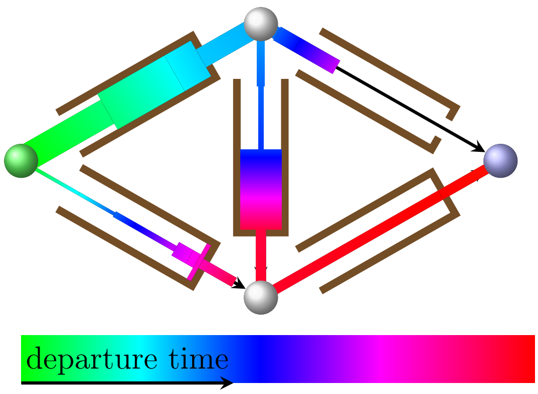
Routing Games over Time

vorgelegt von
Diplom-Wirtschaftsmathematiker
Ronald Koch
geboren in Potsdam

Von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften

genehmigte Dissertation



Promotionauschuss:

Vorsitzende: Prof. Dr. Gitta Kutyniok
Berichter: Prof. Dr. Martin Skutella
Berichter: Prof. Dr. José R. Correa

Tag der wissenschaftlichen Aussprache: 30.04.2012

Berlin 2012

D 83

Diese Arbeit widme ich meinen Großeltern
Omi Toni, Oma Hilde, Opa Arthur und Opa Klaus.

Contents

Chapter 1

Introduction

”YES, finally managed!“ you think. After working hard for a long time, you finally have holiday. In a few days you are going on a trip, and in horror you think of packing your bags. But as always you postpone this task hoping that your partner will feel responsible for it. Instead, you plan the next step first and cogitate about arriving at the airport on time so that your family and you will not miss the flight.

Since you never want to waste time, you look for the fastest route to the airport. Usually, you would rely on the car’s navigation system. Unfortunately, your car has a transmission failure, and so you have borrowed a friend’s old car without a navigation system. While you are printing out the suggested shortest route from an online route planner, you have a sudden brainwave. What happens, if many other people take this route at the same time as you? After all, the high travel season has just started and this is a shortest route to the airport. So if everybody else decides to drive along this route, a traffic jam arises for this very reason.

Hence, you deem yourself as a smart person, pat yourself on the back and search for another path which should not be used as frequently. At best, you will arrive at the airport at the earliest possible point in time provided all other travelers take the shortest path. After half an hour you identify a promising candidate for such a route. But at the same time, a second brainwave hits you. What happens, if other people are just as smart as you? Then these people also want to avoid the traffic jam on the potentially shortest path and will perhaps also use the new route you found. So even on the new route a traffic jam may occur.

But now you are smart enough not to repeat the same process and not to look for a third route to the airport as this would end up with the same conclusions as before and you would never be satisfied with one of the constructed routes. So you are a little bit disappointed and decide to use the route to the airport which was suggested by the online route planner. Further, in order not to miss your flight, you will start significantly earlier. But you still feel queasy.

You cannot stop thinking about the problem. You believe, the primary goal of every person is to reach the airport as fast as possible. Unfortunately, at which time a person arrives at the airport along some route depends on the paths chosen by the other travelers. Now you are confused: ”The best route I could use depends on the best routes of the others and vice versa. So which route should I take and which routes can be taken by the other people?“ Thus, the question is whether or not each person can chose a path in such a way that each path becomes a best one. As you are a hobby mathematician, you observe that such a routing is stable as no person has an incentive to switch his path. In other words, every single person is satisfied with his chosen path because he cannot improve his arrival time by traveling along another path under the assumption that each other person remains on his path. Hence, the entire traffic system is at equilibrium.

But what does this help? You observe that such equilibria can be used to model traffic behavior as you believe that everybody behaves somehow egoistically. Hence, you could chose a fast route with respect to an equilibrium, and maybe, this route works also well in practice. Of course, for this an equilibrium must exist, and it would be even more useful if exactly one equilibrium exists. On the other hand, if you know that exactly one equilibrium exists, you still have to construct it in a reasonable amount of time. And as a discrete mathematician, you know that an efficient computation can be much harder to find than proving uniqueness.

Be that as it may, right now you are quite busy with the preparation of your trip. But from time to time, you think about this topic. You notice that almost everybody you know, except the nice friend who has been borrowing you the old car, has a navigation system. If it were possible to interconnect all navigation systems, one could get a nearly complete picture of the current traffic. In particular, one gets the information at which time somebody wants to go to which place. Using this information, it would be possible to compute best routes. Even more, one could compute an optimal routing in the sense that the number of travelers who have been reaching their destination is maximized. From an abstract point of view, this sounds like a social optimum as, on average, this is the best for all people. However, you discard this idea at the same instant as you have the feeling that this causes long unattractive routes for some of the travelers. And if you would be such a traveler, you would not follow the path from your navigation system. Nevertheless, you ask yourself how far away is an equilibrium from such a social optimum. Maybe, one does not lose so much if everybody behaves egoistically.

At this point in time you sort your questions and decide to make a literature review after your trip. You want to find answers concerning existence, uniqueness, computability, and performance of equilibria.

So the day of your trip has come, and luckily, your partner has packed all bags. You depart from home early in the morning. At the beginning everything is fine, but with increasing time, the roads become more and more congested. In the end you arrive at the airport half an hour later than suggested by the online route planner but still on time to catch your flight. In fact, there is still enough time to drink a coffee at the airport. You use this time and ask yourself why the online route planner failed in computing the right transit time. You guess that the route planner stores a transit time for each street section, which is based on the speed limit and the length of that street section. Of course, such a transit time is a good approximation on the average time needed for traveling along this street section, but only if it is not congested too much. In fact, if more cars want to travel along a street than the street can handle, a traffic jam builds up. You imagine an increasing number of cars in a queue at the end of the particular street section. This means intuitively that a car wanting to travel along this section, first drives at the speed limit until it arrives at the end of the queue and then waits until it reaches the end of the street section. Hence, the actual transit time of a car consists of the transit time plus the current waiting time. For now you are proud of yourself and think that this is the right model for traffic behavior which can be used when considering equilibria.

But suddenly, while you are thinking, you get a nudge from your partner as it is time to check-in. You see all the queues at the check-in counters and the ones in front of the security channels. You are searching for the best possibility to pass the security check as fast as possible. ”Oh“, you think, realizing that the model explaining travel behavior could also work in this case.

Everything goes well. You pass the check-in and the security check, and finally, all of you enter the airplane. Now you have 10 hours of flying ahead of you. You enjoy the perspective of this, but the topic still fascinates you. In school you heard something about the historical problem of the seven bridges of Königsberg, which searches for a round tour through Königsberg where each bridge is crossed exactly once, only. Euler [22] showed in 1741 that such a tour does not exists and his approach is the basis for graph theory. You remember, a graph consists of nodes and edges where an edge always connects two nodes. ”Graphs are a good tool for representing the traffic network“, you think – the nodes are junction, and the edges are street sections connecting two junctions. Furthermore, for each edge you should have a real number determining the low congested transit time of the corresponding street section. In addition, a capacity should bound the number of cars leaving an edge over one time unit, i.e., the rate of leaving cars, and if more cars want to leave this edge, a waiting queue builds up at the end of this edge.

Absorbed in thought, some background noise catches your attention, and you recognize the voice of your partner. Still trying to remember the words you can already hear your partner saying: ”Are you not listening to me, once again?“ Luckily, you manage to reconstruct the words: ”Hello, in which clouds do you have your head?“ So you decide to turn your entire attention to your family for the rest of this flight.

It seems that the airplane arrives at your destination on time, but then an announcement sounds through the loud speakers saying that the airplane is put in a holding pattern. Immediately, your thoughts are back in the world of equilibria. You ask yourself if you can use the idea of equilibria for determining good air lanes. As before, graphs could be used to model all possible air lanes. But of course, you cannot take the model of traffic behavior as it is not a good idea for an airplane to wait in midair. Rather, you need a model which forbids waiting after departure and before arrival. In this sense, your routing model should somehow be direct ensuring that, after takeoff, an airplane can fly nonstop through all intermediate nodes of the chosen air lane. However, you have the feeling that an optimal routing would be the better alternative in this scenario. But of course, even better would be an equilibrium which is optimal.

Suddenly, you notice the increase of pressure on your ears. The airplane lands safely, and it is time to enjoy your holiday. However, already on your first day you take out your iPad and decide to search for some answers to your questions. Starting in the area of graph theory, you hit rather fast on an article of Ford and Fulkerson [26] on flows. Like water in a pipeline system, flow traverses the network from a source node to a sink node. ”Of course, that’s it!“ Now you think of flows as flows of cars or flows of airplanes. Unfortunately, you miss the time dimension. These flows represent only steady states arising out of, for example, average considerations over a long period of time or constant flow rates. Luckily, Ford and Fulkerson also study so-called flows over time in [27], where flow rates may vary over time and also transit times are included in the model. Unfortunately, in their model all transit times are constant and flow-independent. After a short period of time of time you discover the article [89] of Vickrey. Surprisingly, his deterministic queuing model represents traffic behavior exactly in the same manner as you claimed. At this point, you are satisfied as you have a starting point for a more detailed search at home.

So you try to find out something about equilibria, and at the same point in time you are in the world of games. You are wondering, where the relation between games and equilibria is. Then the scales fall from your eyes – all car drivers can be seen as players which compete for free capacities on streets in order to reach as fast as possible their destinations. Clearly, all players have their own strategies in choosing their routes resulting in a particular routing scenario. And if there is no better route in some scenario, no player is going to change his route meaning that such a scenario represents an equilibrium. You find out that such stable scenarios are called Nash equilibria which goes back to the work of Nash in [62].

At this moment, your family decides that all of you want to go to the beach as it is hot and the sun is shining. You say: "I'll come in a moment." You want to check first if there is any connection between flows and Nash equilibria. You observe that for the steady state case Nash equilibria are considered already in [90] by Wardrop. After further five minutes you find the work of Friesz et al. [29] who formalize the notion of Nash equilibria for flows over time. You are really happy about your search so far. In the meantime, your family is getting angry with you. So your partner takes your iPad away from you and puts it in the hotel safe without giving the code to you. Thus, from now on you have to enjoy your holiday whether you like it or not.

Literature review. Back at home you start the literature review in order to find answers to your questions. You already know that flows over time are an adequate modeling tool for formalizing a traffic system which evolves over time. Nevertheless, you decide to learn something about static flows representing steady state scenarios at first in order to get familiar with this topic. After the introductory work of Ford and Fulkerson summarized in their book [28], network flow theory became an own field in the area of discrete mathematics. So there exist a lot of popular text books like [1, 46, 80] providing you with all the basic knowledge about static flows.

Having a rough idea about static flows, your attention turns to flows over time which sometimes are also called dynamic flows. As you already know, Ford and Fulkerson [27] study flows over time in 1958. In [27] they consider the problem to send as many flow units as possible from a unique source to a unique sink within some given time window. In their classical model the number of flow units being able to enter an edge at the same point in time is bounded by some integral capacity. Further, the time for traversing an edge is flow-independent and determined by some integral transit time. Thinking about your traffic system, each car becomes a flow unit in their model.

However, seeing the huge number of cars using the streets of your hometown every day, you ask yourself whether or not it is really necessary to model every single car. You recognize that the resulting model would be too large to be computable. For that reason, you search for a continuous version of this discrete flow over time model where the discrete structure of flow units is somehow smoothened. Lucky as you are, you immediately find the recent survey article [82] on flows over time written by Skutella. Here the flow is given by time-varying rates at which the flow enters the edges.

Reading [82] you are somewhat surprised that there exist flows over time which not only maximize the amount of flow that has arrived at the sink until a single point in time but do this simultaneously for all points in time. For discrete scenarios, the existence of such earliest arrival flows is observed by Gale [32] already in 1959. For the continuous setting Philpott [71] does this in 1990. Putting this in the context of analyzing traffic behavior, earliest arrival flows seem to be the best candidates for a social optimum. As already mentioned, you want to use a social optimum in order to measure the performance of a Nash equilibrium. However, something bothers you. In these classical flow over time models the transit times are always constant, i.e, flow-independent. But this contradicts your personal experience. According to that, it takes longer to traverse a street the more congested the street is. So you proceed with your literature review on flows over time concentrating on flow-dependent transit times.

During your literature review you find the survey paper of Mun [61] which gives a comprehensive overview of existing flow over time models. Mainly, there exist three classes of such models – transit time, outflow, and wave models.

Transit time models have their origin in the seminal work of Greenshields [34] in 1933 who describes the speed depending on the current traffic density. In this manner, depending on how the flow evolves over time, a current transit time is assigned to each point in time. Like in classical flow over time models current transit times determine the traveling time for traversing a particular edge. But now these transit times are flow-dependent and may vary over time. An important subclass of such flow models are load-dependent transit times. Here, current transit times depend only on the current load of an edge, i.e., the amount of flow traversing the edge just now.

As the name suggests, the flow behavior in outflow models is characterized by the outflow of an edge, i.e., the flow leaving an edge. Such models are first considered by Merchant and Nemhauser [55] in 1978. In both, transit time and outflow models, it is assumed that after flow enters an edge at some point in time it travels at constant speed through the edge.

This is no longer the case for wave models, which are introduced by Lighthill and Whitham [51] in 1955 and Richards [76] in 1956. Here, flow is modeled as a 1-dimensional compressible fluid. Hence, the speed of a flow particle varies while traversing an edge depending on how the current flow situation is changed.

During your study of general flow models you observe that, mostly, additional assumptions on the flow model are imposed. This is mainly done because of the particular application but also for analyzing reasons. A detailed description of such assumptions you found for example in [14]. One of the most popular assumption is the FiFo principle which ensures that flow which enters an edge earlier also leaves this edge earlier. In outflow models this assumption is even implicitly assumed, in order to obtain well-defined flow models. Other assumptions are the continuity of the transit time functions with respect to time and that current flow behavior does not depend on future flow scenarios.

At this point you decide that you know enough about general flow over time models and turn your attention to specific flow models. You find out that the deterministic queuing model introduced by Vickrey [89] is one of the most popular flow models used in simulating and analyzing traffic behavior.

The idea behind the deterministic queuing model is quite simple. It is assumed that the flow behavior on an edge is determined by a constant free flow transit time and a capacity. If an edge is unused, the free flow transit time denotes the time needed for traversing this edge. Further, the capacity bounds the flow leaving an edge. Thus, a waiting queue builds up at the end of an edge if more flow wants to leave an edge than its capacity allows. Further, it is assumed that the waiting queue has no physical length. So you think of it as a vertical line with the consequence that the free flow transit time also denotes the time for arriving at the end of the waiting queue. Thus, the flow-dependent transit time consists of the constant free flow transit time plus the time-varying waiting time. Besides you find out that the deterministic queuing model where the waiting queue has a physical length is studied by Daganzo [20] and seems to be harder to analyze.

Recalling your literature review, you observe that the deterministic queuing model is nothing else than the classical flow over time model where the storage at a node is moved to the waiting queues of its outgoing edges. Hence, in order to model an air traffic system you need something like the deterministic queuing or the classical flow over time model which forbids waiting. Unfortunately, literature in this area is rather rare. E.g., Fleischer and Skutella show that forbidding storage reduces the performance of a classical flow over time at most by a factor of 2 (see, e.g., [24]). Further, direct routing is also considered in the the area of packet routing. In packet routing problems a predetermined set of packets has to be sent through a given network. In this manner, packet routing can be seen as a kind of discrete flows over time. In this connection, it seems that Symvonis [86] produces in 1996 the earliest work which studies direct routing explicitly.

For the moment, you have learned enough about flows over time, and so you proceed your literature review in the area of game theory. You find out that game theoretic questions are already studied by Waldegrave in 1713. His research on an optimal strategy for the 2-player card game Le Her can be found in the book of Montmort [57]. Nevertheless, game theory is a very young field of mathematics which becomes important since the middle of the 20th century. A comprehensive overview of this topic is provided by popular textbooks like [31, 64, 66].

A nice aspect of game theory is that everyone understands games from early childhood. A game consists of players who compete with each other, i.e., each player searches for the best gaming strategy in order to win the game or, at least, to lose not that clearly. If, in a particular gameplay, no player can improve his situation by switching unilaterally to another strategy, the game is at equilibrium. As you mentioned already, such stable gameplays are called Nash equilibria due to the work of Nash [62]. In particular, he establishes the existence of a Nash equilibrium for arbitrary games (consisting of a finite number of players with a finite number of strategies), which generalizes a work of Neumann [63] from 1928.

You observe that in literature the performance of a Nash equilibrium is an important question which asks, how bad is egoistic behavior compared with collective action. Dubey attacks this question at first in [21], which is a rewrite of an older version of 1978. Since 1999, it has become common practice to measure the quality of Nash equilibrium as the ratio between the social outcome of a Nash equilibrium and a social optimum. This concept is introduced by Koutsoupias and Papadimitriou [47] and this ratio is called price of anarchy due to the work of Papadimitriou [68].

Having in mind your initial motivation for this literature review, you search for articles combining flow with game theory. In such routing games each player has to choose a path from a source to a sink as his strategy. Thus, in a Nash equilibrium no player should arrive earlier at his sink if he switches unilaterally to another path. You find out that in case of static flows there is already a rich literature.

The first scientific work [73] addressing this topic is written by Pigou in 1920. Pigou analyzes the selfish routing behavior on a network consisting of two parallel edges between a source and a sink node. The transit time on an edge depends continuously on the amount of flow traversing this edge. In this manner, a static flow is a Nash equilibrium if the flow-dependent transit times on both edges coincide. Later, in 1952, this observation results in the first of the two popular Wardrop's principles [90] explaining routing behavior in traffic networks. In fact, a static flow is a Nash equilibrium if and only if it satisfies Wardrop's first principle. This is mathematically verified by Beckmann, McGuire, and Winsten [9] in 1956 who also discover several equivalent formulations of static Nash equilibria. In their seminal work [77], Roughgarden and Tardos analyze the price of anarchy. One of their considerable results states that the price of anarchy for linear transit time functions is tightly bounded by $\frac{3}{4}$.

Concluding your literature review you focus on articles combining game theory and flows over time. You find out that already Wardrop [90] recognizes the necessity of time-varying transit times when considering different time periods in a road traffic network. However, it takes twenty years until Yagar [93] come up with an algorithmic treatment on this topic, and it takes another 25 years until the theory of so-called routing games over times becomes a field of active mathematical research. In between, nearly all contributions analyze Nash equilibria on small instances. As you verify by the survey paper [69] of Peeta and Ziliaskopoulos, since the middle of 1980's the number of publications increase rapidly.

There exist mainly two kinds of routing games over time. In *route choice models* each player searches for a best path from some source to some sink as in static scenarios. In *simultaneous departure-time route-choice models* each player has to choose a departure time in addition. In your opinion, the article [29] by Friesz et al. from 1993 can be seen as the groundbreaking work in this area. They formalize the notion of Nash equilibria for flows over time as a variational inequality.

Related to the deterministic queuing model, you find the early work [83] of Smith who proved in 1984 the existence of a Nash equilibrium for a very special case. Recently, this result is improved by Cominetti, Correa, and Larré [18] which holds for arbitrary networks assuming that the flow over time is piecewise constant. Their approach is strongly based on a characterization of Nash equilibria found by Koch and Skutella [44]. Unfortunately, you do not find any contribution analyzing Nash equilibria for the direct flow model.

Finally, your literature review ends up in finding this thesis [42]. You shortly think of Russel's paradox. It is feasible that this thesis contains itself. But then you recognize that, surprisingly, this thesis seems to be as made for you. It answers or at least addresses all of your questions. So you carefully start reading this thesis.

Contribution of this thesis. As you expect by the title, this thesis deals with routing games over time. In fact, the theory of routing games over time is built from scratch. Even the underlying theory of flows over time is completely revised and placed on a higher abstraction level resulting in a quite general concept of flows over time. Based on this general flow over time model, a precise notion of routing games over time is introduced. In this environment, Nash equilibria for the direct flow over time and the deterministic queuing model are characterized and analyzed.

Everything is discussed in single commodity scenarios. That is, all flow has to be sent from one unique source to one unique sink. In addition, only route choice models are considered, i.e., the departure time of each flow particle controlled by some player is known in advance. On the one hand, this is the simplest setting you can imagine. On the other hand, you think that for understanding the very own nature of routing games over time and for developing a rigorous theory, this should be the starting point.

As common in the literature, flows over time are determined via Lebesgue integrable flow rate functions. A flow rate function denotes the rate at which flow passes by some place of interest. The general flow over time model upon which this thesis is based is a transit time model, i.e., the flow behavior on a network is explained via time-varying transit time functions. In contrast to most of the known models, these transit time functions need not to be continuous with respect to time and are rather given via Lebesgue measurable functions. You somehow have the feeling that such an approach also legitimates the usage of *Lebesgue integrable* flow rate functions representing flow behavior.

Another aspect of the general flow over time model is that current edge transit times depend on the entire flow on that edge – in particular, future flow situations are allowed to influence current transit times. From a notational point of view, this makes dynamic scenarios more similar to static scenarios. Another characterization of this general flow over time model is presented via outflows. This characterization generalizes existing outflow models. In fact, outflow and transit time models can be used alternatively for representing flow behavior. Until now this works only under the FiFo assumption.

Since the general flow over time model seems to be intractable for further detailed analysis, the notion of consistent flow models is introduced. Generalizing certain observations from [14], several assumptions on a consistent flow over time are made, including the FiFo principle. In contrast to previous work, the general notion of flows over time allows a mathematical precise formulation of these assumptions. So far, some assumptions are more or less based on intuition. In addition, a notion of continuity is required ensuring that flow behavior is changed only slightly if the corresponding scenarios differ not that much. Surprisingly, you do not find such a notion of continuity in literature although this is one of the main assumptions ensuring the existence of Nash equilibria in the *static* case. Only for load dependent transit time functions, this is implicitly stated as it is mainly assumed that current transit time functions depend continuously on the current load of an edge.

You wonder why the entire approach is presented two times – once for path-based and once for edge-based flows over time. But reading this thesis you get an answer. The reason is that the theory is applied to the direct flow model which requires path-based formulations and also to the deterministic queuing model which requires an edge-based formulation. It is observed that both models are consistent. Finally, discrete aspects of flows over time are briefly discussed. As discrete models are easier to understand, this provides you with some basic intuition.

Hence, your questions concerning the modeling of traffic behavior on air lane or road networks are answered regardless of whether or not the direct flow and the deterministic queuing model are the most adequate models for this task. With respect to the routing behavior arising out of the selfish decisions of the players, routing games over times are discussed. For consistent flow models, a well-motivated definition of Nash equilibria is introduced. Based on this definition, it is shown that Nash equilibria are representable via static flows leading to nice characterizations of Nash equilibria for the direct flow and the deterministic queuing model. It turns out that a Nash equilibrium for the deterministic queuing model can be seen as a sequence of special static flows. These so-called thin flows are also analyzed in detail in this thesis.

Based on the characterization via static flows, the existence of Nash equilibria is established in general for all consistent flow over time models. This existence result is obtained via a limit approach based on a constructive algorithm. This answers your question about the computation of Nash equilibria at least partially. Concerning the uniqueness of Nash equilibria, it is observed that the direct flow model can admit quite different Nash equilibria. However, a good Nash equilibrium on a given network is computable from an earliest arrival flow. As already observed in [18], Nash equilibria for the deterministic queuing model seem to be unique. Further, Nash equilibria are efficiently computable if static thin flows can be found efficiently.

Finally, you also find answers concerning the performance of Nash equilibria. In this thesis four kinds of the price of anarchy are introduced. The evacuation price of anarchy considers the amount of flow which has been arriving at the sink of the network. The working price of anarchy, which can be seen as an average evacuation price of anarchy, includes the aspect that flow units representing employees start working after arriving at the sink. The completion time price of anarchy asks for the time needed to evacuate a given flow value. Finally, the average arrival time price of anarchy measures, as the name suggests, the average arrival time over all flow particles. Note that a social optimum for these prices of anarchy is given by an earliest arrival flow.

For the direct flow model, a nearly complete analysis of the price of anarchy is given. It is shown that, considering the good Nash equilibria, the evacuation price of anarchy is equal to 2. Further, the working price of anarchy is equal to α^2 where α is the unique solution of $2\alpha = \log \frac{\alpha+1}{\alpha-1}$. The completion time price of anarchy evaluates to $\frac{3}{4}$ which bounds also the average arrival time price of anarchy. In contrast, for the deterministic queuing model, only bounds on the price of anarchy are given. However, you believe that these bounds are based on a worst case scenario. In particular, it is shown that the evacuation price of anarchy grows at least linearly with the number of edges whereas the working price of anarchy seems to be constant. Finally, the possibly worst case scenario admits a completion time price of anarchy of $\frac{e}{e-1}$. In addition, for one special case a complete analysis is presented. That is, if each path of the considered network is a shortest one, a Nash flow is already an earliest arrival flow, i.e., all kinds of the price of anarchy are equal to 1.

Organization of this thesis. This thesis is outlined as follows. Notations and preliminary results which are needed for understanding this thesis are summarized in Section 2. The general flow over time model together with well-motivated restrictions are introduced in Chapter 3. Routing games over time are discussed in Chapter 4. In particular, the existence of Nash equilibria are established if the underlying flow over time model is consistent. In Chapter 5, Nash equilibria for the direct flow model are characterized and analyzed. Finally, the deterministic queuing model is discussed in Chapter 7. Since Nash equilibria for the deterministic queuing model are representable via thin flows, Chapter 6 contains a detailed treatment on this topic

However you remark that this thesis contains nearly no reference to the literature. Therefore, you decide to add this introduction and your own comments directly after each chapter in a separate section. There you include some remarks and bring the results into relation to known literature. Surprisingly, you already include some notes although you read this thesis for the first time. ”Mmh,“ you think, ”this seems to be a practical proof of $1 = 2$.“ Nevertheless you decide to send all comments, remarks, and corrections to the author.

Acknowledgments. Thank you for reading this acknowledgment. Without the support of many helpful and likeable people this thesis would not be completed. It is my inner need to thank these persons.

First of all, I want to thank my supervisor, Professor Dr. Martin Skutella, for introducing me into the fascinating area of routing games over time. Thanks for all the valuable discussions about this topic. I am also thankful to Professor Dr. José R. Correa for revising this thesis. Thank you both for all of your patience during the completion of this work.

Special thanks go to Ebrahim Nasrabadi for very fruitful discussions about flows over time. Thank you for improving my understanding of flows over time to a higher level and for all the nice time our families spent together. You are a great person.

Moreover, I want to thank my colleague and friend Britta Peis. Thank you for the delightful office atmosphere and your bright face every day. Say thank to Thorsten who does not let me alone with one of my vice. Also say thank to Emma for first stroking Jonas. There are not such many faithful persons like you.

I would like to thank the COGA-group for the nice working atmosphere. It is a great pleasure to be part of it.

Furthermore, I want say thanks to my parents. Thanks for supporting me such cordially, insightfully, and tolerantly. I also thank my parents-in-law for caring our children that often.

Special thanks also go to the supervisor of my wife, Michael Joswig, and, once more, to Martin Skutella. Without their unselfish support, our family would not be what it is.

Finally and from the bottom of my heart, I thank my beloved wife and our children Jonas and Alina. Thank you for all what makes life worth living. Thank you that you tolerate my mood swings. Without you, this thesis would never be completed.

Thank you all again!

Chapter 2

Basics

This Chapter provides all the basic preliminaries needed for following and understanding this thesis. In Section 2.1 we specify elementary notation. Lebesgue measurable functions and functional analysis are treated in Section 2.2. Especially, this section shall ensure that this thesis is self-contained. For this very reason and since I am not an expert in these mathematical fields, I took my knowledge from WIKIPEDIA and prove everything which I did not find there. In particular, it is possible that results are already well-known. In Section 2.3 we discuss all the basic ingredients of networks which we consider in this thesis. Static routing games just as a short introduction into the area of game theory are presented in Section 2.4. Finally, basic definitions and results concerning classical flows over time are discussed in Section 2.5.

In addition to these preliminaries, Subsection 5.4.3 uses some results from linear algebra, and Chapter 6 requires a deeper understanding of static flows. As these observations are not used anywhere else, we put the corresponding concepts within Subsection 5.4.3 and Chapter 6.

2.1 Notations and Conventions

This section introduces basic symbols and notations which are used consistently throughout this thesis. In addition, some useful results, equations, and inequalities are discussed.

Numbers. The symbol $\mathbb{N} := \{1, 2, \dots\}$ represents the set of natural numbers *without* 0. Integers, rationals, and real numbers are denoted by $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$, respectively. For a subset $R \subseteq \mathbb{R}$ of real numbers, the symbols $R_0 := R \cup \{0\}$ and $R^\infty := R \cup \{\infty\}$ contain in addition to the numbers in R the 0 and ∞, respectively.

For a natural number $n \in \mathbb{N}_0^\infty$ we define $[n] := \{1, \dots, n-1\}$ as the set of the first $n-1$ natural numbers without 0. Thus, the set $[n]_0$ contains exactly n numbers. Here, we use the conventions $[\infty] := \mathbb{N}$ and $[0]_0 := \emptyset$. Further, for a finite or countable set N we let $[N] := [|N|]$, where $|N|$ equals the numbers of elements in N. Hence, N and $[N]_0$ are of equal cardinality.

Below, we state some inequalities. One of the most popular inequalities compares the arithmetic and the geometric mean. It states that for k nonnegative real numbers $a_1, \ldots, a_k$ it holds

$$\frac{a_1 + \cdots + a_k}{k} \;\geq\; \sqrt[k]{a_1 \cdots a_k}\ . \tag{2.1}$$

Another popular inequality is the inequality between the arithmetic and the harmonic mean. It states that for k nonnegative real numbers $a_1, \ldots, a_k$ it holds

$$\begin{aligned} \frac{a_1 + \cdots + a_k}{k} &\;\geq\; \frac{k}{\frac{1}{a_1} + \ldots + \frac{1}{a_k}} \\ \Leftrightarrow \qquad \frac{1}{a_1} + \ldots + \frac{1}{a_k} &\;\geq\; \frac{k^2}{a_1 + \cdots + a_k}\ . \end{aligned} \tag{2.2}$$

Besides, the geometric mean is as least as large as the harmonic mean. The next lemma shows that if the arithmetic mean of arbitrary real numbers is close to the maximum of these numbers, each number must be close to the value of the arithmetic mean.

Lemma 2.1. *Let $a_1, \ldots, a_k \in \mathbb{R}$ be arbitrary real numbers. Further, consider $a \in \mathbb{R}$ and $\epsilon_1, \epsilon_2 > 0$ such that:*

$$\begin{aligned} a_i &\;\leq\; a + \epsilon_1 \qquad\qquad \forall i = 1, \ldots, k \\ \textit{and} \qquad \frac{1}{k}\sum_{i=1}^{k} a_i &\;\geq\; a - 2\epsilon_2\ . \end{aligned}$$

If $(k-1)\epsilon_1 \leq \epsilon$ and $k\epsilon_2 \leq \epsilon$ holds for some $\epsilon > 0$, we get

$$a_i \geq a - 3\epsilon \qquad\qquad \forall i = 1, \ldots, k\ .$$

Proof. Because each a_i is upper bounded by $a + \epsilon_1$ and their arithmetic mean is lower bounded by $a - 2\epsilon_2$, we obtain for all $i \in [k+1]$:

$$a - 2\epsilon_2 \;\leq\; \frac{1}{k}\sum_{j=1}^{k} a_j \;=\; \frac{a_i}{k} + \frac{1}{k}\sum_{j \neq i} a_j \;\leq\; \frac{a_i}{k} + \Big(1 - \frac{1}{k}\Big)(a + \epsilon_1)\ .$$

Rearranging this inequality to a lower bound on a_i and using the bounds on ϵ result for all $i \in [k+1]$ in

$$\begin{aligned} a_i &\;\geq\; k\Big(a - 2\epsilon_2 - \Big(1 - \frac{1}{k}\Big)(a + \epsilon_1)\Big) \\ &\;=\; a - (k-1)\epsilon_1 - 2k\epsilon_2 \\ &\;\geq\; a - 3\epsilon\ . \end{aligned}$$

□

The following remark notes that Lemma 2.1 is generalizable to arbitrary convex combinations of the real numbers $a_1, \ldots, a_k$.

Remark 2.2. Let $a_1, \ldots, a_k \in \mathbb{R}$ be arbitrary real numbers, $\alpha_1, \ldots, \alpha_k > 0$ be such that $\sum_{i=1}^{k} \alpha_i = 1$, and $\alpha_{\min} > 0$ be such that $\alpha_i \geq \alpha_{\min}$ for all $i = 1, \ldots, k$. Further, consider $a \in \mathbb{R}$ and $\epsilon_1, \epsilon_2 > 0$ satisfying

$$\begin{aligned} a_i &\;\leq\; a + \epsilon_1 \qquad\qquad \forall i = 1, \ldots, k \\ \text{and} \qquad \sum_{i=1}^{k} \alpha_i a_i &\;\geq\; a - 2\epsilon_2\ . \end{aligned}$$

If $\epsilon_1 \leq \frac{\alpha_{\min}}{1-\alpha_{\min}} \cdot \epsilon$ and $\epsilon_2 \leq \alpha_{\min} \cdot \epsilon$ holds for some $\epsilon > 0$, we get:

$$a_i \geq a - 3\epsilon \qquad \forall i = 1, \ldots, k\ .$$

Sets and Families. Sets are usually a collection of unique elements meaning that a set contains each element exactly once. In contrast families can contain multiple copies of the same element. Mainly, we use capital letters for denoting sets and families of ordinary elements and calligraphic letters for denoting more complex sets and families. For example, the symbols I and P are used for denoting a real interval and a path of a graph, respectively. So I is a set of numbers and P a family (or sequence) of edges. On the other hand, $\mathcal{P}$ and $\mathcal{F}$ are used for representing a set of paths in a graph and a family of functions, respectively.

