

Transport capacity constraints on the mass transit system: a systemic analysis

Fabien Leurent

Received: 16 April 2010 / Accepted: 13 January 2011 / Published online: 7 May 2011
© The Author(s) 2011. This article is published with open access at SpringerLink.com

Abstract A public passenger transit system is complex, because it involves two types of traffic unit, passengers and vehicles. The system operations require supplying passengers with stations where they can access vehicles and then travel in them between stations. These operations are subject to various capacity constraints, depending on the traffic type and the situation in the system. The article provides a framework for the analysis of a mass transit system, broken down into four subsystems that respectively pertain to the Passenger, the Vehicle, the Station and the Line. Each subsystem consists of several components that are involved in the system processes and interactions. Within this framework, capacity phenomena are identified, described qualitatively and classified into seven broad categories, namely: (i) the vehicle capacity of an infrastructure; (ii) the vehicle fleet; (iii) the passenger capacity of a vehicle; (iv) the passenger capacity of a route; (v) the passenger capacity of a station; (vi) the vehicle storage and movement capacity of a station; (vii) the capacity of a station for interface with personal transport modes. Moreover, it is shown that capacity phenomena can interact in congestion gears, which impair capacity. Lastly, there are complex retroactive effects that involve either network management or demand behaviour (through route choice which depends on quality of service hence on congestion).

Keywords Passenger transit · Capacity constraints · Congestion gear

F. Leurent (✉)
Université Paris Est, Laboratoire Ville Mobilité Transport,
Ecole des Ponts ParisTech,
6-8 avenue Blaise Pascal, Cité Descartes, Champs sur Marne,
77455, Marne la Vallée Cedex 2, France
e-mail: fabien.leurent@enpc.fr

1 Introduction

1.1 Background: capacity in mass transit systems, a complex notion

The purpose of a public passenger transport mode is to carry flows of people large enough to generate economies of scale and scope on one or more line services (i.e. transit routes). In relation to fixed resources and a quality of service objective, transport capacity is the maximum volume of flow that can be handled in standard conditions for a limited period. Since the service requires the combination of an infrastructure, a vehicle and a traffic protocol, capacity depends on the arrangements for each of the components: for instance the size and passenger capacity of a vehicle, or the fleet of vehicles available to run over a period. However, every vehicle needs to complete its route, which limits its availability at each spot. Moreover, in each station the flows of passengers boarding and alighting influence the vehicle's dwelling time, hence the time it takes to complete its route. In other words, flow rates and configuration in space influence local capacity.

1.2 The issue: identifying local capacity in order to plan the network

In planning the mass transit network, the aim is to match the nature and size of the operating resources to the volume of the flow, subject to certain service quality objectives. The volume of flow needs to be anticipated over the long term, since implementing additional infrastructural resources is a process requiring several years of design, consultation and construction.

The capacity of existing or planned means of transportation also needs to be known, so that it can be adjusted to

the likely volume of trip flows. Insufficient capacity not only results in local queues, causing passengers discomfort and wasting their time, but also vehicle delays that give rise to passenger delays and reduce the vehicle's productivity and availability for its next run.

In public transport, lack of capacity is a factor of congestion, which can trigger a disastrous gear when several forms of congestion combine and reinforce each other. In cities, the operator has few options for regulating passenger flows in the short term: it has to provide transport for all passengers, with no booking system and almost no pricing leverage. It therefore needs to ensure that it has sufficient capacity, without excessive overcapacity but with sufficient reserves.

1.3 Objective: to identify capacity constraints

The aim of this paper is to identify, qualify and classify the capacity constraints for a mass transit network. Such a network contains two types of moving body, passengers and vehicles. Each type is subject to constraints in its movements, but also in its storage in a given place. In addition, operations are restricted by the availability of a limited fleet of vehicles—amongst other resources needed to provide the service. We wish to identify the principal capacity constraints, and describe the nature (the resource limited) and form (mechanism, law) of each of those constraints, with specific factors and conditions and specific consequences.

Each capacity constraint combines several factors or conditions, through composition in space: in principle, this combination is local, but it can also be subject to influence from elsewhere, since both vehicles and passengers are mobile.

The paper provides a framework for the analysis of a mass transit system, in which each component is placed in its specific context in interaction with other components. We situate capacity constraints and congestion phenomena within this framework, each with its principal factors and principal effects. The framework thus makes it possible to identify not only the constraints but also their reciprocal influences and their potential to trigger and reinforce each other.

This framework is intended for engineers and planners involved in the operation and planning of public transport systems, for consulting engineers who apply network simulation models in the design of development projects, and for scientists that work out such simulation models.

1.4 Method: a systemic analysis

The stationing and movement of people within and around buildings and other organised spaces are primarily of

interest to architects: that is why Neufert [10] and others (i.e. [8]) determined the main parameters of the use of space by people at rest or in movement.

In traffic science the time aspects, including durations and speeds, are also considered to describe the kinematics of an isolated moving body on an infrastructure, and the interactions between moving bodies, in particular the laws of congestion. Road traffic theory has developed substantially since the 1950s: its engineering applications are covered in the *Highway Capacity Manual*, produced under the aegis of the *Transportation Research Board* (TRB) and regularly updated since its first edition in 1950 [14]. The equivalent manual for public transport is the *Transit Capacity and Quality of Service Manual* (TCQSM), also commissioned by the TRB but first published in 1999, which shows the relative youth of this scientific field. The manual provides a quasi comprehensive description of the technical aspects and operating principles of the different transit modes, and the associated capacity constraints. Aimed at traffic engineers, the manual contains application instances and provides criteria for assessing the performance of an operational technique. In this article, citations in the form [X.Y-Z] refer the reader to chapter Y and page Z in part X of the second edition [15].

