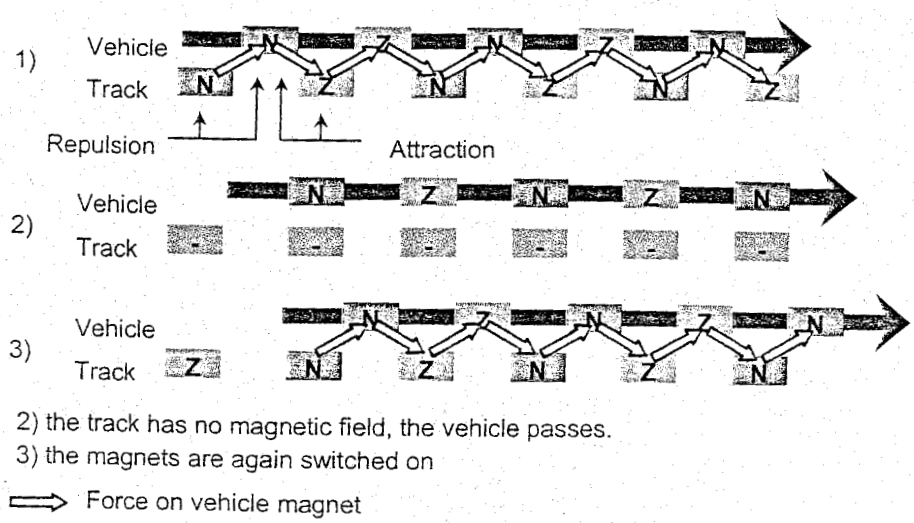


Figure 17.20: Propulsion principle

Propulsion System

The synchronous longstator linear motor of the Transrapid maglev system is used both for propulsion and braking. The function of this non-contact propulsion and braking system can be derived from the functional principle of a rotating electric motor whose stator is cut open and stretched along both sides of the guideway. Instead of a rotary magnetic field, the motor generates an electromagnetic traveling field. The support magnets in the vehicle function as the rotor (excitation portion) of the electric motor. The principle is explained in the Figure 17.24 and Figure 17.25.
In contrast to conventional railway systems, the primary propulsion component of the Transrapid maglev system – the stator packs with three-phase motor winding – are not installed in the vehicle but in the guideway.

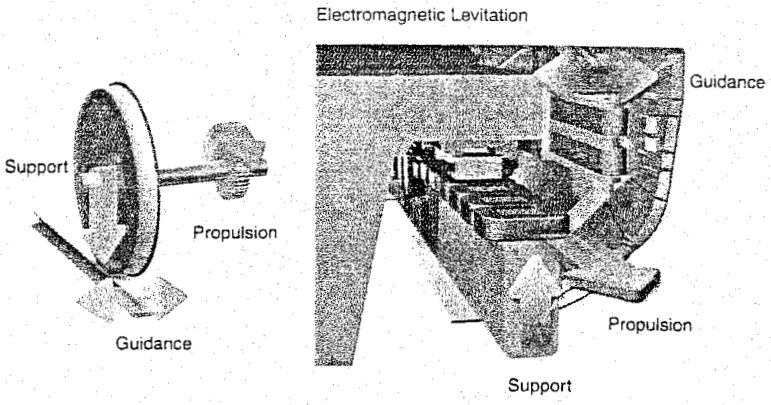


Figure 17.23: Levitation principle Transrapid

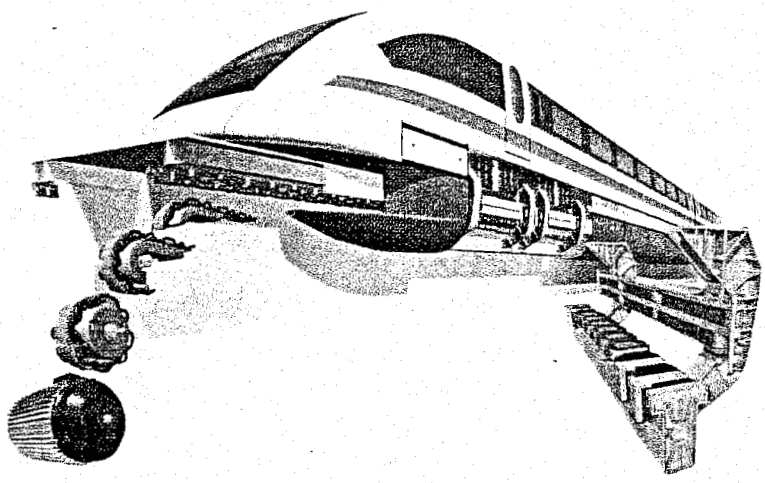


Figure 17.24: Linear Motor principle Transrapid

By supplying alternating current to the three-phase motor winding, an electromagnetic traveling field is generated which moves the vehicle, pulled along by its support magnets which act as the excitation component (see Figure 17.24)

.The speed can be continuously regulated from standstill to full operating speed by varying the frequency of the alternating current. If the direction of the traveling field is reversed, the motor becomes a generator which brakes the vehicle without any contact. The braking energy can be fed back into the public network.

The longstator linear motor in the guideway is divided into individual motor sections which are only supplied with power as the vehicle passes. The location and the installed power of the substations depends on the requirements on the propulsion system. In sections where high thrust is required, e.g. gradients, acceleration, and braking sections, the power of the substations is higher than on level sections which are travelled at constant speed. And because the primary component of the propulsion system is installed in the guideway, Transrapid vehicles need not carry the entire motor power for the peak requirements, as is the case with other types of vehicles. The support and guidance system is supplied with energy without contact via the linear generators integrated into the support magnets. No overhead wires are required for the

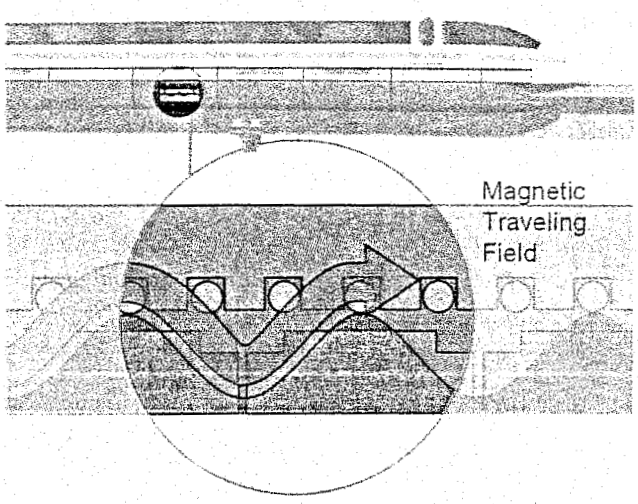


Figure 17.25: Propulsion principle Transrapid

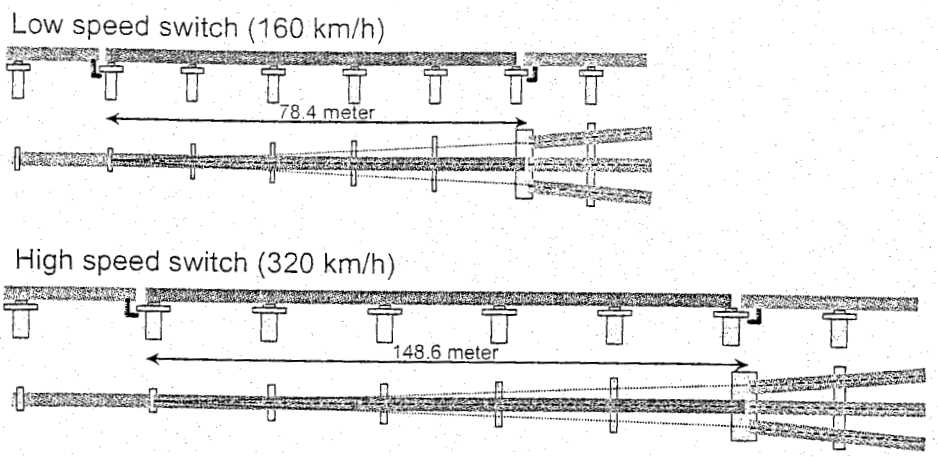


Figure 17.27: Principle of switch structures

Tunnels

With its flexible route alignment parameters, the Transrapid guideway can be adapted to a great extent to the landscape. Therefore, tunnels are seldom necessary, even in hilly and mountainous terrain. Even when they are required, the tunnel cross-sections necessary for the Transrapid are smaller than those of railways. This is due to the smooth, aerodynamic shape of the vehicle and small clearance envelope. Typical tunnel cross-sections for tunnels longer than 150 m and vehicles with 2 to 8 sections are given in Table 17.11.

Speed	250 km/h	400 km/h	450 km/h
Cross section of double track tunnel	70 m^2	180 m^2	225 m^2
Cross section of single track tunnel	36 m^2	85 m^2	120 m^2

Table 17.11: Tunnel cross section values

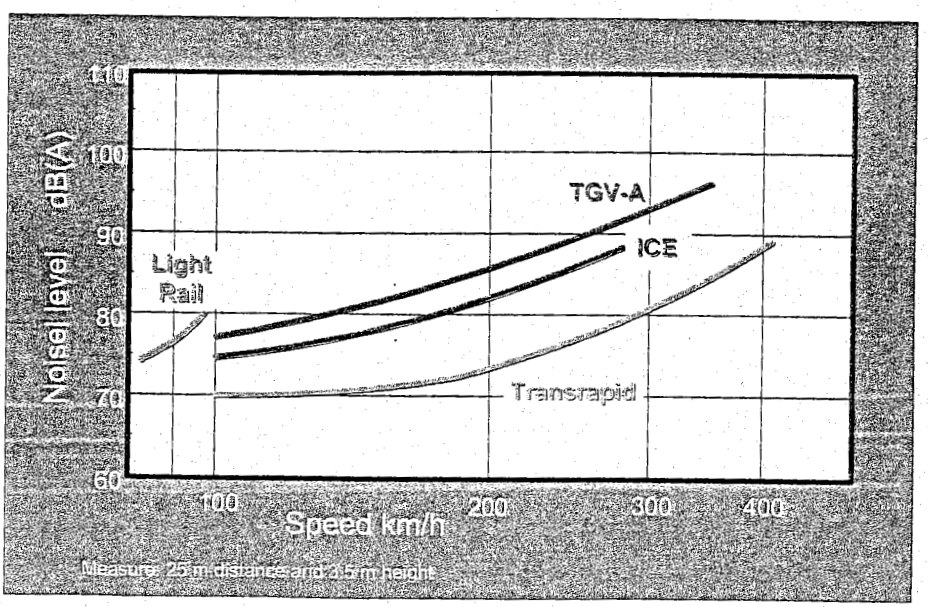


Figure 17.28: Measured noise levels for different systems

Route Alignment Parameters

Two features of the Transrapid maglev system allow it to have extremely favourable route alignment parameters. These are the active guidance of the vehicle along the guideway and the propulsion system (motor) in the guideway. With these, gradients of 10% can be climbed and curves with tight radii and cants of up to 16° can be travelled without difficulty. These favourable parameters also allow the guideway to be flexibly adapted to the landscape without massive earthworks and it is often possible to collocate it with existing traffic routes, thus saving precious land and cost.

Despite the high speeds of up to 500 km/h the safety is relatively great. The vehicle virtually cannot derail because it wraps around its guideway. And since grade crossings are not allowed, nothing can get in its way. Collisions between Transrapid vehicles are also ruled out due to the technical layout of the system and the section-wise switching of the guideway motor. The vehicle and the traveling field of the guideway motor move synchronously, i.e. with the same speed and in the same direction. Additionally, the section of the longstator linear motor in which the vehicle is moving is only switched on as the vehicle passes. Other important elements of the safety concept are:

- automatic train protection
- passive protection equipment to prevent damage to the guideway structure and violation of the vehicle's clearance envelope
- automatic inspection of the guideway
- protection of the passengers during boarding and exit in the stations by means of platform doors and gap bridges

Environmental issues

Compared with other transport systems, the Transrapid is extremely quiet. There is no rolling or propulsion noise. At speeds up to 250 km/h, the Transrapid hovers almost soundless. Due to the non-contact levitation and propulsion technology, the noise emission at speeds above 250 km/h is mainly determined by aerodynamic noise. The pass-by noise levels of the Transrapid have been measured at the Transrapid Test Facility (TVE). The levels for elevated guideway compared with other rail systems are shown in Figure 17.28. For at-grade guideway, the levels are even lower.

The air flow surrounding the Transrapid has been extensively studied at the Transrapid Test Facility (TVE). The aerodynamic effects are comparable to a gentle wind. The measured air flow velocity along the vehicle at a distance of 1 m and a speed of 350 km/h was smaller than 10 km/h.

Characteristic		*TRANSRAPID*	*ICE 3*	*TGV-A*
Speed	max.	500 km/h	330 km/h	300 km/h
Mass / seat		0.6 t	1.1 t	1.0 t
Acceleration time	0-200 km/h	82 s	150 s	170 s
	0-300 km/h	120 s	335 s	345 s
	0-400 km/h	165 s		
	0-500 km/h	225 s		
Acceleration distance	0-200 km/h	2,200 m	5,000 m	
	0-300 km/h	4,900 m	18,900 m	18,500 m
	0-400 km/h	9,300 m		
	0-500 km/h	17,000 m		

Figure 17.29: Performance comparison of different systems

Vibration measurements at the Transrapid Test Facility (TVE) have shown that the vibrations at a distance of 25 m and a speed of 250 km/h, typical in urban areas, are below the human perception threshold. At a distance of 50 m, no vibration is perceivable at any speed.

To conclude this section some comparative figures between some railway systems and the Transrapid are given in Figure 17.29.

18 TRACK MAINTENANCE MANAGEMENT SYSTEMS

18.1 Introduction

In order to meet competition from other modes of transport, there is an increasing demand upon the railways to improve reliability, efficiency, and transit times. The resulting requirements for improvements in speed and axle load mean that the demands made upon the track are becoming more onerous. In order to provide cost-effective track to meet this need in the future, it is essential to be able to improve the methods by which the performance of the track is monitored and to have reliable methods for prediction and planning.

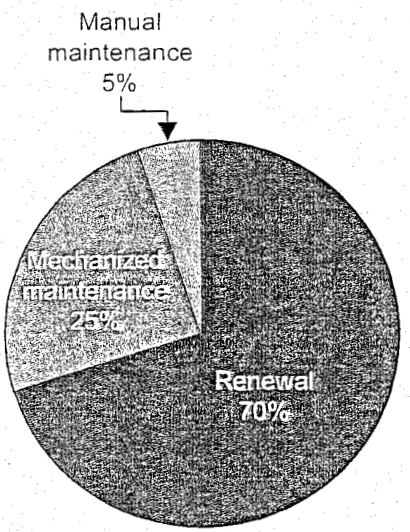


Figure 18.1: Total annual expenditure for maintenance and renewal at NS (price level 2000)

Maintenance and renewal of large railway networks require huge amounts of money. For example, the annual expenditure for NS permanent way (price level 2000) amounts to about EUR (€) 180 million (Figure 18.1). Only 25 % of this value concerns mechanised track maintenance and 5 % manual track maintenance. It is, therefore, obvious that the high expenses are caused by track renewal. To achieve an effective cost reduction the decision makers need to be provided with adequate information. With objective data of this sort, processes become more transparent and, thus, can be better controlled.

As will be explained in more detail in Chapter 16, computer-aided Track Maintenance Management Systems (TMMS) logically represent a constitutive part of Asset Management Systems (AMS). However, until AMS become sufficiently developed and fully accepted as a concept, TMMS will exist more as isolated systems. Even as isolated system, if designed properly, they represent invaluable tools for any track and/or infrastructure manager.

In order to properly manage track maintenance, a vast amount of data is needed. Types of data to be collected for computer-aided TMMS are summarised in Figure 18.2. For an efficient analysis of the track the data had to be divided into segments. In fact, all information is linked through the track segment to which it refers.

This information is by no means restricted to just data from automatic recording systems. Other data, such as from visual inspections, various layout and operating data, data about speed reductions, spot

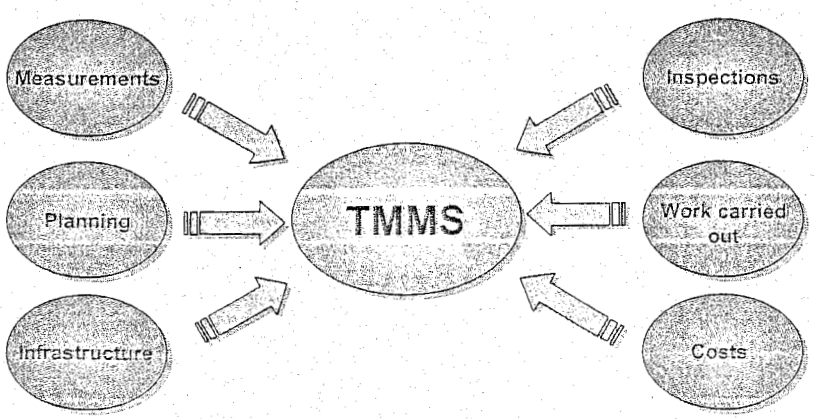


Figure 18.2: Types of data to be collected for computer-aided Track Maintenance Management Systems

maintenance, replacement of track components, and, particularly, maintenance history and costs should be taken into consideration as well.

Segment length could differ from one railway organisation to another, and generally depends on the railways' maintenance policy. However, in most cases this length corresponds to the segment length for which the track recording car produces quality indices. Track recording car data, on the other hand, mostly refer to sections 100 or 200 m long, or alternatively, are stored on a kilometre basis. Finally, the information needs to be aggregated into long "route" sections in order to obtain an assessment of the quality or needs of a total route section between major points on the network.

18.2 Basic data for prediction and planning

Specific data are needed for proper monitoring of the various types of track deterioration. Some basic examples of these data are listed below:

1 *Geometry measurements* (alignment, vertical, twist, crosslevel, gauge, quality indices, number of faults, other user-specific parameters).

2 *Inspections and other measurements*

- General condition;
- Ballast condition (% surface soiling, % pumping, % weedy ballast);
- Fastening condition (% ineffectual fastenings, % loose fastenings);
- Sleeper condition (% bad sleepers, % medium sleepers, sleeper condition);
- Rail failures (number of failures, number of failures outside welding zones, % of defective rails);
- Rail wear (vertical wear of rail head, lateral wear of rail head, chamfered angle of rail head);
- Corrugation (amplitude of corrugation, vertical acceleration).

In addition to the specific data, there are also general data which should in any case be collected for a section of track. These data are referred to as "infra data" and usually consist of:

1 *Layout and operating*

- Curves (start and end km, curve hand, radius, etc.);
- Loads (annual load [MGT], maximum axle load [tons], date from which data is valid, etc.);
- Speeds (speed of freight trains and passenger trains, date from which a speed is valid);
- Gradients (start, end, value).

2 *Superstructure and infrastructure*

- Subgrade (geological condition, various monitored parameters, etc.);
- Ballast (ballast type, date of installation, ballast thickness, etc.);
- Sleepers (sleeper type, were sleepers new when laid, load on sleepers when laid, sleeper

spacing, type of fastening, date of installation);

- Rails (rail type, jointed track or not, weld type, date of installation, were rails new when installed, cumulative tonnage on rails when installed);
- Structures (type of structure, start and end km, code, name);
- Switches and Crossings (type of a switch, code, name, start and end km).

3 *Work history*

- Renewals, grinding and tamping work history (start km, end km, type, cost, etc.);
- Speed restriction history (start and end date of temporary speed restriction, value of reduced speed, cost);
- Spot maintenance history (type, date, cost).

18.3 Track geometry

Of the specific data mentioned above only the track geometry will be discussed here. The whole track system is designed and maintained to provide satisfactory geometry. Renewal decisions are often governed by the geometry. Roughly speaking, renewal is necessary when too many rail failure repairs spoil the geometry, but ballast is also renewed when it can no longer maintain good geometry. Sleepers and fastenings are considered to have failed when the track gauge cannot be maintained. However, the process of determining whether, when, where, and how best to intervene is far more complex.

18.4 Prediction of geometry deterioration

In order to know what the limiting quality is and to decide when maintenance is required, it is necessary to be able to predict geometry deterioration.

Basic principles of track geometry deterioration and assessing track quality for maintenance decisions have been explained in the Chapter 12.

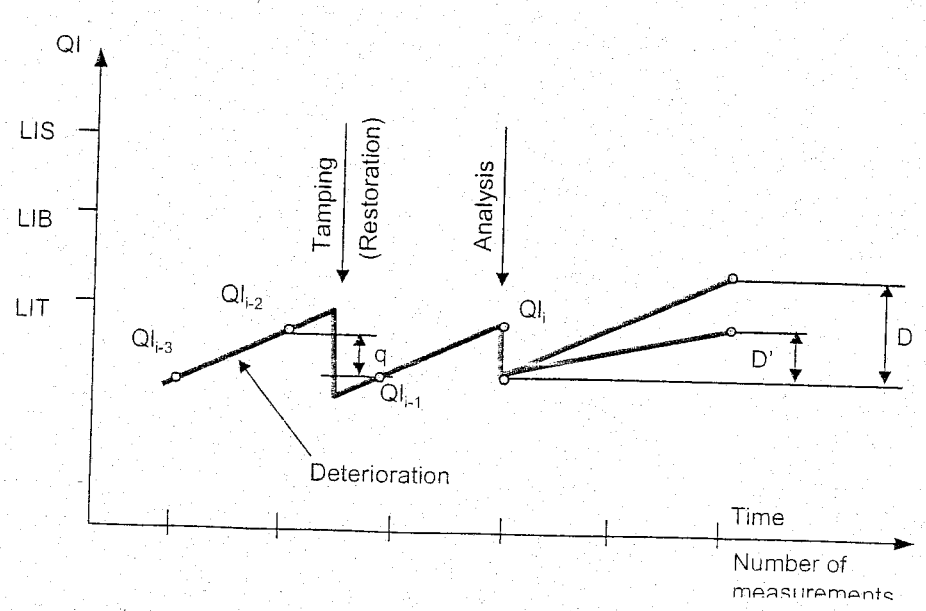


Figure 18.3: rack deterioration and restoration in terms of the Track Quality Index (QI) [64]

As similar sections of track may have very different rates of deterioration and maintenance improvements (restoration), it is necessary to collect geometry data in very short sections for the purpose of optimising maintenance and identifying the influencing factors.These sections may be as short as 200 m. The types of data that are required are as follows:

- Maintenance history;
- Vertical geometry;
- Lateral geometry;
- Short-wave rail geometry;
- Infra data.