As already used, the cardinality of a set A is denoted by $|A|$. That is, $|A|$ is equal to the number of elements in A or to ∞ depending on whether or not A contains a finite number of elements. Further, the power set of A, which contains all subsets of A, is denoted by 2^A implying that the cardinality of 2^A is $2^{|A|}$. Moreover, we use the following definitions in case $A \subseteq \mathbb{R}$ is a subset of the real numbers.

Definition 2.3 (Open, Closed, and Compact Set)**.** A subset $A \subseteq \mathbb{R}$ is *open* if and only if it is the countable union of disjoint open intervals. That is, there exists an $N \in \mathbb{N}_0^\infty$ and real numbers $a_i, b_i \in \mathbb{R}$ for each $i \in [N]_0$ with $a_i < b_i$ for all $i \in [N]_0$ and $b_{i-1} < a_i$ for all $i \in [N]$ such that

$$A = \bigcup_{i=0}^{N} (a_i, b_i)\ .$$

The set A is *closed* if its complement $A^c := \mathbb{R} \setminus A$ is open. If A is additionally bounded, i.e, $A \subseteq [a, b]$ for some $a, b \in \mathbb{R}$, it is called *compact*.

Of course, there exist many other equivalent definitions of open, closed, and compact sets. For example, compact sets can be defined as follows. A set $A \subseteq \mathbb{R}$ is compact if and only if every, possibly uncountable, open cover of A contains a finite cover of A. Here, an open cover is a family of open sets whose union contains A. Further, it is well-known that even the uncountable union of open sets is again an open set. The following definition of a shifted set is used below for defining shifted functions.

Definition 2.4 (Shifted Set)**.** The *shifted set* $A - \tau$ for some $A \subseteq \mathbb{R}$ and some $\tau \in \mathbb{R}$ is defined as

$$A - \tau := \{\theta - \tau \mid \theta \in A\}$$

For explaining Definition 2.4, let $A := \bigcup_{i=0}^{N} (a_i, b_i)$ be an open set and $\tau \in \mathbb{R}_+$ be a nonnegative real number. Then we have $A - \tau = \bigcup_{i=0}^{N} (a_i - \tau, b_i - \tau)$. In this manner, on the real line, $A - \tau$ results from shifting A by τ units to the left.

Null Set, Almost Everywhere, Essentially. Throughout this thesis, especially in this introduction, λ denotes the Lebesgue measure. Informally, λ assigns to subsets of $\mathbb{R}$ its natural size. In this sense, the Lebesgue measure $\lambda(I)$ of an interval $I := (a, b)$ is equal to the length of I, i.e., $\lambda(I) := b - a$. Further, the Lebesgue measure of two or at least countable many pairwise disjoint real subsets should be equal to the sum of the Lebesgue measures of the subsets. Hence, as an open set $O := \bigcup_{i\in\mathbb{N}} I_i$ is the countable union of disjoint open intervals I_i, this means $\lambda(O) := \sum_{i\in\mathbb{N}} \lambda(I_i)$. The Lebesgue measure and Lebesgue measurable sets are treated more formally at the beginning of Section 2.2. Intuitively, a null set $N \subseteq \mathbb{R}$ is a set of size 0, i.e., $\lambda(N) = 0$ should hold for N. This motivates the following definition.

Definition 2.5 (Null Set)**.** A real subset $N \subseteq \mathbb{R}$ is a *null set* if and only if for every $\epsilon > 0$ there exists an open set O containing N with $\lambda(O) \leq \epsilon$.

Definition 2.5 means that a null set can be covered by arbitrary small open sets. Further, it is well-known that the countable union of null sets is again a null set. As a single real number is, of course, a null set, we know that the set of endpoints $\bigcup_{i\in\mathbb{N}}\{a_i, b_i\}$ of an open set $O := \bigcup_{i\in\mathbb{N}}(a_i, b_i)$ is a null set.

When dealing with Lebesgue integrable functions, as it is often the case in this thesis, one also have to deal with logical expressions over the real numbers which hold almost everywhere, i.e., everywhere except on a null set. For example, the integral over an open interval $I \neq \emptyset$ of a Lebesgue integrable non-negative function $f : I \to \mathbb{R}_+$ is equal to 0 if and only if $f(\theta) = 0$ holds for almost all $\theta \in I$. That is, there exists a null set $N \subseteq I$ such that $f(\theta) = 0$ for all $\theta \in I \setminus N$. In this case we also say that $f(\theta) = 0$ holds essentially.

Definition 2.6 (Essentially)**.** Let A_θ be a logical expression for all $\theta \in S \subseteq \mathbb{R}$. We say that A_θ holds *essentially* or for *almost all* $\theta \in S$ if and only if there exists a null set N such that A_θ is valid for all $\theta \in S \setminus N$.

In the following, we derive a sufficient condition for deciding whether a logical expression holds essentially on $\mathbb{R}_+$. For this we need the following lemma.

Lemma 2.7. *Let $S \subseteq \mathbb{R}_+$ be a subset of $\mathbb{R}_+$. Further, for every $\theta \in \mathbb{R}_+$ we have that $\theta \in S$ or there exists an $\epsilon > 0$ such that for almost all $\theta' \in (\theta, \theta + \epsilon)$ it holds $\theta' \subseteq S$. Then S covers almost every point of $\mathbb{R}_+$, i.e., $\mathbb{R}_+ \setminus S$ is a null set.*

Proof. Let $R := \mathbb{R}_+ \setminus S$ be the set of points which are not contained in S. Then for each $\theta \in R$, there exists an $\epsilon_\theta > 0$ such that almost every point of $(\theta, \theta + \epsilon_\theta)$ is contained in S. Defining $S_1 := \bigcup_{\theta\in R}(\theta, \theta + \epsilon_\theta)$, we first show that $S_1 \setminus S$ is a null set.

Since S_1 is open, it is the countable union of disjoint open intervals. Hence, it is enough to show that $I \setminus S$ is a null set for all such disjoint intervals. Let I be such an interval out of the countable union. We show that $[a, b] \setminus S$ is a null set for all $a, b \in I$ with $a < b$. Since $[a, b]$ is compact and $(\theta, \theta + \epsilon_\theta)_{\theta\in R}$ is an open cover of $[a, b]$, there exists a $k \in \mathbb{N}$ and $\theta_1, \ldots, \theta_k \in R$ such that:

$$[a, b] \subseteq \bigcup_{i=1}^{k} (\theta_i, \theta_i + \epsilon_{\theta_i}) \ .$$

Since $(\theta_i, \theta_i + \epsilon_{\theta_i}) \setminus S$ is a null set for each $i = 1, \ldots, k$, also $[a, b] \setminus S$ is a null set. This shows that $S_1 \setminus S$ is a null set.

On the other hand, we know that each $\theta \in R$ coincides with the left boundary of some open interval contained in S_1. Defining $\bar{S}_1 := \bigcup_{k \in \mathbb{N}} [a_k, b_k)$ assuming that $S_1 = \bigcup_{k \in \mathbb{N}} (a_k, b_k)$ is the countable union of the disjoint open intervals (a_k, b_k), this shows $R \subseteq \bar{S}_1 \setminus S$ because $R \cap S = \emptyset$ by definition. Further, $\bar{S}_1 \setminus S$ is contained in $(\bar{S}_1 \setminus S_1) \cup (S_1 \setminus S)$. Since $\bar{S}_1 \setminus S_1 = \{a_1, a_2, \ldots\}$ is countable, it is also a null set. Therefore, $R = \mathbb{R}_+ \setminus S$ is contained in a null set, and hence, it is itself a null set. □

The previous Lemma 2.7 can be used to establish a lemma where a statement has to be proven almost everywhere on $\mathbb{R}_+$.

Corollary 2.8. *Let A_θ be a logical statement for all $\theta \in \mathbb{R}_+$. Assume that for each $\theta \geq 0$ the statement A_θ holds or there exists an $\epsilon > 0$ such that $A_{\theta'}$ holds for almost every $\theta' \in (\theta, \theta + \epsilon)$. Then A_θ is essentially valid, i.e., it holds for almost all $\theta \in \mathbb{R}_+$.*

Proof. Let $S \subseteq \mathbb{R}_+$ be the set such that A_θ holds for every $\theta \in S$. Then S satisfies the assumptions of Lemma 2.7, and hence, covers almost every point of $\mathbb{R}_+$. □

Functions. In this thesis, we mainly consider Lebesgue integrable and absolutely continuous functions. It is well-known that these two kinds of functions are strongly related. Let $I := [a, b]$ be a closed interval, i.e., $a, b \in \mathbb{R}$ with $a \leq b$. For a Lebesgue integrable function $f : I \to \mathbb{R}$, the function $F : I \to \mathbb{R}$ defined by $F(\theta) := \int_a^\theta f(\vartheta)\, d\vartheta$ for all $\theta \in [a, b]$ is absolutely continuous. And vice versa, for every absolutely continuous function $F : I \to \mathbb{R}_+$ with $F(a) = 0$, there exists a Lebesgue integrable function $f : I \to \mathbb{R}$ such that $F(\theta) := \int_a^\theta f(\vartheta)\, d\vartheta$ holds for all $\theta \in [a, b]$. In this case f is called *derivative* of F, and F is called *antiderivative* of f. In Section 2.2 we discuss these functions more in detail.

Throughout this thesis, we denote Lebesgue integrable functions with small letters and absolutely continuous functions with capital letters such that corresponding functions are denoted with the same letter. Moreover, given a small letter denoting a Lebesgue integrable function, we implicitly assume that the corresponding capital letter denotes the corresponding absolutely continuous function and vice versa.

Most functions we are working with are nonnegative functions $\mathbb{R}_+ \to \mathbb{R}_+$ defined on the nonnegative real line. Nevertheless, we always assume that the domain of each function is the entire real line. Thus, we actually work with functions from $\mathbb{R}$ to $\mathbb{R}$. Here, we implicitly assume that a given function $f : A \to B$ with $A, B \subseteq \mathbb{R}$ can be extended to a function $\tilde{f} : \mathbb{R} \to \mathbb{R}$ in an obvious manner. Of course, if we would use $\tilde{f}$ instead of f in some approach, $\tilde{f}$ should show the same behavior as f. For example, if $f : A \to B$ is a Lebesgue integrable function, we actually work with $\tilde{f} : \mathbb{R} \to \mathbb{R}$ defined by

$$\tilde{f}(\theta) := \begin{cases} f(\theta) & \text{if } \theta \in A \\ 0 & \text{otherwise} \end{cases}.$$

Note that by the convention described above, this already defines $\tilde{F}$ for an absolutely continuous function $F : A \to B$. It means that $\tilde{F}$ must be the absolutely continuous function corresponding to $\tilde{f}$, where f is the Lebesgue integrable function with respect to F.

Further, we consider the set of functions $f : \mathbb{R} \to \mathbb{R}$ as a vector space over $\mathbb{R}$. In particular, this ensures that we have an addition and a scalar multiplication. That is:

$$\begin{aligned} & \big(f+g\big)(\theta) &:=& \; f(\theta)+g(\theta) && \forall f,g:\mathbb{R}\to\mathbb{R},\ \theta\in\mathbb{R} \\ \text{and} \quad & \big(\alpha\cdot f\big)(\theta) &:=& \; \alpha\cdot f(\theta) && \forall f:\mathbb{R}\to\mathbb{R},\ \alpha,\theta\in\mathbb{R}\ . \end{aligned}$$

In the same manner, we define the minimum of two functions by

$$\min\{f,g\}(\theta) \;:=\; \min\big\{f(\theta),g(\theta)\big\} \qquad \forall f,g:\mathbb{R}\to\mathbb{R},\ \theta\in\mathbb{R}\ .$$

Besides, we also use the following notations for comparing functions. Let $A \subseteq \mathbb{R}$ be a set of real numbers, $f, g : A \to \mathbb{R}$ be two function defined on A and $\alpha \in \mathbb{R}$ be a real number. Then we write:

$$\begin{aligned} & f<g && \Leftrightarrow && f(\theta)<g(\theta) && \forall\theta\in A\ , \\ & f=g && \Leftrightarrow && f(\theta)=g(\theta) && \forall\theta\in A\ , \\ \text{and} \quad & f=\alpha && \Leftrightarrow && f(\theta)=\alpha && \forall\theta\in A\ . \end{aligned}$$

If f and g are Lebesgue integrable functions, the right hand sides of the equivalences has to hold only for almost all $\theta \in A$. The following basic definitions are often used in this thesis.

Definition 2.9 (Restricted Function)**.** Let $f : A \to B$ be a (Lebesgue integrable) function for some $A, B \subseteq \mathbb{R}$ and $\tilde{A} \subseteq A$ be a subset of A. Then the *restricted function* $f|_{\tilde{A}} : \tilde{A} \to B$ is defined as

$$f|_{\tilde{A}}(\theta) := f(\theta) \qquad\qquad \forall\theta\in\tilde{A}\ .$$

Note that by convention f just as $f|_{\tilde{A}}$ are defined on the entire real line. In particular, this ensures that $f|_{\tilde{A}}$ equals 0 outside of $\tilde{A}$ as expected by the meaning of a restriction. In contrast, the restriction $F|_{\tilde{A}}$ of an absolutely continuous function $F : A \to B$ is in general not equal to 0 outside of $\tilde{A}$. Even it may also differ from F over $\tilde{A}$. Rather, by convention $F|_{\tilde{A}}$ is the antiderivative of $f|_{\tilde{A}}$, where f is the derivative of F.

Definition 2.10 (Shifted Function)**.** Let $f : A \to B$ be a function for some sets $A, B \subseteq \mathbb{R}$, and let $\tau \in \mathbb{R}$ be a real number. Then the *shifted function* $f - \tau : (A + \tau) \to B$ is defined as

$$\big(f-\tau\big)(\theta) := f(\theta-\tau) \qquad\qquad \forall\theta\in A+\tau\ .$$

So if $\tau > 0$ ($\tau < 0$) then the graph of f is shifted to the right (left) in order to obtain the graph of $f - \tau$.

Note that it holds $-(f-\tau) = (-f)-\tau$. Hence, $-(f-\tau)-\tau$ is equal to $-f-2\tau$ and *not* to $-f$. It worth to note that shifting an absolutely continuous function works in the same manner as the shifting a Lebesgue integrable function. The reason for that is that $F - \tau$ is the antiderivative of $f - \tau$ if and only if F is the antiderivative of f.

Shifting a restricted function and restricting a shifted function works as follows: Let f be an arbitrary (Lebesgue integrable) function, $A \subseteq \mathbb{R}$ and $\tau \in \mathbb{R}$. Then we have for all $\theta \in \mathbb{R}$

$$\begin{aligned} \big(f|_A - \tau\big)(\theta) = f|_A(\theta - \tau) &= \begin{cases} f(\theta-\tau) & \text{if } \theta - \tau \in A \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} f(\theta-\tau) & \text{if } \theta \in A + \tau \\ 0 & \text{otherwise} \end{cases} \\ \text{and} \qquad \big(f - \tau\big)\big|_A(\theta) &= \begin{cases} \big(f-\tau\big)(\theta) & \text{if } \theta \in A \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} f(\theta-\tau) & \text{if } \theta \in A \\ 0 & \text{otherwise} \end{cases} . \end{aligned}$$

This shows that $f|_A - \tau = (f - \tau)|_{A+\tau}$ and $(f - \tau)|_A = f|_{A-\tau} - \tau$ are valid.

Mainly in examples, we use the following functions. The symbol id denotes the *identity function id* : $\mathbb{R}_+ \to \mathbb{R}_+$ on $\mathbb{R}_+$, i.e., $\mathrm{id}(\theta) := \theta$ for all $\theta \in \mathbb{R}_+$. Note that the identity function is both Lebesgue integrable and absolutely continuous. Hence, we could be in a conflict if we consider restrictions of the identity function. However, as we only use restrictions of the kind $\mathrm{id}|_{[0,\theta]}$, we avoid such conflicts. Further, for a real set $A \in \mathbb{R}_+$, the *characteristic function* $\chi_A : \mathbb{R} \to \{0, 1\}$ of A is defined by

$$\chi_A(\theta) := \begin{cases} 1 & \text{if } \theta \in A \\ 0 & \text{otherwise} \end{cases} .$$

Characteristic functions are also called indicator functions in the mathematical society. It is well-known that characteristic function are Lebesgue integrable if and only if A is a Lebesgue measurable set.

For some function $f : \mathbb{R} \to \mathbb{R}$ the symbol f^{-1} denotes a set-valued function called the inverse function of f. A set valued function $\mathbb{R} \to 2^{\mathbb{R}}$ assigns to each real number a real subset. The inverse function is defined as

$$a \mapsto \big\{\theta \mid f(\theta = a)\big\} .$$

The composition $f \circ g : \mathbb{R} \to \mathbb{R}$ of two functions $f, g : \mathbb{R} \to \mathbb{R}$ is defined as $\theta \mapsto f(g(\theta))$. If $g : \mathbb{R} \to 2^{\mathbb{R}}$ is a set-valued function such that f behaves constant on each image of g then the composition of $f \circ g$ is well-defined and used in this sense throughout this thesis. E.g., for each surjective function $f : \mathbb{R} \to \mathbb{R}$, which attains every real number, the composition $f \circ f^{-1}$ always equals the identity function.

2.2 Lebesgue Measurable Functions and Functional Analysis

This section imparts basic knowledge concerning Lebesgue measurable and absolutely continuous functions which is necessary for understanding this thesis.

Further, we consider sequences of such functions and their properties. As already mentioned, I am not an expert in these mathematical disciplines. Therefore, every result which is not contained in WIKIPEDIA is proven ensuring that this thesis is self-contained. Before we start considering Lebesgue measurable functions, we introduce the Lebesgue measure as this measure is inseparably associated with the theory of Lebesgue measurable functions. This Theory is introduced by Lebesgue in his famous thesis [49] in 1902. For more background knowledge of the topic presented in this section, we refer to popular textbooks like [78, 79].

As mentioned in the previous section, the Lebesgue measure assigns to each suitable real subset its natural size. In the following, we define the Lebesgue measure via THE Borel measure on $\mathbb{R}$ which we denote by λ for the moment. The set of all measurable sets with respect to λ is called Borel σ-algebra. The Borel σ-algebra is the smallest set containing all real intervals which is closed under the countable union and the complement. The Borel measure of a real interval I with end points $a \leq b$ is set to its length, i.e., $\lambda(I) := b - a$. Then THE Borel measure is completely defined by

$$\lambda\Big(\bigcup_{i=0}^{\infty} A_i\Big) \;:=\; \sum_{i=0}^{\infty} \lambda(A_i)$$

for each family $(A_i)_{i\in\mathbb{N}_0}$ of pairwise disjoint measurable sets $A_i \subseteq \mathbb{R}$. In particular, this ensures $\lambda(\emptyset) = 0$ and $\lambda(A) \geq 0$ for all measurable sets A. As in Section 2.1, a measurable set $N \subseteq \mathbb{R}$ is called null set if $\lambda(A) = 0$.

Intuitively, each subset of a null set should have a measure of 0. Unfortunately, there exist subsets of null sets which are not measurable. The Lebesgue measure and THE Borel measure differ exactly in these sets. That is, the set of all Lebesgue measurable sets equals the smallest set containing all intervals and all subsets of null sets which is closed under the countable union and the complement. In this sense, the Lebesgue measure makes THE Borel measure complete.

However, for the purpose of this thesis it is neither important that the Lebesgue measure and THE Borel measure differ by null sets nor that there exist nonmeasurable sets. In the following, the symbol λ denotes the Lebesgue measure and $\mathcal{B}(A)$ the set of all measurable sets contained in some measurable set A. In case $A := \mathbb{R}_+$ we simply write $\mathcal{B}$ instead of $\mathcal{B}(\mathbb{R}_+)$. The following proposition shows well-known properties of the Lebesgue measure.

Proposition 2.11. *Let $\lambda : \mathcal{B} \to \mathbb{R}_+$ be the Lebesgue measure and $(A_i)_{i\in\mathbb{N}_0}$ be a family of measurable sets $A_i \in \mathcal{B}$ for all $i \in I$.*

(i) The Lebesgue measure is countably subadditive, i.e.:

$$\lambda\Big(\bigcup_{i=0}^{\infty} A_i\Big) \leq \sum_{i=0}^{\infty} \lambda(A_i) \; .$$

(ii) The Lebesgue measure is continuous from above, i.e., if $A_i \subseteq A_{i-1}$ holds for all $i \in \mathbb{N}$ and at least one A_i has finite measure then

$$\lambda\Big(\bigcap_{i=0}^{\infty} A_i\Big) = \lim_{i\to\infty} \lambda(A_i) \; .$$

(iii) The Lebesgue measure is outer regular, i.e., for all $A \in \mathcal{B}$ we have

$$\lambda(A) = \inf\big\{\lambda(O) \bigm| O \text{ is open and } A \subseteq O\big\} .$$

In the following, we introduce several classes of real-valued functions which are mainly used for defining flows over time. As already mentioned, we primary consider nonnegative functions from $\mathbb{R}_+$. In this thesis, time-dependent transit time functions are represented via Lebesgue measurable functions.

Definition 2.12 (Lebesgue Measurable Function)**.** A function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is *Lebesgue measurable* if and only if the preimage of every measurable set is measurable, i.e., $f^{-1}(A) \in \mathcal{B}$ for all $A \in \mathcal{B}$.

Lebesgue integrable functions are used as flow rate functions which show how the rate is changed over time at which flow passes by some point.

Definition 2.13 (Lebesgue Integrable Function)**.** Let $A \in \mathcal{B}$ be a measurable set. A Lebesgue measurable function $f : A \to \mathbb{R}_+$ is called Lebesgue integrable if and only if the Lebesgue integral $\int_A f(\vartheta)\, d\vartheta$ is finite.

A function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is called *locally Lebesgue integrable* if and only if the restricted function $f|_K$ is Lebesgue integrable for each compact set $K \subseteq \mathbb{R}_+$.

We assume that the reader is somehow familiar with the notion of the Lebesgue integral. Nevertheless, we discuss some basics in the following. The Lebesgue integral is constructed upon the very essential assumption

$$\int_{\mathbb{R}_+} \chi_A(\vartheta)\, d\vartheta \;:=\; \lambda(A) \qquad\qquad \forall A \in \mathcal{B} .$$

The following proposition summarizes well-known facts of the Lebesgue integral.

Proposition 2.14. *Well-known properties of the Lebesgue integral are:*

(i) The Lebesgue integral is monotone, i.e., for any two locally Lebesgue integrable functions f and g with $f \leq g$, we have $\int_{\mathbb{R}_+} f(\vartheta)\, d\vartheta \leq \int_{\mathbb{R}_+} g(\vartheta)\, d\vartheta$.

(ii) Monotone Convergence Theorem: For a sequence $(f_i)_{i\in\mathbb{N}_0}$ of nonnegative locally Lebesgue integrable functions with $f_{i-1} \leq f_i$ for all $i \in \mathbb{N}$, we have

$$\lim_{i\to\infty} \int_{\mathbb{R}_+} f_i(\vartheta)\, d\vartheta \;=\; \int_{\mathbb{R}_+} \lim_{i\to\infty} f_i(\vartheta)\, d\vartheta .$$

(iii) Dominated Convergence Theorem: Let $(f_i)_{i\in\mathbb{N}_0}$ be a sequence of nonnegative locally Lebesgue integrable functions converging pointwise to some f. If there exist a locally Lebesgue integrable function g satisfying $f_i \leq g$ for all $i \in \mathbb{N}_0$, we have

$$\lim_{i\to\infty} \int_{\mathbb{R}_+} f_i(\vartheta)\, d\vartheta \;=\; \int_{\mathbb{R}_+} f(\vartheta)\, d\vartheta .$$

Although Lebesgue integrable flow rate functions are commonly used for representing flows behavior, we rather use absolutely continuous functions for defining flows over time. In fact, absolutely continuous functions stand for cumulative flow functions. A value of a cumulative flow function determines the current amount of flow which has been passing by some point of interest.

Definition 2.15 (Absolutely continuous)**.** A function $F : I \to \mathbb{R}_+$ is called *absolutely continuous* on some real interval $I \subseteq \mathbb{R}_+$ if and only if, for all $\epsilon > 0$, there exists a $\delta > 0$ such that every finite sequence $((a_i, b_i))_{i=1,\dots,k}$ of pairwise disjoint sub-intervals of I satisfies

$$\sum_{i=1}^{k} |b_i - a_i| < \delta \qquad \Rightarrow \qquad \sum_{i=1}^{k} \big|F(b_i) - F(a_i)\big| < \epsilon \ .$$

A function $F : \mathbb{R}_+ \to \mathbb{R}_+$ is called locally absolutely continuous if and only if F is absolutely continuous on each compact interval I.

Throughout this thesis we mainly deal with locally Lebesgue integrable and locally absolutely continuous functions from $\mathbb{R}_+$ to $\mathbb{R}_+$. Therefore, we usually omit the term *locally* and assume implicitly that the desired property holds locally – except in this section.

Intuitively, the derivative of some cumulative flow function should coincide with the corresponding flow rate function. Vice versa, the antiderivative of some flow rate function should coincide with the corresponding cumulative flow function. This insight is known as the fundamental theorem of calculus in some general variant.

Proposition 2.16 (Fundamental Theorem of Calculus)**.** *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally Lebesgue integrable function. Then the function $F : \mathbb{R}_+ \to \mathbb{R}_+$ defined by*

$$\theta \mapsto \int_0^\theta f(\vartheta)\, d\vartheta$$

is nondecreasing and locally absolutely continuous with $F(0) = 0$.

Vice versa let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing, locally absolutely continuous function with $F(0) = 0$. Then there exists a locally Lebesgue integrable function $f : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying

$$F(\theta) = \int_0^\theta f(\vartheta)\, d\vartheta \qquad\qquad \forall \theta \in \mathbb{R}_+ \ .$$

In both cases F is differentiable almost everywhere on the nonnegative real line $\mathbb{R}_+$, and it holds $\frac{dF}{d\theta}(\theta) = f(\theta)$ for almost all $\theta \in \mathbb{R}_+$. In this manner, F is the antiderivative *of f, and f is the* derivative *of F.*

In the following, we prove some results and introduce some definitions for the function classes defined above. There are two criteria for putting them here. Either the corresponding object is such basic or technical that it would impair the readability or it is used at several points of this thesis. Because of the idea behind cumulative flow rate functions, these functions must be nondecreasing as cumulative flow only adds up. The following proposition deals with such nondecreasing cumulative flow functions.

Proposition 2.17. *Let $F_1 : \mathbb{R}_+ \to \mathbb{R}_+$ and $F_2 : \mathbb{R}_+ \to \mathbb{R}_+$ be two nondecreasing functions such that their difference $F_2 - F_1$ is nondecreasing. Then F_1 is locally absolutely continuous if this holds for F_2.*

Proof. The following inequality holds for all $a, b \in \mathbb{R}_+$ with $a \leq b$. The "="-sign holds as we only add artificial zeros. The "$\leq$" follows from the assumption that $F_2 - F_1$ is nondecreasing.

$$\begin{aligned} F_1(b) - F_1(a) &= F_1(b) - F_2(b) + F_2(b) - F_2(a) + F_2(a) - F_1(a) \\ &\leq F_2(b) - F_2(a) \end{aligned} \tag{2.3}$$

In order to prove that F_1 is locally absolutely continuous, consider some interval $I \subseteq \mathbb{R}_+$, and fix an $\epsilon > 0$. Since F_2 is locally absolutely continuous, there exists an $\delta > 0$ such that every finite sequence $((a_i, b_i))_{i=1,\ldots,k}$ of pairwise disjoint sub-intervals of I satisfies

$$\sum_{i=1}^{k} |b_i - a_i| < \delta \qquad \Rightarrow \qquad \sum_{i=1}^{k} \big|F_2(b_i) - F_2(a_i)\big| < \epsilon \ .$$

Because of (2.3) this shows

$$\sum_{i=1}^{k} |b_i - a_i| < \delta \qquad \Rightarrow \qquad \sum_{i=1}^{k} \big|F_1(b_i) - F_1(a_i)\big| < \epsilon$$

as F_1 and F_2 are nondecreasing by assumption. This shows that F_1 is locally absolutely continuous. □

The following proposition approximates integrals of value 0 and is directly implied by the construction of the Lebesgue integral.

Proposition 2.18. *Let $A \in \mathcal{B}$ be a measurable set and $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally Lebesgue integrable function which is essentially zero on A, i.e., $\int_A f(\vartheta)\, d\vartheta = 0$. Then, for all $\epsilon > 0$, there exists an open set O such that*

$$\int_O f(\vartheta)\, d\vartheta \leq \epsilon \ .$$

Proof. This proposition is a direct corollary of the dominated convergence theorem (see Proposition 2.14(iii)). Since the Lebesgue measure is outer regular (see Proposition 2.11(iii)), there exists a sequence $(O_i)_i \in \mathbb{N}_0$ with $O_i \subseteq O_{i-1}$ such that $\lim_{i\to\infty} \lambda(O_i) = \lambda(A)$. Hence, the sequence $(f|_{O_i})_{i\in\mathbb{N}_0}$ is bounded by $f|_{O_0}$ and converges almost everywhere to $f|_A$ if i goes to ∞. The dominated convergence theorem states that

$$\lim_{i\to\infty} \int_K f|_{O_i}(\vartheta) = \int_K f|_A(\vartheta)$$

for all compact sets $K \subseteq \mathbb{R}_+$. Hence, for all $\epsilon \in (0,1)$ and for all $k \in \mathbb{N}$, there exists an i_k such that:

$$\int_{[k-1,k+1]} f|_{O_{i_k}}(\vartheta) - \int_K f|_A(\vartheta) \leq \epsilon^k \ .$$

Choosing $O := \bigcup_{k\in\mathbb{N}} O_{i_k} \cap (k-1, k+1)$, we know that O is open, contains A, and satisfies

$$\int_{\mathbb{R}_+} f|_O(\vartheta) - \int_{\mathbb{R}_+} f|_A(\vartheta) \leq \frac{\epsilon}{1-\epsilon} \ .$$

As $\frac{\epsilon}{1-\epsilon}$ goes to 0 if ϵ tends to 0, this finishes the proof. □

Throughout this thesis it is important to identify the points in time at which flow is distributed leading to the notion of the support.

Definition 2.19 (Support)**.** Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally Lebesgue integrable function. The *support* $\operatorname{supp}(f)$ of f is defined as the set of points where f is nonnegative, i.e.,

$$\operatorname{supp}(f) := \left\{\theta \in \mathbb{R}_+ \;\middle|\; f(\theta) > 0\right\} .$$

The *support* $\operatorname{supp}(F)$ of a locally absolutely continuous, nondecreasing function $F : \mathbb{R}_+ \to \mathbb{R}_+$ is defined by $\operatorname{supp}(F) := \operatorname{supp}(f)$ where f is a derivative of F. A set $S \in \mathcal{B}$ is called *F-supporting* if and only if $\operatorname{supp}(F) \setminus S$ is a null set. Further, a set $N \in \mathcal{B}$ is called *F-null set* if and only if $N \cap \operatorname{supp}(F)$ is a null set.

Note that the support of a locally absolutely continuous nondecreasing function is only defined up to a null set. Nevertheless, this causes no conflict for the definitions of an F-supporting and an F-null set. In addition, the following proposition shows that, in general, this is not a problem.