Our own method is theoretical, seeking to identify the generic concepts underlying the forms specific to a given mode or system and thereby to enable a particular system to be analysed more concisely and effectively. By way of such abstraction, it is hoped to put traffic engineers' knowledge into a more scientific shape, thereby facilitating the work of researchers modelling public transit (PT) networks. The network modelling literature makes another reference source, particularly with regard to the modelling of traffic assignment to routes on a PT network and for the analysis of certain capacity constraints.

To sum up, a framework is provided for the analysis of a public passenger transit system, situating capacity constraints within it. Our analysis is qualitative, characterising the “mechanisms” in terms of factors, specific conditions and the consequences associated with them, but providing no mathematical formulation. Certain formulations are provided in the TCQS Manual, others in the traffic modelling literature, but this is a scientific field that still remains to mature.

1.5 Structure

The body of the paper consists of three main parts, followed by a conclusion. Section 2 provides the framework of analysis for a mass transit system, identifying four subsystems relating respectively to: the Passenger, the Vehicle, the Station and the Line. In Section 3, seven categories of capacity constraint are described: (i) the

vehicle capacity of an infrastructure; (ii) the vehicle fleet; (iii) the passenger capacity of a vehicle; (iv) the passenger capacity of a route; (v) the passenger capacity of a station; (vi) the vehicle movement and storage capacity of a station; (vii) the capacity of a station for interface with individual modes of transport. Section 4 covers the composition of resources in the operation of public transport, interactions between capacity constraints and the feedbacks by users or the operator on congestion mechanisms.

2 Systemic analysis of a mass transit system

A mass transit system is made up of technical resources, notably infrastructures and vehicles, which are “components” involved in operation of the service. In addition, the passenger impacts on the system, firstly in an aggregate way in the form of flows, and secondly in a disaggregate way as an individual actor possessing intentions and a pattern of travelling behaviour.

2.1 System overview

Let us analyse public passenger transport as a system composed of subsystems relating respectively to the Passenger, the Vehicle, the Station, the Line and Global Management. The *Global Management* subsystem covers the rules for access and use of the system by passengers; pricing and payment controls; various technical or legal rules, notably relating to operating safety, power supply, working conditions, etc. Rules like this are sophisticated, established for a medium-term timeframe and at a high decision-making level, distinct from day-to-day operating concerns. That is why this subsystem is excluded from the analysis below.

The *Passenger* subsystem covers two aspects: the individual aspect, either physical—relating to a person’s use of space and movement—or economic, relating to trip purposes and travel behaviours; and the aggregate aspect, the macroscopic level of passenger flows, considered in their spatial and temporal structure and also in terms of physical or economic categories.

The *Vehicle* subsystem includes the vehicle’s “external” aspects, i.e. its characteristics in terms of use of space and kinematic behaviour, and the vehicle’s “internal” aspects as a passenger container, with carriage capacity, internal arrangements and access systems for passenger entry and exit.

The *Station* subsystem performs two functions—respectively interface and storage—firstly, for passengers and secondly, for vehicles. With regard to passengers, the interface function links a particular service line (route) to all the other transport services (another route or another mode), whilst the storage function concentrates travellers

within vehicles or waiting areas prior to vehicle loading. As regards the vehicles, the interface function separates the upstream and downstream parts of the route, hinging around the unloading of alighting passengers and the loading of boarding passengers; the storage function pertains to the division of platforms and tracks within the station between vehicles on the basis of their real-time allocation (operational service or temporary immobilisation).

The *Line* subsystem consists of: (i) the vehicle movement infrastructure with right-of-way systems and possibly interactions with the exterior; (ii) the organisation of service routes on the line, a route being a subset of stations served in a predetermined order; (iii) the day-to-day operation of each route, with the runs of the vehicles allocated to it: the specific characteristics of the vehicles give aggregate properties to the routes they cover.

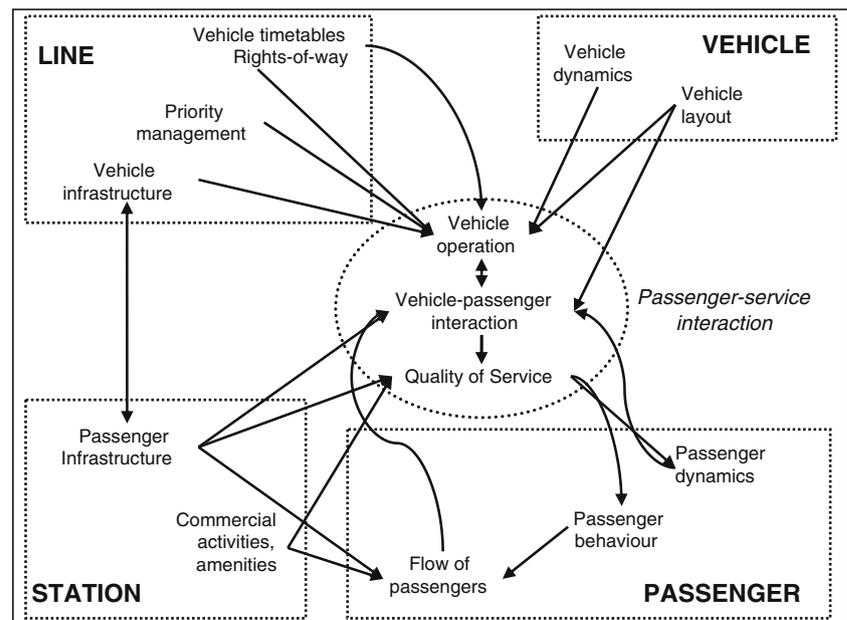
Figure 1 summarises the four subsystems and shows their main interactions. In the TCQSM, capacity factors are divided into Vehicle Characteristics, Right of Way Characteristics, Operating Characteristics, Passenger Traffic Characteristics, Street Traffic Characteristics and Methods of Headway Control (gap between two consecutive vehicles on a route) [1.3–18]. Our framework of analysis is more structured owing to the identification of sub-systems and their components; furthermore, it is abstract enough to cover public transit submodes in a generic way, whereas the manual divides public transport modes into bus, rail or ferry services. Lastly, it includes the economic dimension that pertains to passengers.