Vertical and lateral geometry is usually measured by track recording cars like EURAILSCOUT (see Section 16.9. They enable the standard deviations to be calculated, which has been shown to be useful for predictive purposes. In some cases vehicle reactions calculated from recorded geometry are used to assess track quality. When measurements have been made over two or more tamping and lining cycles, average values for both the "deterioration rate" and the "restoration rate" (improvement from maintenance) can be found for each section (Figure 18.3). Examples of such analyses can be found in literature.

18.5 The basics of the analysis principle

The basics of the analysis principle

A general explanation of the diagnostic principle can be found in Figure 18.4. The basic idea of the diagnosis principle is that the behaviour of the track geometry of a certain segment, expressed by means of a certain geometry parameter, is monitored in time and, thus, captured.[143]

The thick green line shows the hypothetical deterioration of the track geometry if the track had been left to deteriorate without any maintenance input. On this line three phases can be distinguished: the first phase, often referred to as "youth", occurs immediately upon (re)construction or completion of a renewal work and characterises rapid and substantial deterioration caused by the initial settlements

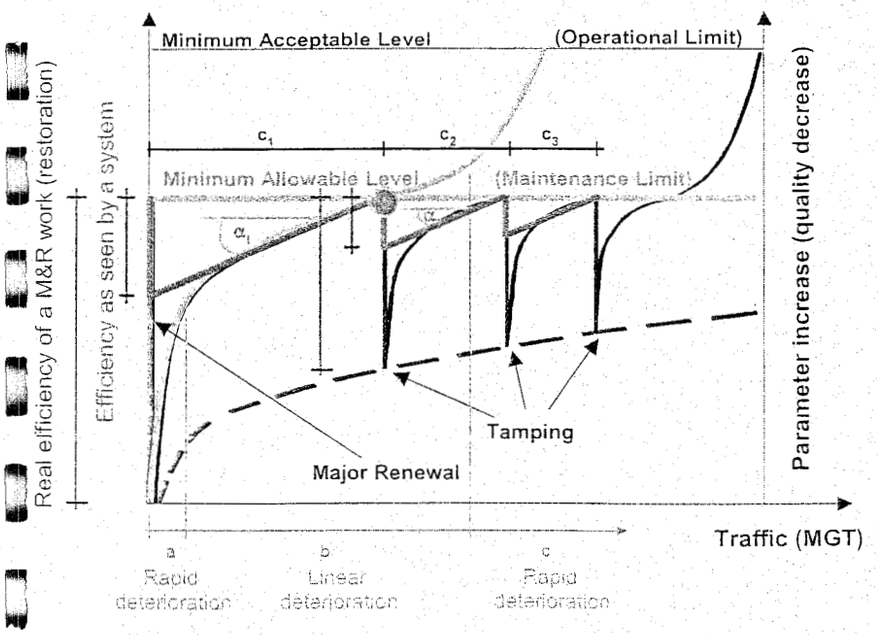


Figure 18.4: Analysis Principle applied on a hypothetical track geometry deterioration

of the track (Figure 18.4 marked with "a"). This period is also highly unpredictable and differs considerably from one track section to another. It is, therefore, very hard to model. This is why this period is usually disregarded by discarding any geometry measurements performed within this period from analyses. The length of this period is fortunately quite short which diminishes the consequences of omission.

The second phase, which occurs once the track has been sufficiently stabilised, shows a more or less linear deterioration pattern. This kind of behaviour occurs during most of the track's life-time and this is exactly the period on which most of the analyses are based (Figure 18.4, marked with "b").

The third period occurs in the latter part of the track's lifetime and is characterised by a more and more rapid deterioration which eventually even takes on a more or less exponential form (Figure 18.4 , marked with "c"). Normally, this is a situation which should never be allowed to happen under any circumstances, as it could effect the safety of traffic as well. This is avoided by applying certain appropriate M&R works at a much earlier stage, i.e. there is always a preset maintenance threshold value (horizontal purple line) which, when reached, triggers certain M&R activity.

Based on this concept, measurement data within the normal (linear) part of the track behaviour are analysed. By doing this the track geometry behaviour is "captured" by calculating the trend line through the measured points and extrapolating it. Provided that the maintenance threshold has been set (purple horizontal line), the moment (or the tonnage) when this extrapolated line will reach the threshold is calculated, marking the moment when certain M&R activity, e.g. tamping, should be performed.

However, this is just one aspect of the analysis. After tamping has been performed, the parameter value (normally standard deviation) drops abruptly, i.e. the quality increases. In Figure 18.4 this is represented by the vertical distance marked with red (real parameter drops), or the blue vertical distance and vertical drops marked by the thick blue saw-tooth line (simulated parameter drops).
After the quality has been improved by tamping, the deterioration process will start again. However, several things will change over time as the track grows older. The first thing that changes is the efficiency of tamping, i.e. the intensity of the "vertical drop". This can be observed by looking at the line showing the "real behaviour" of the track (red vertical dimensioning line) or by looking at the simulated linear line (thick blue saw-tooth-like one). But it becomes even more clear from the dashed red line which shows this change over time.

Another thing that changes is the "deterioration rate", i.e. the slope of the line defined by measured points (represented by green angle marked with α_1 & α_2). Both of these two events have their impact on the required tamping frequency which becomes higher and higher, i.e. the time period between two tamping works (tamping cycle) becomes shorter and shorter. Eventually, the tamping frequency becomes so high that it no longer makes sense to tamp. However, action definitely needs to be taken, e.g. some other M&R activity like ballast renewal.

This kind of logic was also adopted by the ECOTRACK system which determines the optimal work to be done on a certain track section, i.e. for the decision-making process incorporated in ECOTRACK.

The global idea is to analyse the track elements' condition from as many aspects as possible. This is why a track condition database should be quite extensive as explained earlier. The goal is to enable the track manager to see the "big picture", i.e. to simultaneously display all kinds of information that can influence the condition of the track, in order to search for the real cause of certain track problems and reach decisions about the best possible remedial actions. This decision making can be performed either manually, displaying and overlaying all sorts of information, or automatically using the predefined decision rules.

Unfortunately, despite the wish to see all of these data displayed in one place in a user-friendly manner, track managers, in fact, usually do not possess most of the data. A good computer-aided decision support system should therefore be able to work with as little data as provided. Luckily, nowadays, there is hardly any railway organisation that does not at least possess geometry measure-

ments data. Even with only this data computer-aided DSS should already be able to reach very sound decisions. Knowledge about the other data and rules that process them could then help the track managers to see what kind of analyses could be performed with them and what kind of decisions could be made. Ultimately, this could then serve as a motive for the track managers to start collecting and using this data in order to reach even better decisions in future.

18.6 Monitoring system for wheel defects

An important source of noise and vibration or even damage is the presence of wheel thread defects like wheel flats, oval wheels, etc.

To detect such sources a monitoring system could be devised (Figure 18.5) in the track, which in case of a wheel defect records the wheel number, the type end the level of the defect.

Large exceedences of force levels are reported immediately to the responsible authorities.

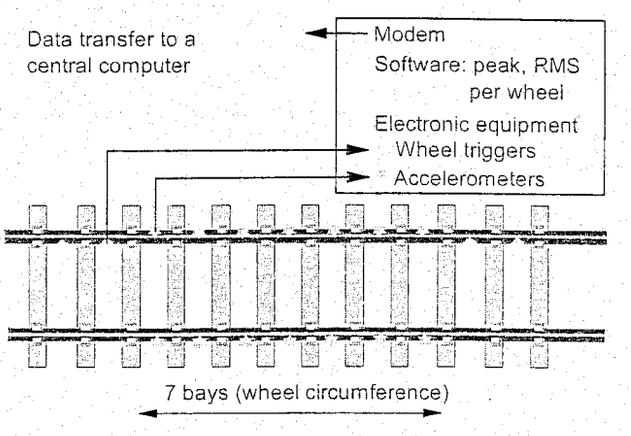


Figure 18.5: Monitoring system for detecting wheel defects

18.7 Rational rail management

Over the years the rail managers (tramway and railway companies) showed a growing need to perform rail management more rationally and more efficiently for the following reasons:

- To be less dependent on the individual know-how of co-workers;
- To create optimal working conditions regarding business economy;
- To bear responsibility and to report to the management.

Therefore a platform of experts from rail companies in the Netherlands and the TU Delft was created to look at the feasibility of an integral rational rail management system (in analogy to the earlier development of the rational road paving system).

The following definition was formulated:

Rational rail management is aimed at the objective assessment of the qualitative and quantitative state of the rail infrastructure, after which, based on a system of objective rules and standards, decisions may be taken regarding the maintenance and the renewal of rail infrastructure.

18.8 ECOTRACK

18.8.1 Introduction

A major development in the field of rational track management was the ECOTRACK project. The UIC (Union International des Chemins de Fer) gave ERRI (European Rail Research Institute) the task to develop a decision support system in order to manage track maintenance and renewal (M&R) in an optimal way. Its goal was to increase the productivity of permanent way. The system was supposed to be based on modern diagnostic methods including expert systems for engineering (analysis of track

sections, geometry, and equipment) and for medium and long-term management (planning and optimization of available resources allocation).

The development of the ECOTRACK system took place within ERRI's D187 Committee between 1991 and 1998, and some 24 European Railways took part in various phases of its development, see Figure 18.6.

18.8.2 Overview

The costs of permanent way and its maintenance are very substantial and form a large part of the total infrastructure expenditure. Any reduction of these costs has a significant impact on the overall efficiency of infrastructure management. It is, therefore, of primary importance to those responsible for the state of the permanent way and its maintenance that it is kept at a given quality level for the period required at as low cost as possible.

The process of determining whether, when, where, and how to intervene, deciding on an optimum allocation of resources, and minimizing costs is very complex because different track sections tend to behave differently under the effects of loading. It is also complex because decision processes for M&R works are closely interrelated technically and economically, and decision making for M&R plans is based on a large quantity of technical and economic information, extensive knowledge, and, above all, experience.

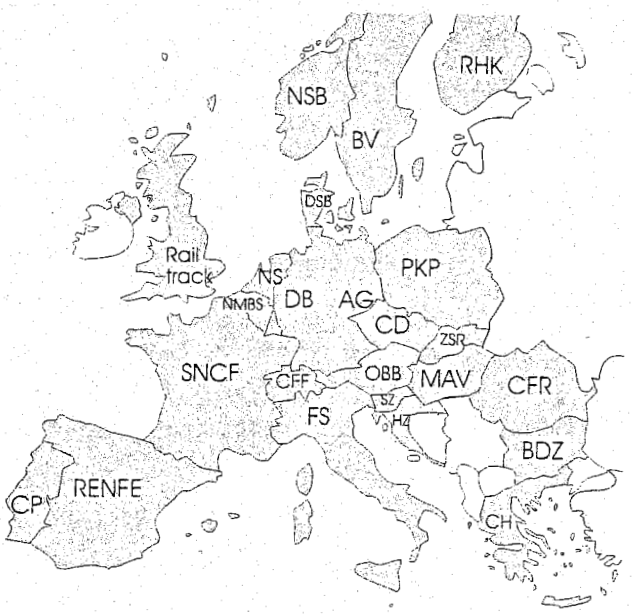


Figure 18.6: Networks participating in the development of ECOTRACK

ECOTRACK software enables the decision makers of railway operators to plan expenditure and the financial consequences of M&R policies.

The general idea of ECOTRACK development was to solve the most complex link in the track management sequence of operations - making decisions about the M&R Work Plan (see Figure 18.7). In order to achieve that, complex interrelationships between the in-situ condition to which the track is subjected (e.g. annual tonnage, axle loads, speeds), the state of the track geometry and track material, and various M&R works input had to be resolved and formulated (Figure 18.8). These interrelationships are referred to as "rules". Some of these rules represent the deterioration process, some represent the restoration process as a consequence of the M&R works, and some represent experience rules based on the decades of M&R practice on various European railways.[141].

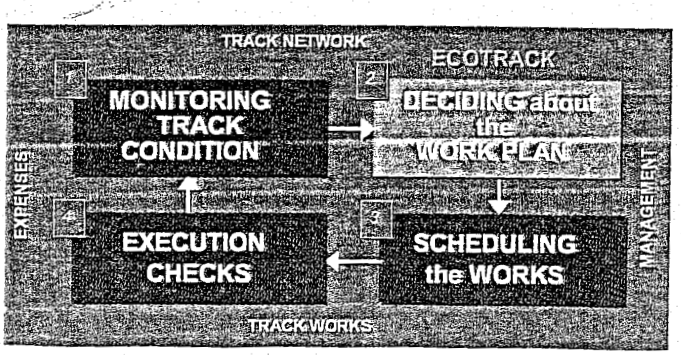


Figure 18.7: ECOTRACK - General concept

Level 5 - Overall network management

This level provides a set of tools required by the planner to achieve the optimal management of M&R works over the whole network. This level includes powerful database tools like showing various thematic maps (Figure 18.12). In this case, Figure 18.12 shows the speed distribution of passenger trains on the Belgium network (NMBS). This tool can be successfully used to present various information to both the higher management levels and the public. Furthermore, there is a "Track material display" tool (Figure 18.13). This is an extremely useful tool that allows the user to see all the track material and its most important parameters in one place in a very logical and user-friendly manner. Also, by clicking on each of the track elements that are displayed, the user receives additional information concerning that track element.

The option also exists to present various statistical analyses, e.g. the distribution and history of M&R works performed in the past (Figure 18.14), which gives the user an instant insight into how much work has been done on a certain line or a piece of the track network. There are in total some 26 parameters that could be statistically analysed and presented to the user. The one shown in Figure 18.15 presents the change in a certain quality parameter, in this case standard deviation of the alignment values, over the years.

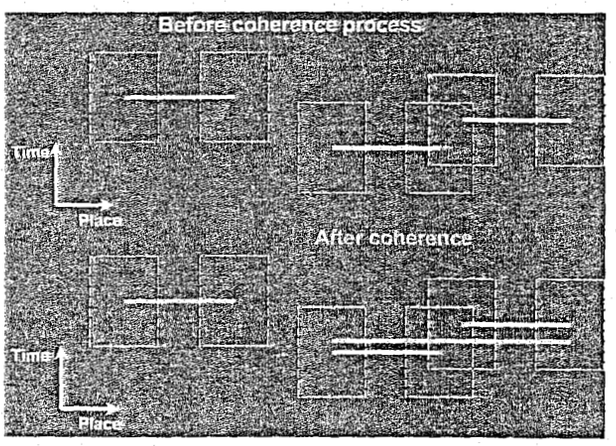


Figure 18.10: Coherence Principle

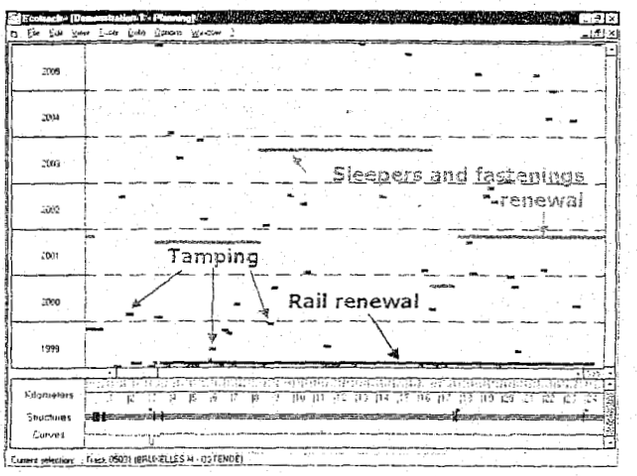


Figure 18.11: ECOTRACK - Work PLan

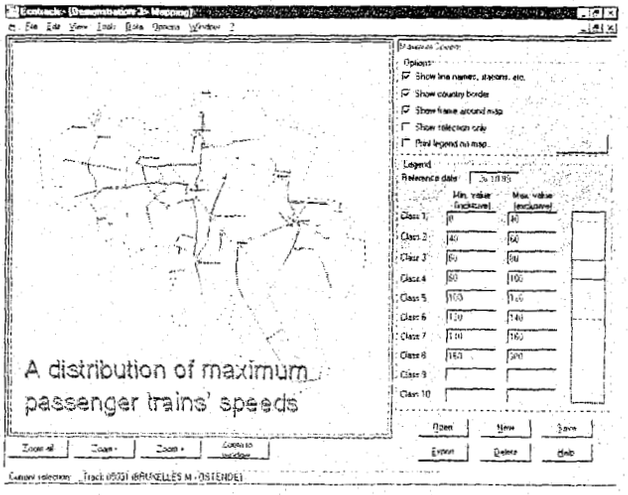


Figure 18.12: Thematic Maps

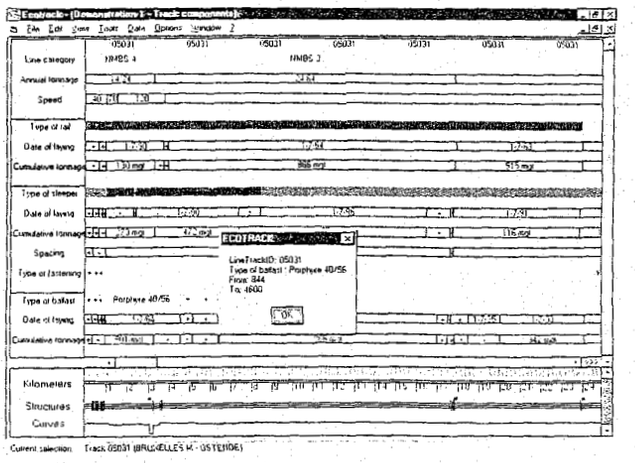


Figure 18.13: Track Material Display

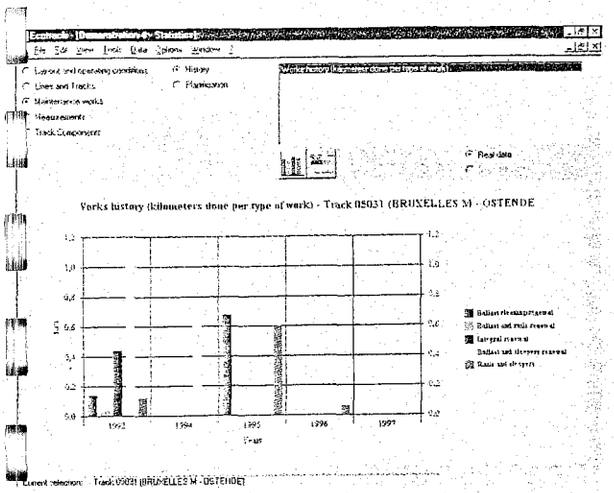


Figure 18.14: Statistics (M&R work history)

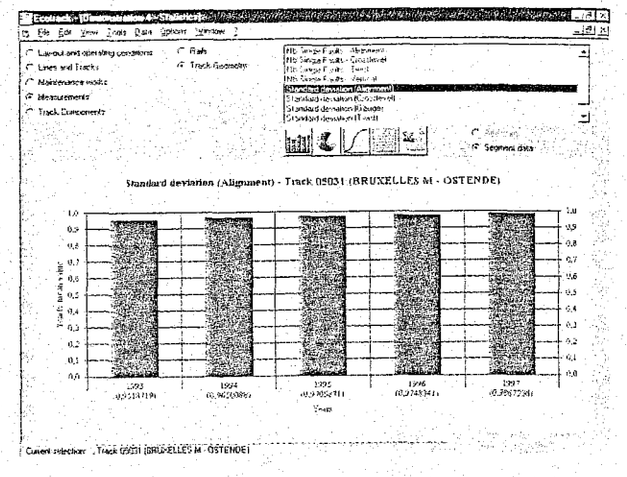


Figure 18.15: Statistics (measurements)

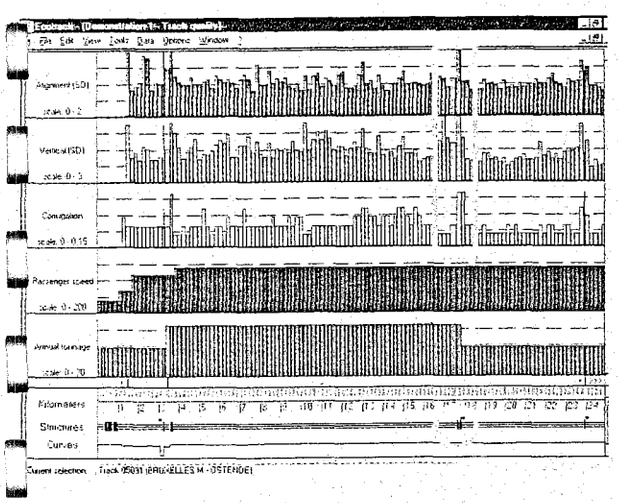


Figure 18.16: Track Quality Display

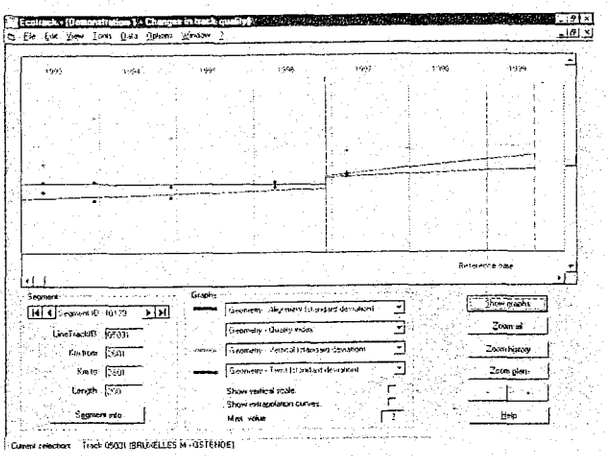


Figure 18.17: Track Quality Evolution Display

This information, for example when coupled with the information on the work history or the budget history, could show the infrastructure manager the effect of budget cutbacks made in the past on the overall track quality. He could then, for example, easily see to what extent these cutbacks caused the drop in the overall track quality, or by how much the certain increase in the M&R work input did improve the overall track quality. He could then use these findings for future budget adjustments and negotiations.

Another useful option is the Track Quality Display (Figure 18.16) which offers the user the ability to see the quality of each track segment (usually 200 m, but could be of any length) according to any of the defined track quality parameters. The user can choose which parameters will be displayed and in which order, so that he can overlay different interrelated parameters in order to get a "complete picture". For example, the obvious increase in the parameters' values observed in the right part of the display (marked with the box) was not caused by some unexpected and unexplainable track deterioration, but was rather due to the fact that a set of switches located on this track segment (marked with a circle) caused the higher measurement values on that track section. This, however, could have quickly been inferred if only the measured parameters had been superimposed on the track layout display, which would have allowed the set of switches to be clearly seen. Without the Track Quality Display, the infrastructure manager could have lost a lot of time figuring out this sudden increase in the measurement values.

