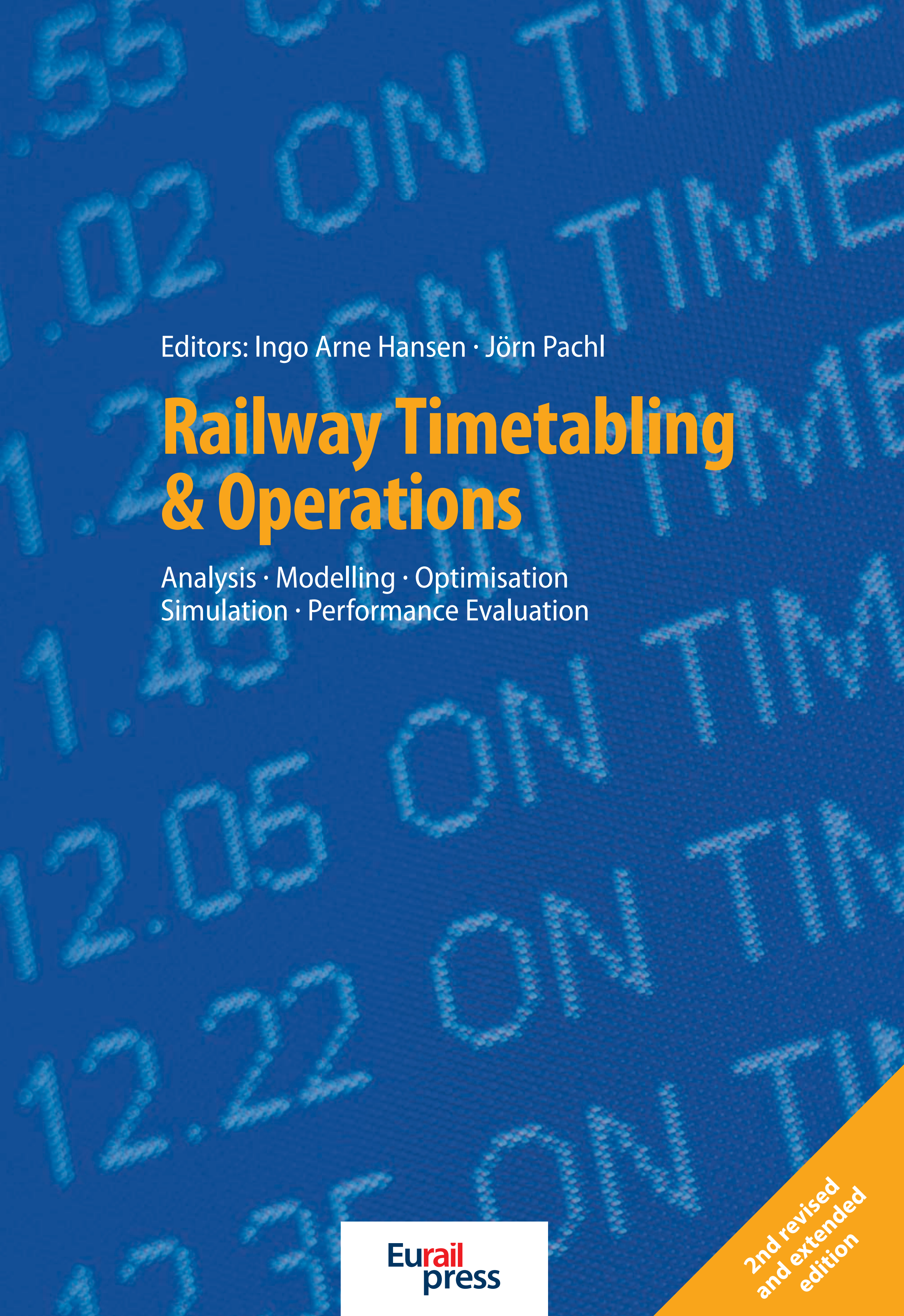


Editors: Ingo Arne Hansen · Jörn Pachl

Railway Timetabling & Operations

Analysis · Modelling · Optimisation
Simulation · Performance Evaluation

2nd revised and extended edition

Eurailpress

Railway Timetabling & Operations

Analysis · Modelling · Optimisation Simulation · Performance Evaluation

Editors
Prof. Dr.-Ing. Ingo A. Hansen, Delft University of Technology
Prof. Dr.-Ing. Jörn Pachl, Technical University of Braunschweig

Authors
Dr.-Ing. Thomas Albrecht, CSC Deutschland Solutions GmbH
Dr. rer. nat. Olaf Brünger, DB Systel GmbH
Dr. Gabrio Caimi, BLS AG
Dr. Francesco Corman, Delft University of Technology
Dr. rer. nat. Elias Dahlhaus, DB Systel GmbH
Dr. Andrea D'Ariano, Roma Tre University
Dr. Rob M. P. Goverde, Delft University of Technology
Prof. Dr.-Ing. Ingo A. Hansen, Delft University of Technology
Dr. Dennis Huisman, Erasmus University Rotterdam
Dr. Pavle Kecman, Delft University of Technology
Prof. Dr. Leo Kroon, Erasmus University Rotterdam
Dr. Gábor Maróti, VU University Amsterdam
Prof. Dr.-Ing. Ullrich Martin, University of Stuttgart
Dr.-Ing. Giorgio Medeossi, University of Trieste
Prof. Dr. Chris Nash, University of Leeds
Prof. Dr.-Ing. Nils Nießen, RWTH Aachen University
Prof. Dr. Dario Pacciarelli, Roma Tre University
Prof. Dr.-Ing. Jörn Pachl, Technical University of Braunschweig
Priv.-Doz. Dr.-Ing. habil. Alfons Radtke, Technical University of Braunschweig
Dr. Andrew Smith, University of Leeds
Dr. Robert Watson, Loughborough University
Dr. Phillip Wheat, University of Leeds
Dr. Jianxin Yuan, Arcadis Nederland BV

Eurail
press

Contents

1 Introduction

Ingo Arne Hansen

Timetabling, capacity analysis, traffic control and performance evaluation are basic elements of railway transport management, which are inherently linked. The function of the timetable in practice, however, is often not closely connected with its execution. Customers, train drivers, conductors, and dispatchers depend on specific sources of information, interpret the timetable and perceive the quality of train service differently. Further, the feasibility of train schedules and operations is subordinated to constraints of the infrastructure, power supply, signalling and safety systems, as well as rolling stock. For trip planning and decision making, railway passengers use live travel information increasingly and book via the internet rather than purchasing tickets at railway stations.