Proposition 2.20. *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally Lebesgue integrable function. Then for an arbitrary Lebesgue measurable set $A \in \mathcal{B}$ we have:*

$$\int_{\mathbb{R}_+} f(\vartheta)\, d\vartheta = 0 \qquad\qquad \Leftrightarrow \qquad\qquad \lambda\big(\operatorname{supp}(f)\big) = 0 .$$

In particular, for an arbitrary measurable set $A \subseteq \mathbb{R}_+$, this implies

$$\int_{A} f(\vartheta)\, d\vartheta = 0 \qquad\qquad \Leftrightarrow \qquad\qquad \lambda\big(A \cap \operatorname{supp}(f)\big) = 0 .$$

Proof. Since $\int_A f(\vartheta)\; d\vartheta = \int_{\mathbb{R}_+} f|_A(\vartheta)\; d\vartheta$ and $\operatorname{supp}(f|_A) = A \cap \operatorname{supp}(f)$ holds by definition, the second equivalence is directly implied by the first equivalence. For proving the first statement, we assume $\lambda(\operatorname{supp}(f)) = 0$ and define $f_i(\theta) := \min\{i, f(\theta)\}$ for all $\theta \in \mathbb{R}_+$ and all $i \in \mathbb{N}_0$. Since the Lebesgue integral is monotone and $0 \leq f_i \leq i \cdot \chi_{\operatorname{supp}(f)}$ holds, we obtain

$$0 \;\leq\; \int_{\mathbb{R}_+} f_i(\vartheta)\; d\vartheta \;\leq\; i \cdot \int_{\mathbb{R}_+} \chi_{\operatorname{supp}(f)}(\vartheta)\; d\vartheta \;=\; i \cdot \lambda\big(\operatorname{supp}(f)\big) \;=\; 0$$

implying $\int_{\mathbb{R}_+} f_i(\vartheta)\; d\vartheta = 0$. Further, the sequence $(f_i)_{i\in\mathbb{N}_0}$ converges pointwise to f and satisfies $f_{i-1} \leq f_i$ for all $i \in \mathbb{N}$. Hence, the monotone convergence theorem shows

$$\int_{\mathbb{R}_+} f(\vartheta)\; d\vartheta \;=\; \lim_{i\to\infty} \int_{\mathbb{R}_+} f_i(\vartheta)\; d\vartheta \;=\; 0 .$$

For proving the other direction, assume $\int_{\mathbb{R}_+} f(\vartheta)\; d\vartheta = 0$, and for all $i \in \mathbb{N}_0$ and all $\theta \in \mathbb{R}_+$, set $f_i(\theta) := \min\{1, i \cdot f(\theta)\}$. Since the Lebesgue integral is monotone and $0 \leq f_i \leq i \cdot f$ holds, we obtain

$$0 \;\leq\; \int_{\mathbb{R}_+} f_i(\vartheta)\; d\vartheta \;\leq\; i \cdot \int_{\mathbb{R}_+} f(\vartheta)\; d\vartheta \;=\; 0$$

implying $\int_{\mathbb{R}_+} f_i(\vartheta)\, d\vartheta = 0$. Further, the sequence $(f_i)_{i\in\mathbb{N}_0}$ converges pointwise to $\chi_{\mathrm{supp}(f)}$ and satisfies $f_{i-1} \leq f_i$ for all $i \in \mathbb{N}$. Hence, the monotone convergence theorem shows

$$\lambda\big(\mathrm{supp}(f)\big) \;=\; \int_{\mathbb{R}_+} \chi_{\mathrm{supp}(f)}(\vartheta)\, d\vartheta \;=\; \lim_{i\to\infty} \int_{\mathbb{R}_+} f_i(\vartheta)\, d\vartheta \;=\; 0\;. \qquad \square$$

The following definition lays the ground for deciding under which circumstances flow is directly sent from one edge to another.

Definition 2.21 (Mutually Singular)**.** Let $A \subseteq \mathbb{R}_+$ and $B \subseteq \mathbb{R}_+$ be two sets of real numbers. Two Lebesgue integrable functions $f_1, f_2 : A \to B$ are called *mutually singular* if an only if $\min\{f_1, f_2\}$ is essentially 0.

Recalling the definition of the support, $\min\{f_1, f_2\}$ is essentially 0 means that the support of $\min\{f_1, f_2\}$ is a null set. In this sense, f_1 and f_2 are mutually singular if and only if $\mathrm{supp}(f_1) \cap \mathrm{supp}(f_2)$ is a null set or, equivalently, if and only if there exist an f_1-supporting and an f_2-supporting set which are disjoint.

The following definition is applied to transit and arrival time functions and, as a conclusion of the subsequent lemma, ensures consistent flow behavior.

Definition 2.22 (Compatible)**.** Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing, absolutely continuous function and $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ be a Lebesgue measurable function. Then ℓ is called *compatible with* F if and only if the preimage with respect to ℓ of every null set is an F-null set.

Lemma 2.23. *Let $F^+ : \mathbb{R}_+ \to \mathbb{R}_+$ be nondecreasing, absolutely continuous and $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ be Lebesgue measurable. Further, let $\lambda(\ell^{-1}(K)) < \infty$ for each compact set $K \subset \mathbb{R}_+$ which means that ℓ is somehow bounded. If ℓ is compatible with F^+ then the function $F^- : \mathbb{R}_+ \to \mathbb{R}_+$ defined by*

$$F^-(\theta) \;:=\; \int_{\ell^{-1}([0,\theta])} f_P^+(\vartheta)\, d\vartheta \qquad\qquad \forall \theta \in \mathbb{R}_+$$

is absolutely continuous.

Proof. Let λ be the Lebesgue measure and $I := [0, \theta]$ be a closed interval for some $\theta \in \mathbb{R}_+$. In the first step we show:

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall A \in \mathcal{B}(I):\quad \lambda(A) \leq \delta \quad\Rightarrow\quad \lambda\big(\ell^{-1}(A) \cap \mathrm{supp}(f^+)\big) \leq \epsilon\;. \tag{2.4}$$

Informally, this means that, with respect to the Lebesgue measure, ℓ^{-1} behaves somehow uniformly continuous on measurable sets. We prove this indirectly. So assume that there exists an $\epsilon \in (0, 1)$ such that

$$\forall \delta > 0\ \exists A \in \mathcal{B}(I):\quad \lambda(A) \leq \delta \quad\text{ and }\quad \lambda\big(\ell^{-1}(A) \cap \mathrm{supp}(f^+)\big) > \epsilon\;. \tag{2.5}$$

For each $i \in \mathbb{N}$, set $\delta_i := \epsilon^i$, and let $A_i \subseteq I$ be a corresponding set satisfying the expression above. Then $B_i := \bigcup_{j=i}^{\infty} A_i$ is measurable as it is the countable union of measurable sets. Further, we know $B_i \subseteq B_{i+1}$ for all $i \in \mathbb{N}$ and, hence, $B^* := \bigcap_{i=1}^{\infty} B_i$ is measurable. Besides, each $\ell^{-1}(A_i)$ is measurable as ℓ is a measurable function. This shows that $\ell^{-1}(B_i) = \bigcup_{j=i}^{\infty} \ell^{-1}(A_i)$ is also measurable. Moreover, we have $\ell^{-1}(B_i) \subseteq \ell^{-1}(B_{i+1})$ for all $i \in \mathbb{N}$, and hence, $\ell^{-1}(B^*) = \bigcap_{i=1}^{\infty} \ell^{-1}(B_i)$ is measurable. Since $A_i \subseteq I$ holds for all $i \in \mathbb{N}$, the set B_i is contained in I for all $i \in \mathbb{N}$. Hence, the measure of each B_i is bounded by $\lambda(I) = \theta$, and the measure of each $\ell^{-1}(B_i)$ is bounded by $\lambda(\ell^{-1}(I)) < \infty$ which is finite by the assumptions of this lemma. Because the Lebesgue measure is continuous from above (see Proposition 2.11(ii)), we obtain from (2.5)

$$\begin{aligned}
\lambda(B^*) \;&=\; \lim_{i\to\infty} \lambda(B_i) \;\leq\; \lim_{i\to\infty} \sum_{j=i}^{\infty} \lambda(A_j) \\
&\leq\; \lim_{i\to\infty} \sum_{j=i}^{\infty} \epsilon^j \;=\; \lim_{i\to\infty} \frac{\epsilon^i}{1-\epsilon} \;=\; 0
\end{aligned}$$

$$\begin{aligned}
\text{and} \quad \lambda\big(\ell^{-1}(B^*) \cap \operatorname{supp}(f^+)\big) \;&=\; \lim_{i\to\infty} \lambda\big(\ell^{-1}(B_i) \cap \operatorname{supp}(f^+)\big) \\
&\geq\; \lim_{i\to\infty} \lambda\big(\ell^{-1}(A_i) \cap \operatorname{supp}(f^+)\big) \;\geq\; \lim_{i\to\infty} \epsilon \;=\; \epsilon \;.
\end{aligned}$$

Thus, B^* is a null set, but $\ell^{-1}(B^*)$ is not an F^+-null set by Definition 2.19. This contradicts the compatibility of ℓ with F^+. Therefore, (2.4) holds.

In order to prove that F^- is absolutely continuous on I, fix another $\epsilon > 0$. Since F^+ is absolutely continuous on I, there exists an $\tilde{\delta} > 0$ such that for all finite families $((\tilde{a}_i, \tilde{b}_i))_{i\in[\tilde{k}]}$, where $\tilde{k} \in \mathbb{N}$, of pairwise disjoint intervals it holds:

$$\sum_{i=1}^{\tilde{k}} \tilde{b}_i - \tilde{a}_i \;\leq\; 2\tilde{\delta} \qquad \Rightarrow \qquad \sum_{i=1}^{\tilde{k}} F^+(\tilde{b}_i) - F^+(\tilde{a}_i) \;\leq\; \epsilon \;. \tag{2.6}$$

From (2.4) we know that there exists an δ such that for all measurable sets $A \in I$ with $\lambda(A) \leq \delta$ we have $\lambda(\ell^{-1}(A) \cap \operatorname{supp}(f^+)) \leq \tilde{\delta}$.

Now let $((a_i, b_i))_{i\in[k]}$, where $k \in \mathbb{N}$, be a finite family of pairwise disjoint intervals for some $k \in \mathbb{N}$ with

$$\lambda\Big(\bigcup_{i=1}^{k} (a_i, b_i)\Big) \;=\; \sum_{i=1}^{k} b_i - a_i \;\leq\; \delta \;.$$

Hence, defining $S := \operatorname{supp}(f^+) \cap \bigcup_{i=1}^{k} \ell^{-1}((a_i, b_i))$, we know $\lambda(S) \leq \tilde{\delta}$. Since the Lebesgue measure of S is approximable from above by open sets (see Proposition 2.11(iii)), there exists an open set $O \subseteq I$ with $S \subseteq O$ and $\lambda(O) \leq 2\tilde{\delta}$. Moreover, we know that each open set is the countable union of pairwise disjoint open intervals. So let $((\tilde{a}_i, \tilde{b}_i))_{i\in\mathbb{N}}$ be a family of pairwise disjoint open intervals with $O = \bigcup_{i\in\mathbb{N}} (\tilde{a}_i, \tilde{b}_i)$. Finally, since F^+ is nondecreasing implying that f^+ is nonnegative, we get from (2.6):

$$\begin{aligned}
\sum_{i=1}^{k} \big|F^-(b_i) - F^-(a_i)\big| \;&=\; \sum_{i=1}^{k} \Big| \int_{\ell^{-1}((a_i,b_i))} f^+(\vartheta)\, d\vartheta \Big| \\
&=\; \sum_{i=1}^{k} \int_{\ell^{-1}((a_i,b_i))} f^+(\vartheta)\, d\vartheta \;=\; \int_S f^+(\vartheta)\, d\vartheta \\
&\leq\; \int_0 f^+(\vartheta)\, d\vartheta \;=\; \lim_{\tilde{k}\to\infty} \sum_{i=1}^{\tilde{k}} F^+(\tilde{b}_i) - F^+(\tilde{a}_i) \\
&\leq\; \epsilon \;.
\end{aligned}$$

This shows that F^- is locally absolutely continuous and the proof is done. $\square$

The following definitions and results play an important role when analyzing the FiFo principle. The monotonicity set of some cumulative flow functions represents the points in time at which a positive amount of flow is distributed.

Definition 2.24 (Monotonicity Set)**.** Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing, continuous function with $F(0) = 0$. The *monotonicity set* S of F is defined as the set of points where F is strictly increasing from left, i.e.,

$$S \;:=\; \big\{\theta \in \mathbb{R}_+ \;\big|\; F(\theta_1) < F(\theta) \;\; \forall \theta_1 < \theta\big\} \;\cup\; \{\theta_0\}\;.$$

Here, $\theta_0 := \max\{\theta \in \mathbb{R}_+ \mid F(\theta) = 0\}$ is the maximal point at which F vanishes. Note that θ_0 exists by the assumptions on F.

The reason for adding θ_0 to the monotonicity set is that F becomes a one-to-one correspondence from S to $F(\mathbb{R}_+)$. Besides, S is an F^+-supporting set. Both properties are formally established in the next proposition.

Proposition 2.25. *Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing, absolutely continuous function with $F(0) = 0$ and S be the monotonicity set of F. Then F is a one-to-one correspondence from S to $F(\mathbb{R}_+)$, and S is F-supporting.*

Proof. In order to prove that F is a one-to-one corresponding from S to $F(\mathbb{R}_+)$, we have to show that for each $\theta \in \mathbb{R}^+$ the set S contains exactly one θ' with $F(\theta) = F(\theta')$, i.e., $|F^{-1}(F(\theta)) \cap S| = 1$ for all $\theta \in \mathbb{R}_+$. Firstly, we show that, for all $\theta \in \mathbb{R}_+$, the set $F^{-1}(F(\theta))$ contains at least one element from S. For $F(\theta) = 0$, this holds trivially as Definition 2.24 and the continuity of F imply $\theta_0 \in S$ and $F(\theta_0) = 0$. So consider a $\theta \in \mathbb{R}_+$ with $F(\theta) > 0$, and define $\theta_1 := \min\{\theta' \mid F(\theta') = F(\theta)\}$. Note that θ_1 exists by the assumptions on F. By the minimality of θ_1, we know $F(\theta') < F(\theta_1)$ for all $\theta' < \theta_1$ implying $\theta_1 \in S$. Since $F(\theta_1) = F(\theta)$ by definition, this shows $|F^{-1}(F(\theta)) \cap S| \geq 1$.

Next, assume that $F^{-1}(F(\theta))$ contains at least two different elements θ_1 and θ_2 from S for some $\theta \in \mathbb{R}_+$. Assuming without loss of generality $\theta_1 < \theta_2$, we obtain $F(\theta_1) < F(\theta_2)$ by the definition of S. Here, we use the fact that θ_0 is the only point in S on which F vanishes. Since $\theta_1, \theta_2 \in F^{-1}(F(\theta))$ implies $F(\theta_1) = F(\theta) = F(\theta_2)$, this is a contradiction. Hence, $|F^{-1}(F(\theta)) \cap S| = 1$ holds for all $\theta \in \mathbb{R}_+$ implying that F is a one-to-one corresponding from S to $F(\mathbb{R}_+)$.

It remains to show that S is an F-supporting set. So let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally Lebesgue integrable function satisfying $\frac{dF}{d\theta}(\theta) = f(\theta)$ for all $\theta \in \mathbb{R}_+ \setminus N$ for some null set N. Note that f exists by Proposition 2.16. By Definition 2.19, it is enough to show $\mathrm{supp}(f) \setminus S \subseteq N$. Consider a nonnegative real number θ which is not contained in S. Because of the definition of S, there exists an $\epsilon > 0$ such that $F(\theta') = F(\theta)$ for all $\theta' \in (\theta - \epsilon, \theta]$ as F is nondecreasing. Hence, the derivative $\frac{dF}{d\theta}(\theta)$ is either equal to 0 or does not exist. This shows that either $\theta \notin \mathrm{supp}(f)$ or $\theta \in N$ holds implying $\mathrm{supp}(f) \setminus S \subseteq N$. Thus, S is an F-supporting set. $\square$

The following definition lays the ground for the FiFo condition which we use in thesis.

Definition 2.26. Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing function. Then a function $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ is called *nondecreasing with respect to F* if and only if there exists an F-supporting set S such that

$$F(\theta_1) < F(\theta_2) \qquad \Rightarrow \qquad \ell(\theta_1) < \ell(\theta_2)$$

holds for $\theta_1, \theta_2 \in S$.

An alternative statement to the previous definition is established in the following Lemma.

Lemma 2.27. *A function $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing with respect to a nondecreasing and locally absolutely continuous function $F : \mathbb{R}_+ \to \mathbb{R}_+$ if and only if there exists an F-supporting set S such that*

$$\theta_1 < \theta_2 \qquad \Leftrightarrow \qquad \ell(\theta_1) < \ell(\theta_2) \tag{2.7}$$

holds for all for all $\theta_1, \theta_2 \in S$. This means, ℓ is strictly nondecreasing on S.

Proof. Firstly, assume that the function $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing with respect to a nondecreasing, absolutely continuous function $F : \mathbb{R}_+ \to \mathbb{R}_+$, and let S_1 be an F-supporting set for which Definition 2.26 holds. Further, let S_2 be the monotonicity set of F. By Proposition 2.25 we know that S is F-supporting implying that $S := S_1 \cap S_2$ is F-supporting. We show that (2.7) is valid over S. Since F is a one-to-one correspondence on S_2, the monotonicity of F is strict on S, i.e.,

$$\theta_1 < \theta_2 \qquad \Leftrightarrow \qquad F(\theta_1) < F(\theta_2)$$

holds for all $\theta_1, \theta_2 \in S$. Hence, Definition 2.26 directly implies $\ell(\theta_1) < \ell(\theta_2)$ for all $\theta_1, \theta_2 \in S$ with $\theta_1 < \theta_2$. On the other hand, $\ell(\theta_1) < \ell(\theta_2)$ for some $\theta_1, \theta_2 \in S$ implies in particular $\ell(\theta_1) \leq \ell(\theta_2)$ and $\theta_1 \neq \theta_2$. By Definition 2.26 this implies $F(\theta_1) \leq F(\theta_2)$ and $\theta_1 \neq \theta_2$. As the monotonicity of F is strict, this shows $F(\theta_1) < F(\theta_2)$ implying $\theta_1 < \theta_2$.

For proving the other direction, let S be an F-supporting set over which (2.7) is satisfied and $\theta_1, \theta_2 \in S$ be such that $F(\theta_1) < F(\theta_2)$ holds. As F is nondecreasing, we know $\theta_1 < \theta_2$ and, hence, $\ell(\theta_1) < \ell(\theta_2)$ by (2.7). This shows that ℓ is nondecreasing with respect to F. □

The following lemma shows in which case a nondecreasing function is nondecreasing with respect to some function F.

Lemma 2.28. *Let $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing function which is compatible with some nondecreasing, absolutely continuous function $F : \mathbb{R} \to \mathbb{R}_+$. Then ℓ is nondecreasing with respect to F.*

Proof. Consider $\theta_1, \theta_2 \in \mathbb{R}_+$ with $F(\theta_1) < F(\theta_2)$, and assume $\ell(\theta_1) \geq \ell(\theta_2)$. Since F is nondecreasing, we know $\theta_1 < \theta_2$ implying $\ell(\theta_1) = \ell(\theta_2)$ as ℓ is nondecreasing as well. This shows $(\theta_1, \theta_2) \subseteq \ell^{-1}(\theta_1)$. For this reason, (θ_1, θ_2) must be an F-null set because ℓ is compatible with F. This is a contradiction because we know

$$\int_{(\theta_1,\theta_2)} f(\vartheta)\, d\vartheta = F(\theta_2) - F(\theta_1) > 0$$

where f is a derivative of F. Therefore, $\ell(\theta_1) < \ell(\theta_2)$ is valid, and ℓ is nondecreasing with respect to F. □

This completes the discussion about inherent properties of functions. In the following, we consider function spaces in order to analyze function sequences. In particular, this lays the ground for discussing continuity aspects of flows over time in this thesis.

Definition 2.29 (Space of (Locally) Bounded Functions)**.** Let $A \subseteq \mathbb{R}$ be some measurable set. The symbol $L_\infty(A)$ denotes the *space of bounded functions* $F : A \to \mathbb{R}$ satisfying

$$\left\|F\right\|_\infty^A \;:=\; \sup_{\theta\in A} \left|F(\theta)\right| < \infty \;.$$

The space $L_\infty^{\mathrm{loc}}(A)$ consists of all locally bounded functions $F : A \to \mathbb{R}_+$, i.e.,

$$\left\|F\right\|_\infty^K \;=\; \sup_{\theta\in K} \left|F(\theta)\right| < \infty$$

holds for all compact sets $K \subseteq A$. For the special case of $A = \mathbb{R}_+$, we set $L_\infty^{\mathrm{loc}} := L_\infty^{\mathrm{loc}}(\mathbb{R}_+)$.

It is well-known that $\|.\|_\infty^A$ is a norm on the vector space $L_\infty(A)$, i.e., $\|.\|_\infty^A$ satisfies

$$\begin{aligned}
&F \neq 0 \quad\Rightarrow\quad \left\|F\right\|_\infty^A > 0 \;, \\
&\left\|a \cdot F\right\|_\infty^A = |a| \cdot \left\|F\right\|_\infty^A \;, \qquad \text{and} \qquad \left\|F+G\right\|_\infty^A \leq \left\|F\right\|_\infty^A + \left\|G\right\|_\infty^A
\end{aligned}$$

for all $a \in \mathbb{R}$ and $F, G \in L_\infty(A)$. Therefore, $\|.\|_\infty^A$ implies a metric on $L_\infty(A)$ given by $(F, G) \mapsto \|F - G\|_\infty^A$ which has to satisfy

$$\begin{aligned}
&\left\|F-G\right\|_\infty^A \geq 0 \;, && \left\|F-G\right\|_\infty^A = 0 \quad\Leftrightarrow\quad F = G \;, \\
&\left\|F-G\right\|_\infty^A = \left\|G-F\right\|_\infty^A \;, \quad \text{and} && \left\|F-H\right\|_\infty^A \leq \left\|F-G\right\|_\infty^A + \left\|G-H\right\|_\infty^A
\end{aligned}$$

for all $F, G, H \in L_\infty(A)$. This metric enables us to consider convergent sequences. The following definition introduces convergence in L_∞. This kind of convergence is commonly known as *uniform convergence*.

Definition 2.30 (Convergence in L_∞)**.** Let $A \subseteq \mathbb{R}$ be a measurable sets. A sequence $(F_i)_{i\in\mathbb{N}}$ of functions $F_i \in L_\infty(A)$ *converges to some* $F \in L_\infty(A)$ if and only if for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$\left\|F - F_i\right\|_\infty^A \leq \epsilon \qquad\qquad \forall i \geq N \;.$$

A sequence $(F_i)_{i\in\mathbb{N}}$ of functions $F_i \in L_\infty^{\mathrm{loc}}$ *converges to some* $F \in L_1^{\mathrm{loc}}$ if and only if for all $\epsilon > 0$ and all compact sets $K \subset \mathbb{R}_+$ there exists an $N \in \mathbb{N}$ such that

$$\left\|F - F_i\right\|_\infty^K \leq \epsilon \qquad\qquad \forall i \geq N \;.$$

That is, $(F_i)_{i\in\mathbb{N}}$ converges in $L_\infty(K)$ to F for all compact sets $K \subset \mathbb{R}_+$.

A *Cauchy sequence* on $L_\infty(A)$ is a sequence $(F_i)_{i\in\mathbb{N}}$ of functions $F_i \in L_\infty(A)$ such that for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$ with

$$\left\|F_j - F_i\right\|_\infty^A \leq \epsilon \qquad\qquad \forall i, j \geq N \;.$$

It is known that the space $L_\infty(A)$ is a Banach space, i.e., a normed vector space in which every Cauchy sequence $(F_i)_{i\in\mathbb{N}}$ converges to some $F \in L_\infty(A)$ in $L_\infty(A)$.

As known, continuity of maps can be defined via convergent sequences. That is, a map is continuous if each convergent sequence of the domain implies a convergent sequence in the codomain such that the corresponding limit points are mapped on each other. The subsequent explanation of continuity in L_∞ is used to formalize flow over time models where a small variation in some flow scenario only leads to a slightly different flow behavior.

Let $A, B \subseteq \mathbb{R}$ be two measurable sets. A map $L_\infty(A) \to L_\infty(B)$ denoted by $F^+ \mapsto F^-$ is called *continuous at* $F^+ \in L_\infty(A)$ if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\left\|F^+ - \tilde{F}^+\right\|_\infty^A \leq \delta \qquad \Rightarrow \qquad \left\|F^- - \tilde{F}^-\right\|_\infty^B \leq \epsilon \qquad \forall \tilde{F}^+ \in L_\infty(A) \ .$$

A map $L_\infty^{\text{loc}} \to L_\infty(B)$ denoted by $F^+ \mapsto F^-$ is *continuous at* $F^+ \in L_\infty^{\text{loc}}$ if and only if for all $\epsilon > 0$ there exists a $\delta > 0$ and a compact set $K \subset \mathbb{R}_+$ such that

$$\left\|F^+ - \tilde{F}^+\right\|_\infty^K \leq \delta \qquad \Rightarrow \qquad \left\|F^- - \tilde{F}^-\right\|_\infty^B \leq \epsilon \qquad \forall \tilde{F}^+ \in L_\infty(A) \ .$$

A map $L_\infty^{\text{loc}} \to L_\infty^{\text{loc}}$ denoted by $F^+ \mapsto F^-$ is *continuous at* $F^+ \in L_\infty^{\text{loc}}$ if and only if for all $\epsilon > 0$ and all compact sets K_ϵ there exists a $\delta > 0$ and a compact set $K_\delta \subset \mathbb{R}_+$ such that

$$\left\|F^+ - \tilde{F}^+\right\|_\infty^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \left\|F^- - \tilde{F}^-\right\|_\infty^{K_\epsilon} \leq \epsilon \qquad \forall \tilde{F}^+ \in L_\infty(A) \ .$$

Such maps are called *continuous* if and only if they are continuous over its entire domain.

Throughout this thesis we also work with at most countable families of (locally) bounded functions, and we also need a notion of continuity for maps between such families. Hence, we need a norm for such objects. Let $n \in \mathbb{N}^\infty$ and $(F_i)_{i\in[n]_0}$ be a family of (locally) bounded functions. With $(L_\infty(A))_1^{[n]_0}$ we denote the set of all families $\mathcal{F} := (F_i)_{i\in[n]_0}$ with $F_i \in L_\infty(A)$ for each $i \in [n]_0$ satisfying

$$\left\|\mathcal{F}\right\|_{\infty,1}^A \ := \ \sum_{i\in[n]_0} \left\|F_i\right\|_\infty^A \ < \ \infty \ .$$

With $(L_\infty^{\text{loc}})_1^{[n]_0}$ we denote the set of all families $\mathcal{F} := (F_i)_{i\in[n]_0}$ with $F_i \in L_\infty^{\text{loc}}$ for each $i \in [n]_0$ satisfying

$$\left\|\mathcal{F}\right\|_{\infty,1}^K \ := \ \sum_{i\in[n]_0} \left\|F_i\right\|_\infty^K \ < \ \infty$$

for each compact set $K \subset \mathbb{R}_+$. In this manner, we use the well-studied 1-norm of $\mathbb{R}^{[n]_0}$ in order to construct a norm for $(L_\infty(A))_1^{[n]_0}$ and $(L_\infty^{\text{loc}})_1^{[n]_0}$, respectively. As all norms on $\mathbb{R}^{[n]_0}$ are equivalent if n is finite, we could have taken any norm in this case – we always get the same characterization of convergent sequences and continuous maps. However, if $n = \infty$ holds, this is no longer true, and the 1-norm of $\mathbb{R}^{[n]_0}$ is essential for obtaining well-defined continuous flow over time models.

Beside this, for determining unique flow behavior, we also work with families of maps whose cardinality can be even uncountable. For such families we explain the notion of equicontinuity in the following. We only consider the most general case. So let $J \subseteq \mathbb{R}$ be some set of real numbers and $n, m \in \mathbb{N}^\infty$. A family $(\mathcal{F}^+ \mapsto \mathcal{F}_j^-)_{j\in J}$ of maps $(L_\infty^{\mathrm{loc}})_1^{[n]_0} \to (L_\infty^{\mathrm{loc}})_1^{[m]_0}$ is called *equicontinuous* if and only if for all $\epsilon > 0$ and all compact sets $K_\epsilon \subseteq \mathbb{R}_+$ there exists a $\delta > 0$ and a compact set $K_\delta \subseteq \mathbb{R}_+$ such that for all $j \in J$ we have

$$\left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \left\|\mathcal{F}_j^- - \tilde{\mathcal{F}}_j^-\right\|_{\infty,1}^{K_\epsilon} \leq \epsilon \qquad \forall \tilde{\mathcal{F}}^+ \in (L_\infty^{\mathrm{loc}})_1^{[n]_0} .$$

Note that behind the norms on the left hand side and on the right hand side a sum consisting of n and m terms is hidden, respectively.

In this thesis, flows over time are primary modeled via nondecreasing, locally absolutely continuous functions $F : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $F(0) = 0$. Therefore, we apply these notions of continuity to such functions and not to its locally Lebesgue integrable counterpart. Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ be such a function. Clearly, F is locally bounded and, hence, is contained in the space L_∞^{loc}. To see this, note that a compact set $K \subset \mathbb{R}_+$ is bounded, i.e., there exists a $\theta \in \mathbb{R}_+$ with $K \subseteq [0, \theta]$. Since F is nonnegative and nondecreasing, this shows $\sup_{\theta\in K} |F(\theta')| \leq F(\theta') < \infty$ implying $F \in L_\infty^{\mathrm{loc}}$.

In this sense, the set consisting of nondecreasing, absolutely continuous functions $F : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $F(0) = 0$ can be seen as a subspace of L_∞^{loc}. Note that this subspace is *not* a vector space as the difference of two nonnegative functions is not nonnegative, in general. However, this causes no really problems. In addition, this subspace is also not complete. Nevertheless, it is quite important for the purpose of this thesis that the limit of such functions lies in this subspace if it exists. For that reason, we subsequently identify assumptions under which such a behavior is guaranteed. The Weierstrass M-Test provides a sufficient criterion under which a function series converges uniformly.

Theorem 2.31 (Weierstrass M-Test)**.** *Let A be a set and $(F_k)_{k\in\mathbb{N}}$ be a sequence of functions $F_k : A \to \mathbb{R}$ such that for each $k \in \mathbb{N}$ there exists a nonnegative real number $M_k \in \mathbb{R}_+$ with $\|F_k\|_\infty^A \leq M_k$. If $\sum_{k\in\mathbb{N}} M_k$ converges then $\sum_{k\in\mathbb{N}} F_k$ converges uniformly on A.*

The following well-known theorem shows that the limit of a uniform convergent sequence of continuous functions is also continuous.

Theorem 2.32 (Uniform Limit Theorem)**.** *Let A be a set and $(F_k)_{k\in\mathbb{N}}$ be a sequence of continuous functions $F_k : A \to \mathbb{R}$ converging uniformly to some function $F : A \to \mathbb{R}$. Then F is continuous.*

Similarly to the Uniform Limit Theorem, we are able to prove a convergence result for a series consisting of nondecreasing, locally absolutely continuous functions $\mathbb{R}_+ \to \mathbb{R}_+$ vanishing at 0.

Lemma 2.33. *Let $(F_k)_{k\in\mathbb{N}}$ be a sequence of nondecreasing, locally absolutely continuous functions $F_k : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $F_k(0) = 0$. Assume that the sequence $(\sum_{k=1}^K F_k)_{K\in\mathbb{N}}$ of partial sums converges uniformly to some function $F : \mathbb{R}_+ \to \mathbb{R}$. Then $F = \sum_{k\in\mathbb{N}} F_k$ is nonnegative, nondecreasing, locally absolutely continuous, and satisfies $F(0) = 0$.*

Proof. Firstly, we observe that $F(0) = 0$ holds. Clearly, as each F_k vanishes at 0, each partial sum $\sum_{k=1}^{K} F_k$ does so. Since F is the limit of the partial sums, this shows $F(0) = 0$.

In order to see that F is nondecreasing, note that each partial sum $\sum_{k=1}^{K} F_k$ is nondecreasing as a finite sum of nondecreasing functions. Thus, we get

$$F(\theta_1) \;=\; \lim_{k\to\infty} \sum_{k=1}^{K} F_k(\theta_1) \;\leq\; \lim_{k\to\infty} \sum_{k=1}^{K} F_k(\theta_2) \;=\; F(\theta_2) \;.$$

for all $\theta_1, \theta_2 \in \mathbb{R}_+$ with $\theta_1 \leq \theta_2$ implying that F is nondecreasing.

It remains to prove that F is locally absolutely continuous. For this consider an arbitrary real interval $[\theta_1, \theta_2]$ with $\theta_1, \theta_2 \in \mathbb{R}_+$ and $\theta_1 < \theta_2$, and let $\epsilon > 0$ be some positive real number. Since the sequence $(\sum_{k=1}^{K} F_k)_{K\in\mathbb{N}}$ of partial sums converges to $F : \mathbb{R}_+ \to \mathbb{R}$, there exists a $\tilde{K}$ such that

$$\big|F(\theta_2) - \tilde{F}(\theta_2)\big| \;=\; F(\theta_2) - \tilde{F}(\theta_2) \;\leq\; \epsilon \;, \tag{2.8}$$

where $\tilde{F}$ is set to $\tilde{K}$-th partial sum, i.e., $\tilde{F} := \sum_{k=1}^{\tilde{K}} F_k$. Note that the "="-sign follows from the nonnegativity of the F_k's as this implies that the sequence of partial sums is nondecreasing. Further, we know that $\tilde{F}$ as the finite sum of locally absolutely continuous functions is locally absolutely continuous as well. Hence, there exists a $\delta > 0$ such that every finite sequence $((a_j, b_j))_{j=1,\ldots,i}$ of pairwise disjoint sub-intervals of $[\theta_1, \theta_2]$ with $\sum_{j=1}^{i} |b_j - a_j| < \delta$ satisfies

$$\sum_{j=1}^{i} \big|\tilde{F}(b_j) - \tilde{F}(a_j)\big| \;=\; \sum_{j=1}^{i} \tilde{F}(b_j) - \tilde{F}(a_j) \;<\; \epsilon \;. \tag{2.9}$$

Since each F_k is nondecreasing, we know that $F - \tilde{F}$ is also nondecreasing. Because of (2.8) this implies

$$\sum_{j=1}^{i} \big(F(b_j) - \tilde{F}(b_j)\big) - \big(F(a_j) - \tilde{F}(a_j)\big) \;\leq\; F(\theta_2) - \tilde{F}(\theta_2) \;\leq\; \epsilon \tag{2.10}$$

as the intervals (a_j, b_j) are pairwise disjoint and $F(\theta_1) - \tilde{F}(\theta_1) \geq 0$. Adding (2.9) and (2.10) we obtain

$$\sum_{j=1}^{i} \big|F(b_j) - F(a_j)\big| \;\leq\; 2\epsilon$$

which, finally, shows that F is locally absolutely continuous. $\square$

Uniform convergence means that a sequence of functions converges to its limit evenly over their common domain. As already mentioned, this is the characteristic of convergence in L_∞. In addition, there exists a weaker kind of convergence. This so-called pointwise convergence requires only that a sequence of functions converges to its limit for each point of their domain separately. More precisely, a sequence $(F_k)_{k\in\mathbb{N}}$ of functions $[a, b] \to \mathbb{R}_+$ converges pointwise to some function $F : [a, b] \to \mathbb{R}_+$ if and only if:

$$\forall \epsilon > 0 \quad \forall \theta \in [a, b] \quad \exists K \in \mathbb{N} \quad \forall k \geq K : \qquad \big|F(\theta) - F_k(\theta)\big| \;\leq\; \epsilon \;.$$

Note that uniform convergence is obtained if we swap the second $\forall$-quantifier and the $\exists$-quantifier. The pointwise convergence becomes important in the following. But before note that Lemma 2.33 remains true if we replace the uniform convergence with pointwise convergence.