Finally, perhaps the most useful option is the Track Quality Evolution Display (Figure 18.17). This feature serves as a basis for future M&R works planning. Basically, the parameter measurements performed over a certain amount of time at regular time intervals are used to "capture" the track behaviour. Except for the short period following the performance of M&R works and the period of extreme track deterioration at the end of the track elements' life when rapid and non-linear deterioration can be observed, the track usually takes up a linear deterioration rate.

This rate, or behaviour, is deduced from the consequent measurements and, thus, the future condition can be foreseen, i.e. extrapolated. If the allowable threshold is already defined for this particular parameter, the moment can be calculated when this parameter will cross the threshold. This will be the moment when certain corresponding M&R work should be performed. This process is used as the main basis for scheduling the M&R works. The process of capturing behaviour, defining the threshold, and determining which work type(s) should be performed in which circumstances is referred to as "Rules" in ECOTRACK [142].

These "Rules" represent the core of ECOTRACK's inference engine, or its knowledge base. During the development of ECOTRACK some 173 rules were gathered from the maintenance practices of the aforementioned 24 European Railways. After an extensive and thorough selection process, 54 rules were adopted and included as standard rules in the ECOTRACK knowledge base. However, these rules can always be turned off and on, modified, and added to so that ECOTRACK can fit any specific Railway policy.

Moreover, these rules can only be used if the data, which they are supposed to analyse, are available. This is why, very often, ECOTRACK users only utilise a certain portion of the available rules. On the other hand, the existing rules which are not used at the moment may form an incentive for Railways to start collecting the data for them, so that even more reliable diagnoses and decisions may be had.

Finally, ECOTRACK possesses also an extremely beneficial option to simulate the future conditions that might occur on a certain track network. Here, the user can change the future speeds, annual tonnage and axle loads (Figure 18.18) and see the resulting impact on the future maintenance and renewal costs and track quality.

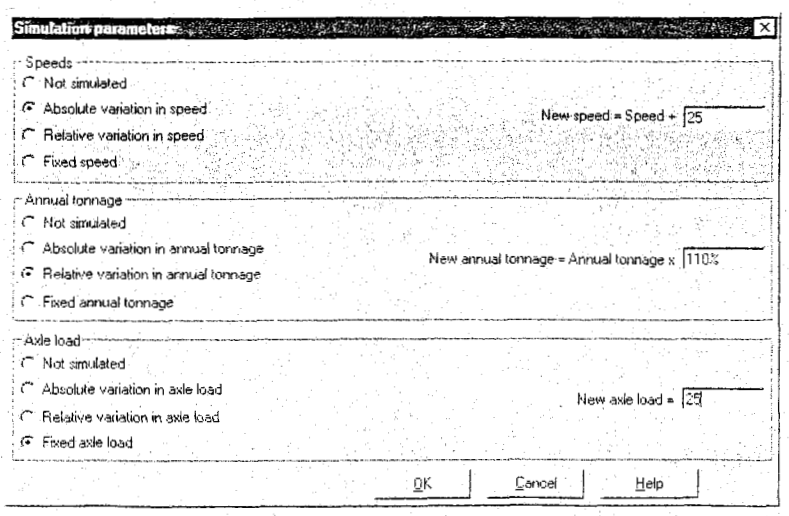


Figure 18.18: Simulating future conditions

18.8.4 Features of the ECOTRACK system

This section is concluded with an overview of the different ECOTRACK features:

- Objectively analyses the track based on the technical and economic aspects;
- Produces an optimal plan for the most important M&R works, i.e. the works that account for most of the track M&R expenditures, several years in advance - it is a tool for strategic M&R management;
- Produces a sound base for an optimal Resource Allocation;
- A modular system with an interface to use the existing database of the individual railways;
- Uses standard rules that can be easily modified to suit each individual railway's practice - it is flexible;
- Instant reference to data on all track components, their level of use, and M&R works history;
- Permits maintenance to be carried out according to a fixed budget and allows decisions to be made on the various priority levels for required maintenance for a number of years;
- Is user friendly and operates in a Windows 95 or NT environment on a PC - it is easily installed and is portable.

19 RAILWAY ASSET MANAGEMENT SYSTEMS

The last two decades of the twentieth century have seen the emergence of computer-aided track maintenance and renewal systems based ever-increasing amount of data. Problems of inaccessibility regarding the required data, especially the data in the digital format suitable for the automatic analyses, which were present during the eighties, were largely overcome during the nineties. This has enabled much better and reliable track condition analyses, and prompted the development of comprehensive condition-based decision support systems for track maintenance and renewal, such as ECOTRACK.

On the other hand, the development of systems like ECOTRACK necessitated further development of more sophisticated and more reliable track deterioration models. This, along with the constantly rising amount of high-quality track condition data available, made it clear that track condition and related maintenance and renewal management could no longer be considered separately from other railway infrastructure elements. Instead, it became obvious that track management had to be coupled to substructure management, track structure management, overhead lines management, etc.

The reasons for that were numerous. Firstly, it was long known that the other railway infrastructure elements, like substructure for example, had very significant influence on the track superstructure behaviour. This was especially true at the locations where persistent and recurring problems with track geometry were observed. However, until the high-quality data were made sufficiently available and until systems like ECOTRACK emerged, this could never be sufficiently quantified. The reason for that was that only by using these systems it would become possible to overlay many different kinds of track condition information at the same time, thereby enabling the condition evaluations to be made for many aspects.

The lack of reliable, efficient, and cost-effective methods for continous and repetitive substructure monitoring additionally hampered the problem. Secondly, ever increasing need and strict regulations regarding track availability, both of which becoming increasingly pronounced during the nineties, made the joint management of all railway infrastructure absolutely necessary. These two reasons, along with the always present strive for higher efficiency and cost reduction, have initiated the appearance of a new management concept, which was usually referred to as Asset Management or Asset Management System (AMS).

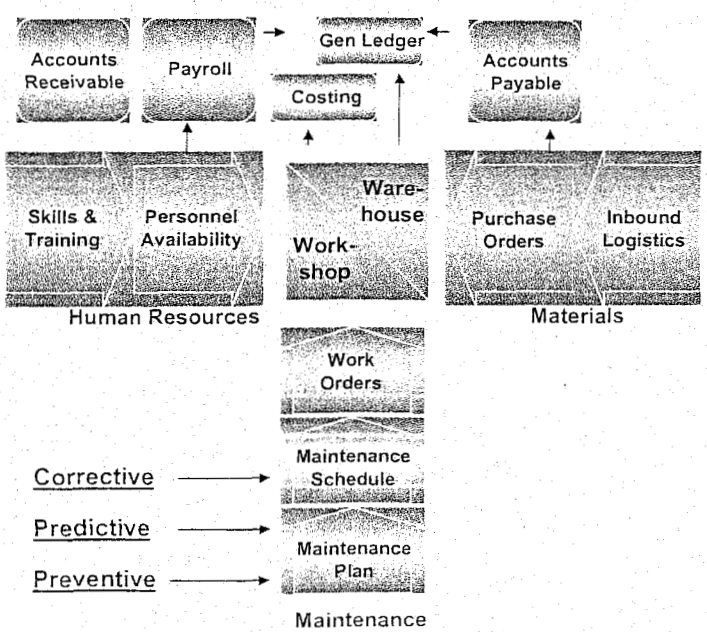


Figure 19.1: Core Elements of an AMS

Within a so defined Railway AMS, previously described (computer aided) track maintenance management system (TMMS) would represent only one of the integrated modules. However, this new position did not limit the abilities of a TMMS. On the contrary, it allowed even broader and more accurate consideration of the problem.

19.1 Railway Asset Management System concept

AMS has gained its importance during the second half of the nineties. In an attempt to achieve a better and more efficient railway system, it was soon discovered that the complexity of interrelations between the management of infrastructure, track availability, rolling stock, work force, maintenance machinery, material, and, above all, budget could not be efficiently resolved without the comprehensive AMS.

At the same time, it should not be forgotten that, for example, European railway networks have a combined length of more than 200.000 km, with a book value estimated at more than EUR 350 billion, and figures of similar significance can also be found in North-America and elsewhere.

This is why a significant rise of interest for such systems worldwide was not surprising at all. Perhaps the best and most recent example comes from the United Kingdom where Railtrack has initiated the development of a large AMS called MIMS, which should serve and be served by not only Railtrack, but also by other maintenance contractors and traffic operating companies. The core elements of AMS are shown in Figure 19.1 [12].

Other basic constitutive parts of an AMS are:

- Railway Assets Locating (usually GIS based);
- Integrated Condition Monitoring and Data Collection;
- M&R Planning and related Decision Support;
- Resource Allocation Optimization (budget, track possession times, labour, machinery, material, etc.);
- Accounting/Reporting;
- (Re-)Investment Planning;
- Overall Data management.

19.2 Development of an AMS

The development of the Railway Infrastructure related part of an AMS is usually divided into two phases. The first phase represents the design and population of the AMS database and definition of the data collection procedures. This, in other words, means the provision of Asset Information. This is an extremely important phase as it serves as a basis for all the other phases and future upgrades. If any flaws are left at this stage, it will have consequences for all following phases.

The second phase represents the formalisation of the deterioration/restoration process(es); i.e. it comprises the actual processing of the data gathered in the first phase. This means the creation of various "rules" that will govern the decision-making process based on the collected data. These processes are characterised by certain parameters, which necessarily have to have certain thresholds such as maintenance, operation, or safety thresholds.

19.3 Railway Assets Locating

Taking a look at the first phase of an AM development in respect to the railway infrastructure, it can be seen that the first important aspect of providing detailed asset information is finding out what the assets are and where they are located. Although it sounds really basic, this has not proven to be an easy task on many railways. Asset documentation was usually paper based, rather old, not regularly updated, inconsistent, and, consequently, unreliable.

The railway infrastructure on many of the world's railways is more than a hundred years old. During these hundred years there have been numerous reconstructions and changes performed both in track layout and track structure and other infrastructure elements. Unfortunately, these changes, especially those performed further in the past, most often escaped being recorded or were poorly documented.

Assets were also usually referred to using milestones which often proved to be wrong or imprecise. These things, along with the retirement of the most experienced engineers who were perhaps the only ones left who knew anything about those changes, necessitated the creation of a consistent and accurate infrastructure location and data collection system. This system had as its aim not only to establish the accurate infrastructure location, but also to capture the existing knowledge and information present among the most experienced engineers while it was still available.

However, since railway networks are often of very large length, ranging from several thousand kilometres up to several tens of thousands of kilometres, surveying of such networks and especially data collection was far from simple. Therefore, new, faster, and more accurate methods were needed. Until now, there have been several methods at different sophistication levels that were able to handle this problem.

19.3.1 Method using ortho-photo technology

The first and perhaps the easiest way to assign the geographical location to the infrastructure elements was by using the ortho-photo technology [140]. The baseline for this technology were the ortho-photo images. Ortho-photo images represent accurate digital raster cartographic products which are obtained after the aerial photographs have been rectified and geo-referenced into a certain orthogonal co-ordinate system.

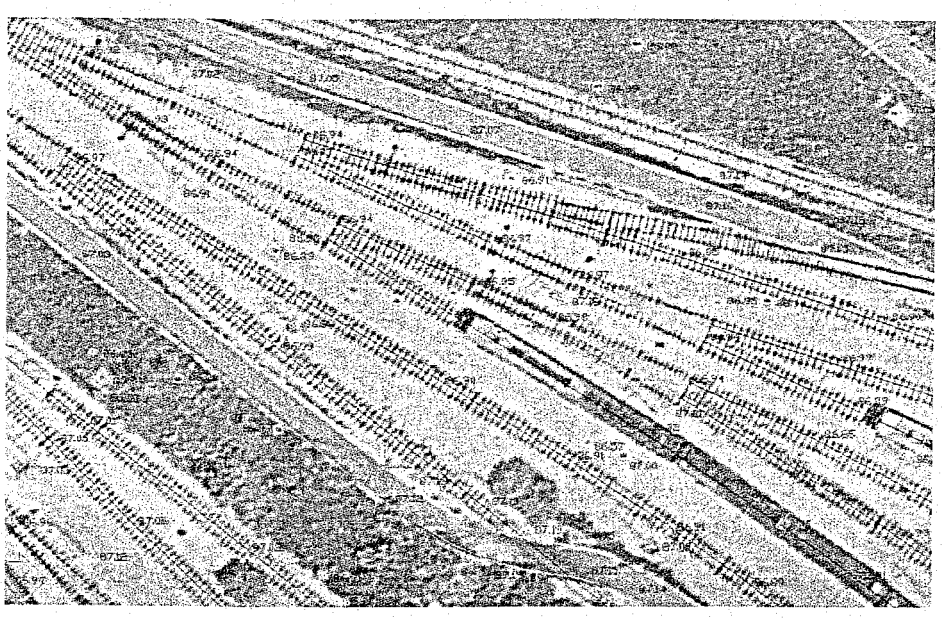


Figure 19.2: Ortho-photo image with overlaid vector track drawing

Rectifying the aerial photograph means removing the deformations of the scanned aerial photographs, both those induced by the camera and those which occurred as a consequence of the irregularities of the terrain. Geo-referencing of an ortho-photo image represents assigning orthogonal co-ordinates to its every pixel. After the ortho-photo images have been created, they are overlaid with the vector images representing the railway infrastructure elements (Figure 19.2). Finally, each of the vector objects is assigned the relevant information.

The essence of this database is the dual referencing system that locates each infrastructure asset along the right-of-way. The system's first frame of reference uses Track Number, Mile post, and Footage locators, based on the mile post system, while the second frame of reference uses Latitude, Longitude, and Elevation (N, E, Z) co-ordinates. This dual system allows the presentation of maintenance management information in both track chart and geographical format as seen in Figure 19.4.

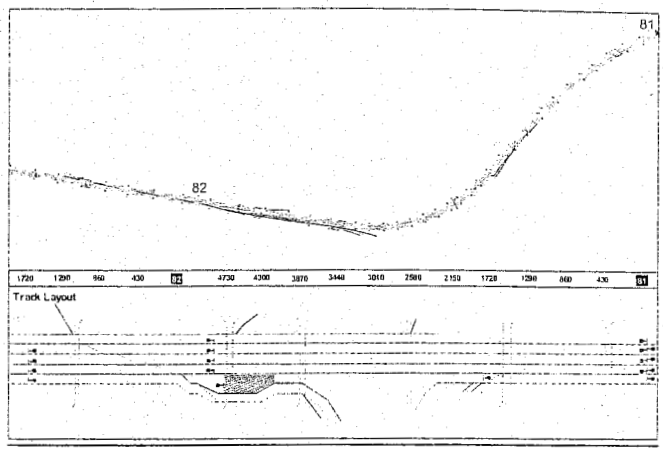


Figure 19.4: Track Chart and Geographic Layout Plot

19.3.3 Video Surveying

A new video surveying system [199] emerged in England in the late nineties. This system was called OMNISurveyor3D and was developed by Omnicom Engineering to be used by one of the largest UK contracting companies, namely Balfour Beatty Rail. OMNISurveyor3D provides a unique method for the high-speed surveying and analysis of complex infrastructures, such as those found on railway networks. It produces a comprehensive relational picture of all visible lineside infrastructure and permanent way geometry which can be maintained in a safe and cost-effective manner. It also enables a 3-dimensional model of the surveyed network and its assets to be created.

The Primary System Components of the OMNISurveyor3D are:

- Positioning Systems:
  - Differential GPS;
  - Inertial Positioning System;
- Digital Imaging:
  - Digital Cameras;
  - Digital Recorders;
  - Multiple Views;
- System Data Logging and Post Processing:
  - Industrial Data Logging Computers;
  - Kalman Filters;
  - Produce MPEG and Position File Library;
- Image Processing:
  - Multimedia PC;
  - Multiple Views 'Point & Identify';
  - 3D Analysis;

- Build Database;
- Printing.

s system allows the network yout to be viewed from differ- angles and measures asset ations in three dimensions, ie population of the existing IS platforms, and exporting h-resolution images of ssets. It also allows identifica- on of the new or existing sets (Figure 19.5) as well as e automatic transfer to asset egisters or asset data bases, ecification of user's details for ch asset, location of access oints in relation to infrastruc- e for maintenance and ewals, etc (Figure 19.6).

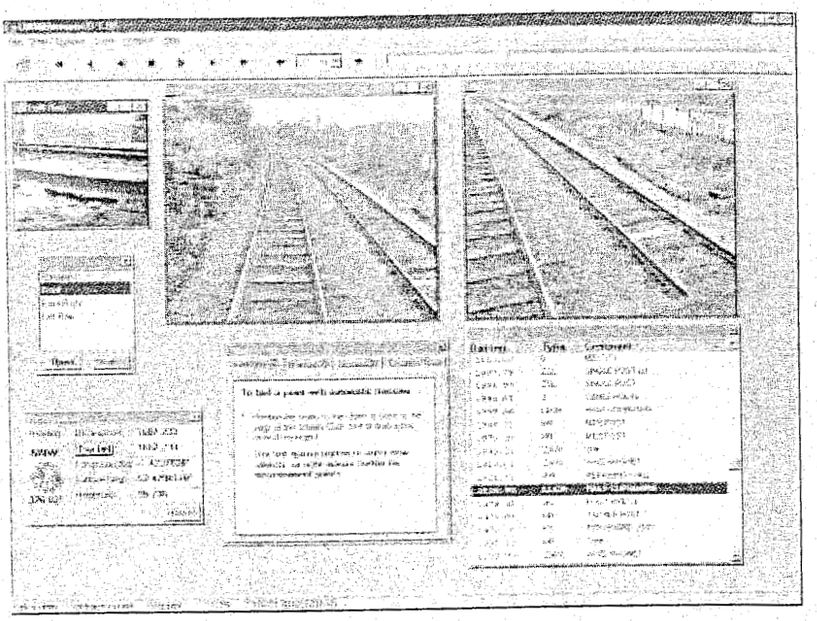


Figure 19.5: Asset Identification in OMNISurveyor3D

his system was found to be ry useful because it was:

- **Safe**: no lineside access was required during the survey and subsequent post processing of the survey data could be completed in an office environment;
- **Fast**: surveys could be undertaken at normal train operating speed;
- **Accurate**: all track and infrastructure could be more accurately surveyed than +/- 2 meters;
- **Re-usable**: images and positioning information were recorded on CD-ROM and were thus easily available for later use or amendment.

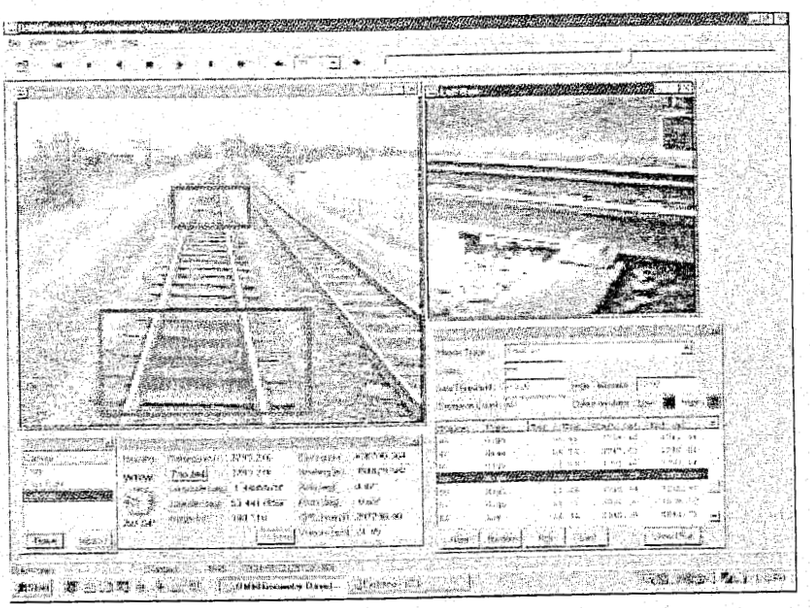


Figure 19.6: Automatic detection and display of locations with exceptions in OMNISurveyor3D

Using the recorded digital mages allowed a permanent isual database of the network to be built up. Moreover, by eviewing these images, further evels of detail about the net- work could be extracted without any need for additional manual site surveys. This basically allowed the network to be revisited and details to be verified without leaving the office.

The distance measuring tools (Figure 19.7) enabled the user to accurately determine the distance between points on the digital images. This meant that the user could measure the distance between trackside objects using a 3D ruler, see the distance between objects in a direct straight line or along the track, and use the measurement data to plan maintenance work, for example, the distance between a trackside cabinet and a signal for cable length, etc.

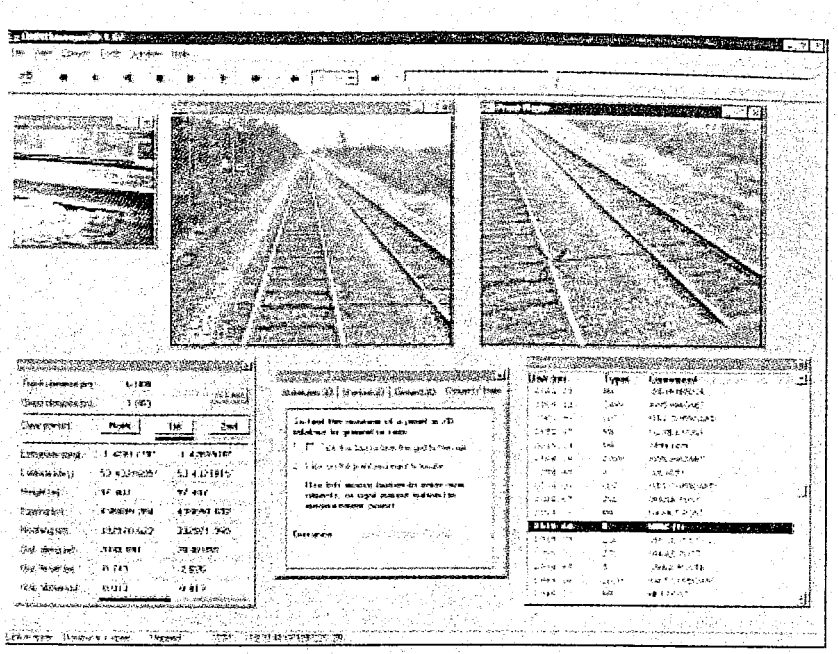


Figure 19.7: Distance measuring tools

Furthermore, the Object Sighting tool (Figure 19.8) enabled the user to place new assets, such as signals and signs, anywhere on the digital images. This tool also enabled the user to create a virtual railway in order to plan different scenarios and view the results, to sight the new object from various driver eye distances in relation to speed and time or distance, to see the line of sight between the driver and the new object and identify obstructions, and, finally, to assess the position and safety of existing signals on the track.