In many railways, however, train drivers and conductors still receive only a printed daily schedule. This will indicate the relevant departure, passing and arrival times, as well as the times and locations for the start, change and end of their duties, and a train radio or mobile phone for communication with the traffic control centres. Dynamic train delay information in the driver's cab has, so far, been introduced only by a limited number of train operators, while actual travel information in passenger coaches and at station platforms is generally made available. During the journey, drivers must observe the signals and cab displays, which indicate the necessary safety information with respect to their movement authority, permitted maximum speed, and in some modern trains the critical target distance. However, continuously updated actual delay information for train drivers is still mostly missing. A more precise and actual traffic information feedback to drivers and conductors would help them to better recognise deviations from schedules, to anticipate conflicts, and to raise operational performance.

The information made available to signallers, dispatchers and network supervisors in traffic control centres has improved significantly. This has been due to the introduction of electronic interlocking, associated with computer screen monitoring of the actual set-up of routes, track occupancy and sectional clearances for each train. Further, time-distance diagrams of the daily scheduled train paths and those actually used have become available at the newer traffic control centres. These are equipped with modern information, telecommunications and computer technology.

Advanced tools for timetable design, simulation and construction have been implemented in the past decades. These enable a much higher quality of timetabling and contribute to more reliable operations. However, the reliability and train punctuality levels achieved by railways, in general, do not give evidence of higher performance – except from Japan. It is hypothesised that the disappointing train punctuality performance of many networks stems only partly from the high traffic volumes and capacity usage. The main cause is an insufficient degree of optimisation, accuracy and robustness of the current practice of timetabling, which misses the feedback of detailed operational performance data.

This book intends to narrow the gap between the theory and practice of railway timetabling, capacity analysis, traffic management and performance evaluation. It does so by compiling the current state-of-the-art in this field and indicating a number of remaining problems and research issues. It is directed at academics, Masters and PhD students, professionals from the railway industry, and also public authorities that tender, contract and supervise railway services. Its aim is to enable readers to better understand the underlying theoretical models and to become acquainted with the actual state of technology.

The authors of the individual chapters are mostly members of the *International Association of Railway Operations Research* (IAROR)[1], founded in 2004, and were selected on the basis of expertise in specific areas. All are involved currently in railway research projects and have published individually in journals or international conference proceedings. The book forms a compendium of the main methodological aspects of railway timetabling, capacity and stability analysis, scheduling, simulation, real-time traffic management and performance evaluation, including some examples of calculation and practical application.

This book is an updated, partly revised and extended version of the previous edition *Railway Timetable & Traffic*, published in 2008. Completely new in this edition are Chapter 9 on macro-micro conflict-free scheduling in highly utilised networks, Chapter 10 Simulation, Chapter 12 Train Delay Prediction, Chapter 13 Rescheduling, and Chapter 15 Economic Performance Assessment.

The book is structured around the elementary steps of railway timetable design, modelling of infrastructure and train operations, and the analysis, optimisation, simulation, and evaluation of timetables and operations respectively. *Pachl* outlines in Chapter 2 the principles and variables of timetable design, and explains the principal methods for scheduling, modelling of train paths, train separation, specific signalling systems, as well as estimation of capacity use. In the following Chapter 3, *Radtke* deals with the models for description of the infrastructure based on graph theory, while showing the differences between node and link oriented, as well as macroscopic and microscopic models.

Chapter 4 contains a detailed explanation by *Brünger* & *Dahlhaus* of the calculation of train running times, depending on infrastructure and rolling stock characteristics, followed by a discussion of standard and advanced deterministic and stochastic calculation methods. *Albrecht* continues in Chapter 5 by describing the practice of energy supply and billing, and then explains the fundamentals and an intelligent tool for energy-efficient train control and driving, depending on the actual state of operations.

The following two Chapters focus on mathematical models for the estimation of waiting times, train delays, and delay propagation in networks. *Nießen* explains in Chapter 6 the fundamentals of queuing theory and its application, while *Goverde* presents in Chapter 7 a Petri net approach by means of the so-called max-plus algebra technique for stability analysis and delay propagation in large network timetables.

The next two Chapters are focusing on current optimisation methods for the design of large network timetables. *Kroon*, *Huisman* & *Maróti* describe in Chapter 8 linear programming approaches for optimal train scheduling in cyclic and acyclic network timetables. In Chapter 9 *Caimi* explains the bi-level approach applied in Switzerland for generating optimal conflict-free schedules in highly utilised networks through macro- and micro scheduling, as well as distinction of condensation and compensation zones.

Watson & *Medeossi* describe in Chapter 10 the role of simulation, discuss the key differentiators of simulation methods and give a comprehensive overview on the development and application of current simulation models for timetable design and train operations.

In Chapter 11, *Yuan* & *Medeossi* present suitable approaches for the detailed statistical analysis of train delays and validation of running and dwell times based on standard track occupation and release data, including goodness-of-fit test and estimates for the distributions and its parameters.

---

[1] www.iaror.org

Kecman & *Goverde* introduce in Chapter 12 a recently developed automatic data-driven traffic state and route conflict prediction model based on train describer, track occupation and clearance records. Chapter 13 consists of a review of the few existing models for the rescheduling of trains in case of service irregularities and disruptions. *Pacciarelli*, *D'Ariano* & *Corman* distinguish centralised from co-ordinated approaches and describe in detail the architecture, main modules and procedures of their real-time rescheduling model Railway traffic Optimisation by Means of Alternative graphs (ROMA).

In Chapter 14 *Martin* discusses the methods and key indicators for performance evaluation of railway operations, capacity use of networks and standard cost-benefit assessment of infrastructure investments. *Wheat*, *Nash* & *Smith* describe in the following Chapter 15 the inputs and outputs for assessing the economic performance of railway systems through productivity analysis, measurement of technical efficiency and cost efficiency.