2.2 The passenger subsystem

Passengers are the elementary moving entities of which traffic flows are made up: at any given moment, passenger trajectories add up locally and form trip flows at every point on the network. Passengers temporarily immobile in a waiting area form a local stock. Passengers temporarily contained within the vehicle also form a stock, either stationary or moving, depending on the status of the vehicle.

Individual passengers have their own specific characteristics: when standing, they occupy a certain portion of space and require a certain distance from other passengers. When moving on foot, they occupy a certain width and possess dynamic characteristics for horizontal or vertical movement, in particular pedestrian “cruise” speed. These characteristics give a passenger flow its macroscopic properties: compressibility, capacity for acceleration and deceleration that governs flow speed, spread of individual speeds and flow viscosity. The properties of pedestrian flows are currently an active field of research (e.g. Hoogendorn and Daamen [7]). For our purposes, it is enough that the static and dynamic characteristics of pedestrians determine the individual load (or occupation,

Fig. 1 Components and interactions in PT operation and use



consumption) that each of them requires on a capacity resource, whether a storage space or a transport element. Passenger trajectories generate the spatial and temporal structure of flows during an analysis period, in particular with respect to each network element and each origin-destination pair. This structure determines the load on the transport means, infrastructure section or vehicle or pedestrian element. In turn, this load determines the local state of congestion and influences quality of service. Finally, quality of service is experienced by the individual passenger, primarily in time spent, discomfort and reliability. These aspects constitute an important criterion in determining the passengers' travel choices, in particular the choice of route and that of departure time.

Throughout a trip, a passenger uses certain elements of the network, forming a sequential composition. In his choice of path, he considers a set of alternative routes he might use, thus relating them in a kind of parallel composition. These possibilities for choice determine the local structure of trip flows within the network, and therefore have a feedback effect on local traffic loads and on quality of service.

So the economic choices made by passengers generate local traffic demand and therefore the demand for capacity on a local resource.

2.3 The vehicle subsystem

Vehicles are moving entities that can be simple or composite, through the assembly of "carriages" into a train. Their movement is the fundamental principle of public

transport, replacing the movement of the individual traveller, who remains immobile (or almost immobile) within the vehicle.

The vehicle has *individual characteristics*: it occupies a certain portion of space; its dimensions restrain its trajectory; its gross mass, power unit and braking systems determine its capacity for acceleration and deceleration, and therefore also its cruising speed.

The vehicle is purported to *contain passengers*. It has a cabin that provides seating and standing space. The dimensions and internal layout of the cabin determine the maximum number of passengers the vehicle can carry, in standard passenger units for a conventional maximum density of people standing (typically 4 people per square metre).

Passenger comfort depends on compartment layout and occupancy, temperature and humidity regulation, sound insulation, vibration control and the smoothness of the trajectory, amongst other factors.

Finally, the vehicle has one or more doors whose width and accessibility dictate the number of travellers able to move through them simultaneously and the time taken to do so: the positioning of the doors relative to the station platform determines the vehicle's exchange capacity per unit of contact time.

2.4 The station subsystem

The ultimate purpose of a station is to give passengers access to vehicles, for both boarding and alighting, for transfer to or from another service route or another mode of transport [2.5–41]. It must therefore allow storage space

and internal circulation space for both passengers and transit vehicles.

Passenger flows require the provision of levels of equipment that match the intensity of the flows, from a simple post, a shelter or a single building, to a complex of buildings for a large station. The buildings contain areas and seats where people wait to board, possibly fare and information kiosks and passenger service areas, whether utilitarian (toilets, left luggage, restaurants) or more commercial (newsagents, cultural products, clothes, groceries etc.). Globally, stations contain areas allocated to different activities, some accessible to passengers, others reserved for personnel. Some activities are necessary for trip making, such as waiting on the platform or moving around the station. Others are useful or profitable in offering passengers opportunities and helping them make better use of their waiting time. As in a vehicle cabin, the layout of the space and control of the surroundings influence the traveller's comfort. In addition, signage and static or dynamic information contribute to quality of service.