On top of that, the Kinematic Envelope tool (Figure 19.9) enabled the user to place an overlay of a vehicle's kinematic envelope onto the video. This meant that the user could see how the envelope for a particular vehicle would run on the track with real track geometry. It also meant that the user could detect potential collisions between the vehicle and objects on the video such as bridges and signals, or load multiple envelopes and switch between them to assess the difference between the vehicles.

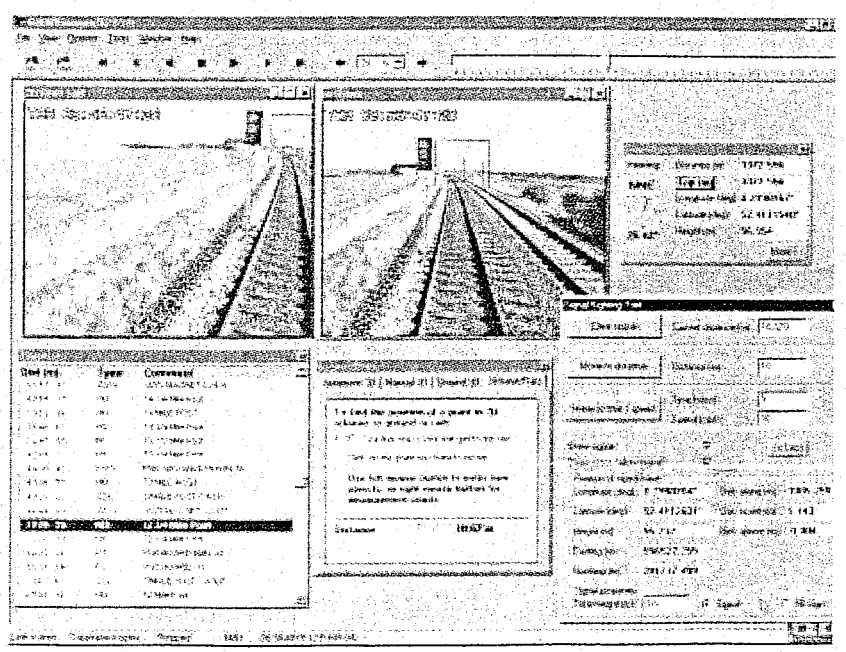


Figure 19.8: Object Sighting tool

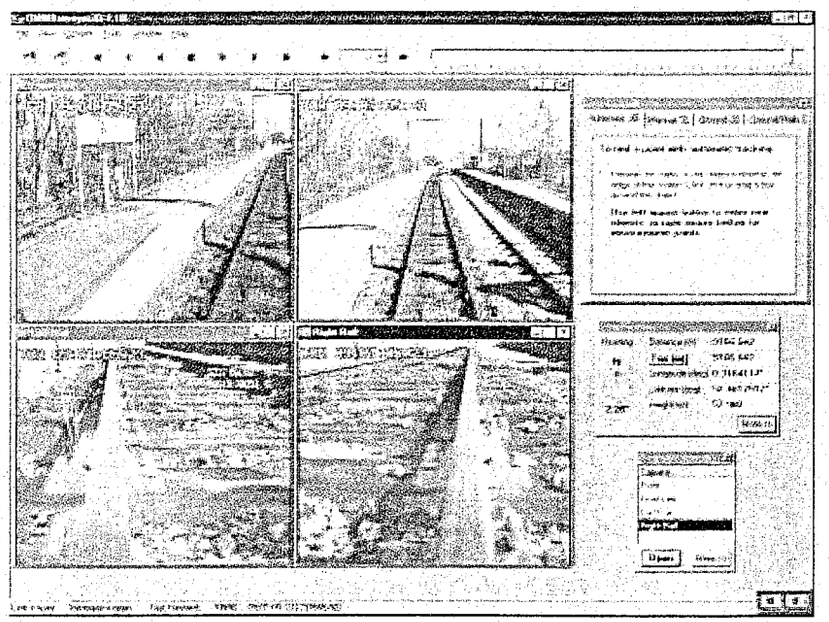


Figure 19.9: Kinematic envelope tool

19.3.4 Method using Satellite Imagery

Another emerging and promising technology to be used for the purpose of locating railway infrastructure elements and data collection is satellite technology [140]. Satellite images (Figure 19.10) have recently not only become more available and inexpensive for civilian purposes, but have also become increasingly accurate. This accuracy is still constantly being improved. Perhaps the maximal capabilities of this technology regarding positional accuracy, currently ranging 1-2 m, are still limited to use in the railway domain. However, as its capabilities significantly improve, which will most certainly happen, this method may soon emerge as the most convenient and cost-effective technology to be used

Figure 19.10: A satellite image

19.4 Integrating a Railway Asset Management System

As mentioned previously, a Railway AMS should include and combine all kinds of specialised monitoring, data collection, and decision support systems. This should be the case for track, but also for all other railway infrastructure elements like bridges, switches and crossings, overhead lines, level crossings, tunnels, culverts, etc. Some of these systems will be given special consideration later on in the text. AMS should also incorporate issues like environmental and hazard management and emergency response systems.

Once integrated, an AMS should serve the needs of all the parties which are in anyway connected to the railway system, such as the infrastructure owners, railway contracting or traffic operating companies, or any other. They all should make use of the AMS, extract the data needed for their everyday or long-term strategic purposes, and also feed the appropriate data back to the system.

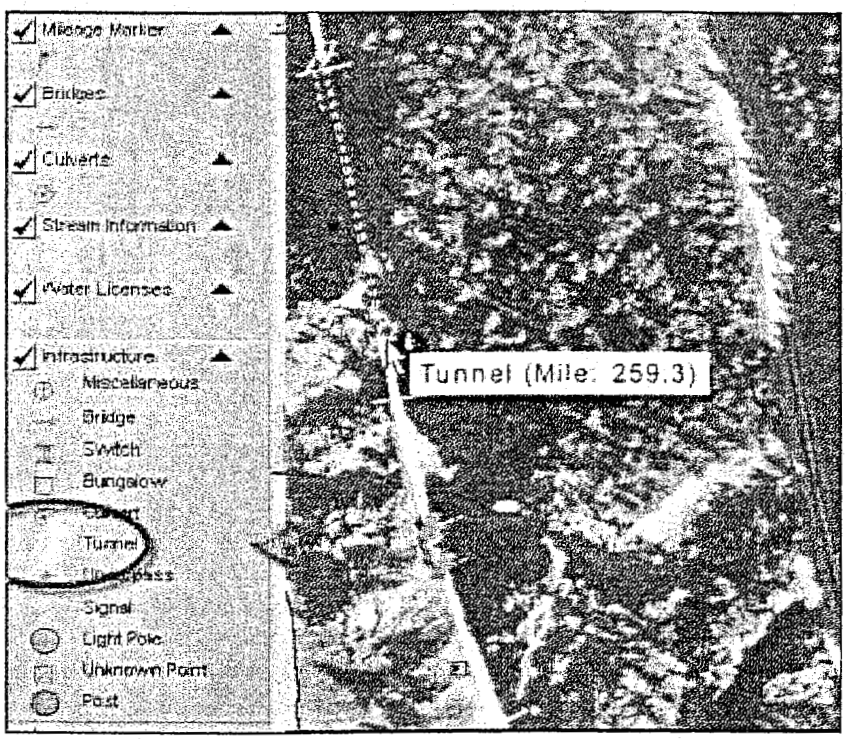


Figure 19.11: GIS background layout of an AMS [62]

The final "look" of the system and some of its features could be similar to those presented in the Figure 19.11, Figure 19.12, and Figure 19.13. This last Figure presents its overall structure. The final "look" should allow for seamless integration between geographical mapping, database management, and multimedia technologies in order to efficiently manage data in today's railway environment. It should also involve such technologies as the Internet to provide worldwide access to all railway information, and enable centralised data management to be performed quickly and easily. It should handle issues like passing information back and forth between various users and their departments, updating data bases, and integrating digital information among the users.

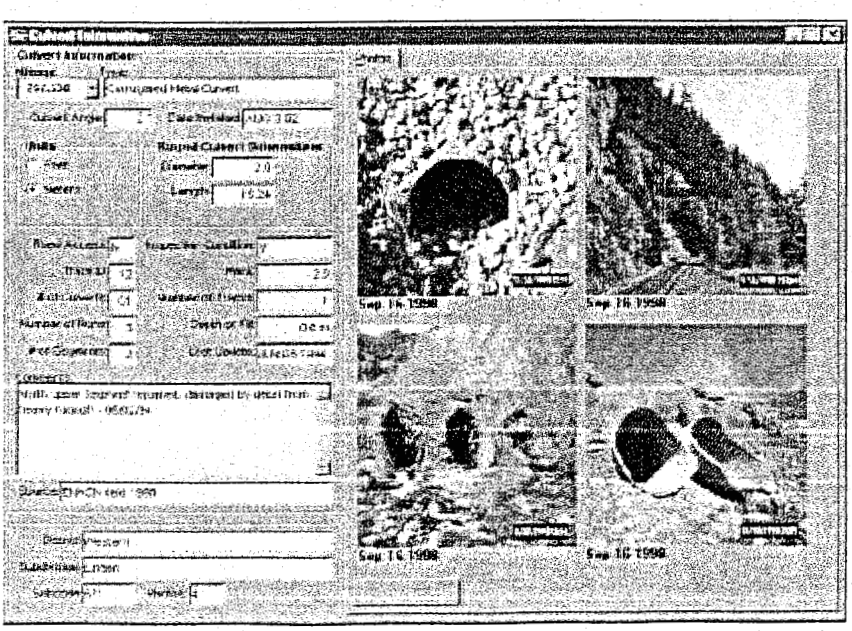


Figure 19.12: GIS background layout of an AMS (2) [62]

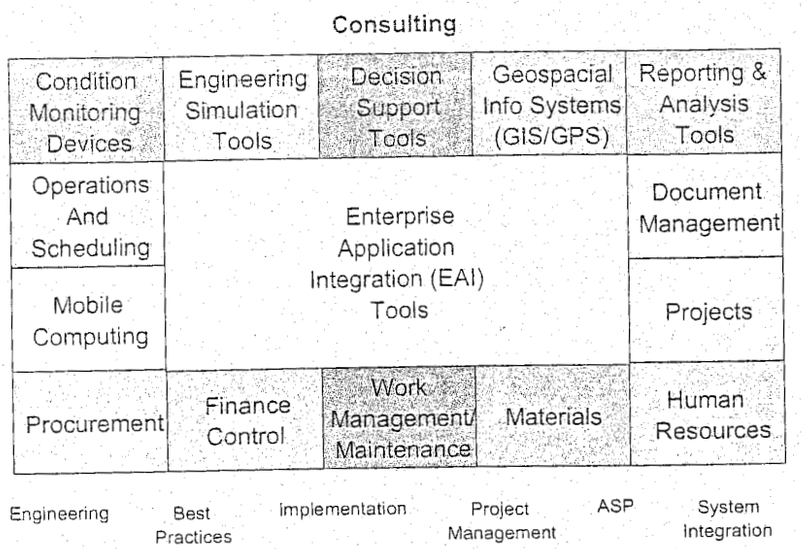


Figure 19.13: Business Performance Framework of an AMS [13]

In an ideal case, AMS should consider all the phases of the assets life cycle from the Planning to the Disposal phase (see Figure 19.14). Of course, in most cases, this can only be achieved with the newly built lines which design documentation is all in place and available, or with lines to be built in future. However, it does not mean that the older lines should not be treated the same way and that all available documentation should not be collected and incorporated into AMS.

19.5 AMS subsystems

An integrated AMS should incorporate several subsystems. All of those subsystems must communicate to each other, exchange input and output data. AMS should serve as a cohesion factor between the subsystems, enabling their proper communication, data exchange and synchronising their work and output. With their perhaps the most important output being M&R work plans, AMS should further optimise those plans in terms of resource allocation, especially regarding track availability, and take care of budgeting and work ordering. Some of the subsystems are stated below, and their position within an AMS presented on the Figure 19.15 [13].

- Track Management System (e.g. ECOTRACK);
- Switches and Crossings Management System;
- Substructure and Drainage Management System;
- Catenary Management System;
- Signalling Equipment Management System.

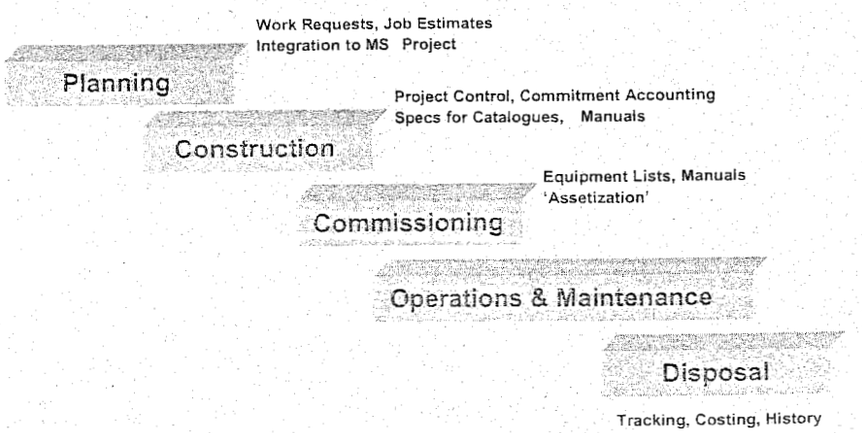


Figure 19.14: Full Asset Life Cycle Management [12]

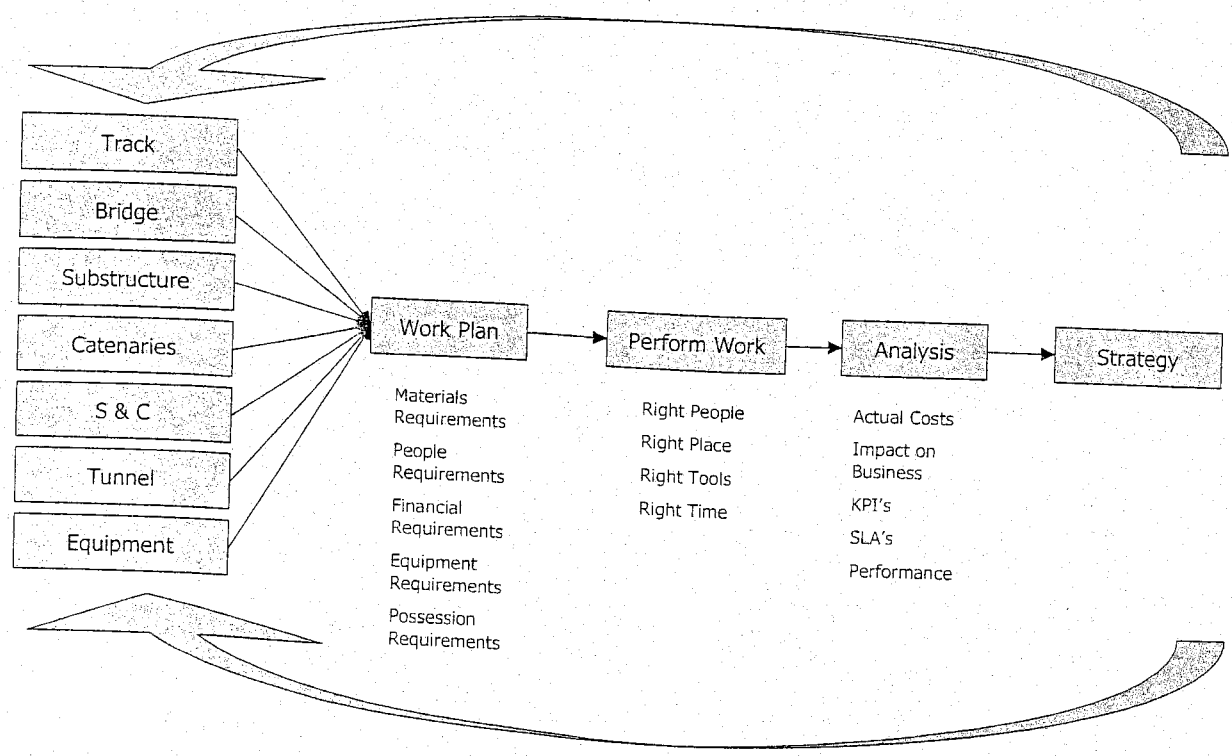


Figure 19.15: Integrated Railway Asset Management

- Bridge Management System
- Tunnel Management System
- Rolling Stock Management System
- Facilities Management System

All of those systems should have in common the condition diagnostic principle, based on extensive and smart monitoring. Some of the most important existing and prospective monitoring techniques are presented in Chapter 16, with the emphasis, of course, being given on the track monitoring.

20 LIFE CYCLE COST ANALYSIS

20.1 Life Cycle Costing

In the 1990s the setting for infrastructure management changed for many railways, especially in Europe. EU directive 91/440 requires that a separate organisation has to provide the railway infrastructure in order to create profit-driven transport operators and transparent cost accounting of infrastructure maintenance and operations [94]. The railway restructuring leads to the introduction of user charging and performance regimes. The operators have to pay access charges for using the railway assets, while the infrastructure manager has to pay penalties in case of unplanned disruption. Decisions in design and maintenance have to be based on estimates of availability, reliability and maintenance costs in order to minimise the total (long-term) costs of ownership for the infrastructure owner.

In the Dutch railway sector three 'change programmes' are initiated since the mid 90s in order to deal with the changed management conditions:

1. *'Life cycle management'* (LCM) aims at the realisation of a systematic approach to underpin and optimize investments in new construction, maintenance and renewal. Costs of ownership, including penalties for track possessions, have to be analysed for a period of 50 years. Since last year the regional maintenance planning staff is obliged to identify the feasible investment and life-lengthening maintenance solutions and to quantify their assumptions on investment and maintenance costs using a special computer application [198].

2. *'Performance-based contracts'* are being introduced for the maintenance and incident management. Contractors with approved quality control systems can acquire this type of contract for periods of 5 years. Their efforts will increasingly be monitored based on agreed performance indicators [254].

3. *'Maintenance window scheduling'* is triggered by more stringent safety demands for maintenance works. Maintenance and renewals will be clustered in periodic maintenance windows (sometimes called slots): the Ministry of Transport will not allow maintenance during operations [110].

The necessity of an Asset Management System (AMS) as described in Chapter 19 for supporting the track maintenance seems to be obvious. For adequate planning insight in the relationships between transport volumes, infrastructure quality, maintenance efforts and availability and reliability *in the long term* is a prerequisite. A Life Cycle Costing approach is presented in this chapter that is able to help develop the AMS concept. Since the implementation of the AMS is in many railways only partly realised, the way to deal with lacking and unreliable data is covered as well in the examples being:

- Appraisal of track designs for the HSL South for an international consortium;
- Revision of track maintenance policy on the Dutch conventional network.

In this Section the general principles of Life Cycle Costing (LCC) are introduced. Section 20.2 contains an outline of a computer application used for life cycle cost analysis of railway track at TU Delft. Two studies are presented shortly in Section 20.3.

20.1.1 Life Cycle Costing principles

Since railway infrastructure, and especially the railway track, has a long life span and investments are very costly, decision makers have to consider the long-term cost impacts in the construction, maintenance and transportation processes. A preventive maintenance regime can for instance postpone renewals and reduce traffic disruptions.

However, it will require more planned track possessions. In order to develop an optimal maintenance strategy the life cycle costs should be analysed. Many questions have to be answered, such as:

- What is the maintainability and reliability of a track design?
- Which cost increases are caused by increased traffic intensities and loads?
- What are the costs implied for train operations and/or the penalties allocated in case of availability-reducing incidents?
- Which construction and maintenance risks have to be covered?
- Which amount of preventive maintenance is needed?
- How to trade off disruption against work efficiency in planning track possessions?

Life Cycle Costing (LCC) is an economic method that supports the decision makers in finding systematically the answers to this type of questions. LCC is defined as an economic assessment of an item, system, or facility and competing design alternatives considering all significant costs over the economic life, expressed in terms of equivalent currency units [256].

The life cycle costs of the railway track can be distinguished in two ways:

1. *Tangible versus intangible costs*. Tangible costs are paid 'out-of-pocket', such as the costs of construction and maintenance (labour, materials and machines). Quality loss, reduction in transport services, a reduced safety and comfort level and noise nuisance cause the intangible, 'hidden costs', of which the exact cost level is not known.

2. *Initial (capital) costs versus running costs*. The initial costs are the costs made for acquisition and installation or construction. The running costs are made during the operational period of the railway. A further distinction can be made between 'annual costs', such as inspections and small maintenance, and 'intermittent costs', such as major overhaul and renewals. The initial capital costs and the running costs for the first year are called the 'full year effect costs'.

Since the cash flows occur during different operational years of the railway line, they are discounted to the year, in which the decision is being made ('year zero'). Usually the real interest rate is used. With this discount rate the interest payments and incomes during the period of construction and operation are estimated.

Three criteria are used for comparing the investment and maintenance strategies:

1. The *net present value (NPV)* is the sum of all discounted cash flows. In the LCC method it concerns mostly costs. The larger the total present value (TPV) of the life cycle costs is, the more unattractive the investment is compared to alternative investment or maintenance strategies.

2. The *internal rate of return (IRR)* shows the profitability of an investment compared to alternative investment or maintenance strategies.

3. The *annual equivalent or annuity (ANN)* is the sum of interest and amortisation, which has to be paid every year to finance an investment or maintenance strategy. The annuity is therefore the most meaningful indicator. With the annuity projects with different life spans can be compared.

The annuity of any investment and maintenance strategy can easily be calculated with the formulas presented in Table 20.1. In Box 1, on the next page, a simple example of the annuity calculation is presented.

Nr.	Formula	Explanation
(20.1)	$TPV = \sum_{y=0}^{n} \frac{C_t}{(1+i)^t}$	The total present value *TPV* is the sum of the discounted costs *C* during all the years (*y*) analysed. Year *n* is the last year, i.e. the time horizon of the analysis. The interest rate applied is *i*.
(20.2)	$v_{y,i} = \frac{1}{(1+i)^y}$	This multiplication factor (*v*) is used to convert a cost of 1 Euro in year *t* into its present value using an interest rate *i*. Used for intermittent costs.
(20.3)	$a_{n,i} = \frac{1 - \frac{1}{(1+i)^n}}{i}$	This multiplication factor (*a*) is used to convert a cost of 1 Euro occurring in all the years analysed into its present value using an interest rate *i*. Used for (annually) constant running costs.
(20.4)	$ANN = \frac{(1+i)^n \bullet i}{(1+i)^n - 1} \bullet TPV$	This formula is used to convert the total present value of the investment or maintenance strategy into the annual equivalent or annuity (*ANN*). This is the fee needed to finance the strategy during the years analysed.