Finally, the Editors of this book, *Hansen* and *Pachl*, derive the main conclusions concerning the current state-of-knowledge on railway timetabling and traffic management from the preceding Chapters. They then summarise the challenges and main topics for future research. Their aim is to encourage readers of the book to contribute further to a better comprehension of railway processes and to improve the attractiveness and efficiency of train services for railway customers.

The approaches discussed cover the main current state-of-the-art theories and analytical, combinatorial optimisation, simulation and economic models for railway timetabling and operations world-wide. This book is a German-Dutch-British-Italian co-production, where the academic and professional quality of railway operations research and education, and the application of railway technology, are amongst the best in the world. The Editors hope that researchers and professionals from any country will enjoy the content and use it for improving the analysis, modelling and performance evaluation of railway timetabling and operations on their own.

2 Timetable Design Principles

Jörn Pachl

2.1 The Purpose of Scheduling

Scheduling performs the following functions:

- It co-ordinates the train paths in the planning process for optimum use of the infrastructure.
- It ensures the predictability of train traffic.
- It produces timetable data for passenger information.
- It is essential for traffic control, locomotive and rolling stock usage and crew scheduling.

In passenger operations, scheduling is essential to provide predictable travel times for passengers. The safe operation of trains also requires the co-ordination of slots. In freight operations, the situation is different. Due to changing demands from the shippers, establishing long term timetables is often not possible. Some freight operators abolished timetables completely and run their trains entirely on-demand with priority-based train dispatching. Train crews just get a “timetable” that does not contain any times, but only the data required for safe train driving (e. g. speeds, braking conditions). This is typical for North American railways but also for a number of mining railways in other parts of the world. European freight operators are also facing changing demands. In the past, this problem was solved by establishing on-demand paths in the yearly timetable. These paths could be used to run scheduled freight trains on a daily basis.

Today, computer-based scheduling allows the principle of pre-constructed on-demand paths to be replaced by running an increasing number of freight trains as extra trains. However, running “extra” does not mean running “unscheduled”. An extra train has to be scheduled as accurately as a regular train, but on a more dynamic basis. On many railways, a freight train operating company can order a train path for an extra train just a few hours in advance. Compared with completely unscheduled operation, this kind of flexible scheduling will lead to a much higher degree of predictability of traffic. On lines with a mixed traffic of freight and passenger trains, it will also ensure the required quality of service.

On railway networks that are operated on an open access basis, scheduling is not just a planning procedure but also the commercial interface between infrastructure operators and train operating companies. There, the scheduled train path is the product sold by the infrastructure operator to the train operating company concerned. Scheduling means to coordinate the train paths ordered by competing train operating companies. Assigning a train path to a train operating company means to sell the right to run a train on that path under specified operating conditions.

2.2 Basic Terms of Railway Operation

For the explanation of scheduling aspects, an understanding of the basic terms of railway operation is required. Unfortunately, railways worldwide use very different terminologies which may cause misunderstandings in international use. There is especially a big difference between British and North American railway terms. This textbook is based mainly on the terms used in *UIC* Code 406 (*UIC*, 2004) which is neither pure British nor pure American. However, at some points, significant differences between the terms of *UIC* Code 406 and the terms used in practical operation of British and North American railways are mentioned. For more detailed explanation of the fundamentals of railway operation see *Pachl* (2009).

Organisations

The effect of EU Directives has been to separate the operational and infrastructure sides of the railway businesses, in the EU but also in some other countries. This does not, of itself, require separation of ownership. It has, however, introduced the concept of a track authority with a network management, which grants access to operators of passenger and/or freight traffic on payment of a fee. With rules which may be determined differently in the various countries concerned, it has to grant track access as may be demanded by qualified and licenced operators, to the extent that this is practical. Hence, efficient timetabling has become of even more major importance in recent years.

The term open access needs to be treated with caution, since it also requires suitable train paths to be available. In time, that may result in consequential changes to the timetable already in place and hence to the services of existing operators. That in turn may have an effect on the contractual arrangements between those operators and the track authority. Alternatively, should conditions so dictate, the request may have to be refused.

Train operators may be classified generally as Train Operating Companies (TOCs), though it should be noted that in Britain this term does not include freight operators.

Tracks, Routes and Stations

Tracks are the routeways of a railway system. A railway network in *UIC* terms consists of nodes connected by links. In *UIC* Code 406 the links between nodes are called lines. Consecutive lines and nodes as a whole, between a defined source and target, are called a route (in a second meaning, the term route is also used for a route through an interlocking or point zone). In its most simple form, a node is just a junction where trains can transfer from one line to another. More complex nodes are extended station areas in which overtaking, crossing, or direction reversal of trains is possible. A terminal is an assemblage of facilities provided at a terminus or intermediate point of a route for the purpose of shunting, assembling, sorting, and marshalling trains. A typical freight terminal consists of track groups for arrival and departure of trains, a shunting district with yard facilities for sorting and storing of wagons, and connecting tracks to industrial sidings. A typical passenger terminal consists of a group of platform tracks and some yard facilities for storing passenger vehicles. In large terminals, there are also often servicing facilities for the rolling stock. Large terminals are often located at the intersection or junction of different routes.

In Britain, the term track is mainly used when talking about the track itself, or the permanent way. In a station, the term “platform” usually includes reference to the track on which the train runs at that point. Thus a train is said to be standing at platform 2, not track 2.

In railway operation, tracks are often referred to as lines. For operational purposes, tracks are divided into two classes. Running lines are the tracks on which trains move through the network. Running lines must be equipped with signalling appliances required for regular train movements. A loop is a running line that is used for passing and overtaking trains. Sidings are lines used for assembling trains, storing vehicles and trains, loading and unloading and similar purposes, but not for regular train movements. A track layout consisting of several sidings is called a yard. In North America, a line is not just a track but may consist of several parallel tracks. Running lines are called main tracks, and loops are called sidings. In a literal translation, the term main track is also used by many railways worldwide, e.g. in Continental Europe. Since the British use of the term line differs from *UIC* Code 406 and may lead to misunderstandings, the term main track is preferred against the term running line in the sections which follow.

In *UIC* Code 406, the term station is only used for the points of a network where overtaking, crossing, or direction reversal is possible. Many railways refer to all places designated as stopping places by name in the timetable as stations. A passenger station is that with one or more platforms for scheduled stops of passenger trains. It is not necessarily associated with a station in the sense of *UIC* Code 406.