Concluding this section, we establish a lemma which is essential for proving the existence of Nash equilibria in quite general routing games over times. For this we need the following two well-known theorems. Note that we only state specialized versions of these theorems such that they can be obviously used in the subsequent lemma. In fact, the real statement of each of these theorems is much stronger and much more general.

Theorem 2.34 (Helly's Selection Theorem, [36]). *Let $(F_k)_{k\in\mathbb{N}}$ be a sequence of nondecreasing functions $\mathbb{R}_+ \to [a,b]$ for some $a,b \in \mathbb{R}$. Then there exists a subsequence $(F_k)_{k\in J}$ converging pointwise to some function F^*.*

Theorem 2.35 (Arzelá-Ascoli Theorem, [5, 6]). *Let $(F_k)_{k\in\mathbb{N}}$ be a sequence of continuous functions $[0,T] \to \mathbb{R}_+$ for some given $T \in \mathbb{R}_+$. Then $(F_k)_{k\in\mathbb{N}}$ converges uniformly if and only if it is equicontinuous and converges pointwise.*

The following lemma establishes a sufficient condition under which a sequence of nondecreasing functions contains a subsequent converging uniformly to some absolutely continuous, nondecreasing function which vanishes at 0.

Lemma 2.36. *For some $T \in \mathbb{R}_+$ let $D : [0,T] \to \mathbb{R}_+$ be an absolutely continuous, nondecreasing function satisfying $D(0) = 0$. Further, let $(F_k)_{k\in\mathbb{N}}$ be a sequence of nondecreasing functions $[0,T] \to \mathbb{R}_+$ such that, for all $k \in \mathbb{N}$,*

(i) the equation $F_k(0) = 0$ is valid and

(ii) the function $D - F_k$ is nondecreasing.

Then there exists a set $J \subseteq \mathbb{N}$ such that the subsequence $(F_k)_{k\in J}$ converges uniformly to some absolutely continuous, nondecreasing function $F^* : [0,T] \to \mathbb{R}_+$ satisfying $F^*(0) = 0$.

Proof. Firstly, (ii) implies that each F_k is absolutely continuous because of Proposition 2.17. Further, because of (i) and (ii) we obtain $F_k(T) \leq D(T)$ for all $k \in \mathbb{N}$. As each function F_k is nondecreasing, this shows that each function F_k takes only values in $[0, D(T)]$. Thus, Helly's Selection Theorem 2.34 ensures the existence of a countable subset $J \subseteq \mathbb{N}$ such that the subsequence $(F_k)_{k\in J}$ converges pointwise to some F^* which, clearly, satisfies $F^*(0) = 0$ because of (i).

Let $J = \{i_k | k \in \mathbb{N}\}$ be such that $i_k < i_{k+1}$ for all $k \in \mathbb{N}$. As each F_k is nondecreasing, we obtain for all $\theta_1, \theta_2 \in \mathbb{R}_+$ with $\theta_1 \leq \theta_2$:

$$F^*(\theta_1) \;=\; \lim_{k\to\infty} F_{i_k}(\theta_1) \;\leq\; \lim_{k\to\infty} F_{i_k}(\theta_2) \;=\; F^*(\theta_2) \;.$$

This shows that F^* is nondecreasing. Further, because of (ii) we know for all $\theta_1, \theta_2 \in \mathbb{R}_+$ with $\theta_1 \leq \theta_2$:

$$\begin{aligned} D(\theta_1) - F^*(\theta_1) \;&=\; \lim_{k\to\infty} D(\theta_1) - F_{i_k}(\theta_1) \\ &\leq\; \lim_{k\to\infty} D(\theta_2) - F_{i_k}(\theta_2) \;=\; D(\theta_2) - F^*(\theta_2) \;. \end{aligned}$$

This shows that $D - F^*$ is nondecreasing. Thus, F^* is absolutely continuous by Proposition 2.17.

Finally, we show that the subsequence $(F_k)_{k \in J}$ converges uniformly to F^*. Because of the Arzelá-Ascoli Theorem 2.35 it is enough to show that $(F_k)_{k \in \mathbb{N}}$ is equicontinuous, i.e., for all $\theta \in [0, T]$ and all $\epsilon > 0$ there exists an δ such that

$$|\theta - \tilde{\theta}| \leq \delta \qquad \Rightarrow \qquad |F_k(\theta) - F_k(\tilde{\theta})| \leq \epsilon \qquad \forall \tilde{\theta} \in [0, T],\ k \in \mathbb{N}\ .$$

So fix some $\theta \in [0, T]$ and some $\epsilon > 0$. Since D is continuous, we know that there exists a $\delta > 0$ such that $|D(\theta) - D(\tilde{\theta})| \leq \epsilon$ holds for all $\tilde{\theta} \in [0, T]$ with $|\theta - \tilde{\theta}| \leq \delta$. Because of (ii) this implies $|F_k(\theta) - F_k(\tilde{\theta})| \leq \epsilon$ for all $\tilde{\theta} \in [0, T]$ with $|\theta - \tilde{\theta}| \leq \delta$ for all $k \in \mathbb{N}$. This shows that $(F_k)_{k \in \mathbb{N}}$ and, in particular, $(F_k)_{k \in J}$ are equicontinuous. Hence, the subsequence $(F_k)_{k \in J}$ converges uniformly to F^*. □

2.3 Networks

The basic mathematical object with which we try to build a copy of the real world are networks. Networks are widely used for studying questions which arise, e.g., in logistic, telecommunication, traffic simulation, and sociology. For our usage a network contains only information which are relevant for traffic on real world routing systems like telecommunication and transportation networks. That is, a network is abstract in the sense that it does not reflect where a particular link or street is physically placed on earth. Rather, it shows how links or streets are connected, and how traffic behaves on them. Further, we limit ourselves to routing systems where traffic originates at one particular source and wants to arrive at one particular sink. That is, we only consider single-commodity networks.

In this section we discuss all components of a network which are relevant for the scope of this thesis. We start with notions related to directed graphs and, in particular, explain what we mean by a path. Subsequently, we explain transit times, capacities, and supplies. Finally, we shortly introduce static flows. For a deeper treatment of static flow theory we refer to Section 6.1.

Throughout this thesis, the symbol $\mathcal{N}$ is reserved for denoting some network. Further, it should be clear from the context which of the following components are part of the particular network. In addition, it should also be clear how a certain component is given and what the intuition behind this specification is.

Graphs. The most important part, upon which almost every routing network is built, are graphs. They are used for modeling the basic infrastructure of a routing network. A graph $G := (V, E)$ consists of a finite set V of nodes like junctions or terminals which are related via a finite set E of edges representing streets or links. An edge $e \in E$ connects some node $v \in V$ to some node $w \in V$. In this case we call $v =: \text{tail}(e)$ the tail of e and $w =: \text{head}(e)$ the head of e and write $e = vw$. Also, we say that e leaves v and enters w. Further, we allow loops which are edges vv from some node v to itself and multiple parallel edges. That is, G can contain several edges from v to w. In order to distinguish different edges $e = vw$, we formally think of E as the set of symbols e and interpret the equality sign as a reference to the node pair (v, w). In terms of graph theory, since e is directed from v to w, the edge e is a directed edge, and hence, G is a directed graph. Since we always work with directed graphs, we mainly omit the word "directed" from now on.

We denote by $s \in V$ the source node where all traffic starts traversing the network. Further, we use $t \in V$ for denoting the common destination of all traffic.

Let $V' \subseteq V$ be a subset of the nodes. The set of all edges leaving V' are denoted by $\delta^+(V) := \{e = vw \in E \mid v \in V', w \notin V'\}$. Similarly, the set of all edges entering V' are determined by $\delta^-(V) := \{e = vw \in E \mid v \notin V', w \in V'\}$. We call V' a (static) v-w-cut if $v \in V'$ and $w \notin V'$. Further, we say that an edge e crosses the cut forwards and backwards if $e \in \delta^+(V')$ and $e \in \delta^-(V')$, respectively. If $V' := \{v\}$ consists only of a single node, we write $\delta^+(v)$ and $\delta^-(v)$ instead of $\delta^+(\{v\})$ and $\delta^-(\{v\})$. Moreover, if the underlying graph G is not clear from the context, we write $\delta^+_G(V')$ and $\delta^-_G(V')$ for identifying G. Also note that we use the symbol δ for denoting a small positive real number. But since δ comes without a superscript in this case, there should be no confusion.

A directed path P on a directed graph G is given by a finite edge sequence $(e_1, \dots, e_k)$, $k \in \mathbb{N}$ where the head of some edge equals the tail of the next edge, i.e., $\text{head}(e_i) = \text{tail}(e_{i+1})$ for all $i \in [k]$. Further, we use $|P| := k$ to denote the number of edges which traffic has to traverse when traveling along P. The symbol $e_i^P := e_i$ stands for the i-th edge of P, and similarly, $v_i^P := \text{head}(e_i)$ stands for the i-th node of P, where the start node v_0^P of P is not counted. Motivated by intuition, if $e = e_i^P$ for some edge $e \in E$, we say that P traverses e, P visits e, or e is contained in P (at position i). Similar terminology is used in case $v = v_i^P$ for some node $v \in V$. Note that P can visit edges and nodes more than once. If $v = \text{tail}(e_1)$ and $w = \text{head}(e_k)$, we call P a v-w-path. If $\text{tail}(e_1) = \text{head}(e_k)$, we call P a cycle. Further, a path P and a cycle C are called simple if no intermediate node is revisited, i.e., $\{v_0^P, \dots, v_k^P\} = k + 1$ and $\{v_0^C, \dots, v_{k-1}^C\} = k$, respectively. Finally, if a graph contains no loop and no parallel edges, we also use the shorter notation $v_0^P v_1^P \dots v_k^P$ for denoting the path P.

The set of all s-t-paths is denoted by $\mathcal{P}$. Note that, in general, $\mathcal{P}$ is infinite but countable because nodes just as edges can be revisited. Nevertheless, during this thesis, we impose very weak restrictions whether or not a certain s-t-path is contained in $\mathcal{P}$. These restriction allow us to see $\mathcal{P}$ as a "locally finite" set. What this means becomes partially clear below.

Transit Times. A crucial feature of flows over time is the incorporation of the time dimension. In particular, we assume that traffic does not travel instantaneously through the network. Rather, we assume that it takes time to traverse the network. This is modeled via transit times which are denoted by τ. In the following, we briefly discuss two different aspects of transit times.

A simple way for including transit times is to assign a constant value $\tau_e \in \mathbb{R}_+$ to each edge. This means that if some flow enters a certain edge e at a time θ, it reaches the head of e at time $\theta + \tau_e$. Hence, if a flow particle chooses an s-t-paths P, it enters the i-th node v_i^P after $\tau_{P,i} := \sum_{j=1}^{i} \tau_{e_j^P}$ time units. Thus, it takes $\tau_P := \sum_{i=1}^{|P|} \tau_e$ time units for traversing P. In case of constant transit times, τ_P is called path transit time. When dealing with constant transit times, we assume (if not mentioned otherwise) that a path P does not contain a cycle of zero transit time. Since transit times are nonnegative by definition, $\mathcal{P}$ can be seen as a "locally finite" set in the following sense: The number of s-t-paths P with $\tau_P \leq \tau_{\max}$ is finite for all $\tau_{\max} \in \mathbb{R}_+$.

However, in real world application transit times are not constant but vary over time. In fact, they usually depend on the flow situation. In this case τ_e is given via a transit time function $\tau_e : \mathbb{R}_+ \to \mathbb{R}_+$. Intuitively, if a flow particle enters an edge e at time θ, it arrives at its head at time $\theta + \tau_e(\theta)$. Hence, if a flow particle originates at s at time θ, the points in time $\ell_{P,i}(\theta)$ at which this particle enters the i-th node of some path P are recursively computable by

$$\ell_{P,0}(\theta) := \theta \qquad \text{and} \qquad \ell_{P,i}(\theta) := \ell_{P,i-1}(\theta) + \tau_{e_i^P}\left(\ell_{P,i-1}(\theta)\right) .$$

So the time needed until such a flow particle arrives at the i-th node of P equals $\tau_{P,i}(\theta) := \ell_{P,i}(\theta) - \theta$. Furthermore, the transit time $\tau_P(\theta)$ of P at time θ amounts to $\ell_{P,|P|}(\theta) - \theta$. Note that $\tau_P(\theta) := \sum_{i=1}^{|P|} \tau_{e_i^P}(\ell_{P,i-1}(\theta))$ holds in this case.

So far, we have considered an edge-based model of transit times, i.e., path transit times arise out of given transit times on the edges. In this thesis, we also consider models where transit times are only given path-based because no equivalent edge-based formulation exists. Hence, transit times occurring in a network can be of four different kinds, i.e., constant or time varying and path-based or edge-based.

Capacities. The time varying nature of transit times can be caused by capacities. Capacities are used to bound the flow which is allowed to traverse a certain edge of a graph. So they represent properties like the number of lanes of a street or the bandwidth of a link.

The capacity of an edge e is denoted by u_e. Usually, we assume that capacities are time varying but independent on the flow. Nevertheless, at several points of this thesis we only consider constant transit times. So u_e is typically a Lebesgue integrable function $\mathbb{R}_+ \to \mathbb{R}_+$ but can also stand for a real number. The value $u_e(\theta)$ bounds the rate at which flow is able to leave e at time θ. In case of constant transit times, we use u_P for defining the minimum capacity on a particular path P, i.e.,

$$u_P := \min\{u_{e_i^P} + \tau_{P,i} \mid 1 \leq i \leq |P|\} .$$

Provided that flow must not wait while traversing P, the value $u_P(\theta)$ bounds the rate at which such flow can enter P at time θ.

It is worth to mention that capacities serve primarily *not* as a feasibility criterion for flows over time but are rather used for defining flow-dependent transit times out of constant transit times. For example, considering the direct flow model (see Section 5.1), flow is only allowed to enter a certain path P if there is free capacity on P. So if at least one edge of P is used up to its capacity, flow must wait in front of P until it can traverse P directly without waiting. Or, regarding the deterministic queuing model (see Section 7.2), if more flow wants to traverse an edge than this edge can handle, a waiting queue is built at the end of this edge. In both cases, the flow-dependent transit time results out of the waiting time plus the constant (path or edge) transit time. Another aspect arising from these two examples is that constant transit times can be interpreted as free flow transit times, i.e., as the transit time of an edge or a path in case it is free of flow.

Supply. In a Nash equilibrium each flow particle tries to travel along a shortest path. For deciding whether or not a given flow over time is a Nash equilibrium, we must be able to follow a flow particle through the network. In particular, we have to know at which point in time a certain flow particle originates. This is modeled via a Lebesgue integrable function $d : \mathbb{R}_+ \to \mathbb{R}_+$ which we call supply rate. The value $d(\theta)$ equals the rate at which flow originates at s at time θ. Similarly, $D(\theta) := \int_0^\theta d(\vartheta)\, d\vartheta$ equals the amount of flow which has arrived at the source s until time θ. For the sake of completeness, we note that d is also used for denoting a constant supply rate which is modeled via a positive real number.

So far, we concentrated on networks for flows over time. The final paragraph deals with static networks containing the basic information for computing static flows. The term *static* refers to the fact that the time dimension is not considered.

Static Networks. Usually, a static network (G, u, s, t) consists of a directed graph $G := (V, E)$, real edge capacities $u_e \in \mathbb{R}_+$, a source s, and a sink t. As for flows over time, a real supply $d \in \mathbb{R}_+$ may also be given. In addition, we use a generalized concept which allows an arbitrary number of sources and sinks as follows. More precisely, we have a value $d_v \in \mathbb{R}$ for each node $v \in V$. If $d_v > 0$ then v is a source and d_v the corresponding supply. On the other hand, if $d_v < 0$ then v is a sink and $-d_v$ is called the demand of v. Since a flow should be able to serve the entire demand with the available supply, we require

$$\sum_{v \in V} d_v = 0$$

in such scenarios. It is worth to mention that every source can potentially serve every sink. In this sense, we still have one commodity which has to be distributed over the network. Summarizing, a static network may also be given as a tuple (G, u, d).

A static flow x on such a network (G, u, d) is given by a family $x := (x_e)_{e \in E}$ of real flow values $x_e \in \mathbb{R}_+$ on the edges. A flow x is called b-flow if and only if it satisfies the following flow conservation constraints

$$\sum_{e \in \delta^+(v)} x_e - \sum_{e \in \delta^+(v)} x_e \;=\; d_v \qquad\qquad \forall v \in V \;.$$

The terminology b-flow refers to the fact that d is also called node balance in literature and, for that reason, is replaced with a b. If the node balance of a node is 0, the corresponding flow conservation constraint requires that the total amount of flow entering a node also leaves this node. The capacity u_e of an edge e mainly bounds the flow on e, i.e., a feasible flow x has to satisfy $x_e \leq u_e$ for all $e \in E$. But we note already here that, in this thesis, this is not the primary interpretation of capacities.

If d_v vanishes for all nodes $v \in V$ except for some $s, t \in V$, a flow x is called s-t-flow. Here, we assume $d_s \geq 0$ and call d_s the flow value of x. Beside this edge-based notion, a static s-t-flow can also be stated path-based. In this case, a flow value $x_P \in \mathbb{R}_+$ is assigned to each simple s-t-path $P \in \mathcal{P}$. Edge-and path-based flows are strongly related via the equation

$$x_e = \sum_{P|e\in E(P)} x_P \ .$$

This equation shows how to transform a path-based flow into an edge-based flow. Vice versa, each edge-based flow is decomposable into a path-based flow and some flow which is sent along cycles.

2.4 Game Theory and Static Routing Games

This section imparts the basic ideas behind game theory and, in particular, behind static routing games. We start by presenting the prisoner's dilemma which is probably the most frequently used example when introducing game theory. Based on this example, we briefly discuss the notion of an atomic game consisting of a finite number of players. Subsequently, we turn our attention to static nonatomic routing games. We exemplarily describe selfish routing behavior using Pigou's example and formalize important aspects of this example.

Example 2.37 (Prisoners Dilemma.). Assume two criminals rob a bank and are arrested by the police. Unfortunately, the police is not able to prove that the two guys have really robbed the bank. Instead, they could only report unauthorized possession of weapons which would imply an imprisonment of three years. On the other hand, a conviction due to bank robbery would result in an imprisonment of ten years. So the police decides to make the following offer by individual surveys of the two criminals such that they cannot cooperate:

- If both confess the bank robbery, each of them gets an imprisonment of seven years.
- If exactly one guy confesses the bank robbery and simultaneously incriminate the other one, he gets an imprisonment of two years and the other guy gets an imprisonment of ten years.
- If no one confesses the bank robbery, both get an imprisonment of three years due to the unauthorized possession of weapons.

What is the behavior of the two criminals – do they confess or not? Let us consider one of the two criminals. Since criminals are, in fact, really egoistic, he figures out what will be the best strategy for him depending on the choice of the other. If the other guy confesses, he would confess as this reduces the imprisonment from ten to seven years. If the other guy does not confess, he would also confess as this reduces again the imprisonment, but now from three to two years. For these reasons, he decides to confess. As the scenario is symmetric, the same considerations remain valid for the other guy. Hence, both guys confess in the end.

Now assume that the police repeats the individual surveys. The only stable behavior of the two criminals is that both confess. In each other behavior of the two criminals at least one of them can reduce its imprisonment by switching from nonconfessing to confessing. From an external view, also the following aspect is interesting. If no one confesses they get a total imprisonment of six years, which is best possible. But in contrast, the egoistic behavior leads to a total imprisonment of 14 years which is the worst obtainable result.

In the following we formalize the prisoner's dilemma leading to a notion of atomic routing games. The term *atomic* means that such a game consists of a finite number $1, \ldots, K$ of players for some $K \in \mathbb{N}$ as it is the case for the prisoner's dilemma. Here, we have two players, namely the two criminals. Each player can choose a strategy out of a given set $\mathcal{S}$. For the prisoner's dilemma these strategies are confessing or nonconfessing. A particular run of the game consists of the chosen strategies of all players which is representable as an element of $\mathcal{S}^K$. Such an element is also called strategy profile. Depending on the strategy profile and his chosen strategies, a particular penalty is assigned to each player. In the case of the prisoner's dilemma these are imprisonments. The amount of the penalty is given by so-called penalty functions $\ell_S : \mathcal{S}^K \to \mathbb{R}$ for all $S \in \mathcal{S}$. Thus, the penalty of some player i in a particular run $(S_1, \ldots, S_K) \in \mathcal{S}^K$ of the game is $\ell_{S_i}(S_1, \ldots, S_K)$. This is the entire setting of a game. As each player shares the same set of strategies and the same penalty functions, such a game is called symmetric.

If all players behave egoistically, a single player would change his strategy in a repeated run of this game if this decreases the penalty under the assumption that each other player remains his strategy. Thus, a particular strategy profile is stable if no player has an incentive to switch his strategy unilaterally. Such stable strategy profiles are called Nash equilibria. In this sense, a strategy profile $S := (S_1, \ldots, S_K) \in \mathcal{S}^K$ is a Nash equilibrium if and only if

$$\ell_{S_i}(S) \;=\; \min_{\tilde{S}_i \in \mathcal{S}} \; \ell_{\tilde{S}_i}(S_1, \ldots, S_{i-1}, \tilde{S}_i, S_{i+1}, \ldots, S_K) \qquad \forall i \in [K+1] \; .$$

On the other hand, assume that all players cooperate and only have in mind their social welfare. In this case a run of this game results in a strategy profile where the total penalty of all players is minimal. Such a strategy profile $S \in \mathcal{S}^K$ is called social optimum and satisfies

$$\operatorname{val}(S) \;:=\; \sum_{i=1}^{K} \ell_{S_i}(S) \;=\; \min_{\tilde{S} \in \mathcal{S}^K} \; \sum_{i=1}^{K} \ell_{\tilde{S}_i}(\tilde{S}) \; .$$

Of course, social welfare can also be determined by other objective functions than simply the total sum of the penalties. An interesting question is, how bad is egoistical behavior for the social welfare. Usually, this is measured by the price of anarchy. The price of anarchy ρ for a particular Nash equilibrium S_{N} is given by

$$\rho(S_{\mathrm{N}}) \;:=\; \frac{\operatorname{val}(S_{\mathrm{O}})}{\operatorname{val}(S_{\mathrm{N}})}$$

where S_{O} is a social optimum. This completes our short discussion about the notion of games.

In the following, we briefly introduce static nonatomic routing games. As already mentioned, we start with explaining Pigou's example. But before, we elucidate the term *nonatomic*. Recall that the term *atomic* stands for a finite number of players. However, if the number of players is too huge the corresponding game becomes computationally intractable. In such a case, an aggregated notion of the players is helpful. Having in mind routing games, exemplarily assume that there are k players controlling a flow unit of size $\frac{1}{k}$ each. In case k

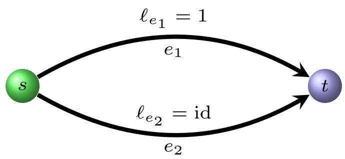


Figure 2.1: The network of Pigou's example.

goes to ∞, the entire flow of size 1 is representable by the interval $[0, 1]$. In this environment, we think of a player as a real number $p \in [0, 1]$ which controls the infinitesimal flow unit at position p. Throughout this thesis an infinitesimal flow unit is called flow particle.

Example 2.38. Pigou's example works on a network consisting of two parallel edges e_1 and e_2 connecting a source s with a sink t (see Figure 2.1). The flow behavior is represented via two arrival time functions $\ell_{e_1}, \ell_{e_2} : \mathbb{R}_+ \to \mathbb{R}_+$ which can also be interpreted as transit time functions. The arrival time function of the upper edge e_1 equals the constant function 1. That is, independent of the respective amount, flow arrives at time 1 at t when traveling along e_1. The arrival time function of e_2 is equal to the identity function. Hence, if flow of value $x \in \mathbb{R}_+$ traverses e_2, it arrives at t at time x. In the following, we consider the behavior of players controlling flow with a total size of 1. Since this setting is symmetric, it does not matter at which position a particular flow particle is placed. In this sense, the number 1 already represents the set of players.

Firstly, we consider the selfish routing which arises out of the egoistic behavior of the players. Clearly, if a positive amount of flow is sent along e_1, the arrival time of e_2 is strictly smaller than 1. Hence, each player which sends his flow particle along e_1 would switch to e_2. This shows that, in a Nash equilibrium, the entire flow is sent along the lower edge e_2.

In order to compute a social optimum, we assume that social welfare is measured via the total load of the network. That is, if flow of value $x \in [0, 1]$ travels along e_2, the total load is given by $(1 - x) \cdot 1 + x \cdot x$. Hence, a social optimum minimizes $x^2 - x + 1 = (x - \frac{1}{2})^2 + \frac{3}{4}$. This shows that, in the social optimum, half of the flow is sent along e_1 and the other half along e_2 which leads to a total load of $\frac{3}{4}$. Considering again the Nash equilibrium, which sends the entire flow along e_2, we see that the total load is equal to 1. Altogether, the price of anarchy for this scenario is $\frac{3}{4}$.

Generalizing Pigou's example, a static nonatomic routing game acts on a network consisting of a directed graph $G := (V, E)$, a source $s \in V$, and a sink $t \in V$. Further, a demand $d \in \mathbb{R}_+$ is given representing the amount of flow which shall be sent through the network. As in Pigou's example, each flow particle is controlled by a player. The strategy set of each player equals the set $\mathcal{P}$ of all s-t-paths. Thus, for each s-t-path P, the decisions of the players result in an amount $x_P \in \mathbb{R}_+$ of flow which is sent along P. Clearly, the respective values x_P have to sum up to d. In this manner, the selfish decisions of the players result in a static path-based s-t-flow x of value d. Moreover, we have a continuous arrival time function ℓ_P for each path $P \in \mathcal{P}$ depending on the entire flow x.

A static flow $x := (x_P)_{P \in \mathcal{P}}$ is a Nash equilibrium or Nash flow if and only if each flow carrying path is a shortest one, i.e., $\ell_P(x) = \min_{P' \in \mathcal{P}} \ell_{P'}(x)$ for all $P \in \mathcal{P}$ with $x_P > 0$. Finally, the total load, which serves as a measure for social welfare, is given by $\sum_{P\in\mathcal{P}} x_P \cdot \ell_P(x)$.

2.5 Classical Flows over Time

This section is devoted to impart a first idea of flows over time. We discuss the classical flow over time model which is introduced by Ford and Fulkerson [27] for a discrete setting. But, we focus on the continuous case. After defining classical flows over time, we shortly discuss the corresponding maximum flow problem. Finally, we turn our attention to so-called earliest arrival flows which can be computed by a successive shortest path algorithm. An earliest arrival is characterized by the fact that it sends the maximum amount of flow into the sink simultaneously until all points in time $T \in \mathbb{R}_+$. We point out that throughout this section, in certain circumstances, a classical flow over time is represented in three different manners – edge-based, path-based, and generally path-based.

Classical flows over time act on a network $(G,\mathcal{U},\mathcal{T},s,t,T)$ consisting of a directed graph G, a source s, and a sink t. Further, a family $\mathcal{U} := (u_e)_{e\in E}$ of real edge capacities $u_e \in \mathbb{R}_+$ is given. A real edge capacity u_e represents a capacity function which is constantly set to u_e. For the purpose of this thesis, we should have in mind that an edge capacity bounds the rate of flow *leaving* an edge. In addition, the flow behavior is also determined by a family $\mathcal{T} := (\tau_e)_{e\in E}$ of constant transit times $\tau_e \in \mathbb{R}_+$ determining the amount of time needed for traversing an edge. More precisely, if a flow particle enters an edge at some point in time $\theta \in \mathbb{R}_+$, it leaves e at time $\theta + \tau_e$. Finally, we are given a time horizon T representing the considered time period. That is, we assume that no flow travels somewhere through the network outside the time interval $[0,T]$.

On such a network, a flow over time is given as a family $\mathcal{F} := (f_e)_{e\in E}$ of Lebesgue integrable functions $f_e : [0,T] \to \mathbb{R}_+$. The value $f_e(\theta)$ stands for the rate at which flow leaves an edge e at a certain time $\theta \in [0,T]$. Since transit times are constant, $f_e(\theta + \tau_e)$ equals the rate at which flow enters e at time θ. A feasible flow over time $\mathcal{F} := (f_e)_{e\in E}$ has to obey edge capacities, i.e.,

$$f_e(\theta) \;\le\; u_e \qquad\qquad \forall e \in E,\;\; \theta \in [0,T]\;.$$

In addition, $\mathcal{F}$ has to satisfy some kind of flow conservation. We require that flow which leaves a certain node v at some point in time $\theta \in [0,T]$ must enter this node before. As storage at nodes is allowed, this results in the condition

$$\sum_{e\in\delta^-(v)} \int_0^\theta f_e(\vartheta)\, d\vartheta \ge \sum_{e\in\delta^+(v)} \int_0^\theta f_e(\vartheta + \tau_e)\, d\vartheta \qquad \forall v \in V\backslash\{s\}, \theta \in [0,T]. \quad (2.11)$$

Clearly, as s is the source of all flow particles, this condition must not be required for s. If equality holds in (2.11) for each $v \in V \setminus \{s,t\}$, no flow is stored at intermediate nodes. Hence, requiring equality in (2.11) for all intermediates nodes forbids storage. In this case we call the flow conservation constraint strict. Below in this section we see that forbidding storage does not decrease the potential of optimizing.

Finally, we have to ensure that no flow remains in the network after time T. This means, no flow is allowed to traverse an edge after time T and no flow is allowed to be stored at some node after time T. This is formulated as

$$\sum_{e\in\delta^-(v)} \int_0^T f_e(\vartheta)\, d\vartheta \;=\; \sum_{e\in\delta^+(v)} \int_0^T f_e(\vartheta+\tau_e)\, d\vartheta \qquad \forall v \in V\setminus\{s,t\}\ .$$

The value $\operatorname{val}(\mathcal{F})(T)$ of $\mathcal{F}$ equals the net inflow of t until time T, i.e.,

$$\operatorname{val}(\mathcal{F})(T) \;:=\; \sum_{e\in\delta^-(t)} \int_0^T f_e(\vartheta)\, d\vartheta - \sum_{e\in\delta^+(t)} \int_0^T f_e(\vartheta+\tau_e)\, d\vartheta\ .$$

In the following, we show how to maximize $\operatorname{val}(\mathcal{F})(T)$ for a given time horizon T. For this we need the notion of temporally repeated flows, which are a special kind of classical flows over time. Temporally repeated flows are path-based and constructed from a feasible static flow on $(G,\mathcal{U},s,t)$. Recall that edge capacities are given by real numbers, and hence, $(G,\mathcal{U},s,t)$ is interpretable as a static network. Let $(x_P)_{P\in\mathcal{P}}$ be a static path-based flow obeying edge capacities. The idea behind temporally repeated flows is the following. For a particular s-t-path P, interpret the value x_P as a constant rate at which flow enters P as long as this flow arrives at t before time T. Further, assume that this flow traverses P directly without waiting at intermediate nodes. In this manner, the path-based representation of the corresponding temporally repeated flow $\mathcal{F} = (f_P)_{P\in\mathcal{P}}$ is given by

$$f_P(\theta) \;:=\; x_P \qquad\qquad \forall P\in\mathcal{P},\ \theta\in[0,T-\tau_P]\ .$$

Since waiting at intermediate nodes does not occur, the edge-based representation $(f_e)_{e\in E}$ of $\mathcal{F}$ is computable as

$$f_e \;:=\; \sum_{P,i|e_i^P=e} f_P - \tau_{P,i} \qquad\qquad \forall e\in E$$

where $\tau_{P,i} := \sum_{j=1}^{i} \tau_{e_i^P}$ is the time needed for arriving at the end of the i-th edge of P after departing at s. Note that, for this definition, it is essential that waiting does not occur. Clearly, by this definition, $\mathcal{F}$ satisfies the strict flow conservation constraint. In addition, the flow rate on an edge is always bounded by x_P. This shows that $\mathcal{F}$ satisfies the capacity constraint. Also by the definition of $\mathcal{F}$, no flow remains in the network after time T. Hence, $\mathcal{F}$ is a feasible classical flow over time.

Recall that we want to maximize $\operatorname{val}(\mathcal{F})(T)$. Already Ford and Fulkerson observe in [27] the following. If the static flow $(x_P)_{P\in\mathcal{P}}$ minimizes the load $\sum_{P\in\mathcal{P}} x_P\cdot\tau_P$ then the corresponding temporally repeated flow maximizes $\operatorname{val}(\mathcal{F})(T)$. But minimizing $\sum_{P\in\mathcal{P}} x_P\cdot\tau_P$ results in a static minimum cost flow problem which is efficiently solvable. Hence, we are able to construct a classical flow over time maximizing $\operatorname{val}(\mathcal{F})(T)$ efficiently in this manner.