Table 20.1: Formulas applied in life cycle cost analysis

For the ex-ante assessment of cost impacts during the construction, maintenance and operation of the railway, input data has to be collected from various sources[97]:

- *Supplier information.* Information from suppliers can be used for composing the likely construction costs, especially for innovative designs. A lot of data from previous tenders is available at the Construction Department of the infrastructure manager.

- *Maintenance history.* The history of infrastructure maintenance, failures and quality for the different track assets should be known at the Maintenance Department. Besides, the Maintenance Department can provide information on cost prices and productivity rates for maintenance and renewals in relation to the duration of track possessions. Finally, the history of the traffic loads on the track can be obtained there as well.

- *Simulation and test data.* For innovative asset types the usage of data from simulation models and laboratory, and if possible, field tests is unavoidable.

- *Expert judgement.* Finally, in most cases judgements from maintenance experts are one of the most important sources. Especially, since the maintenance records of the infrastructure manager are often (still) poorly usable.

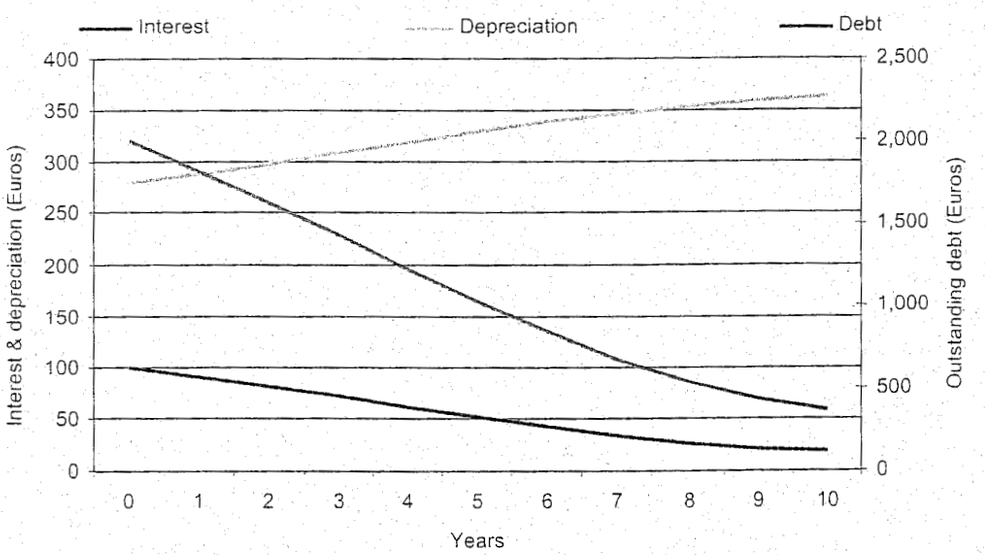
Box 1: Illustration of a simple life cycle cost plan

Consider the following example: an investment in a track component is being made, which will cost 2.000 Euros and will be maintenance-free during the first year. In the following four years the maintenance will cost 100 Euros; from the fifth year on the costs will increase annually with 20%. After eleven years the asset will be replaced and a new cycle starts, which is not considered here.

Using a 5% interest rate a depreciation plan can be made based on the annuity method. First the total present value is calculated using formula (20.1) and eventually with the help of formulas (20.2) and (20.3). The annuity, which has to be paid during eleven years, is calculated with formula (20.4). The annuity is the basis for the depreciation plan. The investment will lead to a debt of 2.000 Euros, which will cost 100 Euros interest for a rate of 5%. This means that in the first year 281 Euros remain available for depreciation of the track asset. Etc.

Year	Cost	Present value	Debt (Jan 1st)	Annuity	Interest	Depreciation
0	2,000	2,000	2,000	381	100	281
1	100	95	1,819	381	91	290
2	100	91	1,629	381	81	299
3	100	86	1,430	381	72	309
4	100	82	1,221	381	61	320
5	120	94	1,021	381	51	330
6	144	107	835	381	42	339
7	173	123	669	381	33	347
8	207	140	529	381	26	354
9	249	160	424	381	21	360
10	299	183	363	381	18	363
Totals	**3,592**	3,163	0 *(yr 11)*	4,189	597	**3,592**

Depreciation plan based on annuity method



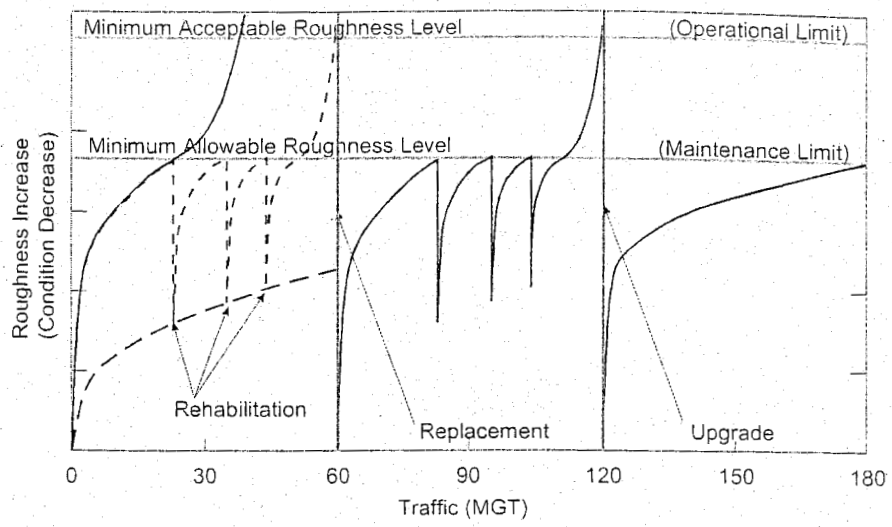


Figure 20.1: The typical saw-tooth curve of track quality decline

The most uncertainty in the life cycle cost analysis is the forecasting of track quality decline. A typical decline pattern is depicted in Figure 20.1. The decline depends on all kinds of factors, such as the initial quality of construction, the quality of the substructure (e.g. settlements and crushing of the ballast bed), and the loads on the track. Generally the effectiveness of maintenance works diminishes after a couple of times and larger overhaul or complete renewal is required. Data from an Asset Management System, like ECOTRACK and TOSMA, provide insight in the actual decline rates and the effectiveness of maintenance activities on specific lines [223], [56].

Due to the unavailability of historic data and indisputable renewal thresholds, uncertainty should be well dealt with in any life cycle cost analysis. Two methods to include uncertainty are [97]:

- *Sensitivity analysis.* In a sensitivity analysis systematically the input values are varied, for instance with percentages of 10%, 20% and 30%, in order to test the robustness of the outcomes. With the sensitivity analysis the most sensitive input parameters can easily be identified by the deviation in the annuity. It is important to get a sound estimation on the likely value for these parameters.

- *Monte Carlo simulation.* The disadvantage of a sensitivity analysis is that usually only one parameter at a time can be varied. A Monte Carlo simulation is a more advanced method, in which all factors at the same time can be varied. It requires more detailed input: not only a most likely value but also a minimum and a maximum value for an input parameter has to be provided. With the help of a random generator a statistical (normal) distribution of the annuity can be calculated, which gives an insight in the chances on a specific annuity. For a life cycle cost analysis this method requires too much effort, except if large amounts of money are involved (construction projects).

In the next paragraph the structure of a Decision Support System is described, which supports the life cycle cost analysis of track designs and maintenance and renewal (M&R) strategies. The DSS provides a format for systematic data collection and has the ability to perform scenario and sensitivity analyses.

20.2 Track Life Cycle Cost DSS

Many factors influence the costs of railway track. In the first place the quality of the track asset (lifetime expectancy and failure chances) and the maintenance strategy applied (such as the amount of preventive maintenance, the stock policy for spare parts and the number of incident repair teams). Further, a number of external factors influence the cost level [301]:

- *Construction and maintenance conditions*, e.g. market prices for labour, materials and machines, safety and noise regulations, the performance quality, and the duration of maintenance windows;
- *The financial environment*, e.g. budget limits, budget allocation between investments and maintenance, the (performance) payment regime, and interest rates;
- *The physical environment*, e.g. quality of the substructure, alignment, climate, and site accessibility;
- *The transportation concept*, e.g. lay-out of the tracks, distribution of train speeds and axle-loads, quality of the rolling stock, traffic intensities, and braking and acceleration capabilities.

To deal with the variety of cost factors an MS Excel based Decision Support System (DSS) has been developed at TU Delft, which has been applied in five studies since 1997 [302], [303], [304], [305], [306]. It has an interface with screens for input, sensitivity analysis and output and a hidden model-base; a DSS chauffeur operates the system.

The most important feature of the DSS model-base is the relationship established between track *use* and *maintenance* (see Figure 20.2). The infrastructure use determines to a large extent the maintenance need (e.g. traffic intensities, train weights, axle loads), while it also determines the impact of track failures (e.g. cumulative train delays and cancellations) and the slots available for maintenance. The output delivered by the DSS is the *annual performance fee*, which is the annuity to be paid by the infrastructure owner for the construction and maintenance of the tracks and switches. This *annual performance fee* is one of the most decisive indicators for the decision-makers. However, also intermediate output, such as asset availability and number of work shifts, is shown. The calculation takes place in a number of steps, which are described below. The figures are only meant to illustrate the calculation steps.

1. Timetable analysis

The starting point of the estimation is an analysis of the transportation forecasts, the so-called *'reference timetable'*. This timetable contains the expected number of trains detailed to different train types. It includes a growth scenario and data on train weights and number of train sets. A (notional) tonnage can be calculated and, besides, the annual scheduled journey time is determined. In some performance regimes, such as for the HSL South, the penalty rates are based on an availability level. The annual scheduled journey time is used as a basis for the calculation of the availability level. The infrastructure manager of the Dutch conventional network uses in its analyses a penalty rate simply depending on the number of hours of planned or unplanned track possession. Other information to be extracted from the reference timetable is the regime for maintenance windows and speed restrictions for track failures and maintenance.

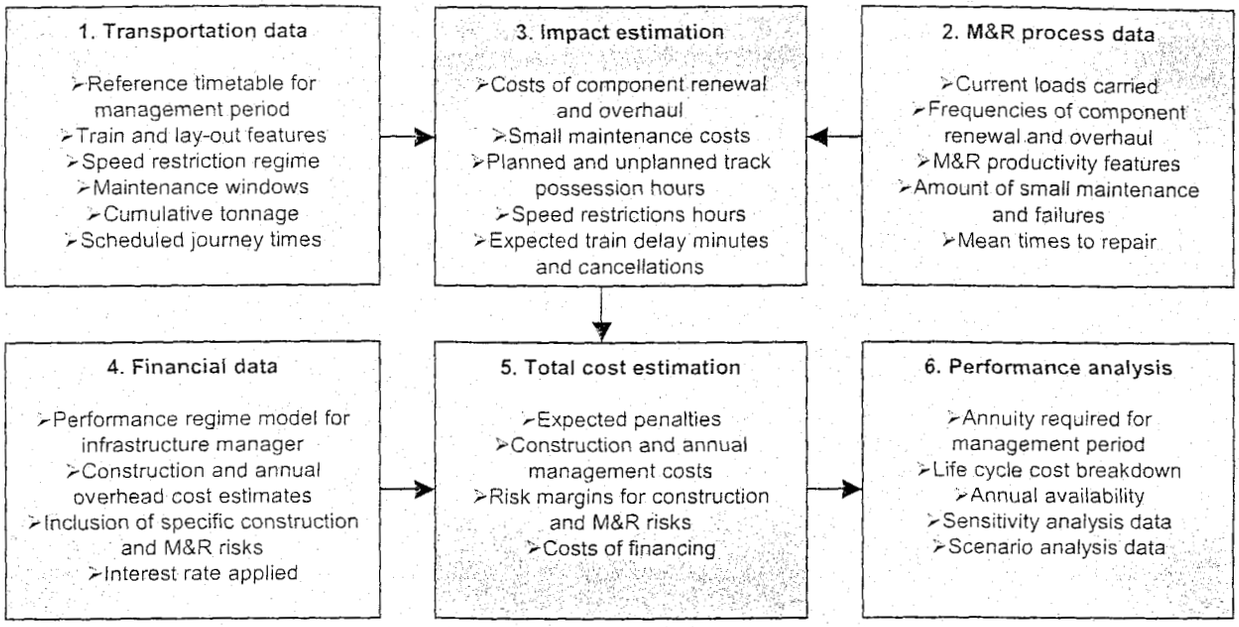


Figure 20.2: Life Cycle Cost DSS model-base

2. Maintenance and renewal analysis

A second starting point is the collection of key data on the maintenance and renewal processes for the track and switch components. First, data on the lifetime expectancy and overhaul and renewal thresholds is needed. Secondly, data on the overhaul and renewal methods is needed, such as the production speed in relation to the duration of a track possession (work efficiency, start-up and finishing time) and the costs per kilometre and per work shift. Besides, data on the maintenance and failure repair process is needed, such as the response and repair time and the annual small maintenance and inspection cost (per ton or per year - per km).
The information should be available in so-called 'Maintenance Concepts' which contain all the information related to the maintenance of a specified asset e.g. for 'a switch 1:9 UIC54 wooden sleepers, depreciation group UIC-3'. If these concepts are not available, a Failure Mode Effect Analysis (FMEA) should be organised. With the FMEA method the knowledge of design and maintenance experts is systematically used to identify risks and develop a maintenance plan [16]. In the DSS a number of specific risks can be selected for innovative track structures, such as the chance on cracking of the concrete supporting bed or early deterioration of elastic materials in case of slab tracks.

3. Impact estimation

By combining the M&R and timetable data the cost impacts during the maintenance management period can be forecasted. First, the volume of renewals and major overhaul, such as tamping and grinding, is scheduled. A number of years for completion of the activity can be included, which depends on the available work capacity and maintenance windows. Secondly, the number of work shifts needed in the specific years is estimated using the data on productivity rates and the duration of maintenance windows (see Figure 20.3). Thirdly, the costs and speed restriction hours due to renewal and overhaul in the specific years is set using the cost rates and speed restriction regime i.e. the number of days with a specified speed limit.

More or less the same calculation takes place for the small maintenance and failure data, except that the estimates are extracted from the FMEA or Maintenance Concepts. The maintenance costs and unplanned track possession time is simply related to the cumulative tonnage or service years of the asset (see Figure 20.4).

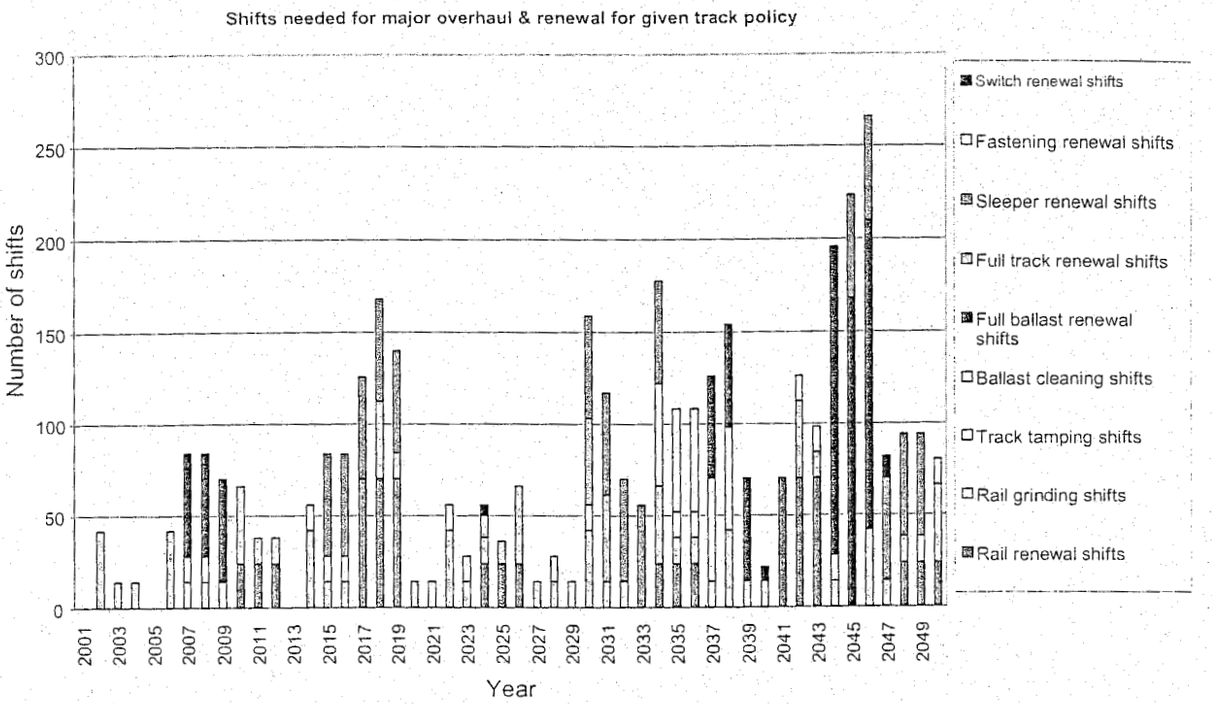


Figure 20.3: Planned work shifts

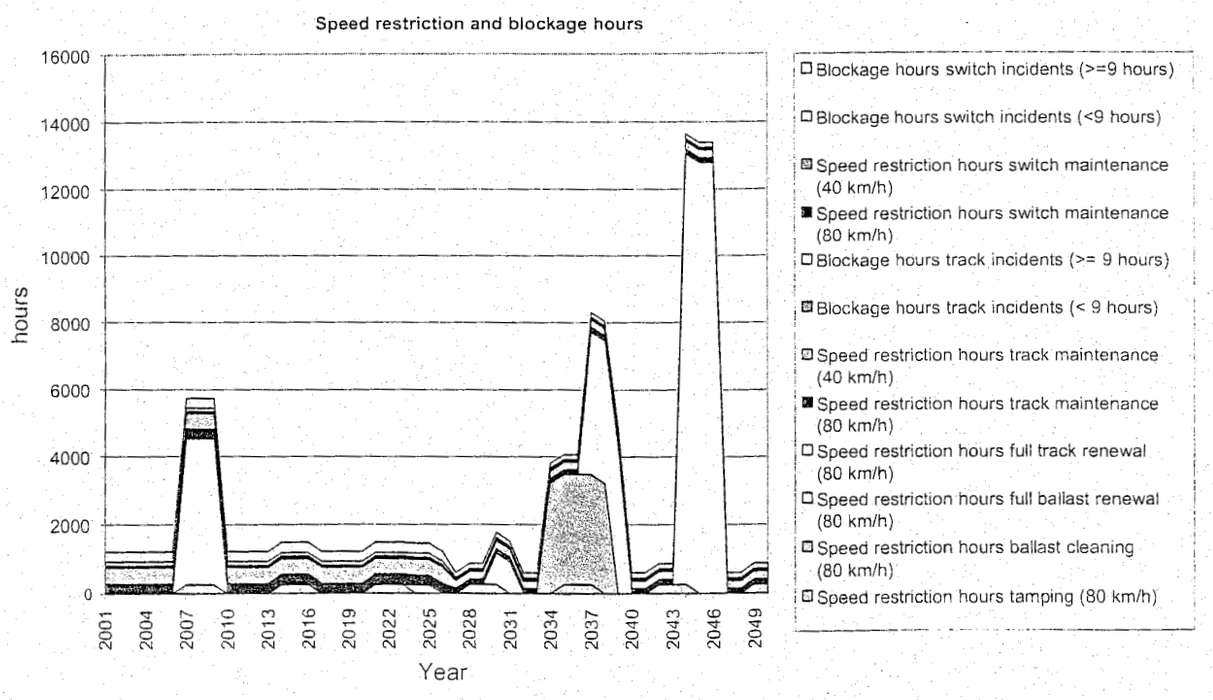


Figure 20.4: Planned and unplanned disruption

Depending on the details of the Performance Regime between the infrastructure manager, transport operators and infrastructure owner, i.e. the central government, a final sub-step is the conversion of speed restriction and possession hours to expected train delay minutes and cancellations. In the DSS a simple estimation model is included, based on the acceleration and deceleration performance of the rolling stock and a number of assumptions on the impact of a speed restriction and full track blockage. A more advanced analysis of the knock-on impacts on later scheduled trains and the use of cross-over switches should be done with a simulation model using for instance Simple++ or RailSys [240].

4. Financial data

The financial data first concerns the construction cost estimate for the track design analysed (in case of new construction) and, eventually, an annual overhead cost for the maintenance organisation. Further, a choice whether to include or exclude the specific risks in the estimation and the interest rate and period of depreciation can be set.
An important part of the financial data is the performance regime, which is the penalty allocation algorithm and penalty rates applied for the infrastructure management. The penalties are used to reflect the extra costs or revenue loss caused for the transport operators in case of failing assets.

5. Total cost estimation

In this step the maintenance and failure expectation is combined with the financial conditions. In this way the expected penalties, costs of construction, maintenance, renewal, and financing over the years are calculated (see Figure 20.5).

6. Performance analysis

Finally, the annual performance fee (the annuity, ANN) needed by the infrastructure manager is calculated based on the cash flows during the management period. Besides, an interface for scenario and sensitivity analysis is provided in order to test the sensitivity of the outcomes to changed conditions. The conditions to be varied are in fact the data from steps 1, 2 and 4.