References made to interchanges between services at stations may, according to context, include those to and from other modes, notably bus. Again, contractual elements may have a bearing on how problems such as late running by any party should be resolved.

Track elements

A turnout is an assembly of rails, movable points, and a frog, which effect the tangential branching of tracks and allows trains or vehicles to run over one track or another. The term turnout is mostly used in civil engineering. In railway operation and signalling, a turnout is usually referred to as a pair of points (points in short), although this term in its original meaning only applies to the part of a turnout where the points are located. On North American railways, this part of a turnout is called a switch, and the point machine is called a switch machine. In North American railway operation and signalling, turnouts are usually referred to as switches.

A crossing is an assembly of rails that enables two tracks to cross at grade. The inner part of a crossing is called a “diamond”. Crossings with a large angle of intersection are constructed rigidly. In case of a small angle of intersection, a crossing may have movable points instead of frogs. Small angle crossings may be equipped with additional points providing a slip connection to permit movements from one track to another. A crossing with a slip connection at one side is called a “single slip”, and a crossing with slip connections at both sides is called a “double slip”. On North American railways, slip crossings are called slip switches.

Derailers are trackside devices that are used to protect train movements against unattended movement of vehicles on converging tracks. An unsafe movement is derailed before it could join the protected route. In the protecting position, a derailing piece is raised over one rail. On many railways, derailers must not be installed in main tracks. Instead of derailers, some railways also use trap points which have much the same effect.

Running and Shunting Movements

Trains consist of a minimum of a locomotive on its own but more generally with coaching stock for passengers, or freight wagons for goods. Other trains may consist of electric or diesel multiple units, and specialist works trains for engineering purposes. The term wagon in this book may or may not include reference to passenger vehicles.

Running movements are locomotives with or without wagons, with authority to operate on a main track under the control of main signals. The operating conditions for a running movement (maximum speed, braking conditions) are specified in the working timetable. On lines with non signal-controlled operation, a running movement needs a movement authority from the operator responsible for controlling traffic on main tracks. Every running movement displays rear end markers (tail lights in Britain) to enable the lineside staff to check the train’s integrity. On European railways, running movements are always scheduled movements.

Shunting movements are those for making up trains, moving vehicles from one track to another, and similar purposes. Shunting movements are accomplished under simplified requirements at a very low speed that allows stopping short of any vehicle or obstruction. Shunting units may enter occupied tracks. Movements in industrial sidings are also carried out as shunting movements. Shunting movements are not scheduled when establishing the timetable, but the need for them and the time necessarily taken must be considered.

Train runs is not a term normally found in Britain; used here, it refers to the complete end-to-end journey undertaken by the train(s) concerned.

In North America, the sharp distinction between running movements and shunting movements does not exist. All movements with authority to run on a main track are called trains. For more details see *Pachl* (2009) and *White* (2003).

Signals and Interlockings

Lineside signals indicate if a movement may enter the section of track beyond the signal. There are different signals for running and shunting movements. A main signal indicates if a running movement may enter a section of line. In a territory with a fixed block system (see Section 2.5.1), lines are divided into block sections for the purpose of safe train separation. A train must not enter a block section until that section has been cleared by the train ahead. In fixed block territory, running movements are authorised by signal indications. Except on lines with a low speed, a signal that authorises a running movement requires an approach aspect at the braking distance in approach to the signal, because the stopping distance is generally longer than the range of vision.

In a territory where the distance between signals does not exceed the braking distance very much, the approach aspect is usually provided by the signal in rear. This is called multiple-section signalling, since a main signal provides information for at least two following block sections. In a territory with very long distances between main signals, distant signals are placed at the braking distance in the approach to a main signal (Fig. 2.1). A distant signal can only provide an approach aspect for the signal ahead but it cannot show a stop aspect. This signalling principle is called one-section signalling, since a main signal can only provide information for one block section.

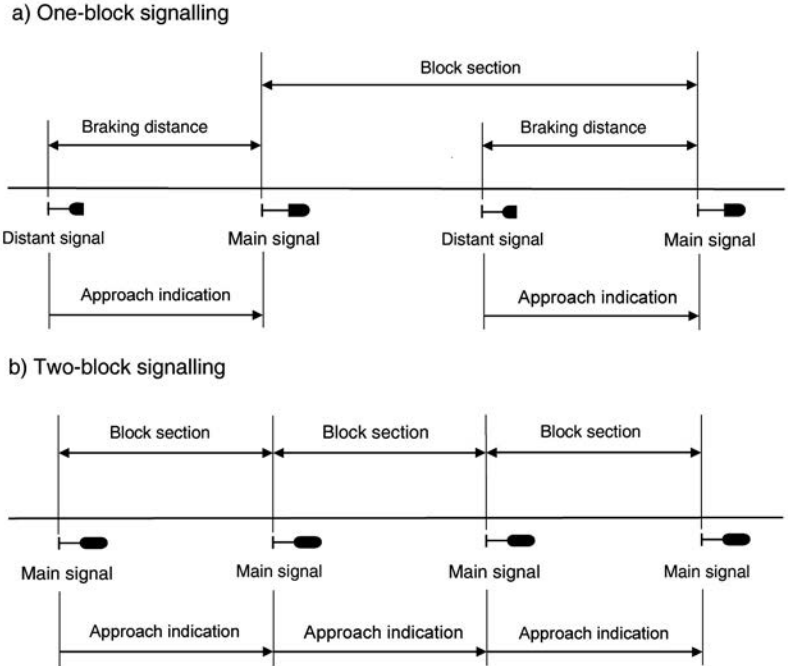


Fig. 2.1: Block Sections limited by Main Signals

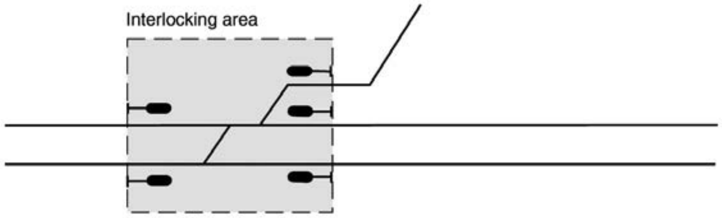
Shunting signals are used to authorise shunting movements and to protect running movements from shunting movements. In a territory where shunting movements may enter main tracks, a shunt aspect is also integrated into the main signal, so that shunting movements may be authorised to pass main signals in the stop position. For shunting signals, an approach aspect is not provided because shunting movements are carried out at a very low speed that allows stopping short of any vehicle or obstruction.