In a station, the concept of circulation capacity varies according to the type of space: waiting areas or corridors for passengers, platforms, lanes or tracks for vehicles. A waiting area can contain passengers in the same way as a vehicle, but in general with more possibilities for interchange with the outside, and conversely exposure to influences such as other movements of passengers or flows of pedestrians. The pedestrian elements under greatest stress are those with different directions and speeds of flow, and also spaces of vertical travel such as stairways and elevators, with particular stress on upwards movement.

Items of equipment involve constraints on capacity, in particular gates, barriers and fare desks. Just as a vehicle's exchange capacity depends on its doors, some capacity constraints combine several parallel elements with similar functions, in particular corridors or stairways leading to platforms.

The elements set aside for vehicles are also subject to capacity constraints. Moreover, in certain stations, passengers access the mass transit system in an individual vehicle, with the concomitant constraints on parking capacity.

To sum up, a station is a collection of spaces and equipment whose individual characteristics place capacity constraints on passengers, mass transit vehicles and even individual vehicles. The configuration of these spaces determines capacities at a higher level, by providing possibilities for the transfer of moving bodies between neighbouring elements.

Finally, the arrival in the station of a large capacity vehicle can generate a flow of relatively high intensity for a short period: this kind of peak flow can saturate corridors

and walkways, and result in transfers or spreads that cause temporary deterioration in certain capacities.

2.5 The line subsystem

Let us analyze the Line subsystem into four components, respectively the vehicle travel infrastructure, route management, route organisation and run scheduling on a route.

The vehicle travel infrastructure follows a linear or almost linear route through space: some lines are defined by a central section with connecting branches, creating a diversity of vehicle routes. At one point on the route, the infrastructure has a number of tracks in each direction. Throughout the route, the material substrate remains the same: road or railway. In the case of the electric railway, the line supplies the vehicle with electricity.

Route management The geometrical characteristics of the route condition the course of a vehicle on the line: the shape, contours and curves (profile in length) and local irregularities (profile in width) condition the movement of the vehicle, its accelerations and decelerations and cruising speed. Beyond the ideal kinetic characteristics of an "isolated" vehicle, the infrastructure is influenced by traffic of different kinds and other disruptive influences. The vehicle's progress can be delayed by other vehicles belonging to the same service, or by other vehicles sharing the lane or tracks on a non segregated infrastructure. Segregated lanes or tracks can be used to exclude vehicles external to the line. However, traffic can also be disrupted by intersections with other flows of traffic, belonging to the same mode or to others (in particular pedestrians), or by disturbances of an external nature (weather, obstructions, etc.): the subway is the mode with the fewest constraints, as it intersects exclusively with passengers boarding and alighting in the station. Route management covers (i) the definition of the route, (ii) interaction management by means of access reservation and intersection management, (iii) the control of each vehicle to match the particular conditions of the run.

Rail transport is characterised by a signalling system purported to maintain safe distances between trains on the track. This system allows trains to travel at high speed in the absence of disruptions.

Route organisation On a public transit line, every vehicle run is allocated to a route, serving a subset of stations along a trajectory on the line. The length of the stops combined with the length of the journey between sections, constitutes the length of the run.

Run scheduling During a run, the vehicle cannot be assigned to another activity, which restricts its availability

for the operator. The latter must also take passenger flows into account in scheduling runs: the headway between two consecutive runs must be long enough for passenger numbers to accumulate, but not too long, because an excessive waiting period would put off some passengers.

2.6 The main interactions between the subsystems

The diagram in Fig. 1 not only shows the four subsystems in the form of blocks, but also their principal interactions, identified by arrows, which indicate:

- The operation of a particular vehicle throughout a run.
- The elementary interaction between the vehicle and the passenger during boarding, alighting and travelling aboard. In particular, the time it takes for a passenger to access the vehicle depends on the respective heights of the vehicle step and the platform, the degree of crowding on each of them, and possibly the need to check tickets inside the vehicle just after boarding, which generates delays.
- The formation of quality of service throughout the passenger's entire journey.

These three complex interactions comprise the entirety of the interaction between the passenger and the service.

3 Seven groups of capacity phenomena

It has been established that a variety of resources are limited in the operation of a mass transit passenger network: the components used for the storage or movement of passengers and vehicles, or the fleet of vehicles available in the short term for the operator.

In order to define the capacity constraints, let us now identify the principal constraints and describe them in terms of the resource limited, factors, specific conditions and consequences. The constraints are grouped based on the type of moving bodies concerned, and on their location in a station or line section. Seven groups of constraints are defined below: (i) the vehicle capacity of an infrastructure; (ii) the vehicle fleet; (iii) the passenger capacity of a vehicle; (iv) the passenger capacity of a route; (v) the passenger capacity of a station; (vi) the vehicle storage and movement capacity of a station; (vii) the capacity of a station for interface with personal transport modes.

3.1 The vehicle movement capacity of the infrastructure

For a section of infrastructure in a given direction, occupancy demand consists of the vehicles to be carried, each corresponding to an elementary time length during

which a given point-wise spot is occupied. Denoting by t_X the average elementary occupancy time per vehicle on a certain route (including a safety margin), and by X the number of such vehicles over a period H , the route demand is $X.t_X$. If a number say Y of other vehicles travels on the section, with an average occupancy time of t_Y per vehicle, then occupancy demand is $\Omega = X.t_X + Y.t_Y$. Other vehicles belong to other routes or other modes.