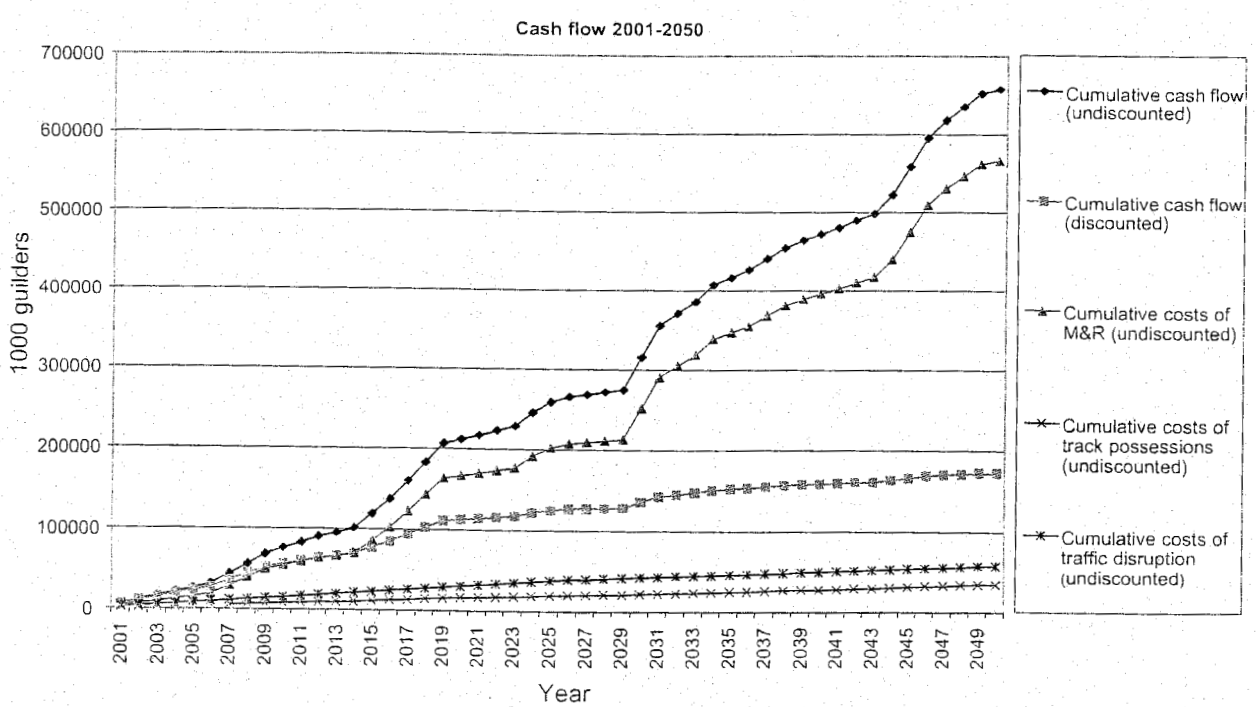


Figure 20.5: Life cycle cost breakdown

In Table 20.2 the basic formulas are shown needed to schedule maintenance and overhaul over the years and to estimate the renewal cost impact for a track or switch component excluding the costs of traffic disruption.

Nr.	Formula	Explanation
(20.5)	$RQ_{y,a} = Q_a \bullet P_{y,a}(T_f \geq TH_a)$	The renewal quantity *RQ* in a specific year (y) for a specific track asset or component (a) is determined. The quantity *Q* is for instance the total track length or number of switches. *P* is the percentage of the asset(s) to be renewed given the fact that the notional tonnage (T_f) has passed the threshold (*TH*).
(20.6)	$S_{y,a} = roundup\left(\frac{(RQ_{y,a})/(PS_a)}{TPP_y - L_a}\right)$	The number of hours needed for the renewal is determined by the production speed (*PS*). The hours have to fit in the track possession periods (TPP) provided. For each TPP also a amount of time lost for production *L* has to be extracted due to preparation and finishing of the shift. The total number of shifts *S* results (whole number).
(20.7) 7	$C_{y,a} = SC_x(TPP_y) \bullet S_{y,a} + RQ_{y,a} \bullet (MC_a - RV_a)$	The costs for asset renewal are calculated by multiplying the number of shifts with the costs per shift *SC* for the given duration of the track possession and by adding the costs of materials (using the unit costs *MC* and the residual value *RV* if a re-use site is available for the replaced materials).
(20.8) 8	$TPV = \sum_{a} \sum_{y=0}^{n} \frac{C_{y,a}}{(1+i)^t}$	The total present value is determined according to formula [20.4].

Table 20.2: Simplified formulas for renewal cost calculation

The formulas from Table 20.2 are simplified; in reality thresholds for renewals can be interdependent: renewals can be harmonised in time and place (clustering of renewals) and components (partial or full track renewal). Besides, it is well possible that a certain renewal activity consists of different subactivities, such as loading and unloading of rails, sleepers, replacement of old rails, tamping, etc. The production speed and preparation time varies for the subactivities, which can impact the total performance for different durations of the track possession. Finally, also the renewal can be spread over a couple of years depending on the renewal length, track possession regime and available work capacity. In the DSS an optimal degree of detail is used depending on the requirements of the life cycle cost study.

In the DSS it is possible to analyse a track composed of assets with various lifetime expectancies - to a manageable level, since the intention is not to replace systems like ECOTRACK.

The DSS assists the decision maker by showing the total cost implications of changing design and maintenance parameters. It is not an expert system, such as ECOTRACK, which means that a team of design and maintenance experts has to provide judgements. A standard data checklist and a process of workshops for data validation support the analysis. The strength of the DSS itself is especially the possibility to investigate a range of scenarios of future conditions in order to test the robustness of the life cycle cost estimates: the impact of different transport and financial conditions can easily be shown. In this way it not only supports the decision making on design and maintenance strategies, but also the negotiations on risk allocation and performance regimes.

Especially, in case of lacking or unreliable maintenance data, this can help decision makers in making the best possible choice under the given uncertainty. Besides, the analysis can be performed quickly. Expert systems alone are usually not fitted to a setting where strategic choices have to be made rapidly and a lot of maintenance data is unavailable.

20.3 Recent studies

In this paragraph two life cycle cost studies in respectively design and maintenance of railway track are presented for illustration. In the first study the DSS has been used for an evaluation of different high-speed track designs for an international consortium (1999) [304]. In the second study the DSS is used to support the revision of the maintenance and renewal policy on the Dutch conventional network (2000 and 2001) [306]. The precise results of the first case are confidential, while the second case is still in a finishing phase. However, the cases illustrate very well the applicability of LCC.

20.3.1 Track design for a high-speed line

The construction of the Dutch high-speed line, HSL South, which is planned to be in service in 2005, is tendered in seven contracts. The most challenging contract is the so-called Infrastructure Provider (IP) Contract. The IP is to construct the rail system: the power supply, signalling and track systems as well as auxiliary (safety) equipment and some sound barriers.

During a period of 25 years the IP has to maintain the assets, after which a guarantee period of 5 years is applicable for backlogs in maintenance. Other contractors will produce the 'settlement-free' concrete substructure. One of the important issues during the Consultation Phase in 1999, has been the selection of a railway track structure, especially since the Dutch State put forward that slab track systems should seriously be considered in the tender [127].

In the engineering team of the Consortium several promising track structures were selected, such as ballasted track, the Rheda system, Shinkansen slab track, a direct fastening system and the Embedded Rails Structure (ERS).

Except for the ballasted track, data on most of the track structures had to be collected through an international investigation and expert judgements. For this reason a Failure Mode Effect Analysis was organised for all the track structures using the ballasted track as a reference. In the FMEA all probable failure causes, modes, measures for prevention and monitoring, and their effects were identified.

For an identification of the effects a distinction was made between immediate failures and loss of quality, which could be restored during the nightly maintenance windows. Failing equipment causes a speed restriction or full track possession; some maintenance activities also require more restoration time than the available maintenance window or require speed restrictions for a number of days. Finally, an expectation on the frequencies of incidents for the given M&R strategy resulted.

During the FMEA the DSS was modified in order to include the Performance Payment Mathematical Algorithm of the HSL South. In the second part of the Consultation Phase the DSS was applied in order to analyse the expected performance and the consequences of for instance different interest rates, construction methods and (longer) maintenance windows. The outcomes were used to support discussions with the State.

An important finding was the sensitivity of the outcomes caused by the penalty regime. If in a specified 'performance period' the train delay minutes increase, the penalty rates increase as well, which causes a progressive trend (see Figure 20.6). This made it also necessary to get insight into the influence of the signalling and power supply systems on the total system availability.

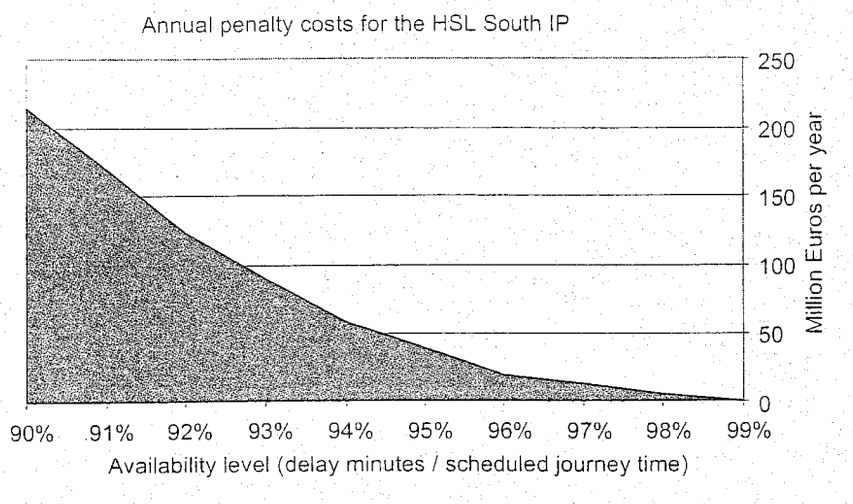


Figure 20.6: Indication of penalty rates (indicative impacts)

Further, the randomness of failures can impact the total amount of penalties. Insight in the distribution over the years of the failures was needed, but this proved to be very hard to get a grip on. The performance regime remained therefore an important issue during the tender.

The output in Figure 20.7 and Figure 20.8 gives an indication for a scenario with ballasted track without a ballast mat and a rather high interest rate. The low construction costs and the large amount of maintenance in the last years of the IP contract prove to have a beneficial impact on the annuity-based performance fee. The absence of a ballast mat was however not assumed to be realistic, which changes the picture a lot. Besides, in a later phase of the tender it was chosen to develop a performance fee, which reflected the actual cost level instead of a 'flat fee' over the years.

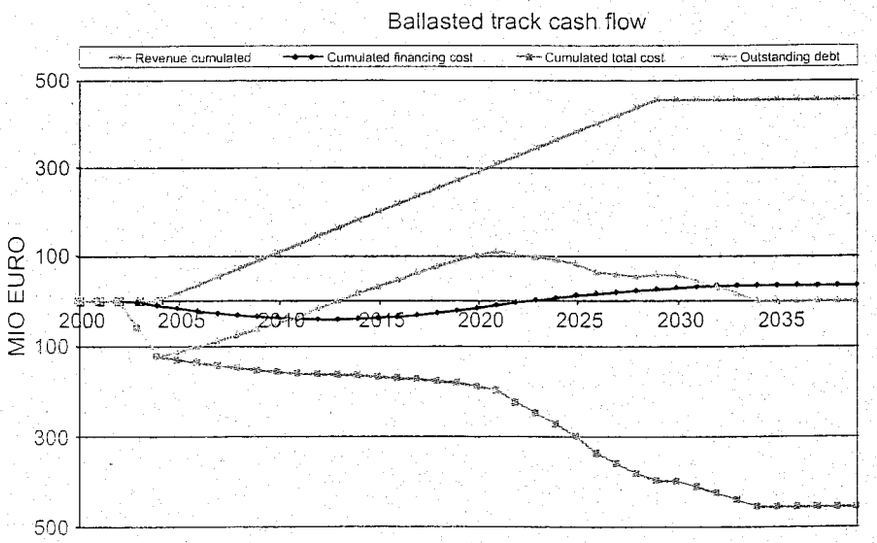


Figure 20.7: Cash flow per track structure (indicative output)

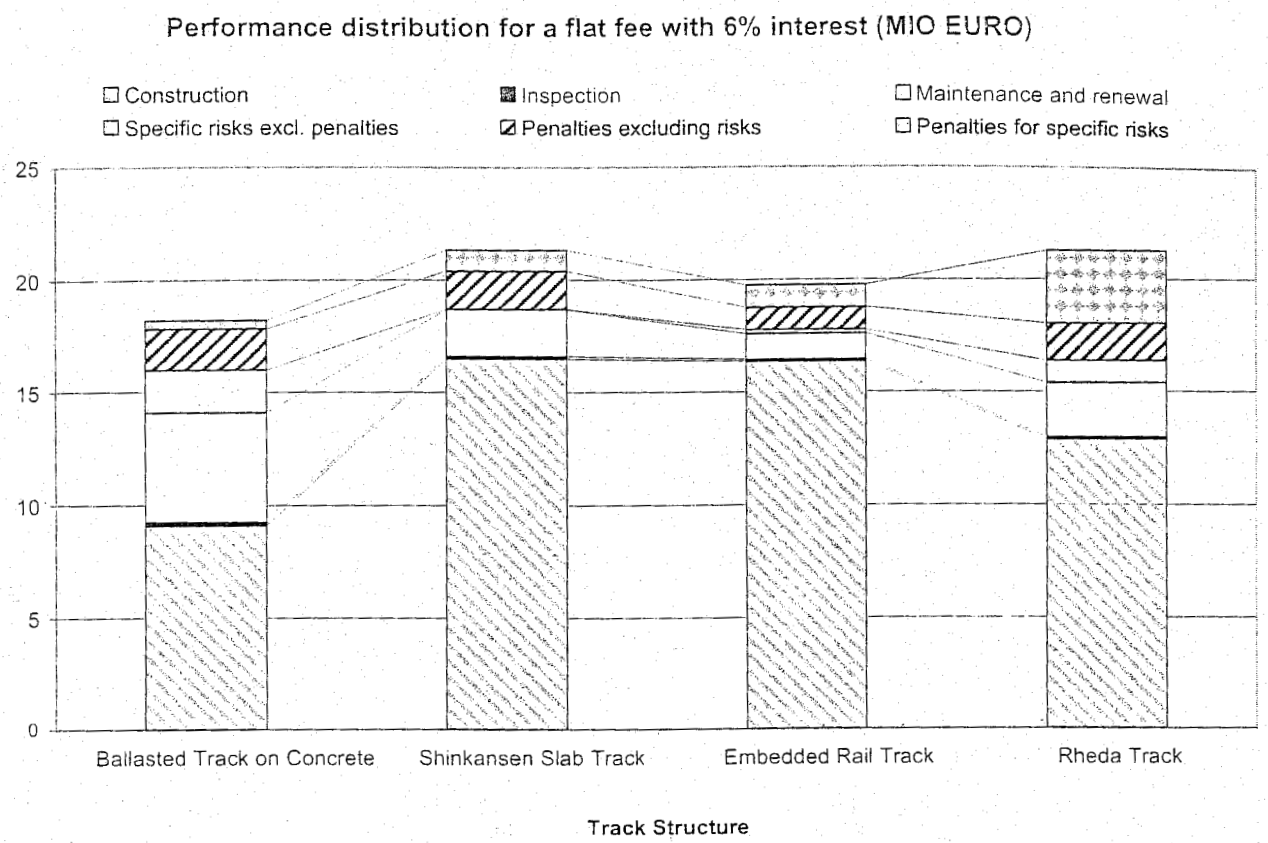


Figure 20.8: Performance fees for one possible scenario (indicative output)

The output proved also to be rather sensitive to changes in the thresholds for renewal and the duration of maintenance windows. For the ballasted track the maintenance proved to increase by the end of the contract period. For the embedded rails the renewal of rails was an issue due to the very short nightly maintenance slots available. Both the stringent performance regime and the short depreciation period (at least for a track structure) were disadvantageous for some of the innovative track structures in the Tender Phase.

20.3.2 M&R strategies for tracks and switches

In 2000 a project named LCM+ Tracks & Switches was initiated by the management of the Dutch infrastructure manager, *Railinfrabeheer*, for the revision of the maintenance strategy for tracks and switches based on an extensive life cycle costing analysis. The immediate reason is the foreseen strong increase in planned renewals in combination with the stringent conditions for track possessions.

A task force from the Headquarters and the four Maintenance Regions realises the revision. For this reason ten pilots have been selected in different parts of the Netherlands that reflect the most important cost areas on the main network and regional lines. Most of the pilots concern renewal projects, which are preliminary planned for 2002. The Track Life Cycle Cost DSS is being used for the central collection of the data and the estimation of life cycle costs for each pilot. The regional track managers are responsible for the maintenance estimates provided, which are discussed in central meetings. The DSS is being used to test the robustness of the outcomes in order to analyse the applicability in other conditions, elsewhere on the conventional network.

The measures for cost-saving which are considered most promising are [306]:

1. *Track design based on required functionality.* In some regional tracks an upgrade to the standard track structure is probably not needed, can be postponed or can be realised with a less expensive design.

2. *Life-lengthening maintenance and overhaul.* Renewals can be postponed in a controlled way through temporary life-lengthening measures. In order to guarantee the long-term quality of the rail network this should only be done for realising the fifth measure.

3. *Partial renewal of tracks and switches.* Another measure to improve the asset utilisation in a controlled way is partial renewal: instead of full track renewals partial renewal of sleepers, ballast and rails can be advantageous.

4. *Instant re-use of tracks and switches in side-tracks.* Especially in construction and renewal projects on the main tracks switches and track components become available for re-use on less loaded side-tracks. This should be realised through a smart, timely planning of the projects in relation to the maintenance needs on the side-tracks and terminals.

5. *Smart combination of renewals.* In case of larger renewal lengths large savings can be realised, especially in combination with the sixth measure.

6. *Smart scheduling of renewals and overhaul.* A longer duration of track possessions (40 to 100 hours) can lead to a drop in renewal costs. State-of-the-art renewal methods can perform the track replacement while on the other track bi-directional traffic services can be provided.

As an example the indicative results for one of the pilots is shown in Figure 20.9. It concerns a switch at a railway yard near Amsterdam.

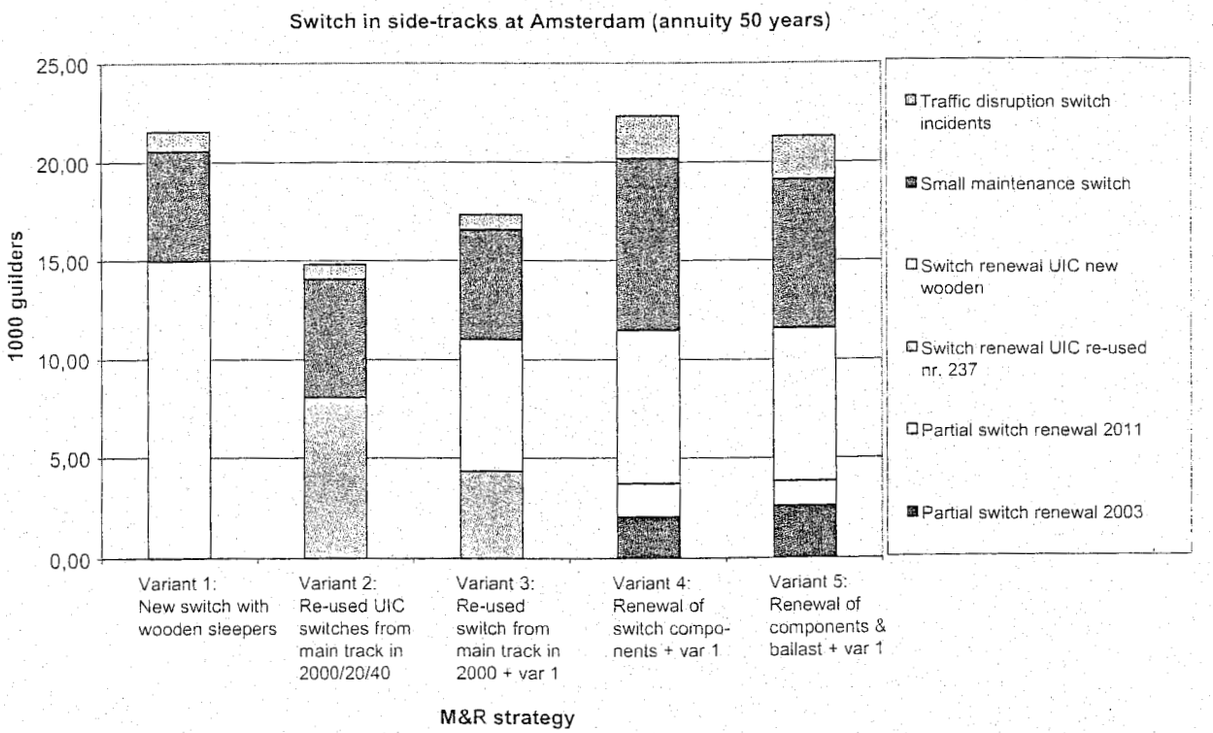


Figure 20.9: Example of an LCM+ life cycle cost analysis (indicative output)

Thanks to the early consideration of the maintenance needs on the railway yard, it has been possible to re-use a switch on the main tracks, which was replaced due to an upgrading project. The advantage of the instant re-use is that all resources are available for the project and the costs of replacement are therefore very low. The prerequisite is, however, that the re-use has to be included in the plans several years in advance. Re-use is not a common policy in the Netherlands any more, since the restructuring of the railways and the transfer of maintenance staffs to the contractors.

20.3.3 Conclusion

Both for design and maintenance of railway track there is a lot of potential to reduce the costs of ownership and the traffic disruption caused by failures. In order to analyse and optimize systematically design and maintenance strategies a life cycle cost analysis with the help of a DSS such as described in Section 20.2 is needed. The principal advantage of the DSS is that the impact of maintenance data and assumptions can be tested systematically. It is very well fitted for supporting a track policy revision with the use of the expertise from the regional maintenance offices.
However, for the further realisation of Life Cycle Cost based management the development of an Asset Management System (AMS) is essential. Maintenance data has long been a neglected area in track maintenance: usually a lot of raw data on track quality is collected, but the data is barely usable due to a number of reasons:

- The maintenance and failure history is not available in database systems
- Track quality data is not collected in central database systems
- The raw data is not processed to usable quality indicators
- The data is spread over different parts of the railway company or even different organisations (after the railway restructuring)

For improving the strategic life cycle costing studies, such as described in Section 20.3.1 and Section 20.3.2, and for a further implementation of life cycle cost based management on a tactical level the development of an AMS is essential. The whole chain of track data collection, data storage and processing has to be structured in order to provide reliable management information on costs of ownership, reliability and availability. An expert system, such as ECOTRACK, is an important element of this chain, since it has the ability to collect and process the data in a single system using objective thresholds for track renewal. A plain implementation of ECOTRACK will however not solve the problem: the life cycle costing method must be embedded in the organisation of the Maintenance Regions. Support from top management and co-operation with the Regional Maintenance Offices will be inevitable.

REFERENCES

[1] Achenbach J.D.: 'Wave propagation in Elastic Solids', Elsevier Science Publishers b.v., Amsterdam, 1973.

[2] Alias J.: 'Le Rail', Eyrolles, Paris, 1987.

[3] Alias J.: 'La voie ferrée', Eyrolles, Paris, 1977.

[4] Ando K. and Miura S.: 'Twenty years' experience on Slab Track', RTRI vol. 35 No. 1, 2, 1994.

[5] ASTM Standard E 399 74: 'Standard method of test for plane strain fracture toughness of metallic materials'.

[6] ASTM Standard E 647 81: 'Standard method for constant load amplitude fatigue crack growth rates above 10^{-8} m/cycle'.

[7] Bachmann H. and Ammann W.: 'Vibrations in structures induced by man and machine', Structural Engineering Documents 3e, IABSE AIPC -IVBH, ETH Hoeggerberg, Zürich, Switzerland.

[8] Bachmann H. and Foege T.: 'Rheda 2000 - Erfahrungen aus Einbau und Verarbeitung', Eisenbahn Ingenieur, 9, 2000.