On North American railways, there is no differentiation between main and shunting signals because of different operating rules.

Interlocking is a description of the arrangement by which points and signals are electrically or otherwise interconnected in a way that each movement follows the other in a proper and safe sequence. Signalled routes for trains on main tracks are usually interlocked. Signals that govern train movements through an interlocking may be called interlocking signals. An interlocking signal can also limit a block section. The points and signals within the interlocking are controlled either by a local interlocking station or from a remote control centre. Local interlocking stations are called signalboxes or signal cabins on railways that follow British principles, and interlocking towers in North America. The main signals between controlled interlockings are often called intermediate block signals. On most railways, the signals that limit an interlocking are called home signals.

There are two basic signal arrangements at an interlocking (Fig. 2.2). First, there are interlockings without consecutive interlocking signals within the home signal limits. The home signals provide the authority to run through the entire interlocking into the next block section. Second, there are interlockings with consecutive interlocking signals. The section of line within the home signal which is limited by consecutive interlocking signals is often called a station track, which is separated from the sections of the so-called open line. Although many

a) Interlocking without consecutive interlocking signals



b) Interlocking with consecutive interlocking signals

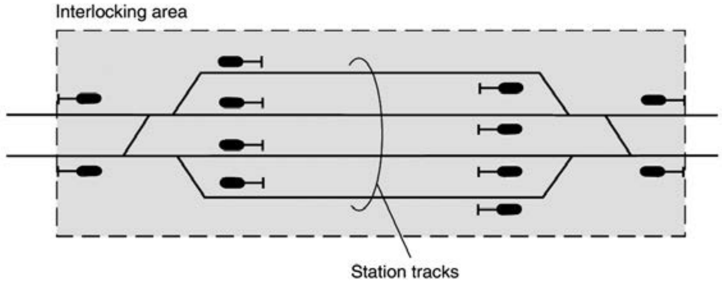


Fig. 2.2: Different Signal Arrangements at Interlockings

railways do not refer to these station track sections as block sections, they have the same effect on train separation as a block section. Different from block sections of the open line, station tracks may be used for overtaking, crossing, and direction reversal of trains. A station platform for scheduled stops of passenger trains is not necessarily associated with a station track which need not have a platform. There are also stations with their associated platforms on the open line, which do not necessarily imply any provision of pointwork or the ability to reverse trains within the signalling system.

On North American railways, interlockings with consecutive interlocking signals are not common. The reason is that the distinction between station tracks and tracks of the open line does not exist in the North American rules. In modern British signalling centres, there is also no longer any distinction between station tracks and sections of the open line. However, from the viewpoint of signalling, there is still a difference between track sections on which trains are protected by controlled interlocking signals and track sections on which trains are controlled by automatic block signals.

2.3 Diagramming Traffic

On most railways, traffic diagrams are used both as the basis of all planning of railway traffic and also as essential documents for the control of the current operation. The only exceptions are the North American railways where traffic diagrams (there called “stringline graphs”) are used for capacity analysis and in very early stages of operation planning, while in current operation tabular sheets are preferred.

The amount of traffic on a line is described in form of a time-distance diagram that consists of a time axis and a station axis. At every station, there is a line parallel to the time axis. This location represents the point where scheduled times for that station apply. Train movements are represented by train paths (time-distance graphs) with a train description inscribed on them. At the intersections of train paths and station lines, the actual time is marked by minute numbers (Fig. 2.3). Traffic diagrams may be shown with a horizontal or with a vertical station

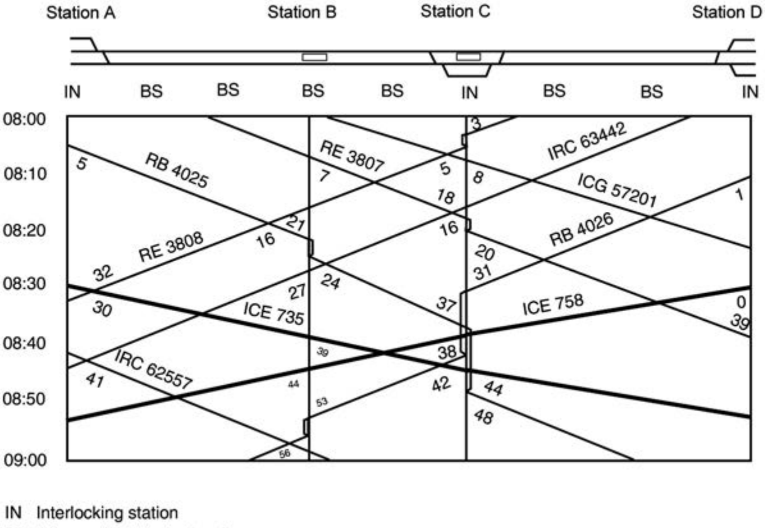


Fig. 2.3: Principle of a Traffic Diagram (Double Track Line)

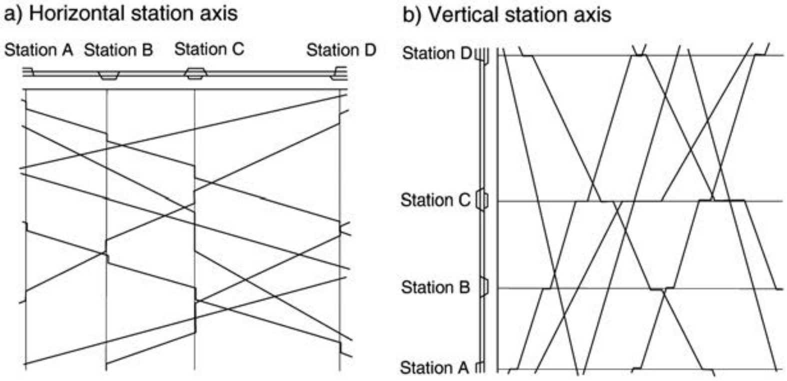


Fig. 2.4: Traffic Diagram with Horizontal and Vertical Station Axis

axis (Fig. 2.4). Information of the two types of diagrams is of the same value. However, for different railways, it is a matter of tradition as to which is used. In addition to the traffic diagram of a line, additional diagrams are used to describe the traffic inside stations or interlockings where different tracks are available per direction. In such a station traffic diagram, every track is represented by a straight line parallel to the time axis.