One possibility for solving minimum cost flow problems is provided by the so-called Successive Shortest Path algorithm. This algorithm mainly acts on residual networks. Given a *static* network $(G,\mathcal{U},\mathcal{T},s,t)$, the graph $G^r := (V,E^r)$ of the residual network contains in addition to G a backward edge $\overleftarrow{e} := wv$ for each original edge $e = vw\in E$ with transit time $\tau^r_{\overleftarrow{e}} := -\tau_e$. The residual capacities $\mathcal{U}^r := (u^r_e)_{e\in E^r}$ depend on a given feasible static flow $x := (x_e)_{e\in E}$.

For an edge $e \in E$, the residual capacity of e and $\overleftarrow{e}$ are set to $u^r_e := u_e - x_e$ and $u^r_{\overleftarrow{e}} := x_e$, respectively. The network $(G^r, \mathcal{U}^r, \mathcal{T}^r, s, t)$ is called residual network of x. Using a feasible static s-t-flow $x^r := (x_e)_{e \in E^r}$ on the residual network, we are able to augment x. It is well-known that the augmented flow, defined by $x^a := (x_e + x^r_e - x^r_{\overleftarrow{e}})_{e \in E}$, is a feasible static s-t-flow on the original network and the value of x^a equals the sum of the values of x and x^r.

The Static Successive Shortest Path algorithm works iteratively as follows (see also Step (1)-(4) of the algorithm presented below). In each iteration, the residual network of the current flow x is constructed. With respect to the residual transit times, a shortest path $P \in \mathcal{P}^r$ which is able to carry flow is computed, i.e., $u^r_P := \min_{i=1,\dots,|P|} u^r_{e^P_i} > 0$ holds in particular. The current flow x is augmented with the flow x^r which sends a value of u^r_P along P. Finally, the augmented flow x^a is passed to the next iteration.

There are three interesting aspects which are founded in the functionality of the Static Successive Shortest Path algorithm. Firstly, the current flow x always minimizes the load $\sum_{P \in \mathcal{P}} x_P \cdot \tau_P$ among all s-t-flows of the same value. Secondly, the transit times of the computed shortest paths are nondecreasing. Thirdly, the current flow x is given by a path-based family $(x_P)_{P \in \mathcal{P}^r}$ containing the path-based family of all previously computed flows. Recalling the construction of the temporally repeated flow maximizing $\mathrm{val}(\mathcal{F})(T)$ for some given T, this suggests that the final flow returned by the Static Successive Shortest Path algorithm encodes a classical flow over time which maximizes $\mathrm{val}(\mathcal{F})(T)$ for all $T \in \mathbb{R}_+$ simultaneously. And, in fact, this is the case. Notice that, as already mentioned, such a classical flow over time is called earliest arrival flow.

Before summarizing the observations above in an algorithm for computing an earliest arrival flow, we provide some intuition for the notion $x := (x_P)_{P \in \mathcal{P}^r}$ and its consequences for the corresponding classical flow over time. It follows directly from the definition of the augmented flow that the edge-based representation $(x_e)_{e \in E}$ of x is given by

$$x_e \;:=\; \sum_{P,i|e^P_i = e} x_P - \sum_{P,i|e^P_i = \overleftarrow{e}} x_P \qquad \forall e \in E\ .$$

Thus, if flow is sent along a backward edge $\overleftarrow{e}$, the flow on e is reduced. Further, the definition of the residual capacities and the augmentation ensure that x_e is a feasible flow value, i.e., satisfies $0 \leq x_e \leq u_e$. In particular, that flow is always augmented along *shortest* paths does not matter for this observation. Next, consider the corresponding classical flow over time $\mathcal{F} := (f_P)_{P \in \mathcal{P}^r}$ which is given by $f_P(\theta) := x_P$ for all $P \in \mathcal{P}^r$ and $\theta \in \mathbb{R}_+$. The meaning of $\mathcal{F}$ coincides with the meaning of a temporally repeated flow. That is, the value x_P is interpreted as a constant rate at which flow enters P. Further, we assume that this flow traverses P directly without waiting at intermediate nodes. Note that, as the residual transit time of backward edges is negative, flow travels back in time on such edges. Further, flow on backward edges reduces the flow on the original edges. Since the flow rate on an original edge is measured at its head, we have to measure the flow rate on backward edges at its tail. This leads to the following edge-based representation $(f_e)_{e \in E}$ of $\mathcal{F}$:

$$f_e \;:=\; \sum_{P,i|e^P_i = e} (f_P + \tau^r_{P,i}) - \sum_{P,i|e^P_i = \overleftarrow{e}} (f_P + \tau^r_{P,i-1}) \qquad \forall e \in E\ .$$

The feasibility of $\mathcal{F}$ requires that, whenever flow is sent along backward edges, flow must be simultaneously sent along the corresponding forward edge. Otherwise, the flow rate would become negative. For proving this, it is essential that the SUCCESSIVE SHORTEST PATH algorithm always sends flow along shortest paths. Note that similar considerations must also be made for the flow on original edges.

A feasible flow $\mathcal{F}$ which is representable as $(f_P)_{P\in\mathcal{P}^r}$ is called generalized temporally repeated flow. Furthermore, we call the respective representation $(f_P)_{P\in\mathcal{P}^r}$ general path-based flow over time or general path decomposition. With this terminology we state the CLASSICAL SUCCESSIVE SHORTEST PATH algorithm which computes an earliest arrival flow.

CLASSICAL SUCCESSIVE SHORTEST PATH ALGORITHM (CSSP)

Input: A network $\mathcal{N} := (G,\mathcal{U},\mathcal{T},s,t)$ consisting of a directed graph G, a family $\mathcal{U}$ of real edge capacities, a family $\mathcal{T}$ of real edge transit times, a source s, and a sink t.

Output: A general path-based flow over time $\mathcal{F} := (f_P)_{P\in\mathcal{P}^r}$ representing a classical earliest arrival flow.

(1) Let $x := (x_P)_{P\in\mathcal{P}^r}$ be the zero flow, i.e., set $x_P := 0$ for all $P \in \mathcal{P}^r$.

(2) Compute the residual network $(G^r,\mathcal{U}^r,\mathcal{T}^r,s,t)$ with respect to x.

(3) Let $P \in \mathcal{P}^r$ be a shortest s-t-path beeing able to carry flow, i.e., $u^r_P > 0$. If no such path exists go to (5).

(4) Set $x_P := u^r_P$ and go to (2).

(5) Set $f_P(\theta) := x_P$ for all $P \in \mathcal{P}^r$ and all $\theta \in \mathbb{R}_+$.

We conclude this section with one remark concerning the correctness of this algorithm. Usually, this algorithm is interrupted if the transit time of the computed shortest path exceeds some given time horizon T. It is known that the corresponding output of the CSSP algorithm is an (restricted) earliest arrival flow until time T, i.e., it maximizes $\text{val}(\mathcal{F})(\theta)$ simultaneously for all times $\theta \in [0,T]$. Note that, defining $(f_P)_{P\in\mathcal{P}^r}$ as the unrestricted earliest arrival flow, the restricted classical earliest arrival flow $(\tilde{f}_P)_{P\in\mathcal{P}^r}$ is given by

$$\tilde{f}_P \;=\; f_P|_{[0,T-\tau^r_P]} \qquad\qquad \forall P \in \mathcal{P}^r \;.$$

2.6 Your Comments

"Well, now I have the basic knowledge for understanding this thesis. A little bit of many sometimes quite different fields of mathematics", you think. You need knowledge about analysis including basic insights in the theory of Lebesgue measurable functions and functional analysis. Moreover, intuition behind game theory, in particular behind classical flows over time and routing games, are also advantageous when reading this thesis. Finally, some stuff of linear algebra and a deeper understanding of static flow theory are also needed.

Section 2.1 introduces the basic notations used throughout this thesis. Most of the notations are standard and commonly used in mathematics. Only the shifting of functions takes getting used to. Running over the pages of this thesis, you observe that this notation is occurs a lot of times. So you decide to engrain this particular notation.

Concerning Section 2.2, you find out that both the Lebesgue measure as well as the Lebesgue integral are introduced by Lebesgue in 1902 in his great thesis [50]. You guess that most of the properties of the Lebesgue measure and the Lebesgue integral are contained therein. For learning more about this subject you find a lot of good textbooks like [78, 79] concerning the theory of real-valued functions on the real line. Since for some of the well-known basic theorems there is no citation, you decide to send the references of the corresponding articles to the author. Unfortunately, you do not succeed in finding all of these references until now. Further, you agree with the author that some of the pursuing results should be known. This holds, in particular, for Lemma 2.23 which proves the absolutely continuity of some function F^-. As you know something about measure theory, you observe that the assumption in Lemma 2.23 directly implies that F^- is the distribution function of some measure which is absolutely continuous with respect to the Lebesgue measure. The well-known definition concerning distribution functions of measures shows that F^- is absolutely continuous in terms of Definition 2.15. But you also find that introducing even measure theory in this thesis is not a good idea. Further, you have the feeling that the monotonicity set and the support (see Definition 2.19 and 2.24) coincide essentially. So far, you are not able to verify this neither personally nor by literature.

Since you have no comment on Section 2.3 about networks, you focus on Section 2.4. You find out that the prisoner's dilemma in the presented form is introduced by Tucker in 1950 during a talk in front of psychologist at Stanford University. For this Tucker fill with life a more abstract version of this game by Flood and Dresher in the same year.

As already mentioned in this section, the presented atomic game is a symmetric game. Usually, the strategy set and the set of penalty functions can differ from one player to another. Further, the presented definition of a Nash equilibrium results in a purely deterministic run of the game as the behavior of one players is determined by exactly one strategy. Such Nash equilibria are called pure in literature, and it is well-known that not all atomic games admit such a pure Nash equilibrium. On the other hand, the behavior of a player can also be given as a probability distribution over his set of strategies. That is, during a run of a game a player chooses a strategy with a certain probability. In this manner, the penalty functions are used for defining the expected penalty of one player. Thus, in a Nash equilibrium no player shall have the incentive to change its probability distribution for improving his expected penalty. That such Nash equilibria exists, for all atomic games, is established in [62] by Nash in 1951. Besides, it may be interesting that Nash was a Ph.D. student of Tucker.

Concerning nonatomic routing games, Pigou discusses his example in 1920 (see [73]). Three decades later, in 1952, Wardrop [90] formalizes the notion of static routing games. In particular, he states two guidelines explaining traffic behavior known as Wardrop's principles. The first principle defines Nash equilibria and the second social optima. Formalizing the notion of static routing games, Beckmann, McGuire, and Winsten [9] lays the ground for the theory of static routing games. In his seminal work [77] on the price of anarchy, Roughgarden and Tardos show that the price of anarchy of $\frac{3}{4}$, which holds for Pigou's example, extends to arbitrary static routing games with linear transit time functions. The price of anarchy for arbitrary classes of transit time functions is analyzed, e.g., by Correa, Schulz, and Stier-Moses [19] in 2004.

Basic knowledge about flows over time is imparted in Section 2.5. As already mentioned in the Introduction, the discrete counterpart is introduced by Ford and Fulkerson [27] in 1958. Already in this article Ford and Fulkerson show that maximal classical flows over time are efficiently computable using a static minimum cost flow algorithm. As Ford and Fulkerson does this for a discrete setting, Anderson and Philpott [3] observe this for the continuous model.

The existence of earliest arrival flows are established by Gale [32] for the discrete case and by Philpott [71] for the continuous case. The Classical Successive Shortest Path algorithm is due to Wilkinson [91] and Minieka [56] for discrete scenarios whereas Fleischer and Tardos [25] analyze this algorithm for continuous scenarios. Besides, it is worth to know that classical earliest arrival flows serve as optima for four different kinds of objective functions:

- the earliest arrival property, i.e., maximizing the flow in the sink.
- maximizing the average rate at which flow arrives at the sink.
- minimizing the time which a given amount of flow need to travel to the sink.
- the average travel time of all flow particles.

Except for the second objective function, these aspects of classical earliest arrival flows are observed by Jarvis and Ratliff [40].

Chapter 3

Flows over Time

In this chapter we introduce general flow over time models. There are two popular classes of flows over time which are built upon either continuous or discrete time representations. In discrete models flow is given via flow units having no physical dimensions but containing a positive amount of flow. In particular, if a flow unit enters an edge at some point in time, the entire amount of flow enters this edge at this time. For example, one can think of a flow of cars on a traffic network in this context. In contrast, a flow over time in a continuous model is given via flow rates. So it consists of infinitesimal (very small) flow units which we call *flow particles*. Intuitively, a flow particle can stand for a single bit (a '0' or '1') which traverses a digital communication network.

In literature there exist several models representing the flow behavior on a network. In order to analyze such flow models, additional assumptions motivated by real-world applications were imposed on the model; sometimes implicitly, sometimes explicitly. In this sense, each flow model has its own theory and, for translating definitions and results to other flow models, complicated transformations are necessary in general. However, the question arises, whether one could generalize the notion of flow over times, to cover most of the popular flow models. The advantage of such a general flow model is obvious. Each basic assumption and each obtained result hold clearly for each covered flow model.

It lies in the nature of flows over time that intuition is a very good tool for discussing this topic. Unfortunately, continuous flows over time are mainly modeled via Lebesgue integrable functions and it lies as well in the nature of Lebesgue integrable functions that intuition sometimes fails. So it is one ambition of this chapter to built a mathematically well-founded basis for the theory of flows over time.

In this chapter we present a very general continuous flow over time model which imposes only very weak assumptions on the participating mathematical objects. In fact, many established flow over time models are covered by this model. In Section 3.1 we develop a path-based representation of flows over time and, along the same lines, we present an edge-based formulation in Section 3.2. The main contribution of these two sections is to characterize equivalent flow models. The more intuitive discrete time models are briefly discussed in Section 3.7.

Based on these general flows over time, we impose natural assumptions on flow models in Section 3.3 leading to the notion of consistent flow models. A large class of existing flow over time models fall into this category including the direct flow model and the deterministic queuing model studied in Chapter 5 and 7, respectively.

The basic feature on which flow models are defined in Section 3.1 and 3.2 are flow-dependent transit time functions. In Section 3.4 we present well-suited transit times for consistent flow models having game theoretic aspects in mind. As in static flow theory, there exists a strong relation between path- and edge-based flows over time leading to the dynamic network loading and the path decomposition problem. We briefly consider both problems in Section 3.5. In Section 3.6 we analyze a shortest path problem which is of particular importance when defining Nash equilibria on consistent edge-based flow over time models. Finally, in Section 3.8 we generalize the notion of classical flows over time to arbitrary time varying capacities. In particular, we define a residual network for this setting leading to another flow over time representation.

Throughout this chapter and subsequently, the symbol $\mathcal{F}$ is reserved for denoting a flow over time independently on the underlying representation. In fact, $\mathcal{F}$ stands simultaneously for any representation and is, therefore, seen as a top-level reference of a particular flow over time. Further, due to the applications in this thesis, we only consider flows over time with one commodity. More precisely, all flow traverses a given network from a unique source s to a unique sink t.

3.1 Path-based Formulation

In this section we define path-based flows over time. We strictly distinguish between a flow over time and an underlying flow model. Roughly speaking, given a set $\mathcal{P}$ of s-t-paths, we assign an inflow F_P^+ to each path $P \in \mathcal{P}$. This inflow shows how flow enters a certain path and the family of these inflows build the flow over time. How the assigned flow traverses a path is determined by a flow model. This works as follows. Depending on the entire family of inflows, the flow model computes transit times for each path $P \in \mathcal{P}$ and for all points in time $\theta \in \mathbb{R}_+$. If a flow particle enters a certain path, the current transit time determines the amount of time needed for traversing this path in order to arrive at t. In particular, the transit time of a path at a certain point in time can depend on the inflow of other paths as well as on past, present, and future flow situations. Using transit times, we identify equivalent flow models and define the outflow F_P^- of some path P. An outflow F_P^- represents how flow particles traveling along P arrive at t. Based on these definitions, we characterize consistent flow models in Section 3.3.

Subsequently, we deal only with $\mathcal{P}$ and not with a graph G. So in fact, $\mathcal{P}$ can have any meaning. However, it is always a good idea to have in mind that $\mathcal{P}$ equals the set of s-t-paths of some graph G implying, in particular, that $\mathcal{P}$ is at most countable. In Section 2.5 flows over time are modeled via Lebesgue integrable flow rate functions. Instead, we define flows over time using antiderivatives of such flow rate functions. That is, F_P^+ represents the cumulative amount of flow entering a path P until a particular point in time. On the one hand, this results in an easier notion because we avoid integrals whenever possible. On the other hand, this enables us to discuss various extensions of flows over time including discrete time models without changing the notion.

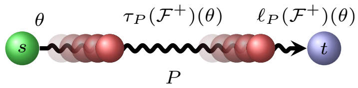


Figure 3.1: A flow particle entering an s-t-path $P \in \mathcal{P}$ at time θ arrives after $\tau_P(\mathcal{F}^+)(\theta)$ time units at the sink t at time $\ell_P(\mathcal{F}^+)(\theta) = \theta + \tau_P(\mathcal{F}^+)(\theta)$.

Definition 3.1 (Path-based Flow over Time)**.** We call a family $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$ *path-based flow over time* if, for all paths $P \in \mathcal{P}$, the functions $F_P^+ : \mathbb{R}_+ \to \mathbb{R}_+$ are nondecreasing, absolutely continuous and satisfy $F_P^+(0) = 0$. For a path P the function F_P^+ is called *path flow* or *inflow function.*

For a point in time $\theta \in \mathbb{R}_+$, we define the *restricted flow* $\mathcal{F}^+|_{\leq\theta}$ until time θ as $\mathcal{F}^+|_{\leq\theta} := (F_P^+|_{\leq\theta})_{P\in\mathcal{P}}$ where $F_P^+|_{\leq\theta} := F_P^+|_{[0,\theta]}$ for all $P \in \mathcal{P}$.

We write $\mathcal{F}^+ := (F_P^+)_{P\in\bar{\mathcal{P}}}$ for some $\bar{\mathcal{P}} \subseteq \mathcal{P}$ if $F_P^+ = 0$ for all $P \in \mathcal{P} \setminus \bar{\mathcal{P}}$. Further, if the underlying path of a path flow is not clear, we use the notation $\mathcal{F}^+ := ((F_P^+, P))_{P\in\bar{\mathcal{P}}}$. If a flow only consists of one path flow, we usually omit the outer brackets and simply write F_P^+ or (F_P^+, P). So not surprisingly, the path flow of any path itself is interpreted as a flow over time.

For a point in time $\theta \in \mathbb{R}_+$, the value $F_P^+(\theta)$ equals the total amount of flow that has been assigned to P until time θ. Since each path flow F_P^+ is absolutely continuous, it is the antiderivative of some Lebesgue integrable function $f_P^+ : \mathbb{R}_+ \to \mathbb{R}_+$. That is, we have $\frac{dF_P^+}{d\theta}(\theta) = f_P^+(\theta)$ for almost all $\theta \in \mathbb{R}_+$ and $F_P^+(\theta) = \int_0^\theta f_P^+(\vartheta)\, d\vartheta$ for all $\theta \in \mathbb{R}_+$. Thus, the flow $\mathcal{F}^+$ is completely defined by $(f_P^+)_{P\in\mathcal{P}}$ and we also use the notation $\mathcal{F}^+ := (f_P^+)_{P\in\mathcal{P}}$. Intuitively, for a point in time $\theta \in \mathbb{R}_+$ the value $f_P^+(\theta)$ equals the rate at which flow enters P at time θ. So $f_P^+(\theta)$ can be seen as the number of flow particles (infinitesimal flow units) departing from s along P at time θ.

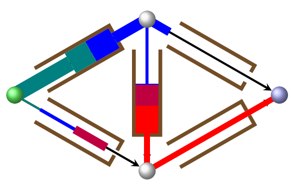
So far, $\mathcal{F}^+$ only shows how flow is assigned to the paths. How flow traverses a chosen path is defined via a flow model. As already mentioned, a flow model is defined via flow-dependent transit times. Exemplarily, Figure 3.1 shows the flow behavior represented via a flow model.

Definition 3.2 (Flow over Time Model)**.** Depending on the network structure, a *flow over time model* defines a family $\mathcal{T} := (\tau_P)_{P\in\mathcal{P}}$ of maps τ_P, $P \in \mathcal{P}$ assigning a locally bounded, Lebesgue measurable function $\tau_P(\mathcal{F}^+) : \mathbb{R}_+ \to \mathbb{R}_+$ to each flow over time $\mathcal{F}^+$. The map τ_P is called *transit time* of $P \in \mathcal{P}$ and $\tau_P(\mathcal{F}^+)$ is called *transit time function* with respect to $\mathcal{F}^+$. The value $\tau_P(\mathcal{F}^+)(\theta)$ is called *current transit time.*

A corresponding family $\mathcal{L} := (\ell_P)_{P\in\mathcal{P}}$ of *arrival times* is given by an *arrival time function* $\ell_P(\mathcal{F}^+) : \mathbb{R}_+ \to \mathbb{R}_+$ defined as $\ell_P(\mathcal{F}^+)(\theta) := \theta + \tau_P(\mathcal{F}^+)(\theta)$ for each path $P \in \mathcal{P}$. The value $\ell_P(\mathcal{F}^+)(\theta)$ is called *current arrival time.*

In a *continuous* flow over time model every transit time function must only be defined almost everywhere on $\mathbb{R}_+$. Further, we require that each transit time function $\tau_P(\mathcal{F}^+)$ is compatible with F_P^+. That is, the preimage of every null set with respect to $\tau_P(\mathcal{F}^+)$ is an F_P^+-null set for each $P \in \mathcal{P}$.

If the flow $\mathcal{F}^+$ is clear from context, we use $\tau_P(\theta)$ and $\ell_P(\theta)$ for denoting current transit and current arrival times, respectively. Furthermore, we use τ_P and ℓ_P for transit and arrival time functions, respectively, if there is no name conflict with the corresponding maps.



Note that a transit time function $\tau_P(\mathcal{F}^+)$ depends on the entire flow $\mathcal{F}^+$ and not only on the inflow F_P^+ of the underlying path P. This also shows that, in principle, current transit times can depend on past, present, and future flow situation. Also we do not require that the family $\mathcal{T}$ of transit times is given explicitly. So there can be highly nontrivial dependencies between different maps τ_P. Why we impose the compatible-condition on transit time functions in continuous models becomes clear below. In particular, it ensures that a positive amount of flow, which enters a path *continuously* over time, does not arrive at t at the same *discrete* point in time.

In the following, we define equivalent flow over time models. Intuitively, if flow is assigned to some s-t-path P at a certain point in time θ, it arrives at t at time $\theta + \tau_P(\mathcal{F}^+)(\theta) = \ell_P(\mathcal{F}^+)(\theta)$. Hence, we could require that two equivalent flow models must lead to the same family $(\ell_P(\mathcal{F}^+))_{P\in\mathcal{P}}$ of arrival time functions for each flow $\mathcal{F}^+$. However, if no flow is assigned to some path P, it makes no difference whether or not the current arrival times coincide.

Definition 3.3 (Equivalent Flow over Time Models)**.** Let τ_1 and τ_2 be transit time functions for some path $P \in \mathcal{P}$ and let $\Theta := \{\theta \in \mathbb{R}_+ \mid \tau_1(\theta) \neq \tau_2(\theta)\}$ be the set of points in time where the two transit time functions differ. Then τ_1 and τ_2 are called *equivalent with respect to some path flow F_P^+* if and only if essentially no flow is assigned to P over the set Θ, i.e., $\int_\Theta f_P^+(\vartheta)\, d\vartheta = 0$.

Let $\mathcal{T}_1 := (\tau_P^1)_{P\in\mathcal{P}}$ and $\mathcal{T}_2 := (\tau_P^2)_{P\in\mathcal{P}}$ be the families of transit times of two flow models. Then these two flow models are called *equivalent* if, for all flows over time $\mathcal{F}^+$ and all paths $P \in \mathcal{P}$, the transit time functions $\tau_P^1(\mathcal{F}^+)$ and $\tau_P^2(\mathcal{F}^+)$ are equivalent with respect to F_P^+.

Informally, Definition 3.3 requires that whenever a positive amount of flow is assigned to some path P then the corresponding current transit times coincide. Note that replacing transit times with arrival times does not change the meaning of Definition 3.3. In this manner, the term *equivalent arrival times* becomes well-defined.

So far, we have defined path-based flow models but we do not know how the flow arrives at the sink. In particular, for a path P we are interested in defining an *outflow function* F_P^- with the following meaning. Given a path-based flow $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$ and an arrival time function $\ell_P(\mathcal{F}^+)$, the value $F_P^-(\theta)$ shall be equal to the amount of flow which has arrived at t along P until time θ. Or equivalently, $F_P^-(\theta)$ results out of the flow particles arriving at the sink t over the time interval $[0,\theta]$. Hence, we have to amount the flow entering P over $\ell_P(\mathcal{F}^+)^{-1}([0,\theta])$. Of course, this shows that the outflow function depends on $\mathcal{F}^+$. In this sense, we think of F_P^- as a map defining an outflow function $F_P^-(\mathcal{F}^+) : \mathbb{R}_+ \to \mathbb{R}_+$. This leads to:

$$F_P^-(\mathcal{F}^+)(\theta) \;:=\; \int_{\ell_P(\mathcal{F}^+)^{-1}([0,\theta])} f_P^+(\vartheta)\, d\vartheta \qquad \forall \theta \in \mathbb{R}_+ \;. \tag{3.1}$$

Since $\ell_P(\mathcal{F}^+)$ is Lebesgue measurable, $\ell_P^{-1}([0,\theta])$ is a Lebesgue measurable set implying that $F_P^-(\mathcal{F}^+)(\theta)$ exists for all points in time $\theta \in \mathbb{R}_+$. Further, the nonnegativity of f_P^+ ensures that $F_P^-(\mathcal{F}^+)$ is nonnegative as well. In the following, we often denote the outflow function simply by F_P^- instead of $F_P^-(\mathcal{F}^+)$ if $\mathcal{F}^+$ is clear from context.

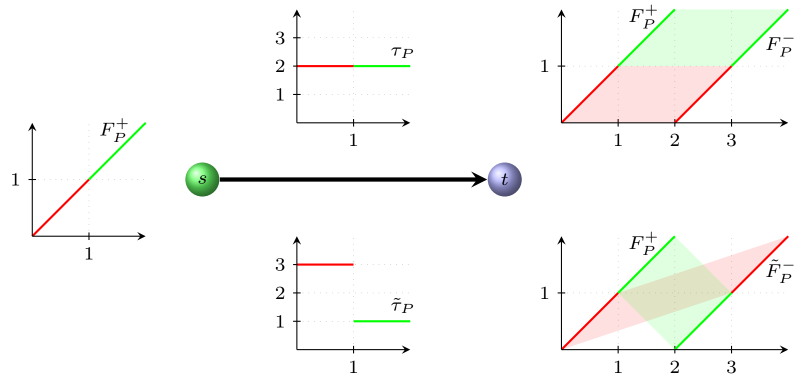


Figure 3.2: The network and the flow behavior of Example 3.5. The red color corresponds to flow entering $P := st$ over the time interval $[0, 1)$ and the green color corresponds to flow entering P over the time interval $[1, 2)$. In the upper scenario no flow particle overtakes any other flow particle whereas in the lower scenario any green flow particle overtakes all red flow particles. Although the flow behavior in both scenarios is completely different, the outflow functions coincide.

Motivated by real world applications, it is worth to know how much flow enters t over all s-t-paths. A flow pattern shows how the particles of a flow over time arrive at the sink t. Intuitively, it could be interpreted as the value over time of a flow over time.

Definition 3.4 (Flow pattern)**.** Let $\mathcal{F}^+$ be a path-based flow over time and, for each path P, let F_P^- be the outflow function of P. Then the *flow pattern* $\mathrm{val}(\mathcal{F}^+) : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as the sum of the outflow functions over all paths P, i.e., $\mathrm{val}(\mathcal{F}^+) := \sum_{P \in \mathcal{P}} F_P^-$.

The value $\mathrm{val}(\mathcal{F}^+)(\theta)$ equals the amount of flow which has arrived at the sink t until time θ. Hence, the flow pattern is of particular importance in evacuation scenarios. In such a case, $\mathrm{val}(\mathcal{F}^+)(\theta)$ represents the number of people reaching the safe point t until time θ. Of course, if θ determines the emergency time, it is worth to identify a flow over time $\mathcal{F}^+$ which maximizes $\mathrm{val}(\mathcal{F}^+)(\theta)$. In this sense, the flow pattern serves as an obvious objective function for flows over time.

Before we conclude this section by establishing basic properties of F_P^-, we derive another characterization of equivalent flow models using outflow functions. The following example motivates the presented approach.

Example 3.5. Consider the simple network shown in Figure 3.2 consisting only of one edge st and let P be the unique s-t-path containing this edge. Next we consider two flow models where the transit time functions τ_P and $\tilde{\tau}_P$ are flow-independent. The first model assigns a constant transit time τ_P of value 2 to P, i.e., $\tau_P(\theta) := 2$ for all $\theta \in \mathbb{R}_+$. In contrast, the transit time $\tilde{\tau}_P$ of P in the second flow model equals 3 over the time interval $[0, 1)$ and 1 over $[1, 2)$, i.e., $\tilde{\tau}_P := 3\chi_{[0,1)} + 1\chi_{[1,2)}$.

Assume that flow enters P at a rate of 1 over the time interval $[0, 2)$, i.e., $f_P^+ := \chi_{[0,2)}$. Hence, concerning the first model, flow arrives at t at a rate of 1 over the time interval $[2, 4)$ implying $f_P^- = \chi_{[2,4)}$. Considering the second model, flow behaves as follows. As flow assigned to P over the time interval $[0, 1)$ arrives at the sink three time units later, this flow enters t at a rate of 1 over the interval $[3, 4)$. Similarly, flow entering P over $[1, 2)$ arrives at t at a rate of 1 over $[2, 3)$. Thus, a flow particle which starts traversing P over $[1, 2)$ overtakes any flow particle entering P over $[0, 1)$. Summarizing, for the second flow model we obtain $\tilde{f}_P^- := \chi_{[3,4)} + \chi_{[2,3)} = \chi_{[2,4)}$. Since this shows $f_P^- = \tilde{f}_P^-$, we are not able to distinguish between these two flow models if we only consider the entire outflow of P.

Example 3.5 shows that we lose information if we consider the entire outflow. On the other hand, the two flow models are recognizable as different if we consider the outflow which is caused by the flow particles originating until time 1. For this restricted flow the outflow rate functions are $\chi_{[2,3)}$ for the first model and $\chi_{[3,4)}$ for the second model, which are of course different. Formalizing this observation for an arbitrary flow over time $\mathcal{F}^+$, we define the *restricted outflow function* $F_P^-|_{\leq\theta}(\mathcal{F}^+) : \mathbb{R}_+ \to \mathbb{R}_+$ of P until time θ as

$$F_P^-|_{\leq\theta}(\mathcal{F}^+)(\theta') \;:=\; \int_{\ell_P(\mathcal{F}^+)^{-1}([0,\theta'])} f_P^+|_{[0,\theta]}(\vartheta)\, d\vartheta \qquad \forall \theta' \in \mathbb{R}_+ \;. \tag{3.2}$$

The main difference to the definition of the entire outflow function in (3.1) is that the integral is taken over the restricted inflow rate $f_P^+|_{[0,\theta]}$ instead of f^+. In particular, $\ell_P(\mathcal{F}^+)$ is still the arrival time function of P with respect to $\mathcal{F}^+$ and *not* with respect to the restricted flow $\mathcal{F}^+|_{\leq\theta}$. Further, note that, in general, the restricted outflow function $F_P^-|_{\leq\theta}(\mathcal{F}^+)$ is not equal to $F_P^-(\mathcal{F}^+)$ restricted to $[0, \theta]$ as it is the case for the restricted inflow $F_P^+|_{\leq\theta}$. Using restricted outflows, equivalent flow models are characterizable as follows.

Lemma 3.6. *Let F^+ be the flow on some path $P \in \mathcal{P}$ and let ℓ and $\tilde{\ell}$ be two arrival time functions. The outflow function on P with respect to ℓ and $\tilde{\ell}$ is denoted by F^- and $\tilde{F}^-$. Then ℓ and $\tilde{\ell}$ are equivalent if and only if*

$$F^-|_{\leq\theta} \;=\; \tilde{F}^-|_{\leq\theta} \tag{3.3}$$

holds for all points in time $\theta \in \mathbb{R}_+$.