[9] Baluch H.: 'Evaluations of the durability of rail fastenings to sleepers using a model of extremes', Rail International/Schienen der Welt, December 1979.

[10] Barthelemy J.-F.M. and Haftka, R.T.: 'Approximation Concept for Optimum Structural Design - a Review', Structural Optimization, Volume 5, pp. 129-144, 1993.

[11] Baus A. and Lieurade H.P.: 'Endurance en torsion et résistance à la fissuration par fatigue de trois nuances d'aciers à rails', Revue de Métallurgie, May 1975.

[12] Baxter P.: 'Information Technologies for Asset Management', Asset Management For Railway Infrastructure Conference - Brussels, Belgium, 28 February - 1 March 2000.

[13] Baxter P.: 'Using New Technologies to Improve Asset Management', Asset Management For Railway Infrastructure Conference - Workshop A - London, UK, 1 December 2000.

[14] Becker S. and Lier K. H.: 'Bewertung und Varianten-vergleich von Bauarten der Festen Fahrbahn', Eisenbahn Ingenieur, 2, 1999.

[15] Bendat J.S. and Piersol A.G.: 'Random Data: Analysis and Measurement Procedures', Wiley-Interscience, New York, 1971.

[16] Bendat J.S.: 'Statistical Errors in Measurements of Coherence Functions and Input/Output Quantities', Journal of Sound and Vibration, Vol.59, No 3, pp. 405, 1978.

[17] Bendat J.S. and Piersol A.G.: 'Engineering of Correlation and Spectral Analysis', Wiley-Interscience, New York, 1980.

[18] Bienzeisler H., Schmedders H. and Wick K.: 'Moderne Schienenerzeugung bei der Thyssen Stahl AG', Thyssen Technische Berichte, Volume 1, 1988.

[19] Bijl F.: 'Buckling force in gapless track calculated in a non linear manner', De Ingenieur, September 1964.

[20] Boer H. de, et. al.: 'Naturharte bainitische Schienen mit hoher Zugfestigkeit', Stahl und Eisen 115, Nr. 2, pp. 93-98.

[21] Bos T.J.: 'Fracture resistance of rails', Internal NS report CTO/3/L/10.110/126, February 1986.

[22] Bos J.A.: 'Low Noise Track', Rail International/Schienen der Welt, Vol.30 No.1, pp. 17-22, 1999.

[23] Bos J.A.: 'Deck Track: Foundation for the railways of the Future', Holland Railconsult.

[24] Bosch R.A. van den: 'Strategies for permanent way maintenance by NS', Proceedings of ÖVG Conference, Salzburg, Austria, 26-28 September 2000.

[25] Box G.E.P. and Draper N.R.: 'Emperical Model-Building and Response Surfaces', John Wiley & Sons, New York, 1987.

[26] Broek D.: 'Elementary engineering fracture mechanics', Noordhoff International publishing', Leiden, 1974.

[27] Bucholz K.H., Lipius J.M. and Schwendt L.: 'Anforderungen des Hochgeschwindigkeitsverkehrs der ICE als geeignetes Zugsystem', Thyssen Technische Berichte, Volume 1, 1988.

[28] CTT: 'Kan het nog stiller?', ICES-project, eindrapport Stiller Treinverkeer, 1999.

[29] CEN: European Committee for Standardization 'Railway applications – Track – Testing methods for fastening systems prEN13146', Draft version, 1998.

[30] CEN: European Committee for Standardization 'Railway applications – Track – Concrete sleepers and bearers prEN13230', Draft version, 1998.

[31] CEN: European Committee for Standardization 'Railway applications – Track – Performance requirements for fastening systems prEN13481', Draft version, 1999.

[32] CEN: European Committee for Standardization 'Railway applications – Track – Rail prEN13674', Draft version, 1999.

[33] Clark R.A., Dean P.A., Elkins J.A. and Newton S.G.: 'An investigation into the dynamic effects of railway vehicles running on corrugated rails', Journal of Engineering Mechanics, 24, 1982.

[34] Clark M., McCann D.M. and Forde M.C., 'Infra-red inspection of trackbed ballast', Railway Engineering 2000 Conference, London, UK, July 2000.

[35] Clough R.W. and Penzien J.: 'Dynamics of Structures', Mc Graw-Hill.

[36] Clouth AG: 'Technical data USM1000'.

[37] Cope G.H.: 'British railway track, Design, construction and maintenance', The Permanent Way Institution, Echo Press, England, 1993.

[38] Cost and Risk Optimisation of Track Maintenance and Renewal, Track MR 2001 Conference, London, UK, 11-13 July 2001.

[39] Cox S.J. and Grassie S.L.: 'Understanding Dynamics as an Aid to Developing Track', Pandrol International Ltd, 1985.

[40] Darr E.: 'Instandhaltung der Festen Fahrbahn', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[41] Darr E.: 'Konstruktion, Bauarten, Gleislagestabilität, Instandhaltung und Systemvergleich der Festen Fahrbahn', Symposium 'Feste Fahrbahn - Mechanisierte Modellierung, Betriebserfahrung und Akustik', Berlin, 4-5 November 1999.

[42] Darr E. and Fiebig W.: 'Feste Fahrbahn', Edition VDEI, Tetzlaff Verlag, Germany, 1999.

[43] Dean F.E., Ahlbeck D.R., Harrison H.D. and Tuten J.M.: 'Effect of the Tie Pad Stiffness on the Impact Loading of Concrete Ties', Heavy Haul Conference, 1982.

[44] Deroche R.Y., Bourdon Y. and Faessel A.: 'Straightening and stress relieving of rails by stretching', Railway Gazette International, March 1983.

[45] Desvallees J., Faessel A., Gouth G. and Mennel G.: 'Universal rolling of rails - state of the art', Unimetal, March 1987.

[46] Deutsche Bahn AG: 'Anforderungskatalog zum Bau der Festen Fahrbahn', 1995.

[47] Deutsche Bahn AG: 'Technische Linie TL918235'.

[48] Deutsche Bahn AG: 'Schalltechnische Untersuchungen Kontinuierliche Schienenlagerung – Sytem Edilon auf der freien Strecke – Versuchsstrecke bei Deurne / Holland', by: Obermeyer Plan und Beraten Projectnumber 10781-10/G.017, 1995.

[49] Devanathan R. and Clayton P.: 'Rolling-sliding wear behaviour of three bainite steels', International Conference of the ASME, Orlando, USA, pp. 91-100, 1991.

[50] Dieterman H. and Metrikine A.V.: 'Steady-state displacements of a beam on an elastic half-space due to a uniformly moving constant load', European Journal of Mechanics A/Solids, Volume 16, pp. 295-306, 1997.

[51] DIN: Deutsches Institut für Normung 'DIN4150 Erschütterungen im Bauwesen, Teil 1-3', 1999.

[52] Dörr J.: 'Der unendliche, federende gebettete Balken unter dem Einfluss einer gleichförmig bewegente Last', Ingenieur Archiv, Volume 14, pp. 167-192, 1943.

[53] Dubourg and Kalker J.J.: 'Crack behaviour under rolling contact fatigue', In: Kalker, Cannon und Orringer (Hrsg.), Rail quality and maintenance for modern railway operation, Kluwer Academic Publ., Dordrecht, pp. 373-384, 1993.

[54] Duškov M.: 'EPS as a Light-weight Sub-base Material in Pavement Structures', Ph.D.Thesis, TU-Delft, ISBN 90-9010660-X, June 1997.

[55] Dutch Ministry of Housing, Land-uses Planning and the Environment (VROM): 'Calculation of rail-borne noise', ISBN 90-346-06880, 1985.

[56] Ebersohn W. and Ruppert J.C. Jr.: 'Implementing a Railway Infrastructure Maintenance System'.

[57] Ebersohn W. and Selig E.T., 1994.

[58] Edilon b.v.: 'Product specifications embedded rails with UIC54 for Netherlands Railways'.

[59] Eisenmann J. and Mattner L.: 'Auswirkung der Oberbaukonstruktion auf die Schotter- und Untergrundbeanspruchung', Der Eisenbahningenieur, pp. 99-107, 1984.

[60] Eisenmann J.: 'Oberbauforschung - Oberbautechnik, Stand und Weiterentwicklung', ETR, 10, 1985.

[61] Eisenmann J. and Mattner L.: 'Gleisverwerfung', Technical University Munich, Heft 52, 1988.

[62] Enkon Information Systems Inc. web site: http://www.enkon.com/

[63] ERRI D182 rp3: 'Unified Assessment criteria for Ballast Quality and Methods for Assessing the Ballast Condition in the Track, Determining the criteria for ballast durability using triaxial tests', Utrecht, September 1994.

[64] ERRI D187 dt299: 'Decision Support System for Track Maintenance and Renewal - ECOTRACK', Utrecht, April 1994.

[65] ERRI D202: 'Lateral Resistance Test', European Rail Research Institute, Utrecht, ERRI D 202/WG3, Draft Report BR research Limited, report RR-TCE-81, November 1996.

[66] ERRI D229 rp2:'Rail Defect Management, State-of-the-Art Report', UIC WEC Joint Research Project, Utrecht, 24 February 1999.

[67] Eschenauer H.A.: 'Multidisciplinary engineering strategies for product and process development - modelling, simulation, optimization, applications', In: V.V. Toropov (Ed.) Engineering Design Optimization, Product and Process Improvement. Proc. of the Ist ASMO UK/ISSMO Conference on Engineering Design Optimization, Ilkley, West Yorkshire, UK, 8-9 July 1999.

[68] Esveld C.: 'Track resistance measurements on Eindhoven- Venlo', Is8/CE16, January 1978.

[69] Esveld C.: 'Spectral analysis of track geometry for assessing the performance of maintenance machines', ORE DT77/D117, March 1978.

[70] Esveld C.: 'Comparison between Theoretical and Actual Transfer Functions of Track Maintenance Machines', Doctoral Thesis, May 1978.

[71] Esveld C.: 'The performance of tamping and lining machines', Rail International/Schienen der Welt, December 1979.

[72] Esveld C.: 'A numerical Butterworth filter for digital data processing, with applications to the analysis of track geometry', Rail International/Schienen der Welt, January 1980.

[73] Esveld C.: 'Track stiffness measurements using an adapted tamping machine', Rail International/ Schienen der Welt, February 1980.

[74] Esveld C.: 'Einige Erfahrungen mit dem dynamischen Gleisstabilisator (DTS) bei den NS', Glasers Annalen, November 1980.

[75] Esveld C., Kosieradzki W. and Peerboom A.: 'Vehicle/track interaction: Theory and Applications of the Multiple Input Single Output method', ORE DT136, 1982.

[76] Esveld C.: 'Vertical rail geometry', Rail International/Schienen der Welt, February 1983.

[77] Esveld C.: 'Automatic curve corrections using DRIVER', ORE Colloquium 'Use of micro processors in the diagnostic systems of railway equipment', Budapest, May 1984.

[78] Esveld C.: 'Q-NORM, Quality rating of track geometry', Rail International/Schienen der Welt, June 1984.

[79] Esveld C.: 'MISO: Application of random signal analysis to vehicle track interaction', Rail International/ Schienen der Welt, August/September 1984.

[80] Esveld C.: 'BMS: The new NS track recording system', Rail International/Schienen der Welt, June 1985.

[81] Esveld C.: 'Measuring and rectifying rail roughness and bad welds', Third International Heavy Haul Railway Conference, Vancouver, Canada, October 1986.

[82] Esveld C.: 'Three years experience with modern weld treatment', Rail International/Schienen der Welt, December 1986.

[83] Esveld C.: 'An objective track maintenance strategy', Rail Engineering International, 2, 1987.

[84] Esveld C.: 'A high speed recording system for rail irregularities', Rail International/Schienen der Welt, July 1987.

[85] Esveld C.: 'BINCO, an information system for track maintenance and renewal', Computers in Railway Management, Computational Mechanics Publications/ Springer Verlag, 1987.

[86] Esveld C.: 'Reinforced sleeper bed, facts and figures', Rail Engineering International, 3, 1987.

[87] Esveld C.: 'A modern track maintenance strategy', Nordic Seminar on Track Technology, Finse, Norway, February 1988.

[88] Esveld C., Jourdain A., Kaess G. and Shenton M.J.: 'Historic data on track geometry in relation to maintenance', Rail Engineering International, 2, 1988.

[89] Esveld C. and van Hengstum L.: 'Track Stability in Tight Curves', Rail International/Schienen der Welt, January 1989.

[90] Esveld C.: 'Computer-Aided Maintenance and Renewal of Track', 4th International Heavy Haul Conference, Brisbane, Australia, September 1989.

[91] Esveld C. and Kok, A.W.M.: 'Interaction between Moving Vehicles and Railway Track at High Speed', Rail Engineering International, Volume 27, 3, pp. 14-16, 1998.

[92] Esveld C.: 'Innovation for the control of infrastructure maintenance', Rail International/Schienen der Welt, February 1997.

[93] Euronorm: 'Railway applications - Ride comfort for passengers - Measurement and evaluation, NVN-ENV 12299', 1999.

[94] European Commission DG VII, 'A strategy for revitalising the community's railways', (White Paper), Brussels, 1997.

[95] Fastenrath F.: 'Die Eisenbahnschiene', Chapter 2: J. Eisenmann: 'Die Schiene als Träger und Fahrbahn', Verlag W. Ernst, Berlin, 1977.

[96] Fillipov A.P.: 'Steady State vibrations of an infinite beam on an elastic half space subjected to a moving load', Izvestija Akademii Nauk SSSR Mehanika I Mashinostroenie, Volume 6, pp. 97-105, 1961.

[97] Flanagan R. and Norman G.: 'Life cycle costing for construction', Reading, 1983.

[98] Frederick C.O.: 'Vibrations in ground: Railway induced ground vibrations', Rail International/Schienen der Welt, October 1987.

[99] Frýba L.: 'Vibration of solids and structures under moving loads', Noordhoff International Publishing, Groningen, 1972.

[100] Fung Y.C.: 'Foundations of solid mechanics', Prentice-Hall inc, Englewood Cliffs, New Jersey, 1965.

[101] Garg K. and Dukkipati R.V.: 'Dynamics of Railway Vehicle Systems', Academic Press, 1984.

[102] Göbel C., Leiberenz K. and Richter F.: 'The Railway Formation', DB-Fachbuch Band 8/20.

[103] Grassie S.L., Gregory R.W., Harrison D. and Johnson K.L.: 'The dynamic response of railway track to high frequency vertical excitation', Journal of Engineering Mechanics, 24, 1982.

[104] Grassie S.L. and Kalousek J.: 'Rolling contact fatigue of rails: characteristics, causes and treatments', In: Proceedings 6th International Heavy Haul Conference, Cape Town, South Africa, pp. 381-404, 1997.

[105] Groenhuysen F.J.: 'Construction, maintenance and renewal of tracks', lecture notes e61, TU-Delft, February 1987.

[106] Hald J. and Madsen K.: 'Combined LP and Quasi-Newton Methods for Nonlinear L1 Optimization', SIAM J. Numer. Anal., Vol. 20, pp. 68-80, 1985.

[107] Halvorson W.G. and Bendat J.S.: 'Noise Source Identification Using Coherent Output Power Spectra', Sound and Vibrating, Vol. 9, No. 8, pp.15, 1975.

[108] Harmsen E.J.M.: 'Developments in track technology', special issue of the Dutch Permanent Way Institution, 1974.

[109] Haupt H.J.: 'An integrated System for Rail Track Inspection, Monitoring and Maintenance Control', European Railway Review, Volume 6, Issue 4, 2000.

[110] Hazelaar A.H.M. et al.: 'Veilig Werken aan de Infra', NS-Railinfrabeheer, 1997.

[111] Heller W.: 'Herstellung, Eigenschaften und Betriebsverhalten von Schienenstählen', In: 'Die Eisenbahnschiene', Wilhelm Ernst & Sohn, 1977.

[112] Heller W., Körfer E. and Schmedders H.: 'Rail Technology and Rail Development in Germany (FRG)', Krupp, Klöckner, Thyssen.

[113] Heller W., Schweitzer R., Weber L.: 'Modern Developments in Rail Steel Metallurgy and Production', Krupp international report 38442.

[114] Heller W., Schweitzer R.: 'Untersuchungen und betriebliche Erprobung neuartiger naturharter Schienenstähle', Techn. Mitt. Krupp Werksber., pp. 79-88, October 1979.

[115] Hengstum L. van: 'A finite element analysis of track stability', Rail Engineering International, 4, 1987.

[116] Heppding H.G.: 'Zur Berechnung nichtlinearer Verbund- probleme, dargestellt am Beispiel der Längskraftermittlung Eisenbahnbrücke', Der Bauingenieur 62, 1987.

[117] Hetényi M.: 'Beams on Elastic Foundations', University of Michigan, 1971.

[118] Heynen G.W.J.: 'Railway induced vibrations', Internal NS-CTO report, November 1987.

[119] Hillig J.: 'Erdbautechnische Anforderungen an eine schotterlose Fahrbahn', Eisenbahn Ingenieur, 5, 1994.

[120] Hilliges D.: 'Anwendung der Festen Fahrbahn / Bauart Rheda Entwicklung der Bauart DYWIDAG-Rheda', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[121] Hofman C. et al.: 'Erprobung hochfester naturharter Schienen auf der Gotthard-Strecke', Eisenbahntechnische Rundschau Vol. 38, No. 12, pp. 775-781, 1989.

[122] Holtzendorff K.: 'Ursachen von Schottersetzung und deren Modellierung', IFV Symposium 'Schotteroberbau - Mechanisierte Modellierung, Laborversuche und die Praxis', Berlin, 1999.

[123] Holzinger R. and Fritz D.: 'Entwicklung moderner Hochleistungsweichen zur Wahrung der Zukunftschancen der Bahn', ETR-Eisenbahntechnische Rundschau, Vol. 39, No. 1-2, pp. 71-78, January-February 1990.

[124] Horsnell M.R. and Adam C.H.: 'Geotechnical investigations for railway maintenance and renewal', Railway Engineering 2000 Conference, London, UK, July 2000.

[125] Horvath J.S.: 'Geofoam geosynthetic', Horvath Engineering, P.C., New York, USA, July 1995.

[126] Holzinger H. and Schreiner U.: 'Feste Fahrbahn Bögl - Generationswechsel im Gleisoberbau', Eisenbahn Ingenieur, 9, 2000.

[127] HSL-South Project Organisation, 'Ballastless Track for the Netherlands High-Speed line', Conference for tederers, Schiphol, the Netherlands, 1999.

[128] Hunt G.A.: 'Analysis of requirements for railway construction on soft ground', Technical Report LR TM 031, British Rail Research, London, UK, 1994.

[129] Hunt H.E.M. and Newland D.E.: 'Theoretical Investigation of Track Dynamic Settlement', 1995.

[130] Ichinose H. et al.: 'An investigation on contact fatigue and wear resistance behaviour in rail steels', Rail Steel Symposium, Paper I.3, 1978.

[131] IMC Geophysics Ltd., 'Ballast Scan'.

[132] Infundo GmbH: 'Prüfung des FF-systems Infundo mit kontinuierlich eingebetteter Schiene', by: TU München, Bericht 1645(a), 1996.

[133] Infundo GmbH: 'Prüfung des Systems Infundo mit kontinuierlich eingebetteter Schiene S41', by: TU München, Bericht 1677, 1997.

[134] ISO: International Organization for Standardization 'ISO2631 Mechanical vibration and shock - ISO-2631-1 Mechanical vibration and shock - Evaluation of human exposure to whole-body vibration, Part 1-2', 1997.

[135] Jackson P.D., Northmore K.J., Meldrum P.I., Gunn D.A. and Hallam J.R.: 'Non-invasive moisture monitoring within an earth embankment - a precursor to failure', Railway Engineering 2000 Conference, London, UK, July 2000.

[136] Janin G.: 'Maintaining track geometry decision making for levelling and lining', French Railway Review, Vol. 1, No. 1, 1983.

[137] Janssens, M.H.A.: 'ICES-STV Effect van de spoorstaafvorm op de geluidproductie van betonplatenspoor', TNO-TPD report 97-0108, 1997.

[138] Jarmulak J. and Giling J.M.: 'CBR systems for interpretations of eddy-current and ultrasonic NDT data', In: Applications and Innovations in Intelligent Systems VII, Proceedings of the 19th International Conference on Knowledge Based Systems and Applied Artificial Intelligence, Springer, Cambridge, England, UK, pp. 102-118, 1999.

[139] Jenkins H.H., Stephenson J.E., Clayton G.A., Morland G.W. and Lyon D.: 'The effect of track and vehicle parameters on wheel/rail vertical dynamic forces', Railway Engineering Journal, January 1974.

[140] Jovanovic S. and Nestorov I.: 'System for location and data-collection about railway infrastructure elements', First Symposium on Condition Assessment, Maintenance and Remedy of the Civil Engineering Objects, Tara, Yugoslavia, December 1999.

[141] Jovanovic S. and Korpanec I.: 'Ecotrack Proves Its Worth', International Railway Journal and Rapid Transit Review (official web site), August 2000.

[142] Jovanovic S. and Zaalberg H.: 'ECOTRACK: Two years of experience', Rail International/Schienen der Welt, April 2000.

[143] Jovanovic S. and Pearce M.: 'ECOTRACK: An overview of the Systems's Functionality and Implememtation to Date', American Railway Engineering and Maintenance-of-Way Association, 2000 Annual Conference, Dallas, Texas, USA, 10-13 September 2000.

[144] Jovanovic S. and Korpanec I.: 'ECOTRACK: The Implementation Phase', Rail Technology Magazine, ISSN 1471 0668, pp.19-22, February-March 2001.

[145] Jovanovic S. and Esveld C.: 'ECOTRACK: An objective condition-based Decision Support System for long-term track M&R Planning directed towards reduction of Life Cycle Costs', 7th International Heavy Haul Conference, Brisbane, Australia, 10-15 June 2001.

[146] Kaess G.: 'Die Langsstabilitaet von Weichen in geschweis sten Gleisen', ETR, March 1982.

[147] Kaess G. and Ebersbach D.: 'Schnelle Züge, schwere Lasten, was sagt der Oberbau dazu?', ETR, January 1986.

[148] Kaess G.: 'Erfahrungen und Ergebnisse aus dem Einsatz des dynamischen Gleisstabilisators', ETR, October 1987.

[149] Kalker J.J: 'Three-Dimensional Elastic Bodies in Rolling Contact', Volume 2 of 'Solid Mechanics and its Applications', Kluwer Acaedemic Publishers, Dordrecht, 1990.