The track occupancies are marked by stripes which have the description, arrival, and departure times inscribed on them (Fig. 2.5). Small bent pieces at the end of the stripes are suggested time-distance graphs to show the direction of movement. On small layouts, station timetables are often used instead of those more complex station traffic diagrams. On some railways, the occupation of station tracks is integrated in the line traffic diagrams. For smaller intermediate stations, this is sufficient. For larger terminals, a separated station traffic diagram leads to a better overview.

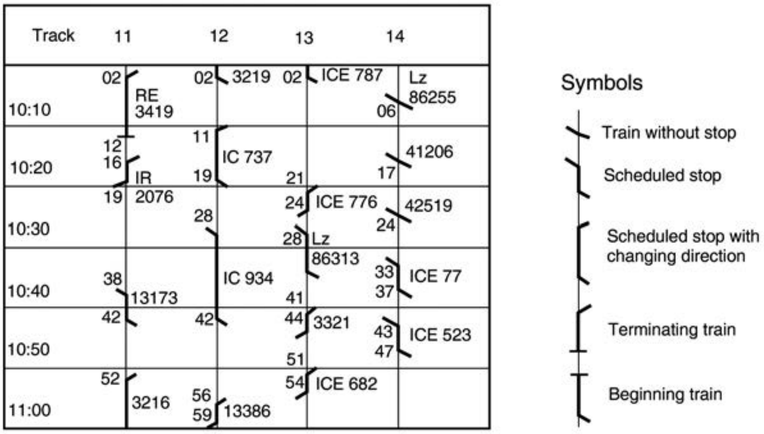


Fig. 2.5: Station Traffic Diagram

2.4 Scheduled Running Time

The scheduled running time of a train consists of the following components:

- The pure running time between scheduled stops
- The dwell time at scheduled stops
- Recovery time
- Scheduled waiting time

The pure running time between scheduled stops is the shortest possible running time as a result from a running time calculation (see Chapter 4). To enable a train to make up small delays, recovery time must be added to the pure running time. There are two kinds of recovery time:

- Regular recovery time
- Special recovery time

The regular recovery time is added to every train path as a percentage of the pure running time. The typical time supplement is 3%–7% on European railways and 6%–8% for passenger trains on North American railways. On some railways, the regular recovery time is evenly spread over the train path while other railways prefer to concentrate the recovery time at the end of the run and at large intermediate terminals (Fig. 2.6). Sometimes, at large terminals, the recovery time is not added to the running time in the approach to the terminal but to the dwell time within it.

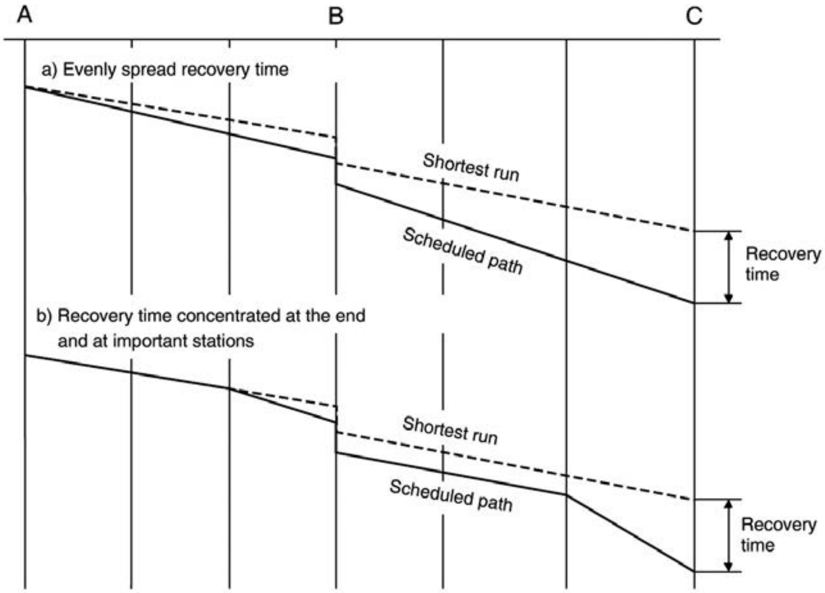


Fig. 2.6: Principles of Adding Recovery Time to a Schedule

The special recovery time is used to compensate the influence of maintenance and construction works and of sections with temporarily bad track conditions. Different from the regular recovery time, it is not added as a percentage of the running time but as a fixed supplement to the running time on the section concerned. Scheduled waiting time is added for scheduling reasons, e.g. to synchronise schedules of different passenger lines at changing points, to synchronise schedules to a cyclic timetable (see Section 2.9), and to wait for a scheduled passing or overtaking. Scheduled waiting times are mostly added to the dwell time of scheduled stops but sometimes also to the running time.

2.5 Modelling Train Paths

A train path describes the usage of the infrastructure for a train movement on a track and in time. As a basic requirement, a timetable must not contain any schedule conflicts between train paths. It is not sufficient to describe a train path just by its time-distance graph. There must also be a model to describe the “time channel” a train movement produces around its time-distance line. A model that describes that time channel very exactly is the so-called blocking time model. The same principle is also well known in analytical models for capacity management, see *Pachl & White* (2003), *Pachl* (2009), *Pachl* (2013).

The idea of the blocking time model was developed in the late 1950s by *O. Happel*, who was at the time the chief scheduling officer in the Frankfurt headquarters of German Federal Railways (Deutsche Bundesbahn), see *Happel* (1959). The term blocking time is a literal translation of the German term “Sperrzeit” which was introduced by Happel. Independently from Happel, a similar method was developed by G. Adler at the former Dresden College of Transportation, see *Adler* (1967). Implementation of that idea in practical railway scheduling lasted several decades until computer-based scheduling systems became available. Use in daily operation required yet more time and computer system development. The German infrastructure operator DB Netz introduced a computer-based scheduling system based on the blocking time model in 1998, see *Brünger* (2000). Beside the German scheduling system, the blocking time model is also used in scheduling systems of other European railways, e. g. in Austria, Denmark, Sweden, the United Kingdom and Switzerland.