Proof. Firstly, assume that ℓ and $\tilde{\ell}$ are equivalent and consider the set of points in time $\Theta := \{\theta \in \mathbb{R}_+ \mid \ell(\theta) \neq \tilde{\ell}(\theta)\}$ where the two arrival time functions differ. This implies $\ell^{-1}([0, \theta'])\backslash\Theta = \tilde{\ell}^{-1}([0, \theta'])\backslash\Theta$. Hence, because of $\int_\Theta f^+(\vartheta)\, d\vartheta = 0$ we obtain from (3.2)

$$\begin{aligned} F^-|_{\leq\theta}(\theta') \;&=\; \int_{\ell^{-1}([0,\theta'])\backslash\Theta} f^+|_{[0,\theta]}(\vartheta)\, d\vartheta \\ &=\; \int_{\tilde{\ell}^{-1}([0,\theta'])\backslash\Theta} f^+|_{[0,\theta]}(\vartheta)\, d\vartheta \;=\; \tilde{F}^-|_{\leq\theta}(\theta') \;. \end{aligned}$$

For proving the other direction, assume the opposite, i.e., $\int_\Theta f^+(\vartheta)\, d\vartheta > 0$. Further, we partition $\Theta := \Theta_1 \dot\cup \Theta_2$ into two sets depending on whether ℓ is greater than $\tilde{\ell}$ or not. More precisely, we set $\Theta_1 := \{\theta \in \mathbb{R}_+ \mid \ell(\theta) > \tilde{\ell}(\theta)\}$ and $\Theta_2 := \{\theta \in \mathbb{R}_+ \mid \ell(\theta) < \tilde{\ell}(\theta)\}$. Then the integral of f^+ over at least one of these two sets must be positive. Without loss of generality we assume $\int_{\Theta_1} f^+(\vartheta)\, d\vartheta > 0$.

Next, we show the existence of a $\theta' \in \mathbb{R}_+$ such that the integral of f^+ over the set $\Theta' := \{\theta \in \mathbb{R}_+ \mid \ell(\theta) > \theta' \geq \tilde{\ell}(\theta)\}$ is positive. Note that $\Theta' \subseteq \Theta_1$ for all $\theta' \in \mathbb{R}$. On the other hand, since between two different real numbers there is also a rational number, we know:

$$\Theta_1 = \bigcup_{\theta' \in \mathbb{Q}} \left\{ \theta \in I \;\middle|\; \ell(\theta) > \theta' \geq \tilde{\ell}(\theta) \right\} .$$

Since $\int_{\Theta_1} f^+(\vartheta)\, d\vartheta > 0$ holds, the integral of f^+ over at least one set of the countable union must be positive as desired. Let θ' be such a point in time and define

$$\epsilon \;:=\; \int_{\Theta'} f^+(\vartheta)\, d\vartheta > 0 .$$

By definition Θ' contains exactly the origination times in Θ_1 for which the current arrival time is less than θ' with respect to $\tilde{\ell}$ and larger than θ' respect to ℓ. That is,

$$\left(\tilde{\ell}^{-1}([0,\theta']) \setminus \ell^{-1}([0,\theta'])\right) \;\cap\; \Theta_1 \;=\; \Theta' .$$

Since $\ell^{-1}([0,\theta']) \cap \Theta_1 \subseteq \tilde{\ell}^{-1}([0,\theta']) \cap \Theta_1$ by the definition of Θ_1, this shows

$$\int_{\tilde{\ell}^{-1}([0,\theta'])\cap\Theta_1} f^+(\vartheta)\, d\vartheta - \int_{\ell^{-1}([0,\theta'])\cap\Theta_1} f^+(\vartheta)\, d\vartheta = \int_{\Theta'} f^+(\vartheta)\, d\vartheta = \epsilon . \tag{3.4}$$

As Θ_1 and Θ_2 are disjoint by definition, we know $\int_{\Theta_1} f^+|_{\Theta_2}(\vartheta)\, d\vartheta = 0$. Hence, Proposition 2.18 ensures the existence of an open set O containing Θ_1 such that

$$\int_O f^+|_{\Theta_2}(\vartheta)\, d\vartheta \leq \frac{\epsilon}{2} . \tag{3.5}$$

Writing $O := \bigcup_{i\in\mathbb{N}}(a_i, b_i)$ as the countable union of disjoint open intervals, this leads to

$$\begin{aligned}
0 &\overset{(3.3)}{=} \sum_{i\in\mathbb{N}} \Big(\big(F^-|_{\leq b_i}(\theta') - \tilde{F}^-|_{\leq b_i}(\theta') \big) - \big(F^-|_{\leq a_i}(\theta') - \tilde{F}^-|_{\leq a_i}(\theta') \big) \Big) \\
&\overset{(3.2)}{=} \sum_{i\in\mathbb{N}} \Big(\int_{\ell^{-1}([0,\theta'])} f^+|_{(a_i,b_i)}(\vartheta)\, d\vartheta - \int_{\tilde{\ell}^{-1}([0,\theta'])} f^+|_{(a_i,b_i)}(\vartheta)\, d\vartheta \Big) \\
&= \int_{\ell^{-1}([0,\theta'])\cap O} f^+(\vartheta)\, d\vartheta - \int_{\tilde{\ell}^{-1}([0,\theta'])\cap O} f^+(\vartheta)\, d\vartheta \\
&= \int_{\ell^{-1}([0,\theta'])\cap O} f^+|_{\Theta_2}(\vartheta)\, d\vartheta - \int_{\tilde{\ell}^{-1}([0,\theta'])\cap O} f^+|_{\Theta_2}(\vartheta)\, d\vartheta \\
&\quad + \int_{\ell^{-1}([0,\theta'])\cap O} f^+|_{\Theta_1}(\vartheta)\, d\vartheta - \int_{\tilde{\ell}^{-1}([0,\theta'])\cap O} f^+|_{\Theta_1}(\vartheta)\, d\vartheta \\
&\overset{(3.5)}{\leq} \frac{\epsilon}{2} + \int_{\ell^{-1}([0,\theta'])} f^+|_{\Theta_1}(\vartheta)\, d\vartheta - \int_{\tilde{\ell}^{-1}([0,\theta'])} f^+|_{\Theta_1}(\vartheta)\, d\vartheta \\
&\overset{(3.4)}{=} \frac{\epsilon}{2} - \epsilon \;<\; 0
\end{aligned}$$

where the third equality sign is implied by the definition of O and the fourth by the definitions of Θ_1 and Θ_2. Further, the nonnegativity of f^+ is also needed for obtaining the "$\leq$"-sign. However, this is a contradiction which finishes the proof. $\square$

Lemma 3.6 shows that flows over time on a network are completely determined by $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-(\mathcal{F}^+))$ where $\mathcal{F}^- = (F_P^-|_{\leq\theta}(\mathcal{F}^+))_{P\in\mathcal{P},\theta\in\mathbb{R}_+}$ is the entire family of restricted outflow functions. Therefore, we use the term *flow over time* also for the whole $\mathcal{F}$ representing, in addition to $\mathcal{F}^+$, the entire flow behavior. But mainly, we use the short notation $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-)$.

We conclude this section by establishing basic properties of the outflow as well as of the restricted outflow. Among others, we show that outflow functions are absolutely continuous. Hence, the outflow of a path can be equivalently defined via Lebesgue integrable outflow rate functions which should be an obvious feature of continuous flow over time models. Here, we use the fact that each transit time function τ_P is compatible with respect to F_P^+ as required in Definition 3.2.

Proposition 3.7. *Let F^+ be the inflow of some path $P \in \mathcal{P}$ and let ℓ an arrival time function. Further, let F^- be the corresponding outflow function. Then the following statements hold.*

(i) The outflow function F^- is bounded by F^+, i.e, $F^+ \geq F^-$ and, vice versa, we have $\lim_{\theta'\to\infty} F^-(\theta') \geq F^+(\theta)$ for all $\theta \in \mathbb{R}_+$.

(ii) The function F^- has the same properties as F^+, i.e., it is nonnegative, absolutely continuous, nondecreasing, and satisfies $F^-(0) = 0$.

Proof. In order to prove (i), consider a point in time $\theta \in \mathbb{R}_+$. Since transit time functions are nonnegative by definition, we know that $\ell(\theta') \geq \theta'$ holds for all $\theta' \in \mathbb{R}_+$ implying $\ell^{-1}([0,\theta]) \subseteq [0,\theta]$. By the definition of F^- in (3.1) this shows

$$F^+(\theta) \;=\; \int_{[0,\theta]} f^+(\vartheta)\, d\vartheta \;\geq\; \int_{\ell^{-1}[0,\theta]} f^+(\vartheta)\, d\vartheta \;=\; F^-(\theta)\;. \tag{3.6}$$

Further, transit time functions are defined to be locally bounded, i.e., there exists an $M \in \mathbb{R}_+$ such that $\ell(\theta') \leq \theta' + M$ holds for all $\theta' \in [0,\theta]$. As this implies $[0,\theta] \subseteq \ell^{-1}([0,\theta+M])$, we obtain from (3.1)

$$F^-(\theta+M) \;\geq\; \int_{[0,\theta]} f^+(\vartheta)\, d\vartheta \;=\; F^+(\theta)\;.$$

Since F^- is nondecreasing as shown in the second part of this proof, this leads to $\lim_{\theta'\to\infty} F^-(\theta') \;\geq\; F^+(\theta)$ for all $\theta \in \mathbb{R}_+$ and statement (i) is established.

For proving (ii) first observe by (3.1) that F^- is nonnegative as f^+ is nonnegative. In addition, this implies $0 \leq F^-(0) \leq F^+(0)$ because of (3.6) and we obtain $F^-(0) = 0$ as we have $F^+(0) = 0$.

To see that F^- is nondecreasing, note that $\ell^{-1}([0,\theta'])\subseteq\ell^{-1}([0,\theta])$ holds for all $\theta > \theta'$. Therefore, we obtain from (3.1)

$$F^-(\theta') \;=\; \int_{\ell^{-1}[0,\theta']} f^+(\vartheta)\, d\vartheta \;\leq\; \int_{\ell^{-1}[0,\theta]} f^+(\vartheta)\, d\vartheta \;=\; F^-(\theta)$$

for all $\theta > \theta'$. It remains to show that F^- is locally absolutely continuous. Since $\ell^{-1}([0,\theta]) \subseteq [0,\theta]$ holds for all $\theta \in \mathbb{R}_+$ as already mentioned above, this follows directly from Lemma 2.23. Note that ℓ is compatible with F^+ by the definition of a flow over time model. $\square$

Comparing the definition of outflow functions with the definition of restricted outflow functions, we see that all statements of Proposition 3.7 remain valid for restricted outflows. In particular, statement (i) ensures that all flow which is assigned to some path until some point in time also leaves this path in finite time. Concluding this section, we show additional basic properties of restricted outflow functions.

Proposition 3.8. *Let F^+ be the inflow of some path $P \in \mathcal{P}$ and let ℓ an arrival time function. Further, let F^- be the corresponding outflow function and $F^-|_{\leq\theta}$ be the restricted outflow function until a point in time $\theta \in \mathbb{R}_+$. Then we have*

$$\begin{aligned} 0 \;\leq\; F^-|_{\leq\theta_2} - F^-|_{\leq\theta_1} \;&\leq\; F^+(\theta_2) - F^+(\theta_1) && \forall \theta_1 \leq \theta_2 \\ \textit{and} \qquad F^-|_{\leq\theta}(\theta') \;&\leq\; \min\bigl\{F^-(\theta'), F^+(\theta)\bigr\} && \forall \theta, \theta' \in \mathbb{R}_+ \\ \textit{and} \qquad F^-|_{\leq\theta}(\theta') \;&=\; F^-(\theta') && \forall \theta \geq \theta' \end{aligned}$$

Proof. Let $\theta' \in \mathbb{R}_+$ be some point in time. To see that the first inequality chain holds, we get from (3.2) and the nonnegativity of f^+

$$\begin{aligned} &F^-|_{\leq\theta_2}(\theta') - F^-|_{\leq\theta_1}(\theta') \\ &\qquad = \int_{\ell_P^{-1}([0,\theta'])} f_P^+|_{[0,\theta_2]}(\vartheta)\, d\vartheta - \int_{\ell_P^{-1}([0,\theta'])} f_P^+|_{[0,\theta_1]}(\vartheta)\, d\vartheta \\ &\qquad = \int_{\ell_P^{-1}([0,\theta'])} f_P^+|_{[\theta_1,\theta_2]}(\vartheta)\, d\vartheta \\ &\qquad \leq \int_{[\theta_1,\theta_2]} f_P^+(\vartheta)\, d\vartheta \;=\; F^+(\theta_2) - F^+(\theta_1)\ . \end{aligned}$$

For verifying the second and the third inequality we observe:

$$F^-|_{\leq\theta}(\theta') \;=\; \int_{\ell_P^{-1}([0,\theta'])} f_P^+|_{[0,\theta]}(\vartheta)\, d\vartheta \;=\; \int_{\ell_P^{-1}([0,\theta'])\cap[0,\theta]} f_P^+(\vartheta)\, d\vartheta\ .$$

Hence, the second inequality follows directly from $F^+(\theta) = \int_{[0,\theta]} f_P^+(\vartheta)\, d\vartheta$ and from $F^-(\theta') = \int_{\ell_P^{-1}([0,\theta'])} f_P^+(\vartheta)\, d\vartheta$ which hold by definition. Further, the nonnegativity of the transit time functions imply $\ell_P^{-1}([0,\theta']) \subseteq [0,\theta]$ if $\theta \geq \theta'$. Recalling (3.1), this proves the third inequality. □

3.2 Edge-based Formulation

In many cases path transit times are not an explicit component of a network but are rather given via transit times on edges. Beside this, edge-based formulations are a basic feature of flow theory. Consequently, we need edge-based flow representations. In this section we define edge-based flows over times along the same lines as we have done this for path-based representations. In order to follow flow particles through the network, it is fundamental to know how flow enters and leaves edges. Therefore and in contrast to the path-based setting, outflow functions are a already basic ingredient of edge-based flows over time.

Definition 3.9 (Edge-based Flow over Time)**.** Let $G := (V, E)$ be a directed graph. A family of pairs $\mathcal{F} := (F_e^+, F_e^-)_{e\in E}$ is called *edge-based flow over time* if, for all edges $e \in E$, the functions $F_e^+ : \mathbb{R}_+ \to \mathbb{R}_+$ and $F_e^- : \mathbb{R}_+ \to \mathbb{R}_+$ are nondecreasing, absolutely continuous and satisfy $F_e^+(0) = 0$ just as $F_e^-(0) = 0$. For an edge e the functions F_e^+ and F_e^- are called *inflow* and *outflow function* of e, respectively. The two families of inflow and outflow functions are denoted by $\mathcal{F}^+ := (F_e^+)_{e\in E}$ and $\mathcal{F}^- := (F_e^-)_{e\in E}$, respectively. Furthermore, we require that any flow particle leaving an edge is immediately assigned to the next edge, i.e.,

$$\sum_{e\in\delta^-(v)} F_e^- = \sum_{e\in\delta^+(v)} F_e^+ \qquad \forall v \in V \setminus \{s,t\} \tag{3.7}$$

where s is the unique source and t is the unique sink of the network. In addition, flow originates only at s and vanishes only at t, i.e., we require that the two functions

$$-\sum_{e\in\delta^-(s)} F_e^- + \sum_{e\in\delta^+(s)} F_e^+ \qquad \text{and} \qquad \sum_{e\in\delta^-(t)} F_e^- - \sum_{e\in\delta^+(t)} F_e^+$$

are nondecreasing.

For an edge $e \in E$ and a point in time $\theta \in \mathbb{R}_+$, we define the *restricted inflow* $F_e^+|_{\le\theta}$ until time θ as $F_e^+|_{\le\theta} := F_e^+|_{[0,\theta]}$.

In contrast to classical flows over times (see Section 2.5), we forbid storage at intermediate nodes by (3.7). But in fact, this is no limitation because we work with arbitrary flow-dependent transit times as explained below. Further, note that in Section 3.1 the symbol $\mathcal{F}^-$ represents the family of all restricted outflow functions of a path-based flow over time whereas in Definition 3.9 the symbol $\mathcal{F}^-$ represents the family of the *non*restricted outflow functions. However, as explained in Section 3.3, this causes no conflicts for subsequent chapters.

For a point in time $\theta \in \mathbb{R}_+$, the value $F_e^+(\theta)$ equals the amount of flow that has been assigned to e until time θ. In contrast, $F_e^-(\theta)$ equals the amount of flow that has left e until time θ. Since F_e^+ and F_e^- are absolutely continuous, they are antiderivatives of Lebesgue integrable functions $f_e^+ : \mathbb{R}_+ \to \mathbb{R}_+$ and $f_e^- : \mathbb{R}_+ \to \mathbb{R}_+$, respectively. That is, $\frac{dF_e^+}{d\theta}(\theta) = f_e^+(\theta)$ and $\frac{dF_e^-}{d\theta}(\theta) = f_e^-(\theta)$ hold for almost all $\theta \in \mathbb{R}_+$ and $F_e^+(\theta) = \int_0^\theta f_e^+(\vartheta)\, d\vartheta$ and $F_e^-(\theta) = \int_0^\theta f_e^-(\vartheta)\, d\vartheta$ hold for all $\theta \in \mathbb{R}_+$. Thus, the flow $\mathcal{F}$ is completely defined by $(f_e^+, f_e^-)_{e\in E}$ and we use the notation $\mathcal{F} := (f_e^+, f_e^-)_{e\in E}$. For a point in time $\theta \in \mathbb{R}_+$, the values $f_e^+(\theta)$ and $f_e^-(\theta)$ equal the rate at which flow is assigned to and at which flow leaves e at time θ, respectively. In particular, this allows us to redefine the flow conservation condition at s and t by

$$\sum_{e\in\delta^-(s)} f_e^- \le \sum_{e\in\delta^+(s)} f_e^+ \qquad \text{and} \qquad \sum_{e\in\delta^-(t)} f_e^- \ge \sum_{e\in\delta^+(t)} f_e^+ \ .$$

For other nodes the flow conservation constraint 3.7 is equivalent to

$$\sum_{e\in\delta^-(v)} f_e^- = \sum_{e\in\delta^+(v)} f_e^+ \qquad \forall v \in V \setminus \{s,t\}\ .$$

So far, the inflow F_e^+ and the outflow F_e^- of some edge e seem to be relatively unrelated. However, they are strongly connected via an underlying edge-based flow over time model. See Figure 3.3 for an illustration.

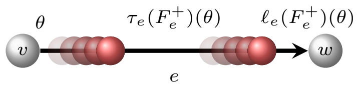


Figure 3.3: A flow particle entering an edge $e = vw$ at time θ arrives after $\tau_e(F_e^+)(\theta)$ time units at the head of e at time $\ell_P(F_e^+)(\theta) = \theta + \tau_e(F_e^+)(\theta)$.

Definition 3.10 (Flow over Time Model)**.** An *edge-based flow over time model* defines a family $\mathcal{T} := (\tau_e)_{e\in E}$ of maps τ_e, $e \in E$ assigning a locally bounded, Lebesgue measurable function $\tau_e(F_e^+) : \mathbb{R}_+ \to \mathbb{R}_+$ to each possible inflow F_e^+. The map τ_e is called *edge transit time* of $e \in E$ and $\tau_e(F_e^+)$ is called *edge transit time function* with respect to F_e^+. The value $\tau_e(F_e^+)(\theta)$ is called *current edge transit time.*

A corresponding family $\mathcal{L} := (\ell_e)_{e\in E}$ of *arrival times* is given by an *arrival time function* $\ell_e(F_e^+) : \mathbb{R}_+ \to \mathbb{R}_+$ which is defined as $\ell_P(F_e^+)(\theta) := \theta + \tau_e(F_e^+)(\theta)$ for each $e \in E$. The value $\ell_e(F_e^+)(\theta)$ is called *current arrival time.*

In a *continuous* flow over time model every transit time function need only be defined almost everywhere on $\mathbb{R}_+$. Further, we require that each transit time function $\tau_e(F_e^+)$ is compatible with F_e^+. That is, the preimage of every null set with respect to $\tau_e(F_e^+)$ is an F_e^+-null set for each edge $e \in E$.

If the inflow F_e^+ is clear from context, we use $\tau_e(\theta)$ and $\ell_e(\theta)$ for denoting the current transit and the current arrival time, respectively. Furthermore, we use τ_e and ℓ_e for the transit and arrival time function, respectively, if there is no name conflict with the corresponding maps.

Note that edge transit times are defined similarly to path transit times with one exception. A path transit time τ_P depends on the entire inflow $\mathcal{F}^+$ whereas the edge transit time τ_e depends only on the inflow F_e^+ on e. From an abstract point of view the flow on an edge can be perceived as a special case of a path flow. But this does not mean that edge-based flow models are easier to handle. As already mentioned, the inflow function F_e^+ and the outflow function F_e^- are linked together via the underlying flow model. As in path-based flow models, we have

$$F_e^-(F_e^+)(\theta) := \int_{\ell_e(F_e^+)^{-1}([0,\theta])} f_e^+(\vartheta)\, d\vartheta \tag{3.8}$$

for the continuous model. But in contrast, this is an explicit constraint which a *feasible* edge-based flow over time on a given edge-based flow over time model has to satisfy. Formally, this constraint could be written as $F_e^- = F_e^-(F_e^+)$ for all edges $e \in E$. Also note that Proposition 3.7(ii) applied to the inflow function of an edge ensures that Definition 3.9 and, in particular, equation (3.7) makes sense for continuous edge-based flow models. Further, (3.8) shows that the outflow of an edge depends on the arrival time of that edge which in turn depends on the inflow. But the inflow itself depends again on the outflow of the predecessor edges. This leads to highly nontrivial dependencies mainly caused by cycles in the underlying graph.

In particular, at a first glance, it is not clear a priory whether or not some flow model supports feasible nonzero flows especially on cycles. This question is of particular importance if we want to decide, whether a path-based flow can be transformed in an equivalent edge-based flow. This problem is known as dynamic network loading problem in literature. The inverse problem, given an edge-based flow, find the path-based representation is called dynamic flow decomposition problem. We briefly consider these two problems in Subsection 3.5 for so-called consistent flow models which are easier to analyze than the very general formulation presented in this section.

Following Section 3.1, equivalent flow models for edge-based representations are defined along the same lines as for path-based flow models.

Definition 3.11 (Equivalent Flow over Time Models)**.** Let τ_1 and τ_2 be transit time functions for some edge $e \in E$ and let $\Theta := \{\theta \in \mathbb{R}_+ \mid \tau_1(\theta) \neq \tau_2(\theta)\}$ be the set of points in time where the two transit time functions differ. Then τ_1 and τ_2 are called *equivalent with respect to some inflow* F_e^+ if and only if essentially no flow is assigned to e over the set Θ, i.e., $\int_\Theta f_e^+(\vartheta)\, d\vartheta = 0$.

Let $\mathcal{T}_1 := (\tau_e^1)_{e\in E}$ and $\mathcal{T}_2 := (\tau_e^2)_{e\in E}$ be the families of transit times of two edge-based flow models. Then these two flow models are called *equivalent* if, for all edges $e \in E$ and for all possible inflows F_e^+, the transit time functions $\tau_e^1(F_e^+)$ and $\tau_e^2(F_e^+)$ are equivalent with respect to F_e^+.

Informally, Definition 3.11 requires that whenever a positive amount of flow enters an edge e, the corresponding current transit times coincide.

Unlike path-based representations the outflow, of an edge is part of the flow over time definition. A flow pattern shows how the particles of a flow over time arrive at the sink t. It could be interpreted as the value over time of a flow over time.

Definition 3.12 (Flow pattern)**.** Let $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-)$ be some edge-based flow over time. Then the flow pattern $\operatorname{val}(\mathcal{F}) := \sum_{e\in\delta^-(t)} F_e^- - \sum_{e\in\delta^+(t)} F_e^+$ is defined as the net inflow in t.

Concluding this section we derive another characterization of equivalent flow models using outflow functions by restating Lemma 3.6 of Section 3.1. For this we need the following definition. The outflow which caused is by the flow particles entering an edge e until time θ is defined by

$$F_e^-(F_e^+)|_{\leq\theta}(\theta') := \int_{\ell_e(F_e^+)^{-1}([0,\theta'])} f_e^+|_{[0,\theta]}(\vartheta)\, d\vartheta\ . \tag{3.9}$$

Note that $\ell_e(F_e^+)$ is still the arrival time function of e with respect to F_e^+ and *not* with respect to the restricted flow $F_e^+|_{\leq\theta}$. Further, the restricted outflow is defined with respect to the flow particles entering e and not with respect to the flow particles originating at s until time θ. Using restricted outflows, equivalent flow models are characterizable as follows.

Lemma 3.13. *Let F_e^+ be the inflow of some edge $e \in E$ and let ℓ and $\tilde{\ell}$ be two arrival time functions. The outflow function on e with respect to ℓ and $\tilde{\ell}$ is denoted by F_e^- and $\tilde{F}_e^-$. Then ℓ and $\tilde{\ell}$ are equivalent if and only if*

$$F_e^-|_{\leq\theta} \;=\; \tilde{F}_e^-|_{\leq\theta}$$

holds for all points in time $\theta \geq 0$.

Proof. This proof is exactly the same as the proof of Lemma 3.6 when replacing P with e. □

As already mentioned, the flow on an edge can be perceived a special case of the flow on a path. Hence, all results and discussions about path-based flows over time are trivially true for the edge-based setting. This is already observed in the proof of Lemma 3.13 and holds in particular for Proposition 3.7 and 3.8. Further, recalling the discussion after Lemma 3.6, Lemma 3.13 shows that an edge-based flow over time on a network is completely determined by $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-(\mathcal{F}^+))$ where $\mathcal{F}^- = (F_e^-|_{\leq\theta}(\mathcal{F}^+))_{e\in E,\theta\in\mathbb{R}_+}$ is the entire family of restricted outflow functions. But in the edge-based setting the symbol $\mathcal{F}^-$ is already reserved for denoting the family of the *non*restricted outflow functions. In this manner, $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-)$ has two meanings depending on whether $\mathcal{F}$ shall represent a flow over time or the flow behavior of the flow particles. However, for consistent flow models this causes no conflict as in this setting the complete flow behavior is already represented by the family of nonrestricted outflow functions. Consistent flow models which provide the basis for subsequent chapters are discussed in the next section.

3.3 Consistent Flow over Time Models

So far, flow models are hardly restricted. However, many real world applications have some basic properties in common which leads to natural limitations of the underlying flow model. We start by discussing the very popular FiFo (**F**irst **i**n **F**irst **o**ut) condition in Subsection 3.3.1. This condition forbids overtaking of flow particles along a path or an edge and enables us to identify nice transit times representing equivalent flow models. However, it turns out that these transit times are not sufficient for defining Nash flows over time in Chapter 4.

Besides FiFo, we study the continuity of flow models in Subsection 3.3.2. It is worth to mention that we do not consider the continuity of transit time or outflow functions with respect to time. Instead, we consider the desired behavior of the flow particles if the inflow is changed slightly.

In Subsection 3.3.3 we define past-oriented flow models. Informally, these are flow models where the current flow behavior depends only on current and past but not on future flow situations. Finally, in Subsection 3.3.4 we impose an assumption on the set $\mathcal{P}$ of s-t-paths in order to avoid pathological flow behavior in path-based flow models. In this case we call $\mathcal{P}$ locally finite.

A flow model which combines all the properties explained in this section is called consistent. In the subsequent sections and chapters of this thesis we always deal with such consistent flow over time models. So the main goal of this section is to explain the following definition.

Definition 3.14 (Consistent Flow over Time Model)**.** A flow over time model (either edge- or path-based) is called *consistent* if it is continuous, satisfies FiFo, and obeys past-orientation. Further, for a consistent path-based flow model the set $\mathcal{P}$ must be locally finite.

All definitions and results in this section are mainly devoted to the flow behavior on a single edge or on a single path. From an abstract point of view, the flow on an edge can be seen as a special case of the flow on a path. This implies that definitions and results for path-based flows over time are directly transferable to the edge-based setting. Therefore, we present our approach only for path-based flow models. However, it is obvious which parts are worth to be restated for edge-based flow models as well.

3.3.1 FiFo - Lower and Upper Arrival Time Functions

Many real world applications are modeled such that overtaking on an edge or on a path is forbidden. Mainly, this is implicitly implied by other simplifying assumptions such as constant speed, increasing transit times, etc., or it is required explicitly. If overtaking is forbidden, flow particles entering an edge or a path earlier also leave this edge or path earlier. The flow behavior in such models is completely described by the family of outflows. In fact, we show that flow models leading to the same family of outflow functions for each given flow over time are equivalent. For this we identify and analyze natural arrival times representing equivalent classes of flow models. Further, these arrival times are used in Subsection 3.4 for defining foresighted arrival times which are even more appropriate for the purposes of this thesis. In fact, they enable us to define Nash equilibria for flows over time in Chapter 4 in a desired manner.

FiFo can be enforced by requiring strictly nondecreasing arrival time functions. The very popular deterministic queuing model (see Chapter 7) does *not* satisfy this assumption. This is verified by the following example motivating the subsequent definition of flow models satisfying FiFo.

Example 3.15. Assume you are in a supermarket and you have all desired products in your shopping cart. The only remaining thing you have to do is to pay. But you see that there is a long waiting queue in front of the cash desk. So you have to make a decision. Either you wait for a moment and look around (for nothing or for seeing what else is available), or you go directly to the end of the waiting queue. In case you look around for nothing, you observe that as long as the last person in the waiting queue remains the same, you always leave the cash desk at the same point in time. Hence, the arrival time remains constant in such a case and is not *strictly* nondecreasing. But while you are thinking, another guy has entered the waiting queue. Now you get angry about your indecision because the time at which you could leave the cash desk has increased. But, surprisingly, the waiting time is decreased. So you are happy and come to the conclusion: Only if a positive amount of flow enters an edge or a path, the arrival time must be strictly increasing.

Example 3.15 motivates the following definition.

Definition 3.16 (FiFo)**.** We say that a path-based flow model *satisfies FiFo (**F**irst **in** **F**irst **out**)* if and only if the arrival time function $\ell_P(\mathcal{F}^+)$ is nondecreasing with respect to F_P^+ for each path P and each path-based flow over time $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$. That is, there exists an F_P^+-supporting set S such that for all $\theta_1, \theta_2 \in S$ we have:

$$F_P^+(\theta_1) < F_P^+(\theta_2) \qquad \Rightarrow \qquad \ell_P(\mathcal{F}^+)(\theta_1) < \ell_P(\mathcal{F}^+)(\theta_2)\ .$$

Intuitively, the FiFo condition ensures that a flow particle entering a path P after a certain point in time θ arrives later than flow particles departing before time θ. Further, flow particles departing before time θ arrive at t as long as the restricted outflow function $F_P^-|_{\leq\theta}$ of P is strictly nondecreasing. Hence, as long as $F_P^-|_{\leq\theta}$ is nondecreasing, no flow particle departing after time θ arrives

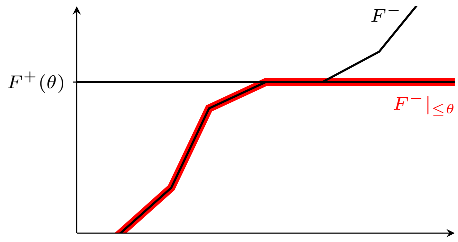


Figure 3.4: Restricted outflow function in a flow model satisfying FiFo.

implying that the outflow F_P^- and $F_P^-|_{\leq\theta}$ coincide for such points in time. On the other hand, if all flow particles departing before time θ have finished their trip along P then $F_P^-|_{\leq\theta}$ must be equal to $F_P^+(\theta)$. Intuitively, this shows that the restricted outflow $F_P^-|_{\leq\theta}(\theta')$ is always equal to $F^-(\theta')$ or to $F^+(\theta)$. By Proposition 3.8 this shows $F_P^-|_{\leq\theta}(\theta') = \min\{F^-(\theta'), F^+(\theta)\}$ which is illustrated in Figure 3.4 and formally proved in the subsequent lemma.

Lemma 3.17. *Consider a continuous flow model and let F^+, F^-, and ℓ be inflow, outflow, and arrival time function of some path P with respect to some flow over time $\mathcal{F}$. Then FiFo is satisfied on P, if and only if*

$$F^-|_{\leq\theta}(\theta') \;=\; \min\bigl\{F^-(\theta'), F^+(\theta)\bigr\} \tag{3.10}$$

holds for all θ and θ' in $\mathbb{R}_+$.

Proof. Assume that FiFo is satisfied, i.e., ℓ is nondecreasing with respect to F^+. By Lemma 2.27 there exists an F^+-supporting set S such that

$$\forall \theta_1, \theta_2 \in S: \qquad \theta_1 < \theta_2 \qquad \Leftrightarrow \qquad \ell(\theta_1) < \ell(\theta_2)\ . \tag{3.11}$$

Next we show that for all $\theta, \theta' \in \mathbb{R}_+$ at least one of the following expressions

$$\ell^{-1}\bigl([0,\theta']\bigr) \cap S \;\subseteq\; [0,\theta] \qquad \text{and} \qquad [0,\theta] \cap S \;\subseteq\; \ell^{-1}\bigl([0,\theta']\bigr) \tag{3.12}$$

is valid. Assume the opposite implying that there exists a $\theta_1 \in \ell^{-1}([0,\theta']) \cap S$ which is not contained in $[0,\theta]$ and a $\theta_2 \in [0,\theta] \cap S$ which is not contained in $\ell^{-1}([0,\theta'])$. Since θ_2 is contained in $[0,\theta']$ and θ_1 not, this shows that $\theta_2 < \theta_1$. On the other hand, we know that $\ell(\theta_1) < \ell(\theta_2)$ as θ_1 is contained in $\ell^{-1}([0,\theta'])$ and θ_2 not. But this is a contradiction to (3.11) as both, θ_1 and θ_2, are contained in S. Hence, at least one expression of (3.12) is valid.