The capacity resource is an area of travel of a certain width allowing movement on f lanes of traffic in the direction concerned: the baseline capacity is $K_0 = f.H$ for the period H . If flow in that section and direction conflicts with flows in other directions, let us say Z vehicles each requiring a time of t_Z on the average, then the initial flow capacity is reduced to $K = K_0 - Z.t_Z$. If the conflict between traffic flows occurs at a regulated junction, then $K = \eta.K_0$ where η is the proportion of the time in which the initial direction has priority on the junction.

The ratio between occupancy demand Ω and allocated capacity K determines the load rate, $\rho = \Omega/K$. Both Ω and K depend on local flows using the infrastructure and the local geometrical conditions: these are the specific conditions of this capacity phenomenon. The general consequence is that the average time taken to cover the section by a vehicle on the route increases with ρ , as do the associated risks of disruption, the energy consumption and the emissions. Whilst the ratio ρ remains well below 1, traffic remains fluid. But if ρ gets close to or temporarily exceeds 1, a queue forms and its size at the time a vehicle arrives determines how long that vehicle will wait (bottleneck model).

3.2 Operating capacity of a route

For a service route z , vehicle usage demand is the desired number of runs, N , multiplied by the service time, say an average of T_z for a round-trip: therefore $\Omega = N.T_z$.

The vehicle time resource during the period H is $K = n.H$, where n is the number of vehicles. The capacity constraints on the fleet of vehicles allocated to the route is therefore $N.T_z \leq n.H$, i.e. in terms of frequency, $N/H \leq n/T_z$.

The composition of the round-trip determines the cycle time T_z , which is the segment running time plus the time spent in stations, together with operating margins for a reserve or recovery time for the vehicle driver.

With a fixed fleet, the constraint limits the frequency of service, which influences the number of passengers boarding and alighting at each station, and therefore dwelling time in the station.

3.3 Passenger capacity of a vehicle

This category includes three capacity constraints, respectively the number of seats, the total containment capacity of

the vehicle and the number of passengers entering/leaving per unit of transfer time.

Seating capacity A vehicle is fitted with a fixed number of seats that defines the seating capacity. Occupancy demand is obviously a quantity of passengers, each of which would in principle induce a unit load (or more in the case of large luggage or passengers who generate avoidance responses). For a train, moreover, the distribution between carriages may be uneven. Other specific conditions relate to the passengers' legs from access to egress points, which dictate both the residual capacity in a station after outgoing passengers have alighted and local boarding demand. On the whole, standing passengers who remain on board have priority to get a seat over those who will board, but some passengers may prefer to remain standing when their egress station is close. The main consequence of the rate of seat use is a degree of discomfort per unit of time spent in the vehicle, relatively low at seating, but higher at standing, particularly when the density of standing passengers is greater.

Total passenger capacity This capacity is the sum of seating and standing capacities inside the vehicle. It reflects a maximum limit tolerated by passengers. In principle, the physical time needed to cover the section is not affected, but the discomfort per unit of time increases with density. In addition, if the vehicle is full, a stock of passengers remains on the platform at least until the next vehicle arrives.

Access capacity The total width of the access doors that open when the vehicle stops at a station determines a passenger capacity per unit of time. The corresponding occupancy demand consists of the number of alighting passengers multiplied by the elementary length of time required to alight, plus the number of boarding passengers multiplied by the elementary length of time required to board. The specific conditions include the presence of any fare checking equipment inside the vehicle, which increases the elementary boarding time; likewise, the difference in height between the vehicle step and the platform adds to the elementary time, as does the proportion of people with reduced mobility.

3.4 Passenger capacity of a route

The passenger capacity of a vehicle induces the passenger capacity of a route by adding up the capacities of the vehicles covering the route during a period H . Demand itself consists of a flow of passengers who arrive at the platform during the period.

If the vehicles are of different sizes or have different loading conditions upstream, or if the headways between two consecutive vehicles are irregular, then the relation between demand and capacity varies over the period.

The *seating capacity* is the sum of the seating capacities of the vehicles covering the route over the period.

The *total passenger capacity* is the sum of the total capacities of the vehicles covering the route over the period. After a relatively long headway, it may be that the vehicle has insufficient capacity to absorb the stock of passengers waiting on the platform to board: those who cannot board experience an additional delay until the arrival of the next vehicle. If several routes use the same platform, their respective stocks of waiting passengers mix and generate mutual obstruction.

Access capacity is the total loading-unloading capacity per unit of exchange time of the vehicles covering the route during the period multiplied by their respective transfer times (dwelling time). Adjusting the dwelling time is an important instrument for regulating passenger flows and maintaining regular gaps between runs, and therefore service reliability.

3.5 Passenger capacity of a station

This group includes the capacities of the elements dedicated to pedestrian movement and to passenger storage.

Pedestrian traffic capacity The passenger capacity of a circulation element is analogous with the vehicle capacity of a section of infrastructure. Conflicts between traffic flows can significantly reduce the capacity available to each one of them. The specific load rate of the traffic flow determines both the speed of flow, and hence the time taken to pass through the element, and the discomfort per unit of time spent. If the load rate exceeds 1 for a certain duration, a queue forms and accumulates upstream.

Pedestrian storage capacity The occupancy demand of a waiting area is the stock of passengers present in that area or in the immediate vicinity, which changes with the arrival of passengers (stock inputs) and with the departure of passengers boarding vehicles (outputs). The capacity of the area is the seating capacity plus the standing capacity, as with the capacity of a vehicle. In addition, alighting passengers need to move through the stock, as need boarding passengers when the waiting area serves several routes (in which case boarding does not trigger a general movement), and sometimes even pedestrians who are not users of the service. These through flows reduce storage capacity and increase the density of waiting passengers. Density influences the discomfort experienced per passenger and per unit of time. Moreover, the size of the stock may prevent certain passengers boarding a vehicle and force them to wait for a later vehicle.