2.5.1 Principles of Train Separation

Correct application of the blocking time model requires a basic understanding of the principles of train separation. In highway traffic, separation of vehicles is effected by the principle of relative braking distance. If one vehicle brakes, the driver of a following vehicle will notice the brake rear lights and start braking. Two vehicles can follow in the distance that equals the difference of the braking distances of the vehicles plus an additional safety distance to ensure sufficient reaction time for the driver of the second car.

In a steel wheel on steel rail system, the coefficient of adhesion is on average eight times less than in highway traffic. As a result, the braking force that can be transmitted from a train to the track is also eight times less than the braking force that can be transmitted from a highway vehicle to the road surface. That leads to braking distances for railway vehicles which may significantly exceed the viewing range of the driver. Thus, train separation by the sight of the driver is only possible when running at a restricted speed (usually no more than 25 km/h). This is only acceptable for shunting movements and for running movements in degraded mode operation. For regular running movements, procedures of train separation are required that work independently from the viewing range of the driver.

The principle used for safe train separation depends on the following criteria:

- how movement authority is transmitted from track to train,
- how the line is released for further traffic behind a train.

If movement authority is only transmitted at discrete points, e.g. at lineside signals, or by written or verbal orders, this will necessarily lead to train separation in a fixed block distance since each movement authority has to cover the entire section up to next point where further authority may be received. This principle is in effect on all lines where trains are governed by fixed lineside signals and protected by an intermittent ATP system. This works for all kinds of intermittent ATP systems including the Level 1 of the European Train Control System ETCS.

On lines where train separation in block distance is used, the track is divided in block sections. A block section may be occupied exclusively by only one train. In a signalled fixed block operation, the block sections are limited by signals which provide movement authority to enter the block section protected by the signals. To clear a signal for a train that is to enter a block section, the following conditions must have been fulfilled:

- The train ahead must have cleared the block section.
- The train ahead must have cleared the overlap beyond the next signal (only on lines where block overlaps are used).
- The train ahead must be protected from following train movements by a stop signal.
- The train must be protected against opposing movements.

On railways where block overlaps are not required, the control length of a signal equals the block section (Fig. 2.7 a). Examples are mainline railways in North America and in Russia. Other railways require a control length of a signal that is longer than the block section (Fig. 2.7 b).

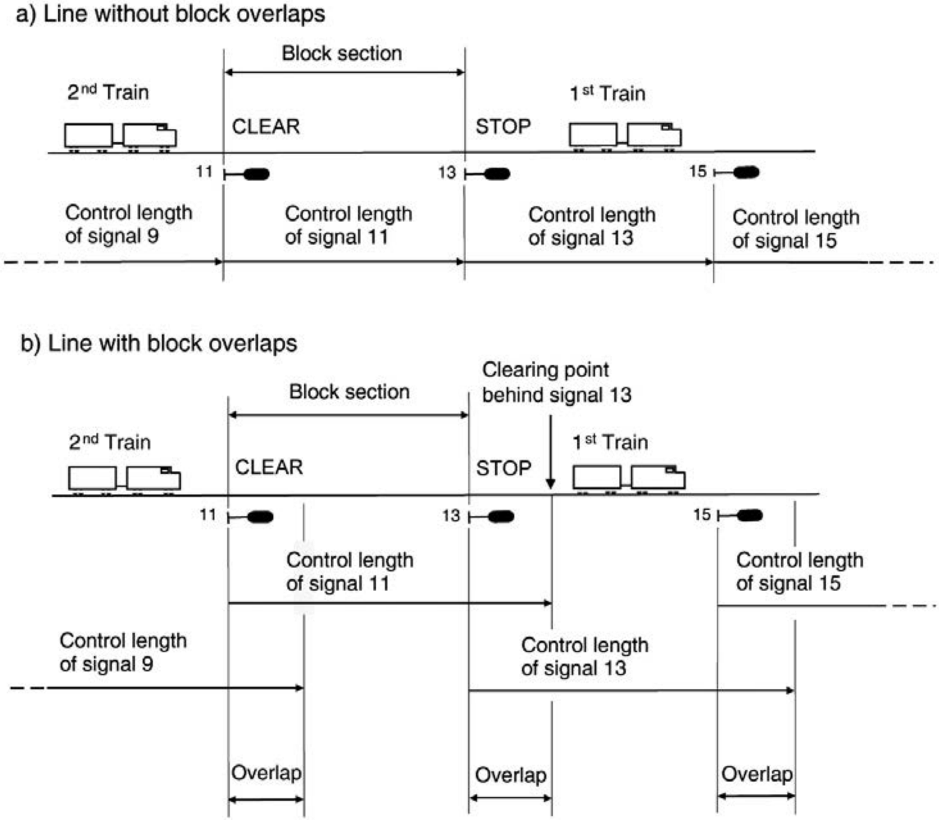


Fig. 2.7: Control Length of Signals in Fixed Block Territory

The difference is called “overlap” because in that area the control length of a signal overlaps with the control length of the next signal. Overlaps are a common safety feature on railways with high density passenger operation. The purpose of overlaps is to protect train movements against trains that are overrunning a stop signal by a short distance due to bad brake handling. In North America, overlaps are typical for subways and heavy rail rapid transit but they are not normally used on standard railways. In Europe, where passenger operation is the backbone of the railway system, overlaps are also frequently used on standard railways. Since a main signal can only be cleared when both the block section and the overlap behind the next main signal are clear, a train that is approaching a stop signal will always have a clear overlap beyond that signal.

On lines where trains are governed continuously by a cab signal system, lineside signals are no longer needed. However, continuous transmission of movement authority is not yet a sufficient criterion to abolish fixed block sections. In addition, the train has to release the track not at fixed intervals but continuously. This requires a permanent train-borne checking of train integrity. Since for traditional railway systems a sufficient solution for that problem has not yet been found, train separation in fixed block distance is still the standard principle for safe train spacing on most railways worldwide.