[150] Kaynia A.M., Madshus C. and Zackrisson P.: 'Ground vibration from high-speed trains: prediction and countermeasure', Journal for Geotechnical and Geo-environmental Engineering, Volume 126, pp. 531-537, 2000.

[151] Kerr A.D.: 'On the stress Analyses of Rails and Ties', AREA Bulletin 659, 1976.

[152] Kerr A.D. and Bathurst L.A.: 'Pads ease track transitions', Railway Track & Structures, ISSN 0033-9016, Vol.96 No.8, pp. 57-64, 2000.

[153] Knothe K. and Yu M.: 'Ansätze zur Modellbildung der Festen Fahrbahn und Unterschiede zur Modellbildung des Schotteroberbaus', IFV symposium 'Feste Fahrbahn - Mechanisierte Modellierung, Betriebserfahrung und Akustik', Berlin, November 1999.

[154] Knote K.: 'Gleisdynamik', Ernst & Sohn, Berlin, ISBN 3-433-01760-3, 2001

[155] Krüger F.: 'Erprobung einer einfachen kontinuierlich elastichen Schienenlagerung auf einer Tunnelstrecke', Verkehr und Techniek, No. 7, pp. 306-312, 1999.

[156] Labra J.J.: 'An axially-stressed railroad track on an elastic continuum subjected to a moving load', Acta Mechanica, Volume 22, pp. 113-129, 1975.

[157] Lamson S.T.: 'A-symmetrical grinding challenges heavy-haul punishment', Railway Track and Structures, February 1984.

[158] Lewis R.B.: 'High Speed Track Recording Coach', Journal of Engineering Mechanics, September 1975.

[159] Lewis R.B. and Richards A.N.: 'A new method for the routine measurement of rail head corrugations', Rail International/Schienen der Welt, February 1986.

[160] Leykauf G. and Mattner L.: 'Feste Fahrbahn mit Asphalttragschicht', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[161] Leykauf G. and Maleki N.: 'Feste Fahrbahn - Entwicklungen im Ausland', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[162] Leykauf G. and Mattner L.: 'Moderne Fahrweg-systeme', ETR, 2-3, 1998.

[163] Lichtberger B.: 'Rechnergestützte Systeme zur Gleislageberichtigung', Vortrag Oslo, March 1990.

[164] Lichtberger B.: 'State of chord measuring using EM-SAT or GPS', EI/Der Eisenbahningenieur, 46, 8, pp. 560-ff, 1995.

[165] Lieb M. and Sudret B.: 'A fast algorithm for soil dynamics calculations by wavelet decomposition', Archive of Applied Mechanics, Volume 68, pp. 147-157, 1998.

[166] Madshus C. and Kaynia A.M.: 'High-speed railway lines on soft ground: dynamic behaviour at critical train speed', Journal of Sound and Vibration, Volume 231, pp. 689-701, 2000.

[167] Man A.P. de: 'Determination of dynamic track properties by means of excitation hammer testing', Rail Engineering International, 4, pp. 8-9, 1996.

[168] Man A.P. de and Moraal J.: 'Mechanical testing of the definite SA42-rail structure in behalf of ICES', Railway Engineering, TU-Delft, report 7-98-111-7, 1998.

[169] Man A.P. de: 'Interim Report Dynatrack: 2 years of research on the dynamic behaviour of railway track structures', Railway Engineering, TU-Delft, report 7-99-111-10, 1999.

[170] Man A.P. de and Moraal J.: 'Behaviour of rail fastenings: Laboratory testing of Direct Fastenings', Railway Engineering, TU-Delft, report 7-00-120-9, 2000.

[171] Maree J.S. and Gräbe P.J.: 'Use of a Dynamic Track Stabiliser to Improve track maintenance and optimisation of track tamping', Proceedings of 6th International Heavy Haul Conference, Capetown, South Africa, June 1979.

[172] Markine V.L., de Man A.P. and Esveld C.: 'A procedure for design and optimization of a railway track structure', Proc. Interactive Conference: Cost Effectiveness and Safety Aspects of Railway Track, Union Internationale des Chemins de Fer, Paris, France, 8-9 December 1998.

[173] Markine V.L.: 'Optimization of the Dynamic Behaviour of Mechanical Systems', Ph.D.Thesis, TU-Delft, Shaker Publishing B.V., ISBN 90-423-0069-8, 1999.

[174] Markine V.L. and Esveld C.: 'Analysis of Lateral Behaviour of a Railway Track Structure Using an Optimization Technique', Railway Engineering, TU-Delft, report 7-99-124-1, 1999.

[175] Markine V.L., de Man A.P., Jovanovic, S. and Esveld C.: 'Modelling and Optimization of an Embedded Rail Structure', Proceedings of Innovations in the design and assessment of railway track, International Conference, Trade Show and Workshops, TU-Delft, The Netherlands, 2-3 December 1999.

[176] Markine V.L. and Esveld C.: 'Determination Of Track Lateral Resistance And Train Speed Limits Using Tamping Machine And Numerical Optimisation', In: M. Forde (ed.) The 3rd International Conference Railway Engineering - 2000 (CD Proceedings), ISBN 0-947644-43-1, London, UK, 5-6 July 2000.

[177] Markine V.L., de Man A.P., Jovanovic, S. and Esveld, C.: 'Multicriteria Optimisation Of Railway Track For High-Speed Lines', In: J. Sienz (ed.) Proceedings of the 2nd ASMO UK / ISSMO Conference on Engineering Design And Optimization, ISBN 0-86076-168-1, Swansea, Wales, UK, pp.149-156, 10-11 July 2000.

[178] Markine V.L., Toropov V.V. and Esveld C.: 'Optimization of Ballastless Railway Track Using Multipoint Approximations: Issue of Domain-Dependent Calculability', Proceedings of the 8th AIAA/USAF/ NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, (CD Proceedings). AIAA Paper 2000-4714, Long Beach, California, 6-8 September 2000.

[179] Massarsch K.R.: 'Isolation of vibrations in soils', Report FIT Franki, Liège, Belgium.

[180] Masumoto H., Sugino K. and Hayashida H.: 'Development of wear resistant and anti-shelling high strength rails in Japan', Paper H.1, Rail Steel Symposium, 1978.

[181] Matten D. and Langhagen K.: 'Feste Fahrbahn im Spannungsfeld von Innovation und Bewährung', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[182] McMichael P.L.: 'Stone blowing from theory to practice', Third International Heavy Haul Conference, Vancouver, Canada, October 1986.

[183] McMichael P.L.: 'Implementing the Stoneblower', Permanent Way Institution Technical Seminr, Instution of Mechanical Engineers, Lodon, October 1995.

[184] Megyeri J.: 'Bewegungsgeometrische Ueberlegungen bei der Entwicklung von Eisenbahnweichen', AET-Archiv fuer Eisenbahntechnik, Vol. 40, pp. 59-63, 1985.

[185] Merkblatt Stahl 241: 'Aluminothermic welding of rails', Beratungsstelle für Stahlverwendung, Düsseldorf, 1983.

[186] Montagné S.: 'Five years experience with French high speed track: achievements and future prospects', Rail International/Schienen der Welt, October 1988.

[187] Moraal J.: 'Testing of Pandrol Rail Fastening Systems', Railway Engineering, TU-Delft, report 7-97-118-2, 1997.

[188] Moraal J. and Bientjes J.W.: 'Pulling force testing of rail fastenings', Railway Engineering, TU-Delft, report 7-97-103-6, 1997.

[189] Moraal J.: 'Laboratory testing on Spanish rail fastening systems', Railway Engineering, TU-Delft, report 7-99-118-7, 1999.

[190] Mörscher J.: 'Anforderungskatalog zum Bau der Feste Fahrbahn', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[191] Moubray J.: 'Reliability-centred maintenance', Butterworth-Heinemann, Oxford, UK, 1997

[217] ORE D156 rp4: 'Studies concerning the measurement and improvement of the level of residual stresses', Utrecht, September 1987.

[218] ORE D161 rp1 : 'General conditions for the study of the evolution of track geometry based on historical information', Utrecht, April 1987.

[219] ORE D161 rp2: 'Assessment of the service life of rails from their defect occurrence', Utrecht, April 1983.

[220] ORE D161 rp4: 'Dynamic vehicle/track interaction phenomena from the point of view of track maintenance', Final report: conclusions and recommendations, Utrecht, September 1988.

[221] ORE D161.1 rp3 : 'The influence of wheel flats on the track with 20 and 22.5 t axle loads', Utrecht, September 1986.

[222] ORE D161.1 rp4 : 'The dynamic effects due to increasing axle loads from 20 to 22.5 t and the estimated increase in maintenance costs', Utrecht, September 1987.

[223] Ohtake T. et al.: 'Practical use of TOSMA on track maintenance work', In: Cost effectiveness and safety aspects of railway track, ERRI and UIC, Paris 1998.

[224] Pahnke U.: 'Konstruktive Geschichtspunkte bei der Gestaltung von Fahrbahnplatten aus Beton für Schienenbahnen', ETR, 4, 1992.

[225] Peterson F.J., Wolf G.P. and Bachinsky G.S., 'The Relationship Among Operating Procedures, Track Geometry, Rail Profile, and Rail Wear Rates', ARM 1999 Rail/Wheel Interface Seminar 5 May 1999.

[226] Pfleiderer AG: 'Rheda 2000 - Die monolitische Feste Fahrbahn mit Direktafbau auf HGT'.

[227] Phoenix AG: 'Subballast mats for track technology'.

[228] Pichler D.: 'Noise and Vibration attenuating measures for modern railway superstructures', In: P.Sas (Ed.) 'Proceedings ISMA25', KU Leuven, ISBN 90-73802-72-5, pp. 755-762, 2000.

[229] Powell M.J.D.: 'A Fast Algorithm for Nonlinearly Constrained Optimization Calculations', In: Proceedings of the 1977 Dundee Conference on Numerical Analysis, Lectures Notes in Mathematics, Springer-Verlag, 630, pp. 144-157, 1978.

[230] Proceedings of the International Conference on Wheel/rail Load and Displacement Measurement Techniques, DOT-TSC-UMTA-82-3, September 1981.

[231] Prud'homme A.: 'Les problèmes que pose, pour la voie, la circulation des rames à grande vitesse', Revue Général des Chemins de Fer, November 1976.

[232] Rasmussen, S. and de Man, A.P.: 'The High Speed Deflectograph and the Hammer Excitation Test: two measurement techniques for determining track performance', Rail Engineering International, Vol.29, Nr.1, pp. 13-16, 2000.

[233] Redi.: 'Cedias - Railway Lines Diagnostic System', By: Sistemi Energetici Integrati sri, Prato, Italy.

[234] REMAIN: 'Modular System for Reliability and Maintainability Management in European Rail Transport, Final Report, 1998.

[235] Richart F.E., Woods R.D. and Hall J.R.: 'Vibrations of soils and foundations', Prentice-Hall, Englewood Cliffs, New Jersey, 1970.

[236] Riessberger K.: 'New Transition Geometry', In: 'High Speed Track Course', By: PAO Delft, the Netherlands, 10-11 June 1997.

[237] Riessberger K.: 'Frame sleeper upgrade ballast track', WCRR Cologne, 2001.

[238] Rivier R.E. and Salchli B.: 'Management information system for track maintenance and renewal', Computers in Railway Management, Computer Mechanics Publications/Springer Verlag, 1987.

[239] Roney M.D.: 'Maintaining Optimal Wheel and Rail Performance', In: 'Guidelines To Best Practices for Heavy Haul Railway Operations: Wheel and Rail Interface Issues', International Heavy Haul Association, Virginia Beach, USA, Part 5, pp. 5-113, May 2001.

[240] Rudolph R.: 'Operational simulation of light rail systems', European Transport Conference, Cambridge, UK, 2000.

[241] Ruiter B. de: 'Noise abatement in metro tunnels, part 1 and 2', Weg en Water, Ed. 1985, No. 5, pp. 9-11, No. 6, pp. 12-14, 1985.

[242] Rump R. von: 'Warum Feste Fahrbahn?', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany 1997.

[243] Sato Y., Fujimori S. and Yoshimura A.: 'Real time digital processing of track irregularities in time of speed-up on Tokaido and Sanyo Shinkansen', ORE Colloquium on experience with the introduction and operation of high and very high speed lines, Arrezo, September 1987.

[244] Sato Y., Morimura T. and Watanabe S.: 'Theoretical analysis and experimental results on track moduli with use of wheel-set drop test', Vehicle System Dynamics, Vol. 24 (supplement), 1995.

[245] Schmit L.A. Jr. and Farshi B.: 'Some Approximation Concepts for Structural Synthesis', AIAA Journal, Volume 12, pp. 692-699, 1974.

[246] Schilder R. and Piereder F.: 'Formation rehabilitation on Austrian Federal Railways - five years of operation experience with the AHM 800 R', Rail Engineering International, April 2000.

[247] Schooleman R.B.: 'Transition between embankments and bridges for High-Speed Lines', Railway Engineering, TU-Delft, report 7-96-110-10, October 1996.

[248] Schreiner H.: 'Feste Fahrbahn - Höhengenaue Asphalt-Tragschichten', Eisenbahn Ingenieur, 3, 1992.

[249] Schweitzer R. and Heller W.: 'Zum Abbrennstumpfschweissen von naturharten Schienen aus Sondergütte Chrom-Mangan-stahl', ETR, 23, 1974.

[250] Schweitzer R. and Huber O.: 'Feinperlitisieren von Weichen', Technische Mitteilungen Krupp 37/3, October 1979.

[251] Selig E.T. and Waters J.M.: 'Track Geotechnology and Substructure Management', Thomas Telford Serivices Ltd, London, 1994.

[252] Shenton M.J.: 'Track standards for high speed trains', ORE Colloquium on experience with the introduction and operation of high and very high speed lines, Arezzo, Italy, September 1987.

[253] Siderius R.M.: 'Feasibility of EPS as a lightweight sub-base material in rail structures', Railway Engineering, TU-Delft, report 7-98-211-8, August 1998.

[254] Snippert G.: 'Dealing with Process Contracting', NS-Railinfrabeheer, Utrecht, 1999.

[255] Stadler W.: 'Multicriteria Optimization in Engineering and in the Science', Plenum Press, New York, London, 1988.

[256] Stephen K.J. and Dellisola A.: 'Life cycle costing for design professionals', McGraw-Hill, USA, 1995.

[257] Suiker A.S.J. and Esveld C.: 'Stiffness transition subjected to an instantaneous moving load passage', In: Proceedings of the Sixth International Heavy Haul Railway Conference, Capetown, South Africa, pp. 1194-1205, 1997.

[258] Suiker A.S.J. and Esveld C.: 'Instantaneous load passage on stratified media', In: Proceedings of the Sixth International Heavy Haul Railway Conference, Capetown, South Africa, April 1997.

[259] Suiker A.S.J., de Borst R. and Esveld C.: 'Critical behaviour of a Timoshenko beam-half plane system under a moving load', Archive of Applied Mechanics, Volume 68, pp. 158-168, 1998.

[260] Suiker A.S.J., Chang C.S., de Borst R. and Esveld C.: 'Surface waves in a stratified half space with enhanced continuum properties. Part I - Formulation of the boundary value problem', European Journal of Mechanics A/Solids, Volume 18, pp. 749-768, 1999.

[261] Suiker A.S.J., Chang C.S., de Borst R. and Esveld C.: 'Surface waves in a stratified half space with enhanced continuum properties. Part II - Analysis of the wave characteristics in regard to high-speed railway tracks', European Journal of Mechanics A/Solids, Volume 18, pp. 769-784, 1999.

[262] Suiker A.S.J., Metrikine A.V. and de Borst R.: 'Dynamic behaviour of a layer of discrete particles. Part I - Analysis of body waves and eigen modes', Journal of Sound and Vibration, Volume 240, No.1, pp. 1-18, 2001.

[263] Suiker A.S.J., Metrikine A.V. and de Borst R.: 'Dynamic behaviour of a layer of discrete particles. Part II - Response to a uniformly moving, harmonically vibrating load', Journal of Sound and Vibration, Volume 240, No.1, pp. 18-39, 2001.

[264] Thompson D.J. and Vincent N.: 'Track dynamic behaviour at high frequencies. Part 1:...', In: Knothe, K. Grassie, S.L. and Elkins, J.A. (Eds.) 'Interaction of railway vehicles with the track and its substructure', ISSN 0042-3114, pp. 86-99, 1994.

[265] Thompson D.J. and Janssens M.H.A.: 'TWINS User Manual', TNO-TPD, report 93-0213, 1993 and (revision) 1996.

[266] Tiflex Ltd: 'Product specifications fc9'.

[267] Timoshenko S.P.: 'Methods of analysis of statical and dynamical stresses in rails', In: Proceedings of the Second International Congress of Applied Mechanics, Zurich, Switzerland, pp. 407-420, 1927.

[268] Timoshenko S., Young D.H. and Weaver W.: 'Vibration Problems in Engineering', John Wiley & Sons, 1954.

[269] Timoshenko S. and Goodier J.N.: 'Theory of Elasticity', McGraw-Hill, New York, Second Edition, 1951.

[270] Timoshenko S.P.: Theory of Elastic Stability. McGraw-Hill, New York, 1936.

[271] Tjaden H.J.: 'A new concept for measuring vertical rail geometry', TU-Delft, June 1986.

[272] Tjaden H.J.: 'Calculation of vehicle reactions from track geometry', Internal NS report, March 1988.

[273] Tjaden H.J.: 'Automatic correction of vertical track geometry', Internal NS report, March 1988.

[274] Toropov V.V.: 'Simulation Approach to Structural Optimization', Structural Optimization, 1, pp. 37-46, 1989.

[275] Toropov V.V., van Keulen F., Markine V.L. and Alvarez L.F.: 'Multipoint Approximations Based on Response Surface Fitting: a Summary of Recent Developments', In: V.V.Toropov (Ed.) Proceedings of the Ist ASMO UK/ISSMO Conference on Engineering Design Optimization, pp. 371-381, ISBN 0-86176-650-4, Ilkley, West Yorkshire, UK, 8-9 July 1999.

[276] Tunna J.: 'Forces due to wheel flats', International Wheelset Conference, September 1988.

[277] Tunna J.M., Hough A. and Waters J.M.: 'Autograph track geometry measuring trolley: measuring vertical and horizontal track geometry profiles', Rail Enginering International, 4, pp. 3-4, 1994.

[278] Ueda M. et al.: 'Development of bainitic rail steel with excellent surface damage resistance', Proceedings IHHA 1999, Moscow, pp. 259-266, 1999.

[279] UIC leaflet 712 R: 'Rail defects', 1979.

[280] UIC leaflet 703 R: 'Design of alignment for tracks run by fast passenger trains', 1988.

[281] Van M.A. and van Dam A.J.P.: 'A finite element model for the calculation of axial track forces', TU-Delft, 1988.

[282] Van M.A.: 'Stability of Continuous Welded Rail Track', Ph.D.Thesis, Delft University Press, ISBN 90-407-1485-1, 1997.

[283] Vermeij, D.J.: 'Vehicle dynamics and passenger comfort related to track geometry in high speed track', Railway Engineering, TU-Delft, report 7-98-121-1, 1998.

[284] Vossloh GmbH: 'Product specifications 1403'.

[285] Vossloh GmbH: 'Product Specifications Feste Fahrbahn System 300', 1997.

[286] Vossloh GmbH: 'Product Specifications Schallabsorber System', 1997.

[287] Watanabe K.: 'Shinkansen track maintenance by closed-loop control', Railway Gazette International, March 1982.

[288] Waters J.M.: 'The FROG track measuring system', International Railway Journal, December 1984.

[289] Waters J.M.: 'Pneumatic Stone Injection - A means of adjusting track level', Fourth International Rail Track and Sleeper Conference, Adelaide, Australia, September 1981.

[290] Weigend M.: 'On-line evaluation of track geometry with the aid of electronic computers in the DB's track recording cars', ETR 31, November 1982.

[291] Weigend M.: 'Gestaltung von Geleiseverbindungen, Elsner's Taschenbuch der Eisenbahntechnik', Tetzlaff Verlag, Darmstadt, 1983.

[292] Wenty R.: 'Basic considerations for mechanical ballast cleaning', Rail Engineering International, February 1990.

[293] Widmann H.: 'Feste Fahrbahn Bauart Züblin', Feste Fahrbahn, Edition ETR, Hestra Verlag, Germany, 1997.

[294] Zaalberg H.: 'Economising Track Renewal and Maintenance with ECOTACK', Proceedings Interactive Conference on Cost effectiveness and safety aspects of railway track, ERRI and UIC, Paris, France, 1998.

[295] Zand J.van 't : 'Track technology', lecture notes e61, TU-Delft, February 1987.

[296] Zand J.van 't : 'Experimental evaluation of dynamic properties of rail pads at lower temperatures', TU-Delft, December 1988.

[297] Zand J.van 't: 'Ballast mats', Railway Engineering, TU-Delft, report 7-91-105-11, 1991.

[298] Zand J.van 't and Moraal J.: 'Ballast Resistance under Three Dimensional Loading', Railway Engineering, TU-Delft, report 7-97-103-4, February 1997.

[299] Zeevenhooven N.: 'Railway Vehicle Technology', lecture notes 170b, TU-Delft, 1987.

[300] Zimmermann H.: 'Die Berechnung des Eisenbahnoberbaues', 3rd edition, Wilhelm Ernst und Sohn, Berlin, 1941.

[301] Zoeteman A.: 'Planning the infrastructure performance of railways? Decision support for track design of a Dutch high-speed link', European Transport Coference, Cambridge, UK, 2000.

[302] Zoeteman A.: 'A financial comparison of track structures: application of life cycle costing for metro track construction in Madrid' (confidential), Comunidad de Madrid, Spain, 1998.

[303] Zoeteman A.: 'Application results of Life Cycle Cost planning on ballasted and embedded rail track structures' (confidential), Strukton Railinfra, Maarssen, The Netherlands, 1998.

[304] Zoeteman A.: 'Report on the Life Cycle Costs of the HSL-Zuid Track System' (confidential), TU-Delft, 2000.

[305] Zoeteman A.: 'Resultsof the Optimisation of the HSL-Zuid Rail System: Estimates on Life Cycle Cost and Availability Levels of the System Components' (confidential), TU-Delft, 2000.

[306] Zoeteman A.: 'State of the art of the Life Cycle Management Plus pilots', NS-Railinfrabeheer, Utrecht, 2000.

[307] Zwarthoed J.M.: 'Monoblock sleeper Betonfabriek Nigtevecht bv', Railway Engineering, TU-Delft, report 7-00-118-5, 2000.

[308] Zwarthoed J.M.: 'Slab Track design: flexural stiffness versus soil improvement', In: Proceedings Railtech 2001 Conference, Utrecht, the Netherlands, 2001.

INDEX