Recalling the definition of restricted outflow functions in (3.2), we know for all $\theta, \theta' \in \mathbb{R}_+$:

$$F^-|_{\leq\theta}(\theta') = \int_{\ell^{-1}([0,\theta'])} f^+|_{[0,\theta]}(\vartheta)\, d\vartheta = \int_{\ell^{-1}([0,\theta'])\cap[0,\theta]} f^+(\vartheta)\, d\vartheta\ .$$

Hence, depending on which relation holds in (3.12), the value $F^-|_{\leq\theta}(\theta')$ is equal to $\int_{\ell^{-1}([0,\theta'])} f^+(\vartheta)\, d\vartheta = F^-(\theta')$ or to $\int_{[0,\theta]} f^+(\vartheta)\, d\vartheta = F^+(\theta)$. Since we know $F^-|_{\leq\theta}(\theta') \leq \min\{F^-(\theta'), F^+(\theta)\}$ by Proposition 3.8, this proves (3.10).

For proving the other direction, consider a $\theta \in \mathbb{R}_+$ and let θ' be minimal with $F^-(\theta') = F^+(\theta)$. Note that such a θ' exists because F^- is continuous and satisfies $F^-(0) \leq F^+(\theta) \leq F^-(\theta + M)$ for some $M \in \mathbb{R}_+$ by Proposition 3.7. Because of (3.10) this shows $F^-|_{\leq\theta}(\theta') = F^-(\theta') = F^+(\theta)$ implying

$$\int_{\ell^{-1}([0,\theta'])} f^+(\vartheta)\, d\vartheta = \int_{\ell^{-1}([0,\theta'])\cap[0,\theta]} f^+(\vartheta)\, d\vartheta = \int_{[0,\theta]} f^+(\vartheta)\, d\vartheta \ .$$

As $\ell^{-1}([0,\theta'])\cap[0,\theta] \subseteq \ell^{-1}([0,\theta'])$, we know that $\ell^{-1}([0,\theta'])\setminus[0,\theta]$ is an F^+-null set. Similarly, we also know that $[0,\theta] \setminus \ell^{-1}([0,\theta'])$ is an F^+-null set. Hence, there exists an F^+-supporting set S_θ such that

$$[0,\theta] \cap S_\theta = \ell^{-1}([0,\theta']) \cap S_\theta \ .$$

Hence, for all $\theta_1 \in S_\theta$ we have $\theta_1 \leq \theta$ if and only if $\ell(\theta_1) \leq \theta'$. Therefore, we get for all $\theta_1, \theta_2 \in S_\theta$

$$\theta_1 \leq \theta < \theta_2 \qquad\qquad \Leftrightarrow \qquad\qquad \ell(\theta_1) \leq \theta' < \ell(\theta_2) \ .$$

Let $S := \bigcap_{\theta\in\mathbb{Q}_+} S_\theta$. Then S is F^+-supporting as it is the countable intersection of F^+-supporting sets. Finally, for all $\theta_1, \theta_2 \in S$ with $\theta_1 < \theta_2$ there exists a $\theta \in \mathbb{Q}_+$ such that $\theta_1 \leq \theta < \theta_2$ and $\theta_1, \theta_2 \in S_\theta$ implying

$$\theta_1 < \theta_2 \qquad\qquad \Leftrightarrow \qquad\qquad \ell(\theta_1) < \ell(\theta_2) \ .$$

Because of Lemma 2.27 this finishes the proof. □

As also suggested by Figure 3.4, Lemma 3.17 shows that the graph of the restricted outflow $F^-|_{\leq\theta}$ until time θ can be constructed as follows. First, follow the graph of the outflow F^- until a value of $F^+(\theta)$ is reached. Then proceed parallel to the time axis. This is possible because F^- is nondecreasing. Further, Lemma 3.17 characterizes the FiFo-principle using outflow functions. In fact, it shows that the flow behavior on a path is completely described by the corresponding inflow and outflow function. Hence, there is no reason for considering restricted outflow functions if the flow model satisfies FiFo. This is verified by the following theorem.

Theorem 3.18. *Let F^+ be the flow on some path $P \in \mathcal{P}$ and let ℓ and $\tilde{\ell}$ be two arrival time functions satisfying FIFO. The outflow on $\mathcal{P}$ with respect to ℓ and $\tilde{\ell}$ is denoted by F^- and $\tilde{F}^-$. Then ℓ and $\tilde{\ell}$ are equivalent if and only if*

$$F^- \ = \ \tilde{F}^- \ . \tag{3.13}$$

Proof. Given a point in time $\theta \in \mathbb{R}_+$, we have $F^-(\theta') = F^-|_{\leq\theta}(\theta')$ for all $\theta' \leq \theta$ by Proposition 3.8. Therefore, increasing θ to ∞, Lemma 3.6 shows that equation (3.13) holds if ℓ and $\tilde{\ell}$ are equivalent.

So assume that (3.13) is valid. Using equation (3.10) of Lemma 3.17 we get for all $\theta \in \mathbb{R}_+$

$$\begin{aligned} F^-|_{\leq\theta}(\theta') \ &= \ \min\{F^-(\theta'), F^+(\theta)\} \\ &= \ \min\{\tilde{F}^-(\theta'), F^+(\theta)\} \ = \ \tilde{F}^-|_{\leq\theta}(\theta') \qquad \forall \theta' \in \mathbb{R}_+ \end{aligned}$$

because of (3.13). Hence, the equivalence of ℓ and $\tilde{\ell}$ follows from Lemma 3.6. □

By Theorem 3.18 flow models satisfying FiFo are completely defined by the family of outflow functions. In the following, we analyze the flow behavior on a given path $P \in \mathcal{P}$ with respect to transit and arrival times. Firstly, we give another characterization of the FiFo-principle in Lemma 3.19, which can be used for identifying equivalent arrival times. Thereafter, we look for a nice representative among such equivalent arrival times.

To get an idea, let F^+ and F^- be the inflow and outflow function of P with respect to some path-based flow over time $\mathcal{F}$, respectively. Consider a flow particle entering P at a certain point in time $\theta \in \mathbb{R}_+$ and let $\ell(\theta)$ be the current arrival time. Intuitively, FiFo ensures that all flow particles entering P before time θ arrive at t before time $\ell(\theta)$ implying $F^+(\theta) \leq F^-(\ell(\theta))$. In addition, FiFo guarantees that no particle departing after time θ arrive before time $\ell(\theta)$ implying $F^+(\theta) \geq F^-(\ell(\theta))$. Hence, intuitively, we get $F^+(\theta) = F^-(\ell(\theta))$. The following lemma shows that the FiFo-principle is, in fact, equivalent to this equation. Furthermore, it shows that a function ℓ satisfying this equation, can be seen as an arrival time function for this flow model. In particular, we have to verify (3.1) for this.

Lemma 3.19. *Consider a path P. Further, let F^+ and F^- be the inflow and the outflow function of P, respectively. If the FiFo principle is satisfied, there exists an F^+-supporting set S such that*

$$F^+(\theta) = F^-\big(\ell(\theta)\big) \qquad\qquad \forall \theta \in S \tag{3.14}$$

where ℓ is the corresponding arrival time function of P computed by the underlying flow model.

On the other hand, if an arbitrary arrival time function ℓ satisfies (3.14)*, FiFo holds with respect to ℓ. In particular, we have that F^- is the outflow function with respect to ℓ defined by* (3.1)*.*

Proof. Since ℓ is nondecreasing with respect to the inflow function F^+, there exists an F^+-supporting set S such that

$$\forall \theta_1, \theta_2 \in S : \qquad \theta_1 < \theta_2 \qquad \Leftrightarrow \qquad \ell(\theta_1) < \ell(\theta_2) \tag{3.15}$$

by Lemma 2.27. We show that (3.14) is satisfied on S. Using the definition of the outflow in (3.1), we know because S is F^+-supporting:

$$F^-\big(\ell(\theta)\big) = \int_{\ell^{-1}([0,\ell(\theta)])} f^+(\vartheta)\, d\vartheta = \int_{\ell^{-1}([0,\ell(\theta)])\cap S} f^+(\vartheta)\, d\vartheta \;.$$

Further, because of (3.15), we know $\ell^{-1}([0,\ell(\theta)]) \cap S = [0,\theta] \cap S$ if $\theta \in S$. Hence, using again that S is F^+-supporting, we obtain for all $\theta \in S$:

$$F^-\big(\ell(\theta)\big) = \int_{[0,\theta]\cap S} f^+(\vartheta)\, d\vartheta = F^+(\theta) \;.$$

The other direction follows directly from the fact that F^+ and F^- are nondecreasing. To see this assume that (3.14) is satisfied for all θ contained in some F^+-supporting set S_1 and let S_2 be the monotonicity set of F which is F^+-supporting by Proposition 2.25. We show that (3.15) holds for the set $S := S_1 \cap S_2$, which is, of course, F^+-supporting. So consider some point in times $\theta_1, \theta_2 \in S$. Since F^+ is strictly increasing on S_2 by Proposition 2.25, we have $\theta_1 < \theta_2$ if and only if $F^+(\theta_1) < F^+(\theta_2)$ is valid. Because of (3.14) we get $F^-\big(\ell(\theta_1)\big) < F^-\big(\ell(\theta_2)\big)$ if and only if $F^+(\theta_1) < F^+(\theta_2)$. This shows $\theta_1 < \theta_2$ if and only if $\ell(\theta_1) < \ell(\theta_2)$ as F^- is nondecreasing. Thus, (3.15) is valid and this part is done by Lemma 2.27.

It remains to show that F^- is the outflow function with respect to ℓ defined by (3.1), i.e., we have

$$F^-(\theta') \;=\; \int_{\ell^{-1}([0,\theta'])} f^+(\vartheta)\, d\vartheta \qquad\qquad \forall \theta' \in \mathbb{R}_+ \,.$$

Consider some $\theta' \in \mathbb{R}_+$ with $0 < F^-(\theta') < F^+(\infty)$ and let $\epsilon > 0$ be a positive real number. Further, let S be an F^+-supporting set such that (3.14) and (3.15) hold. As S is F^+-supporting, there exist $\theta_1, \theta_2 \in S$ such that

$$F^-\big(\ell(\theta_1)\big) \;<\; F^-(\theta') \;<\; F^-\big(\ell(\theta_2)\big) \;\leq\; F^-\big(\ell(\theta_1)\big) + \epsilon \tag{3.16}$$

because of (3.14). Since F^- is nondecreasing, $\ell(\theta_1) < \theta' < \ell(\theta_2)$ holds implying

$$\int_{\ell^{-1}([0,\ell(\theta_1)])} f^+(\vartheta)\, d\vartheta \;\leq \int_{\ell^{-1}([0,\theta'])} f^+(\vartheta)\, d\vartheta \;\leq \int_{\ell^{-1}([0,\ell(\theta_2)])} f^+(\vartheta)\, d\vartheta \,. \tag{3.17}$$

As in the first part of this proof, we know $\ell^{-1}([0, \ell(\theta)]) \cap S = [0, \theta] \cap S$ if $\theta \in S$ because of (3.15) proving

$$F^-\big(\ell(\theta_1)\big) \;=\; \int_{\ell^{-1}([0,\ell(\theta_1)])} f^+(\vartheta)\, d\vartheta \tag{3.18}$$

and

$$F^-\big(\ell(\theta_2)\big) \;=\; \int_{\ell^{-1}([0,\ell(\theta_2)])} f^+(\vartheta)\, d\vartheta \tag{3.19}$$

because of (3.14). Identifying equal terms in (3.16) and (3.17) using (3.18) and (3.19) shows

$$\Big| F^-(\theta') - \int_{\ell^{-1}([0,\theta'])} f^+(\vartheta)\, d\vartheta \Big| \leq \epsilon \,.$$

Finally, let ϵ tend to 0 and equation (3.1) is established for every $\theta' \in \mathbb{R}_+$ satisfying $0 < F^-(\theta') < F^+(\infty)$. If $F^-(\theta')$ is equal to one of its bounds, we use the same lines of arguments as follows. In case of $F^-(\theta') = 0$ and $\theta' > 0$ we set $\theta_1 := 0$ and assume without loss of generality $\ell(0) := 0$. In case of $F^-(\theta') = F^+(\infty)$ we set $\theta_2 := \infty$ implying $\ell(\theta_2) = \infty$. Finally, note that the case $\theta' = 0$ is trivially satisfied because $\ell^{-1}(0) \subseteq \{0\}$. □

Recalling Theorem 3.18, Lemma 3.19 shows that all arrival time functions satisfying (3.14) are equivalent provided that the underlying flow model satisfies FiFo. However, equation (3.14) only has to hold on an F^+-supporting set and is, therefore, a little bit unhandy. The situation would be much better if there would exist an arrival time function satisfying (3.14) for all nonnegative points in time. Identifying such nice representatives among equivalent arrival time functions is the motivation of the following definition which is illustrated in Figure 3.5.

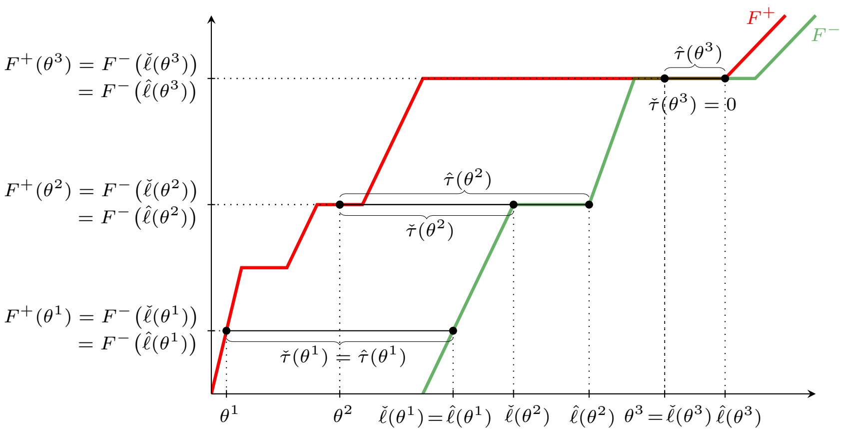


Figure 3.5: Lower and upper arrival time function.

Definition 3.20 (Lower and upper arrival times)**.** Let $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$ be some path-based flow over time of a model satisfying FiFo. For each path $P \in \mathcal{P}$ let F_P^- be the corresponding outflow function. Then the *lower* and the *upper arrival time function* of P are defined by

$$\check{\ell}_P(\theta) := \min\{\vartheta \geq \theta \mid F_P^+(\theta) \leq F_P^-(\vartheta)\} \tag{3.20}$$

and

$$\hat{\ell}_P(\theta) := \max\{\vartheta \mid F_P^+(\theta) \geq F_P^-(\vartheta)\}, \tag{3.21}$$

respectively.

First we observe that, given a path P, current lower and upper arrival times are well-defined. To see this, look at Proposition 3.7 which shows that F_P^- is continuous, and satisfies $F_P^-(0) \leq F_P^+(\theta) \leq F_P^-(\theta + M)$ for some $M \in \mathbb{R}_+$. This implies that current lower arrival times always exist and are finite whereas current upper arrival times can be equal to ∞.

Moreover, the current lower arrival time $\check{\ell}_P(\theta)$ is defined such that $\theta \leq \check{\ell}_P(\theta)$ holds for all $\theta \in \mathbb{R}_+$. For the current upper arrival time, $\theta \leq \hat{\ell}_P(\theta)$ follows from $F_P^+ \geq F^-$. In this sense, the corresponding transit time functions are nonnegative as required by Definition 3.2. Further, since F_P^+ just as F_P^- are continuous, we know

$$F^-\big(\check{\ell}(\theta)\big) = F^+(\theta) \qquad \text{and} \qquad F^-\big(\hat{\ell}(\theta)\big) = F^+(\theta) \qquad \forall \theta \in \mathbb{R}_+ \,. \tag{3.22}$$

Before we analyze lower and upper arrival times, we address the following question. Both $\check{\ell}$ and $\hat{\ell}$ are defined on an existent flow model. Of course, $\check{\ell}$ and $\hat{\ell}$ shall represent this flow model. More precisely, if ℓ is the arrival time of the flow model, we have to show that ℓ, $\check{\ell}$, and $\hat{\ell}$ are pairwise equivalent in terms of Definition 3.3. Because of (3.22), Lemma 3.19 shows that $\check{\ell}$ and $\hat{\ell}$ imply the same outflow functions as ℓ. Hence, all three arrival time functions are equivalent because of Theorem 3.18. We conclude this subsection by presenting some basic properties of $\check{\ell}$ and $\hat{\ell}$.

Using lower and upper arrival times, we are able to identify all equivalent arrival times satisfying (3.14) everywhere on $\mathbb{R}_+$. In addition, the following proposition shows which arrival time functions simplify the notion of Definition 2.26 which provides the basis for the FiFo-definition.

Lemma 3.21. *Consider a flow over time $\mathcal{F}$ satisfying FiFo and let $\check{\ell}$ and $\hat{\ell}$ be the lower and upper arrival time function of some path $P \in \mathcal{P}$, respectively. An arrival time function ℓ satisfies* (3.14) *for all nonnegative points in time if and only if $\check{\ell} \leq \ell \leq \hat{\ell}$. In both cases ℓ is equivalent to $\check{\ell}$ (and $\hat{\ell}$). Moreover, we have*

$$F^-|_{\leq \theta}(\theta') = \begin{cases} F^-(\theta') & \text{if } \theta' < \ell(\theta) \\ F^+(\theta) & \text{if } \theta' \geq \ell(\theta) \end{cases} . \tag{3.23}$$

Proof. Let F^+ and F^- be the inflow and outflow functions of P. Assume that ℓ satisfies equation (3.14) everywhere on $\mathbb{R}_+$, i.e., $F^+(\theta) = F^-(\ell(\theta))$ holds for all $\theta \in \mathbb{R}_+$. Since ℓ is an arrival time function, we know $\ell(\theta) > \theta$ for all $\theta \in \mathbb{R}_+$. Hence, $\check{\ell}(\theta) \leq \ell(\theta) \leq \hat{\ell}(\theta)$ follows for all $\theta \in \mathbb{R}_+$ directly from Definition 3.20.

On the other hand, let ℓ satisfy $\check{\ell}(\theta) \leq \ell(\theta) \leq \hat{\ell}(\theta)$ for some point in time θ. Because of $F^-(\check{\ell}(\theta)) = F^-(\hat{\ell}(\theta))$ we know $F^-(\ell(\theta)) = F^-(\check{\ell}(\theta))$ as F^- is nondecreasing. Hence, by equation (3.22) we obtain $F^+(\theta) = F^-(\ell(\theta))$ for all $\theta \in \mathbb{R}_+$.

Since in both cases we have that $\check{\ell}$ just as ℓ satisfy (3.14), Lemma 3.19 shows that F^- is the outflow function with respect to $\check{\ell}(\theta)$ and with respect to ℓ. Thus, Theorem 3.18 implies that $\check{\ell}$ and ℓ are equivalent.

For proving equation (3.23) we use Lemma 3.17 which states:

$$F^-|_{\leq \theta}(\theta') = \max\{F^-(\theta'), F^+(\theta)\} .$$

Therefore, if $F^-(\theta') < F^+(\theta)$ is valid, this shows that $F^-|_{\leq \theta}(\theta') = F^-(\theta')$ and $F^-(\theta') < F^-(\ell(\theta))$ hold by (3.14). Since we have $\theta' < \ell(\theta)$ in this case, equation (3.23) holds. If $F^-(\theta') = F^+(\theta)$ is valid, (3.23) holds trivially. Finally, for $F^-(\theta') > F^+(\theta)$, we obtain $F^-|_{\leq \theta}(\theta') = F^+(\theta)$ and $F^-(\theta') > F^-(\ell(\theta))$ because of (3.14). Hence, in this case we have $\theta' > \ell(\theta)$ implying (3.23). $\square$

Lemma 3.21 identifies a nice subclass of equivalent arrival times where Definition 3.16 holds everywhere on $\mathbb{R}_+$ and not only on some supporting set. However, a basic intuition behind FiFo is that flow particles departing earlier also arrive earlier. In this sense, FiFo would imply that arrival time functions are nondecreasing. Unfortunately, not every arrival time function identified by Lemma 3.21 is nondecreasing in general. This is especially the case if $\check{\ell}$ is constant over some time interval and differs from $\hat{\ell}$. See, e.g., Figure 3.5 and consider θ_2. It is not hard to see that an arrival time function satisfying Lemma 3.21 need to be nondecreasing around θ_2.

The discussion above shows that we are, mainly, interested in nondecreasing arrival time functions. Note that nondecreasing arrival time functions lead directly to a flow model satisfying FiFo by Lemma 2.28 as they have to be compatible with the corresponding inflow functions. The following proposition shows in particular that lower arrival time functions just as upper arrival time functions are nondecreasing.

Proposition 3.22. *Let $\check{\ell}$ and $\hat{\ell}$ be the lower and upper arrival time function of some path P, respectively. Further, let F^+ and F^- be corresponding inflow and outflow functions, respectively. Then the following statements are valid:*

(i) *No flow arrives at t along P over $[\check{\ell}(\theta), \hat{\ell}(\theta))$ for all $\theta \in \mathbb{R}_+$, i.e., we have $F^-(\check{\ell}(\theta)) = F^-(\hat{\ell}(\theta))$.*

(ii) *The functions $\check{\ell}$ and $\hat{\ell}$ are nondecreasing.*

(iii) *We have $\hat{\ell}(\theta) \geq \lim_{\theta' \searrow \infty} \check{\ell}(\theta)$ implying $\check{\ell}(\theta) \leq \hat{\ell}(\theta)$ for all $\theta \in \mathbb{R}_+$. Moreover, if F^+ is strictly nondecreasing to the right at θ, equality holds, i.e., $\hat{\ell}(\theta) = \lim_{\theta' \searrow \infty} \check{\ell}(\theta)$.*

(iv) *The lower arrival time function is left continuous and the upper arrival time function is right continuous.*

Proof. Statement (i) follows directly from the equations in (3.22). Therefore, we start by proving (ii) for the lower arrival time function. Let $\theta_2 > \theta_1 \geq 0$. Since F^+ is nondecreasing, (3.22) implies $F^-(\check{\ell}(\theta_2)) \geq F^-(\check{\ell}(\theta_1))$. In case of $F^-(\check{\ell}(\theta_2)) > F^-(\check{\ell}(\theta_1))$ we obtain $\check{\ell}(\theta_2) > \check{\ell}(\theta_1)$ because F^- is nondecreasing. So assume $F^-(\check{\ell}(\theta_2)) = F^-(\check{\ell}(\theta_1))$ which shows $F^+(\theta_1) = F^-(\check{\ell}(\theta_2))$ by (3.22). Hence, as (3.20) implies $\check{\ell}(\theta_2) \geq \theta_2 > \theta_1$, we get $\check{\ell}(\theta_2) \geq \check{\ell}(\theta_1)$ because of the minimum in (3.20). Thus, (ii) is proven for $\check{\ell}$. Similar arguments show that $\hat{\ell}$ is also nondecreasing. Therefore, we omit the details.

Consider statement (iii). Because of the maximum in (3.21) it is enough to show $F^-(\lim_{\theta' \searrow \theta} \check{\ell}(\theta')) = F^+(\theta)$. Note that $\lim_{\theta' \searrow \theta} \check{\ell}(\theta')$ exists because $\check{\ell}$ is nondecreasing by (ii). Since F^+ and F^- are continuous, we obtain from (3.22)

$$F^-\big(\lim_{\theta' \searrow \theta} \check{\ell}(\theta')\big) \;=\; \lim_{\theta' \searrow \theta} F^-\big(\check{\ell}(\theta')\big) \;=\; \lim_{\theta' \searrow \theta} F^+(\theta') \;=\; F^+(\theta)$$

which proves the first part of (iii). In order to prove the second part of (iii), let F^+ be nondecreasing to the right at $\theta \in \mathbb{R}_+$. Because of (3.22) this shows for all $\theta' > \theta$:

$$F^-\big(\check{\ell}(\theta')\big) \;=\; F^+(\theta') \;>\; F^+(\theta) \;=\; F^-\big(\hat{\ell}(\theta)\big)\ .$$

Since F^- is nondecreasing, this shows $\check{\ell}(\theta') > \hat{\ell}(\theta)$ implying $\lim_{\theta' \searrow \theta} \check{\ell}(\theta') \geq \hat{\ell}(\theta)$. Thus, the second part of (iii) follows directly from its first part.

For proving the left continuity of $\check{\ell}$ in (iv) we have to show that, for all $\theta > 0$, the equation $\lim_{\theta' \nearrow \theta} \check{\ell}(\theta') = \check{\ell}(\theta)$ is valid. From (ii) we get that $\lim_{\theta' \nearrow \theta} \check{\ell}(\theta')$ exists and that $\lim_{\theta' \nearrow \theta} \check{\ell}(\theta') \leq \check{\ell}(\theta)$ holds. In order to show $\lim_{\theta' \nearrow \theta} \check{\ell}(\theta') \geq \check{\ell}(\theta)$, we consider the following equation chain obtained from (3.22) and the continuity of F^+ and F^-:

$$F^-\big(\lim_{\theta' \nearrow \theta} \check{\ell}(\theta')\big) \;=\; \lim_{\theta' \nearrow \theta} F^-\big(\check{\ell}(\theta')\big) \;=\; \lim_{\theta' \nearrow \theta} F^+(\theta') \;=\; F^+(\theta)\ .$$

Further, the definition of $\check{\ell}$ in (3.20) implies

$$\lim_{\theta' \nearrow \theta} \check{\ell}(\theta') \;\geq\; \lim_{\theta' \nearrow \theta} \theta' \;=\; \theta\ .$$

Hence, from the minimum in (3.20) we get $\lim_{\theta' \nearrow \theta} \check{\ell}(\theta') \geq \check{\ell}(\theta)$. Therefore, $\check{\ell}$ is left continuous. The proof of (iv) for $\hat{\ell}$ follows the same line of arguments and is omitted for this reason. $\square$

Although lower and upper arrival time functions have nice properties and seem to be a good choice as a representative of equivalent flow models, they are not well designed for the purposes of routing games over time. As we will see in Chapter 4, both functions do not reflect the desired behavior of flow particles in a selfish routing system. But based on these arrival time functions, we define and analyze another arrival time function in Section 3.4 which allows us to define Nash flows over time.

3.3.2 Continuity of Flow Models

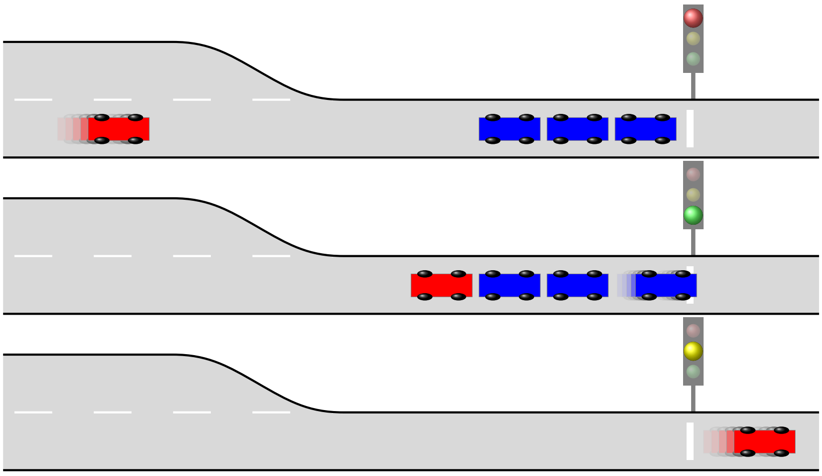
A consistent flow model should be a valid modeling tool for a wide class of applications. In addition to the FiFo principle, it is quite natural that a small perturbation of the inflow of some edge e or of some path P does not lead to a completely different flow behavior on e or P. It would be rather expected that the flow distribution on e or P is changed only slightly. Hence, some kind of continuity is needed.

Since the set $\mathcal{P}$ of s-t-paths is countable in general, a small perturbation of the inflow of each path can lead to a large, even infinitely large deviation of the entire inflow. For this reason we require that the entire inflow $D := \sum_{P\in\mathcal{P}} F_P^+$ must be locally finite for all feasible path-based flows over time $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$, i.e., $D(\theta) < \infty$ for all $\theta \in \mathbb{R}_+$. In particular, this implies that the the sequence of partial sums of $\sum_{P\in\mathcal{P}} F_P^+$ satisfies the assumptions of the Weierstrass M-test on each compact interval. Hence, $\sum_{P\in\mathcal{P}} F_P^+$ converges locally uniform to D. Thus, by Proposition 2.33 the entire inflow D is nondecreasing, locally absolutely continuous and satisfies $D(0) = 0$. Note that all of this is obviously satisfied by edge-based flows over time as the number of outgoing edges of s is finite. Also note that the same holds for the outflow pattern $\mathrm{val}(\mathcal{F}^+)$. Because of Proposition 3.7(i) the outflow pattern val is bounded by D and, for applying the Weierstrass M-test, we can use the same bounds for the outflow functions which we already used for the inflow functions.

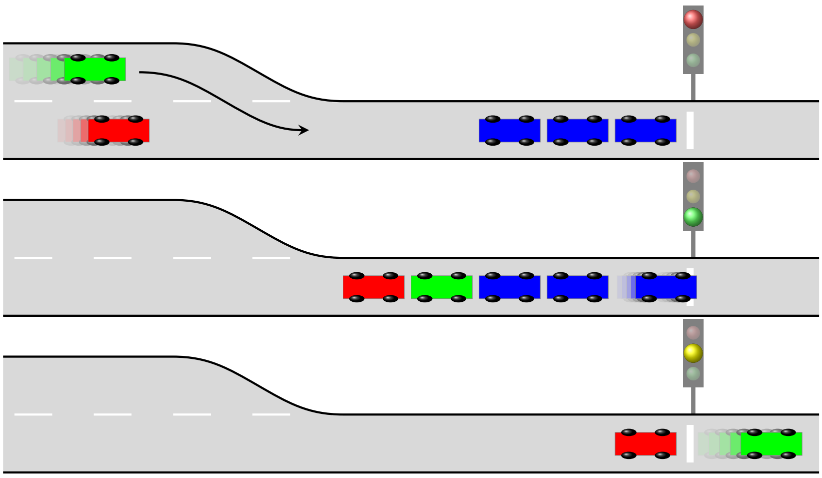
In this section we characterize flow models which are continuous depending on the inflow and show how the notion of continuity can be simplified if the model satisfies FiFo. Be aware that a *continuous* flow over time model refers to a model where each inflow function is describable via Lebesgue integrable flow rate functions. In order to avoid conflicts between these two notions of continuity, we use the term $\mathcal{F}$-continuity for flow models which behave continuous in the inflow. For motivating the subsequent definition of $\mathcal{F}$-continuity, consider the following example depicted in Figure 3.6.

Example 3.23. Let e be a street segment with a traffic light at its head. Assume that there is some red car which traverses e and which has to wait behind some other cars in front of the traffic light because it is red. During the next green phase, the red car is the last car which could cross the traffic light – just in time before the traffic light jumps to red.

Now consider another scenario where an aggressive green car pushes itself into a small gap in front of the red car. This causes that the red car must wait an additional red phase before it is able to pass the traffic light. So the transit time of the red car increases by a constant real number. This shows that it is not a good idea to define $\mathcal{F}$-continuity in terms of transit or arrival times. On the other hand, considering the outflow F_e^- until the red car leaves e, we



(a) Scenario where the red car need not wait for an additional red phase.



(b) The green car causes that the red car has to wait for an additional red phase.

Figure 3.6: Example motivating the definition of $\mathcal{F}$-continuity.

observe that they only differ by one car.

Example 3.23 shows that $\mathcal{F}$-continuity defined in terms of outflow functions seems to be appropriate. Another aspect for preferring outflow functions is that transit and arrival time functions may take arbitrary values if no flow enters e. This motivates the following definition.

Definition 3.24 (Continuity of a Flow Model)**.** A given path-based flow over time model is $\mathcal{F}$-*continuous* if and only if the map $\mathcal{F}^+ \mapsto F_P^-|_{\leq\theta}$ is continuous as a map from $(L_\infty^{\mathrm{loc}})_1^{\mathcal{P}}$ to L_∞^{loc} for all paths P and all times $\theta \in \mathbb{R}_+$ on any network. For a given path-based flow over time $\mathcal{F}^+$, this means that for all $\epsilon > 0$ and all compact sets K_ϵ there exist a $\delta > 0$ and a compact set K_δ such that

$$\left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \left\|F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta}\right\|_{\infty}^{K_\epsilon} \leq \epsilon \ .$$

Consider a sequence $(\mathcal{F}^i)_{i\in\mathbb{N}} := (F_P^i)_{P\in\mathcal{P}}$ of path-based flows over time converging to some path-based flow $\mathcal{F}^* = (F_P^*)_{P\in\mathcal{P}}$. The metric on $(L_\infty^{\text{loc}})_1^{\mathcal{P}}$ ensures that the total net inflow $\sum_{P\in\mathcal{P}} F_P^i$ converges in L_∞^{loc} to $\sum_{P\in\mathcal{P}} F_P^*$ as i goes to ∞. The following lemma implies that this also holds for the corresponding restricted outflow functions if the underlying flow model is $\mathcal{F}$-continuous. Note that for this result the additional assumption $D < \infty$ on the entire inflow D is essential.