3.6 Vehicle capacity in a station

This group comprises four types of constraint.

Vehicle platform dwelling time A vehicle dwelling at a platform spends much longer there than at a standard section point (t_x in the first group). In the event of a disruption, dwelling time may be extended and cause delays to subsequent vehicles, and this delay affects passenger travelling and waiting time.

Platform capacity The number of platforms in the station governs the number of boarding areas for routes. This may force the operator to reduce the frequency on certain routes, especially in large railway stations. When different routes share a single platform, they become mutually dependent and vulnerable to delay, which affects passenger travelling and waiting time.

Weaving constraints Vehicles access station platforms from the tracks, and vice-versa, either directly when the platform is continuous with the track, or via intersection points. The flow capacity of such a point is limited, in the same way as an infrastructure junction point (first group), with the same specific conditions and the same consequences.

Storage of standby vehicles Vehicles that are temporarily not allocated to routes have to be stored, generally in a station where their storage requires space. Their occupancy of space limits the capacity available for ongoing services, especially in rail transport.

3.7 Capacity for interface with individual mechanised modes

A share of the passengers may access public transport by individual mechanised modes, which employ a specific vehicle: private or shared car, taxi or two-wheels. In certain stations, access for such vehicles is provided by dedicated

amenities, including park-and-ride areas where private vehicles can be stored.

In a station, therefore, the interface with individual vehicles involves elements of movement and elements of storage, each with its own specific operation, analogous with the handling of passengers within a station (apart from the fact that an unoccupied vehicle remains stationary).

The *traffic capacity* of an individual vehicle access lane is analogous with the vehicle capacity of an infrastructure (see first group).

The *storage capacity* of an individual vehicle park-and-ride centre is generally limited by a fixed number of slots, identified by ground markings or locking systems for two-wheels.

Taxi operation has certain distinctive features. Often the access route is also used for storage, and its occupancy level depends on the arrival and departure of taxis and passengers. Passenger access to taxis can itself entail specific arrangements, with a waiting area that constitutes a storage element in the same way as for a mass transit route, with specific conditions and similar consequences for passengers.

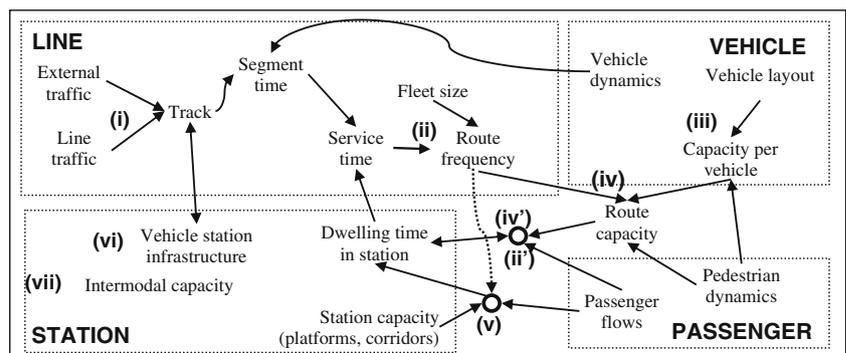
3.8 Graphic summary

Figure 2 summarises the capacity phenomena identified above and situates them within the framework shown in Fig. 1. Each phenomenon is situated in relation to the components of the system. For example, the second group, which relates to service frequency, depends on service time and fleet size. Another example for the fifth group: service frequency on a route influences the size of the stock of waiting passengers, and therefore the interaction between the stock and the capacity of the waiting area.

4 Compositions

A public passenger transport system has been analyzed into four main subsystems, with components in each subsystem. Each capacity phenomenon is an interaction between compo-

Fig. 2 Capacity phenomena on a PT network



nents of the system. This section deals with the compositions that involve capacity phenomena and exercise a feedback effect on them. Simple compositions are addressed first, before moving on to more elaborate ones. Then, “congestion gears” are described—defining such a gear as a combination of capacity phenomena that triggers a reduction in capacity. Lastly, consideration is paid to complex feedback effects triggered by the network users or operator.

4.1 Occupancy in space and time

An area of space, such as a waiting area, constitutes a capacity resource for the elementary forms of occupancy required by entities such as passengers.

A period of time is a capacity resource for elementary periods of occupancy required by entities such as passengers or vehicles. In fact, duration relates to a physical resource: a passing point on an infrastructure or a mass transit vehicle.

To analyse a spatial resource such as a waiting area or section of infrastructure over a period of time, a valuable concept is that of space-time capacity, which is calculated by multiplying the spatial dimension of the resource by the duration of the period. This enables one to look at spatial resources that are larger than a single point in space. An elementary occupancy in space-time is calculated by multiplying the elementary spatial footprint by the length of occupancy of each point: i.e. $L.t_X$ for a moving body travelling through a segment of length L and occupying each point in it for t_X ; or $s_X.t_X$ for an entity of size s_X occupying a space in a waiting area for a period of t_X . Space-time analyses are useful in urban planning [4] and in transport [1, 2, 5, 6, 9].