From the viewpoint of train separation, the essential benefit of a fixed block system with cab signalling compared with a fixed block system with lineside signals is the independence of the cab signals from the approach distance of the lineside signal system, i. e. the distance between the signal at the entrance of the block section and the signal in rear that provides the approach indication. This allows trains to run at higher speeds. That is why cab signalling is today the standard system on all high speed lines, worldwide.

On most railways, cab signalling is combined with a continuous ATP system. The most common systems are the German LZB which uses a cable loop track antenna, systems based on coded track circuits (e. g. in the French TVM system), and systems based on digital radio as used in ETCS Level 2. However, on some railways, e. g. in Russia and North America, cab signalling is also used on conventional lines as a pure signalling system without brake enforcement.

Cab signalling without fixed block sections is generally known as moving block. This term expresses the fact that fixed block sections no longer exist or, in other words, the block length is reduced to zero. Since the train location data is transmitted at short intervals, the time interval for safe train location is part of the minimum headway for two successive trains. Due to the unsolved problem of train-borne checking of train integrity, moving block systems are not yet used in standard railway operation but on a couple of transit railways. However, with respect to future applications, the ETCS specification contains a Level 3 which is based on moving block.

2.5.2 Application of the Blocking Time Model

The blocking time is the total elapsed time a section of track (e. g. a block section, an interlocked route) is allocated exclusively to a train movement and therefore blocked for other trains. Therefore, the blocking time of a track section begins with issuing a train its movement authority for this section (e. g. by clearing a signal). The movement authority must be issued before the train has reached the braking distance in its approach to this section. For example, in signalled operation the train must not yet have passed the signal that gives the approach indication to the signal at the entrance of the section. The blocking time ends after the train has completely left the section and all signalling appliances have been reset to normal position so that movement authority can be issued to another train to enter the same section. Thus, the blocking time of a track section is usually much longer than the time the train occupies the section. In a territory with lineside signals, the blocking time of a block section for a train without a scheduled stop at the entrance of that section consists of the following time intervals (Fig. 2.8):

- The time for clearing the signal
- A certain time for the driver to view the clear aspect at the signal that gives the approach indication to the signal at the entrance of the block section (this can be a block signal or a separate distant signal)
- The approach time between the signal that provides the approach indication and the signal at the entrance of the block section
- The running time between the block signals
- The time to clear the block section and – if required – the overlap with the full length of the train
- The release time to "unlock" the block system

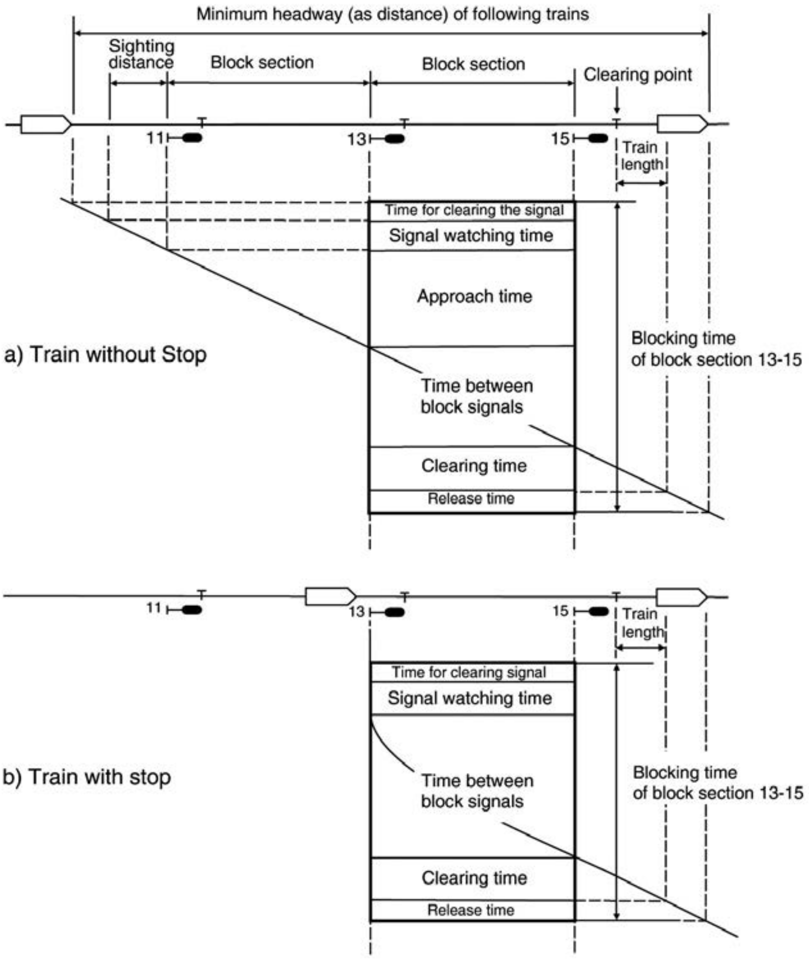


Fig. 2.8: Blocking Time of a Block Section

The performance of many railway networks and the quality of service offered is becoming more and more critical. The main issues to be addressed are the increasing traffic volumes and making the best use of the available capacity, at the same time resolving train scheduling and management problems.

This is an updated, revised and extended edition of 'Railway Timetable & Traffic', published in 2008. It describes the state-of-the-art methods of railway timetabling and optimisation, capacity estimation, train operations analysis and modelling, simulation, rescheduling and performance assessment. The intention is to stimulate their broader application in practice and to highlight current and future research areas. It is directed at academics, Masters and PhD students, as well as professionals from the railway industry. It will also be of interest to the public authorities that tender, monitor and perhaps fund railway service provision. The overall aim is to improve the attractiveness and efficiency of the train services which can be offered to the public.

The key to achieving a higher efficiency and quality of train operations is an awareness of the impact of availability, reliability and robustness of the subsystems on train processes. A deeper insight into the probability of incidents and the propagation of train delays depends on a thorough analysis of real world railway operations and the feedback obtainable. This leads to an optimisation of the timetable and a network-wide improvement in traffic management performance.

This know-how should increase the efficiency of the railway system, making it more attractive for regular, occasional and new customers, and ensure that the railways continue to innovate. They will then be able to make the maximum contribution possible to the transport needs of the future.

ISBN 978-3-7771-0462-1

9 783777 104621