All these occupancies are simple. Elementary occupancies are cumulative, and the resource is similarly calculated by adding together components in both time and space.

4.2 Sequences of factors and consequences

Figure 2 depicts the factors and consequences of capacity phenomena for the seven groups that were identified. The diagram enables one to track sequences between the phenomena and even cycles of dependency. For example, a vehicle’s dwelling time in a station depends on the flow of passengers boarding and alighting in that station; then, the dwelling time influences the vehicle’s cycle time, which influences route frequency; in turn, route frequency influences incoming and outgoing flows per vehicle.

The TCQS Manual [15] presents several combinations of factors respectively composing the capacity of a bus infrastructure [4.1–16], or the passenger capacity of a bus line [4.1–17], or railway capacities, in particular in the guise of flowcharts. This manual is very helpful in providing a more concrete understanding of the details of

capacity constraints, with specific reference to the mode of transport concerned.

In addition, the Manual reveals complex sequences with subtle effects on capacity. Here are some instances:

- The need for a moving body to vacate a transfer point before it can be used by another: the alighting and boarding of passengers depending on the layout of the vehicle doors; “clearance time” between successive buses at a stop; time taken for a bus to rejoin the traffic flow after stopping in another lane; railway switches; the interaction between switches and turnbacks in a rail terminal.
- At the confluence between passengers and vehicles in a station, the respective passenger capacity of a vehicle and of the pedestrian facilities dictate the headway to be maintained between two successive vehicles, and therefore the vehicle capacity of the lines [7.3–20].
- In a station, passengers who use amenities (information desk, ticket machine, etc.) complicate their activities and introduce detours into their route.
- At the terminal of a transit route, on-duty drivers have to take a break; the operator must also provide recovery time to offset fluctuations in runtimes [1.2–13].
- The flow of vehicles on a route can be limited by the number of vehicles available to the operator: this is an economic constraint on vehicle capacity, which also affects the number of carriages in a train and therefore passenger capacity [5.1–8].
- The number of drivers is a constraint on the capacity in terms of the number of vehicles. At a rail terminal, a driver who continues on the return journey must walk the length of the train, which is a constraint on the stopping time, and the condition of the train must be checked at the same time [5.2–17].
- The respective configurations of several transit routes which run along the same infrastructure line, interaction between omnibus services and express services and skip-stop services.

4.3 Congestion gears

A “congestion gear” is an interaction between several capacity phenomena such that they mutually affect each other, resulting in a reduction in capacity. In other words, a gear has devastating and potentially disastrous consequences. In running a system, they need to be avoided and prevented, or at the very least detected and their consequences managed.

A number of congestion gears can be identified, making up the following list—which is likely not complete:

- Elementary occupancies within a single capacity resource combine: their cumulated occupancy generates a

local congestion state. On a circulation element, the local congestion state influences in an increasing way the dynamic elementary occupancy, because the speed of a moving body diminishes with congestion. Since elementary occupancies determine the local congestion state, this state is self-reinforcing: this is the *problem of circulation capacity occupancy*.

- The *problem of dwelling time* for a vehicle in a station has already been described: higher boarding and alighting flows lengthen dwelling time, thereby increasing cycle time, reducing service frequency and in turn increasing the flow of passengers boarding and alighting per vehicle.
- Let us also recall the *local interpenetration* of flows converging on a single passing point in different directions, therefore coming into conflict: a traffic flow in one direction reduces flow capacity in other directions.
- In stations, passenger capacity constraints on a route interact: the denser the stock of passengers waiting on a platform, the more it hinders boarding and alighting passengers. Assuming fixed dwelling time, stock congestion reduces loading and unloading flows, in particular the flows of boarding passengers, which increases the accumulation of passengers on the platform. Similarly, the stock of passengers in the vehicle also obstructs movement into and out of the carriage. These obstructions cause delays to passengers and possibly also to vehicles. Moreover, passenger discomfort per unit of time increases with congestion.
- Congestion that affects moving bodies can propagate in space, when congestion in one element spreads to an adjacent element and reduces its own capacity. This process of propagation occurs in stocks of waiting passengers and in passenger flows, as well as in stocks and flows of vehicles. For a flow of moving bodies, congestion spreads upstream: this phenomenon is called *spillback*.
- In public passenger transit systems, congestion can also spread downstream amongst passengers, through vehicles, i.e. act remotely. The vehicles carry stocks of passengers who become factors of congestion for downstream stations. In addition, a delay generated upstream on a vehicle spreads downstream, itself becoming a factor of congestion.

4.4 Complex feedback effects

In a public transport system, both users and operators are likely to react to traffic conditions. These reactions constitute complex feedback effects.

On the user side, *quality of service determines the path choice*. Every user perceives the quality of service under a

given travel option, in particular with regard to time and discomfort, and chooses their option by trading off quality of service and price factors according to their own preferences, i.e. for personal rather than communal benefit. High levels of local congestion along a route will prompt users to transfer to another route, which distributes traffic between routes since *route choices generate the trip flows*. In principle, this should reduce congestion on the initial route. However, route diversions can actually trigger paradoxical effects: Braess [3] has shown that for certain network configurations and loads, detours by certain users can disadvantage others and adversely affect the global state of the system. For our purposes, it is enough that microeconomic behaviour by the user triggers interactions with local quality of service through the formation of traffic flows.

On the operator side, there are three groups of instruments:

- *Traffic management*: given limited infrastructure and vehicle resources, operators can still manage traffic by directing passenger flows with signage and route guidance and channelling them by regulating capacity. The instruments that can be used to allocate capacity are: adjusting dwelling time per run and per station; regulating traffic on the infrastructure; adjusting runs scheduled on routes; and changing route organisation. Adjustments to run and route schedules include changing the allocation of vehicles to routes, subject to compatibility between the vehicles and the infrastructure lines.
- *Demand management*, by providing information on traffic conditions, also an aspect of traffic management, or by pricing mechanisms. A higher baseline fare will put some people off travelling; less drastically, a spatial adjustment in the fares for different links allows the operator to direct flows by influencing their origin-destination structure on its network. More subtly, applying different fares at different times can reduce congestion by influencing passenger departure times. The stakes are high for an urban public transport network, which experiences high traffic peaks during morning and evening rush hours.
- *Adjustments to network capacity* can be applied to vehicles or infrastructure. On the vehicle side, the operator can adjust the number of vehicles in its fleet or the number of trains in its rolling stock, or the size and layout of its vehicles, to determine passenger capacity. On the infrastructure side, the operator can design the platforms and more broadly all the areas in a station so as to allocate a given amount of space, possibly with several levels, to a range of functions. In addition, the operator can adjust infrastructural capacity by choosing a traffic management system, by setting the number and layout of the tracks, by connecting the tracks to the stations—and sometimes even by config-

uring the lines near a station to accommodate access and transfer requirements.

5 Conclusion

A fairly general framework has been provided for the analysis of a public passenger transport system, in order to show how it functions technically in interaction with the behaviour of the passengers who use it. In this framework, the system is broken down into four subsystems, respectively the Passenger, the Vehicle, the Station and the Line, each subsystem comprising several components. The capacity phenomena have been identified, described qualitatively and classified into seven groups, on the basis of the capacity resource, the type of moving bodies and the place of occurrence. In addition, some sequences of effects have been outlined, including congestion gears and complex feedback effects associated with either passengers or the operator.

The framework is qualitative: it can be used by network operators to identify and reveal a capacity phenomenon and even a congestion gear. However, the framework is primarily intended for network planning, and specifically for the simulation of development schemes and operating plans. This kind of simulation is based on a model for the assignment of traffic to routes on a network (cf. e.g. [11–13]). For the simulation to be relevant, the assignment model must be sensitive to capacity phenomena and their consequences, including the retroactive impact of passenger behaviour. Our framework can be used as a checklist to assess which capacity phenomena are considered in an assignment model and to inspire the development of more sensitive models. It is the author's intention to review in a companion paper the capacity phenomena as represented in network traffic assignment models.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Benz Gregory P (1986) Application of the time-space concept to a transportation terminal waiting and circulation area. In Transportation Research Record 1054, TRB, National Academy Press, Washington, DC
2. Benz GP (1992) Pedestrian time-space concept: a new approach to the planning and design of pedestrian facilities, 2nd edn. Parsons Brinckerhoff, Inc., New York
3. Braess D (1968) Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung* 12, 258–268. English version in Braess D, Nagurney A and Wakolbinger T (2005) On a Paradox of Traffic Planning. *Transportation Science* 39-4:446–450
4. Fruin JJ (1971) Pedestrian planning and design. Metropolitan Association of Urban Designers and Environmental Planners, New York
5. Fruin JJ, Benz GP (1984) Pedestrian time-space concept for analyzing corners and crosswalks. *Transp Res Rec* 959:18–24
6. Guieysse L, Marchand L (1988) “Congestion urbaine et réponse de la planification en agglomération parisienne”, dans Les transports collectifs urbains, actes du colloque des 3–5 mai 1988, Presses de l'ENPC, Paris, pp 111–122
7. Hoogendoorn SP, Daamen W (2005) Applying microscopic pedestrian flow simulation to railway design evaluation in Lisbon. *Transp Res Rec* 1878:83–94
8. Kagan LS et al (1978) *A pedestrian planning procedure manual*, 3 vols. Report NO FHWA-RD-79-45, FHWA-RD-79-46, et FHWA-RD-79-47, US Department of Transportation, Federal Highway Administration, Washington D.C.
9. Marchand Louis (1977) “Qu'est-ce que la mobilité”, *Métropolis*, n° 24-25, pp. 51-54.
10. Neufert E (1951) *EBauentwurfslehree—Eléments des projets de construction: principes fondamentaux, normes et règles concernant la conception, l'exécution, la forme, l'espace nécessaire, les relations spatiales, les mesures des édifices, des locaux, des meubles, des pièces ouvrées avec l'homme qui doit s'en servir*. 10^{ème} édition française, Dunod et Le Moniteur, 2010
11. Ortuzar J. de Dioz, Willumsen LG (1990) *Modeling Transport*. Wiley, London
12. Sheffi Y (1984) *Urban transportation networks*. Prentice Hall, Englewood Cliffs
13. Thomas R (1991) *Traffic assignment techniques*. Avebury Technical, Aldershot
14. TRB (1950, 1965, 1985, 2000, 2010) *Highway Capacity Manual*. See article in http://en.wikipedia.org/wiki/Highway_Capacity_Manual
15. TRB (2003) *Transit capacity and quality of service manual*. Online report prepared for the Transit Cooperative Research Program, available on-line at the following website address: http://gulliver.trb.org/publications/tcrp/tcrp_webdoc_6-a.pdf. First edition 1999