Lemma 3.25. *Consider an $\mathcal{F}$-continuous path-based flow over time model. Then, for each $\theta \in \mathbb{R}_+$, the map $\mathcal{F}^+ \mapsto (F_P^-|_{\leq\theta})_{P\in\mathcal{P}}$ is continuous as a map from $(L_\infty^{\text{loc}})_1^{\mathcal{P}}$ to $(L_\infty^{\text{loc}})_1^{\mathcal{P}}$.*

Proof. Fix some point in time $\theta \in \mathbb{R}_+$ and consider a path-based flow $\mathcal{F}^+$. Further, let $\epsilon > 0$ be some small real number and $K^\epsilon \subset \mathbb{R}_+$ be some compact set. Since we have $D(\theta) := \sum_{P\in\mathcal{P}} F_P^+(\theta) < \infty$, there exists a finite set of paths $\mathcal{P}^\epsilon := \{P_1, \ldots, P_k\} \subseteq \mathcal{P}$ such that

$$D_\epsilon(\theta) := \sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} F_{P_i}^+(\theta) \;\leq\; \epsilon\;.$$

Further, for all path-based flows over time $\tilde{F}^+$ with $\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{\{\theta\}} \leq \epsilon$, we know

$$\left|D_\epsilon(\theta) - \tilde{D}_\epsilon(\theta)\right| \;\leq\; \sum_{P\in\mathcal{P}} \left|F_P^+(\theta) - \tilde{F}_P^+(\theta)\right| \;\leq\; \epsilon \quad \text{implying} \quad \tilde{D}_\epsilon(\theta) \;\leq\; 2\epsilon\;.$$

Since the restricted outflow function of a path P until time θ does not exceed the amount of flow entering P until time θ (see Proposition 3.8), we deduce

$$0 \leq \sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} F_P^-|_{\leq\theta} \leq \epsilon \qquad \text{and} \qquad 0 \leq \sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} \tilde{F}_P^-|_{\leq\theta} \leq 2\epsilon$$

implying

$$\begin{aligned}\sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} \left\|F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta}\right\|_\infty^{K_\epsilon} \;&\leq\; \sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} \left\|F_P^-|_{\leq\theta}\right\|_\infty^{K_\epsilon} + \sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} \left\|\tilde{F}_P^-|_{\leq\theta}\right\|_\infty^{K_\epsilon} \\ &\leq\; 3\epsilon \end{aligned} \tag{3.24}$$

for all path-based flows over time $\tilde{F}^+$ with $\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{\{\theta\}} \leq \epsilon$.

Since the flow model is $\mathcal{F}$-continuous, there exist a $\delta_1 > 0$ and a compact set K_{δ_1} such that

$$\left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{K_{\delta_1}} \leq \delta_1 \qquad \Rightarrow \qquad \left\|F_{P_i}^-|_{\leq\theta} - \tilde{F}_{P_i}^-|_{\leq\theta}\right\|_{\infty,1}^{K_\epsilon} \;\leq\; \frac{\epsilon}{k} \tag{3.25}$$

holds for all $i = 1, \ldots, k$.

Now set $\delta := \min\{\epsilon, \delta_1\}$ and $K_\delta := \{\theta\} \cup K_{\delta_3}$ and let $\tilde{\mathcal{F}}^+$ be a path-based flow over time with $\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{K_\delta} \leq \delta$. Hence, we have

$$\left|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{\{\theta\}} \leq \epsilon \qquad \text{and} \qquad \left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+|\right\|_{\infty,1}^{K_{\delta_1}} \leq \delta_1$$

which implies

$$\sum_{P\in\mathcal{P}} \left\| F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta} \right\|_{\infty,1}^{K_\epsilon}$$

$$\leq \sum_{P\in\mathcal{P}\setminus\mathcal{P}^\epsilon} \left\| F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta} \right\|_{\infty,1}^{K_\epsilon} + \sum_{i=1}^{k} \left\| F_{P_i}^-|_{\leq\theta} - \tilde{F}_{P_i}^-|_{\leq\theta} \right\|_{\infty,1}^{K_\epsilon} \leq 4\epsilon. \quad (3.26)$$

This finishes the proof. □

A particular consequence of Lemma 3.25 is that the flow pattern $\mathrm{val}(\mathcal{F}^+)$ behaves continuously in $\mathcal{F}^+$.

Theorem 3.26. *For an $\mathcal{F}$-continuous flow model, the map $\mathcal{F}^+ \mapsto \mathrm{val}(\mathcal{F}^+)$ is continuous as a map from $(L_\infty^{\mathrm{loc}})_1^{\mathcal{P}}$ to L_∞^{loc}.*

Proof. Consider a path-based flow over time $\mathcal{F}^+$. By Definition 3.4, the equation

$$\mathrm{val}(\mathcal{F}^+)(\theta) = \sum_{P\in\mathcal{P}} F_P^-(\theta) = \sum_{P\in\mathcal{P}} F_P^-|_{\leq\theta}(\theta)$$

holds for each $\theta \in \mathbb{R}_+$ because Proposition 3.8 shows $F_P^-(\theta) = F_P^-|_{\leq\theta}(\theta)$ in particular.

Let $\epsilon > 0$ and $K_\epsilon \subset \mathbb{R}_+$ be a compact set. Since compact sets are bounded, there exists a $\theta \in \mathbb{R}_+$ such that $K_\epsilon \subseteq [0,\theta]$. Applying Lemma 3.25 for this θ, we know that there is a $\delta > 0$ and a compact set $K_\delta \subset \mathbb{R}_+$ such that

$$\left\| \mathcal{F}^+ - \tilde{\mathcal{F}}^+ \right\|_{\infty,1}^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \sum_{P\in\mathcal{P}} \left\| F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta} \right\|_{\infty,1}^{K_\epsilon} \leq \epsilon .$$

But the right hand side of this implication shows directly

$$\begin{aligned} \left\| \mathrm{val}(\mathcal{F}^+) - \mathrm{val}(\mathcal{F}^+ \right\|_1^{K_\epsilon} &= \left\| \sum_{P\in\mathcal{P}} F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta} \right\|_{\infty,1}^{K_\epsilon} \\ &\leq \sum_{P\in\mathcal{P}} \left\| F_P^-|_{\leq\theta} - \tilde{F}_P^-|_{\leq\theta} \right\|_{\infty,1}^{K_\epsilon} \leq \epsilon \end{aligned}$$

which finishes this proof. □

Considering a path P, an $\mathcal{F}$-continuous flow model requires that each restricted outflow function $F_P^-|_{\leq\theta}$ depends continuously on $\mathcal{F}^+$. The following lemma shows that this continuity is somehow evenly in θ. That is, the entire flow behavior on a path P is changed only slightly if the inflow $\mathcal{F}^+$ is changed only slightly.

Lemma 3.27. *Consider an $\mathcal{F}$-continuous path-based flow over time model and let $F^-|_{\leq\theta}$ be the restricted outflow function of a given path P until some point in time $\theta \in \mathbb{R}_+$. Then the family $(\mathcal{F}^+ \mapsto F^-|_{\leq\theta'})_{\theta'\leq\theta}$ of maps is equicontinuous for all $\theta \in \mathbb{R}_+$. That is, for all path-based flows over time $\mathcal{F}^+$, it holds that for each $\epsilon > 0$ and each compact set $K_\epsilon \subset \mathbb{R}_+$ there exist a $\delta > 0$ and a compact set $K_\delta \subset \mathbb{R}_+$ such that*

$$\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \|F^-|_{\leq\theta'} - \tilde{F}^-|_{\leq\theta'}\|_\infty^{K_\epsilon} \leq \epsilon$$

is valid for all nonnegative $\theta' \leq \theta$.

Proof. Consider a path-based flow over time $\mathcal{F}^+$ and let F^+ be the inflow of P. Further, let $\theta \in \mathbb{R}_+$ be some point in time, $\epsilon > 0$ be a small positive real number, and K_ϵ be a compact set. Since F^+ is continuous and locally bounded, there exists a partition $0 =: \theta_1 < \ldots < \theta_k := \theta$ of $[0, \theta]$ such that $F^+(\theta_{i+1}) - F^+(\theta_i) \leq \epsilon$ holds for all $1 \leq i = j-1 < k$. Because of the definition of the supremum norm, all other inflows $\tilde{F}^+$ of P satisfy

$$\left\|F^+ - \tilde{F}^+\right\|_\infty^{[0,\theta]} \leq \epsilon \qquad \Rightarrow \qquad \left|F^+(\theta_i) - \tilde{F}^+(\theta_i)\right| \leq \epsilon$$

for all $i = 1, \ldots, k$. In addition, since the flow model is $\mathcal{F}$-continuous, we know that there exists a $\delta_1 > 0$ and a compact set K_{δ_1} such that

$$\left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{K_{\delta_1}} \leq \delta_1 \qquad \Rightarrow \qquad \left\|F^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta_i}\right\|_\infty^{K_\epsilon} \leq \epsilon \tag{3.27}$$

holds for all $i = 1, \ldots, k$.

Now we define $\delta := \min\{\epsilon, \delta_1\}$ and $K_\delta := [0, \theta] \cup K_{\delta_1}$ and let $\tilde{\mathcal{F}}^+$ be a path-based flow over time with $\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{K_\delta} \leq \delta$. In order to establish this lemma, we show that $\|F^-|_{\leq \theta'} - \tilde{F}^-|_{\leq \theta'}\|_\infty^{K_\epsilon} \leq 5\epsilon$ is valid for all $\theta' \in [0, \theta]$. For this fix a $\theta' \in [0, \theta]$ and let i be such that $\theta_i \leq \theta' \leq \theta_{i+1}$. Since we have $\|F^+ - \tilde{F}^+\|_\infty^{[0,\theta]} \leq \delta_1$ in particular, we obtain out of the definition of the partition of $[0, \theta]$ that

$$\begin{aligned}
F^+(\theta') - F^+(\theta_i) &\leq \epsilon \\
\text{and} \quad \tilde{F}^+(\theta') - \tilde{F}^+(\theta_i) &\leq \tilde{F}^+(\theta_{i+1}) - \tilde{F}^+(\theta_i) \leq F^+(\theta_{i+1}) - F^+(\theta_i) + 2\epsilon \\
&\leq 3\epsilon \, .
\end{aligned}$$

By Proposition 3.8 this shows

$$F^-|_{\leq \theta'} - F^-|_{\leq \theta_i} \leq \epsilon \quad \text{implying} \quad \left\|F^-|_{\leq \theta'} - F^-|_{\leq \theta_i}\right\|_\infty^{K_\epsilon} \leq \epsilon \, , \tag{3.28}$$

$$\tilde{F}^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta'} \leq 3\epsilon \quad \text{implying} \quad \left\|\tilde{F}^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta'}\right\|_\infty^{K_\epsilon} \leq 3\epsilon \, . \tag{3.29}$$

Together with $\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_\infty^{K_{\delta_1}} \leq \delta_1$, we finally obtain

$$\begin{aligned}
&\left\|F^-|_{\leq \theta'} - \tilde{F}^-|_{\leq \theta'}\right\|_\infty^{K_\epsilon} \\
&\leq \left\|F^-|_{\leq \theta'} - F^-|_{\leq \theta_i} \quad + \quad F^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta_i} \quad + \quad \tilde{F}^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta'}\right\|_\infty^{K_\epsilon} \\
&\leq \left\|F^-|_{\leq \theta'} - F^-|_{\leq \theta_i}\right\|_\infty^{K_\epsilon} + \left\|F^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta_i}\right\|_\infty^{K_\epsilon} + \left\|\tilde{F}^-|_{\leq \theta_i} - \tilde{F}^-|_{\leq \theta'}\right\|_\infty^{K_\epsilon} \\
&\leq 5\epsilon
\end{aligned} \tag{3.30}$$

by adding artificial zeros and using (3.28), (3.27), and (3.29) for bounding the first, second, and third summand in line 3, respectively. □

Concluding this general discussion about the $\mathcal{F}$-continuity of path-based flow models, the following theorem shows that the entire flow behavior on a network changes only slightly if the inflow is modified a little bit.

Theorem 3.28. *Consider an $\mathcal{F}$-continuous path-based flow over time model. For a path-based flow over time $\mathcal{F}^+$ let $\mathcal{F}^-|_{\leq \theta} := (F_P^-|_{\leq \theta})_{P \in \mathcal{P}}$ be the family of restricted outflows until some point in time $\theta \in \mathbb{R}_+$. Then the family $(\mathcal{F}^+ \mapsto \mathcal{F}^-|_{\leq \theta'})_{\theta' \leq \theta}$ of maps is equicontinuous for all $\theta \in \mathbb{R}_+$. That is, for*

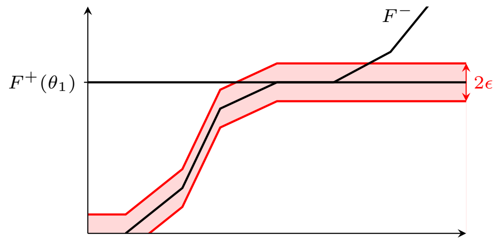


Figure 3.7: If F^- and F^+ are changed by ϵ with respect to the ∞-norm, the light red area shows how the restricted outflow $F^-|_{\leq\theta_1}$ can be changed.

all path-based flows over time $\mathcal{F}^+$, it holds that for each $\epsilon > 0$ and each compact set $K_\epsilon \subset \mathbb{R}_+$ there exist a $\delta > 0$ and a compact set $K_\delta \subset \mathbb{R}_+$ such that

$$\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \|\mathcal{F}^-|_{\leq\theta'} - \tilde{\mathcal{F}}^-|_{\leq\theta'}\|_{\infty,1}^{K_\epsilon} \leq \epsilon$$

is valid for all nonnegative $\theta' \leq \theta$.

Proof. Fix some point in time $\theta \in \mathbb{R}_+$. Following the proof of Lemma 3.25, we observe that inequality (3.24) remains valid for all $\theta' \leq \theta$ because the inflow functions are nondecreasing. Further, the implication (3.25) holds also for all $\theta' \leq \theta$ because of Lemma 3.27. Hence, the final inequality chain is satisfied for all $\theta' \leq \theta$ which finishes of this proof. □

Although Lemma 3.27 shows that the family of restricted outflows of a certain path behaves equicontinuous, we still have to check all restricted outflows in order to verify $\mathcal{F}$-continuity. However, for flow models satisfying the FiFo condition, we are able to give a simpler characterization of $\mathcal{F}$-continuity.

Using Lemma 3.17, we formulate $\mathcal{F}$-continuity of a flow model without the notion of restricted outflow functions. Intuitively, this seems to be obvious because Lemma 3.17 shows that all restricted outflow functions are already defined by corresponding inflow and nonrestricted outflow functions (see Figure 3.7).

Lemma 3.29. *Consider a flow over time model satisfying FiFo. Then this flow model is $\mathcal{F}$-continuous if and only if the map $\mathcal{F}^+ \mapsto F_P^-$ is continuous as a map from $(L_\infty^{\text{loc}})_1^{\mathcal{P}}$ to L_∞^{loc} for all $P \in \mathcal{P}$. For every path P this means that for all path-based flows over time $\mathcal{F}^+$, all $\epsilon > 0$, and all compact sets $K_\epsilon \in \mathbb{R}_+$ there exist a $\delta > 0$ and a compact set K_δ such that*

$$\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\|_{\infty,1}^{K_\delta} \leq \delta \qquad \Rightarrow \qquad \|F_P^- - \tilde{F}_P^-\|_\infty^{K_\epsilon} \leq \epsilon\ .$$

Proof. Consider a path P and let F and F^- be the corresponding inflow and outflow function with respect to some flow over time $\mathcal{F} = (\mathcal{F}^+, \mathcal{F}^-)$, respectively.

First assume that the flow model is $\mathcal{F}$-continuous. Thus, we have to show that the map $\mathcal{F}^+ \mapsto F^-$ is continuous if each map $\mathcal{F}^+ \mapsto F^-|_{\leq\theta}$ with $\theta \in \mathbb{R}_+$ is continuous. This follows directly from Proposition 3.8 as it shows that F^- and $F^-|_{\leq\theta}$ coincide on $[0, \theta]$ for all $\theta \in \mathbb{R}_+$. This implies that the map $\mathcal{F}^+ \mapsto F^-$ is continuous as a map from $(L_\infty^{\text{loc}})_1^{\mathcal{P}}$ to L_∞^{loc} – for a given compact set K_ϵ choose θ such that $K_\epsilon \subseteq [0, \theta]$. Note that we do not need the FiFo assumption for this direction.

For proving the other direction, let $\theta \in \mathbb{R}_+$ be some point in time and assume that the map $\mathcal{F}^+ \mapsto F^-$ is continuous. In order to establish the $\mathcal{F}$-continuity of the flow model, we have to show that the map $\mathcal{F}^+ \mapsto F^-|_{\leq\theta}$ is continuous. Considering an $\epsilon > 0$ and a compact set K_ϵ, there exist a $\delta_1 > 0$ and a compact set K_{δ_1} such that

$$\left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{K_{\delta_1}} \leq \delta_1 \qquad \Rightarrow \qquad \left\|F^- - \tilde{F}^-\right\|_{\infty}^{K_\epsilon} \leq \epsilon \ .$$

In addition, we know

$$\left\|\mathcal{F}^+ - \tilde{\mathcal{F}}^+\right\|_{\infty,1}^{\{\theta\}} \leq \epsilon \qquad \Rightarrow \qquad \left|F^+(\theta) - \tilde{F}^+(\theta)\right| \leq \epsilon \ .$$

Defining $\delta := \min\{\epsilon, \delta_1\}$ and $K_\delta := K_{\delta_1} \cup \{\theta\}$, this leads to

$$\left\|F^-|_{\leq\theta_1} - \tilde{F}^-|_{\leq\theta_1}\right\|_{\infty}^{K_\epsilon} \leq 2\epsilon$$

because

$$\begin{aligned}
\left|F^-|_{\leq\theta}(\theta') - \tilde{F}^-|_{\leq\theta}(\theta')\right| &= \left|\min\{F^-(\theta'), F^+(\theta)\} - \min\{\tilde{F}^-(\theta'), \tilde{F}^+(\theta)\}\right| \\
&= \Big| \quad \min\{F^-(\theta'), F^+(\theta)\} - \min\{F^-(\theta'), \tilde{F}^+(\theta)\} \\
&\qquad + \min\{F^-(\theta'), \tilde{F}^+(\theta)\} - \min\{\tilde{F}^-(\theta'), \tilde{F}^+(\theta)\}\Big| \\
&\leq \left|F^+(\theta) - \tilde{F}^+(\theta)| + |F^-(\theta') - \tilde{F}^-(\theta')\right|
\end{aligned}$$

holds for all $\theta' \in \mathbb{R}_+$. □

Lemma 3.29 shows that $\mathcal{F}$-continuity can be modeled using only outflows instead of the entire family of restricted outflows if the flow model satisfies FiFo. However, for a given $\epsilon > 0$ and a compact set K_ϵ, the set K_δ cannot be bounded a priori. The next subsection identifies a property of flow models which gives a particular answer to this question.

So far, we have discussed $\mathcal{F}$-continuous flow models in terms of outflows. Especially, if the FiFo-principle is satisfied, the flow model is representable via nice arrival times. The following lemma analyzes the behavior of lower and upper arrival time functions if the inflow is changed is slightly.

Lemma 3.30. *Consider an $\mathcal{F}$-continuous flow model which satisfies FiFo and let P be some s-t-path. Then the lower arrival time of P is lower semi-continuous and the upper arrival time of P is upper semi-continuous.*

Proof. Let $(\mathcal{F}_i^+)_{i\in\mathbb{N}}$ be a sequence of flows over time which converges to some flow over time $\mathcal{F}^+$ and let F_i^- and F^- be the outflow functions of P with respect to $\mathcal{F}_i^+$ for $i \in \mathbb{N}$ and $\mathcal{F}^+$, respectively, and F_i^+ and F^+ be the corresponding inflow functions of P. Further, let $\check{\ell}_i$ and $\check{\ell}$ be the lower arrival time functions of P with respect to $\mathcal{F}_i^+$ for $i \in \mathbb{N}$ and $\mathcal{F}^+$, respectively. Similarly, let $\hat{\ell}_i$ and $\hat{\ell}$ denote the corresponding upper arrival time functions of P.

Firstly, we prove that the lower arrival time function is lower semi-continuous at an arbitrary point in time $\theta \in \mathbb{R}_+$, i.e., for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$\check{\ell}_i(\theta) \geq \check{\ell}(\theta) - \epsilon \qquad \forall i \geq N \ .$$

Clearly, if $\check{\ell}(\theta) = \theta$ nothing has to be shown as lower arrival time functions are never less than θ at time θ by definition. So let $\check{\ell}(\theta) > \theta$ be valid in the following and fix an $\epsilon \in (0, \check{\ell}(\theta) - \theta]$. Because of the minimum in the definition of the lower arrival time function (see (3.20)), we know $\delta := F^-(\check{\ell}(\theta)) - F^-(\check{\ell}(\theta) - \epsilon) > 0$. Since $(\mathcal{F}_i^+)_{i \in \mathbb{N}}$ converges to $\mathcal{F}^+$, there exists an N_1 such that

$$F_i^+(\theta) > F^+(\theta) - \frac{\delta}{2} \qquad\qquad \forall i \geq N_1 \ .$$

Further, because the flow model is $\mathcal{F}$-continuous, there exists an $N_2 \in \mathbb{N}$ such that

$$F_i^-\big(\check{\ell}(\theta) - \epsilon\big) < F^-\big(\check{\ell}(\theta) - \epsilon\big) - \frac{\delta}{2} \qquad\qquad \forall i \geq N_2 \ .$$

Defining $N := \max\{N_1, N_2\}$ and using (3.22), we obtain for all $i \geq N$ that

$$\begin{aligned} F_i^-\big(\check{\ell}_i(\theta)\big) \;=\; F_i^+(\theta) \;>\; F^+(\theta) - \frac{\delta}{2} \;&=\; F^-\big(\check{\ell}(\theta)\big) - \frac{\delta}{2} \\ &=\; F^-\big(\check{\ell}(\theta) - \epsilon\big) + \frac{\delta}{2} \;>\; F_i^-\big(\check{\ell}(\theta) - \epsilon\big) \ . \end{aligned}$$

Since each F_i is nondecreasing, we get $\ell_i(\theta) \geq \ell(\theta) - \epsilon$ for all $i \geq N$. This shows that the lower arrival time function is lower semi-continuous. Since the line of arguments for the proof of the upper semi-continuity of the upper arrival time remains quite the same, we omit the details here. □

In contrast to the continuity of the outflows, the semi-continuity of the upper and lower arrival times is not uniform. The reason is that there exists no positive lower bound on δ if, e.g., $\check{\ell}(\theta)$ tends from right to the left boundary of an interval over which the outflow function F^- is constant. If F^- remains constant over an interval I, the current upper and lower arrival times differ over $(F^-)^{-1}(I)$.

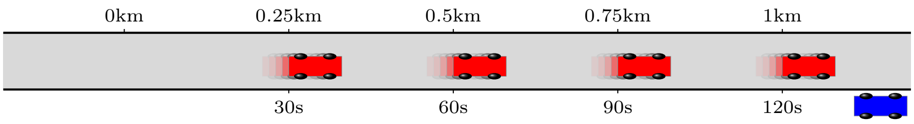
Let us consider the case where current lower and upper arrival times coincide over an open set O. By Proposition 3.22(iv) this is equivalent to the continuity of the *functions* $\check{\ell}$ and $\hat{\ell}$ over O. Further, by Proposition 3.22(iii), Lemma 3.29 shows that $\check{\ell}$ and $\hat{\ell}$ behave continuously in the inflow in this case, i.e., they are continuous as maps from $(L_\infty)_1^{\mathcal{P}}$ to $L_\infty(O)$. Surprisingly, this continuity is locally uniform. To see this, observe that F^- is strictly nondecreasing over $\check{\ell}(O)$. Hence, $\delta = F^-(\check{\ell}(\theta)) - F^-(\check{\ell}(\theta) - \epsilon)$ is strictly positive and continuous for all $\theta \in O$. Thus, the minimum of δ over each compact set $K \subseteq O$ exists and is strictly positive. Using this minimum in the proof, Lemma 3.29 shows that $\check{\ell}$ is uniformly continuous.

3.3.3 Past-Orientation

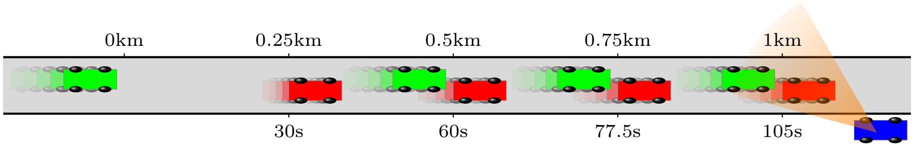
This subsection addresses the question which flow particles can influence the behavior of other flow particles. Informally, past-orientation requires that current transit times are independent on flow particles entering an edge or path later. The following example gives some intuition behind this topic.

Example 3.31. Consider an one-way street e with one lane and let a red car enter e as drawn in Figure 3.8. The length of e amounts to $1km$ and the speed-limit is $30km/h$. Since the driver of the red car has 17 points,[1] she obeys

[1] In Germany you get at least one point if you are $21km/h$ too fast. If you collect 18 points, you lose your driver's license until you pass the so-called Medical Psychological Assessment.



(a) Scenario where the red car is not influenced by other cars.



(b) Scenario where the red car accelerates because of the behavior of the green car.

Figure 3.8: Example explaining past-orientation.

strictly the traffic rules and, hence, would leave this edge after two minutes. Otherwise she may lose her driver's license. However, 30 seconds after the red car, a green car enters e and the driver of this car is really in a hurry. After the red car has covered half of the distance, the green car sticks to the red car bumper-to-bumper and gives wildly gesticulating flashlight. So the red driver feels coerced to accelerate a little bit because she can really good put oneself in the green driver's shoes. This decreases the transit time of the red car to $1min\ 35s$. However, at the end of the street this behavior also causes two shortly sequenced flash lights out of a car parking on the near side.

A closer look at the scenario of Example 3.31 shows that obeying the speed limit would have been better for the red driver. In flow terminology this means that the current flow situation on e, i.e., the transit times of the assigned flow, is independent on flow entering e later. Or equivalently, the current (restricted) outflow only depends on the current (restricted) inflow and not on the future inflow. This leads to the following definition.

Definition 3.32 (Past-orientation)**.** A path-based flow model is *past-oriented* if and only if $F_P^-|_{\leq\theta}(\mathcal{F}^+) = F_P^-(\mathcal{F}^+|_{\leq\theta})$ holds for all points in time $\theta \in \mathbb{R}_+$, on all paths $P \in \mathcal{P}$, and all path-based flow over times $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$.

Definition 3.32 ensures that flow particles originating later cannot influence the current transit time of flow particles originating earlier. Hence, for a flow model satisfying FiFo, past-orientation has the following useful impact. Given the inflow function until a particular point in $\theta > 0$, we know the outflow function of some path P over the time interval $[0, \check{\ell}_P(\theta)]$ where $\check{\ell}_P$ is the lower arrival time function of P. Technically, by Lemma 3.21 this can be written as $F_P^-|_{[0,\check{\ell}_P(\theta)]}(\mathcal{F}^+) = F_P^-|_{\leq\theta}(\mathcal{F}^+) = F_P^-|_{[0,\check{\ell}_P(\theta)]}(\mathcal{F}^+|_{\leq\theta})$.

For a past-oriented flow model we can also simplify the notation of $\mathcal{F}$-continuity. In fact, a past-oriented flow model is $\mathcal{F}$-continuous if and only if each map $\mathcal{F}^+ \to F_P^-|_{\leq\theta}$ is continuous as a map from $(L([0,\theta]))_1^{\mathcal{P}}$ to L_1^{loc}. If, in addition, the flow model also satisfies FiFo, the flow model is $\mathcal{F}$-continuous if and only if each map $\mathcal{F}^+ \to F_P^-$ is continuous as a map from $(L([0,\theta]))_1^{\mathcal{P}}$ to $L([0,\ell_P(\theta)])_1$.

As explained in Section 3.3.1, flow models which satisfy the FiFo-principle are representable by lower and upper arrival times. The following proposition shows that past-orientation carries over to these arrival times in some sense.

Proposition 3.33. *Consider a past-oriented flow model which satisfies FiFo and let $\check{\ell}$ and $\hat{\ell}$ be the lower arrival and upper arrival time of some path $P \in \mathcal{P}$, respectively. Then for a path-based flow over time $\mathcal{F}^+$ we have:*

$$\begin{array}{llcll} & \check{\ell}(\mathcal{F}^+)(\theta') & = & \check{\ell}\big(\mathcal{F}^+|_{\leq\theta}\big)(\theta') & \forall \theta \geq \theta' \\ \text{and} & \hat{\ell}(\mathcal{F}^+)\big(\theta'\big) & = & \hat{\ell}\big(\mathcal{F}^+|_{\leq\theta}\big)(\theta') & \forall \theta, \theta' \text{ with } F^+(\theta) > F^+(\theta') \end{array}$$

where F^+ is the inflow function of P.

Proof. Let F^- be the outflow function of P with respect to $\mathcal{F}^+$ and let $\theta, \theta' \in \mathbb{R}_+$ be two points in time with $\theta' \leq \theta$. By Lemma 3.17 we know

$$F^-|_{\leq\theta}(\mathcal{F}^+)(\vartheta) \;=\; \min\big\{F^-(\mathcal{F}^+)(\vartheta), F^+(\theta)\big\} \tag{3.31}$$

for all $\vartheta \in \mathbb{R}_+$. Since $\theta' \leq \theta$ implies $F^+(\theta') \leq F^+(\theta)$, this shows the equivalence of $F^+(\theta') \leq F^-(\mathcal{F}^+)(\vartheta)$ and $F^+(\theta') \leq F^-|_{\leq\theta}(\mathcal{F}^+)(\vartheta)$ for all $\vartheta \in \mathbb{R}_+$. Using the definition of the lower arrival time function in (3.20), this shows directly

$$\begin{aligned} \check{\ell}(\mathcal{F}^+)(\theta') &= \min\big\{\vartheta \geq \theta' \;\big|\; F^+(\theta') \leq F^-(\mathcal{F}^+)(\vartheta)\big\} \\ &= \min\big\{\vartheta \geq \theta' \;\big|\; F^+(\theta') \leq F^-|_{\leq\theta}(\mathcal{F}^+)(\vartheta)\big\} \\ &= \min\big\{\vartheta \geq \theta' \;\big|\; F^+|_{\leq\theta}(\theta') \leq F^-(\mathcal{F}^+|_{\leq\theta})(\vartheta)\big\} \;=\; \check{\ell}(\mathcal{F}^+|_{\leq\theta})(\theta') \end{aligned}$$

because the flow model is past-oriented. If $F^+(\theta') < F^+(\theta)$ holds then equation (3.31) shows $F^+(\theta') \geq F^-(\mathcal{F}^+)(\vartheta)$ if and only if $F^+(\theta') \geq F^-|_{\leq\theta}(\mathcal{F}^+)(\vartheta)$ for all $\vartheta \in \mathbb{R}_+$. Using the definition of the upper arrival time function in (3.21), this directly shows

$$\begin{aligned} \hat{\ell}(\mathcal{F}^+)(\theta') &= \max\big\{\vartheta \;\big|\; F^+(\theta') \geq F^-(\mathcal{F}^+)(\vartheta)\big\} \\ &= \max\big\{\vartheta \;\big|\; F^+(\theta') \geq F^-|_{\leq\theta}(\mathcal{F}^+)(\vartheta)\big\} \\ &= \max\big\{\vartheta \;\big|\; F^+|_{\leq\theta}(\theta') \geq F^-(\mathcal{F}^+|_{\leq\theta})(\vartheta)\big\} \;=\; \hat{\ell}(\mathcal{F}^+|_{\leq\theta})(\theta') \,. \end{aligned}$$ ☐

Proposition 3.33 shows that lower arrival times are past-oriented in the sense that the restricted inflow until time θ already determines all current lower arrival times until and *at* time θ. Recalling Definition 3.32, this behavior is also required for the outflow. In this sense, upper arrival times are not entirely past-oriented as we always have $\check{\ell}(\mathcal{F}^+|_{\leq\theta})(\theta) = \infty$. Thus, computing a current upper arrival time for some point in time θ requires knowledge about future flow behavior. In fact, we have to know at which point in time the first flow particles departing after time θ arrives at the sink t. This time would be the current upper arrival time at time θ. Especially, if the inflow rate is zero after time θ, this requires knowledge about future behavior of flow particles.

3.3.4 Locally Finiteness of $\mathcal{P}$

This subsection analyzes a property of flow models in order to avoid counter-intuitive flow behavior. This assumption has to hold only for path-based flow models. We require that the subset $\mathcal{P}_{\leq\theta}$ of paths which can be entered by a positive amount of flow until a given point in time θ is finite for all $\theta \in \mathbb{R}_+$. In other words, for each $\theta \in \mathbb{R}_+$ there exists a finite subset $\mathcal{P}' \subseteq \mathcal{P}$ such that for each path-based flow $\mathcal{F}^+ := (F_P)_{P\in\mathcal{P}}$ we have $\{P \in \mathcal{P} \mid F_P(\theta) > 0\} \subseteq \mathcal{P}'$.

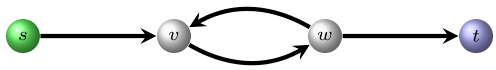


Figure 3.9: A network where each edge has a constant capacity of 1 and a constant transit time of 0.

Definition 3.34. We call $\mathcal{P}$ *locally finite* with respect to a given path-based flow model if and only if the set

$$\mathcal{P}_{\leq\theta} \;:=\; \{P \in \mathcal{P} \mid \exists \mathcal{F}^+ : F_P^+(\theta) > 0\}$$

is finite for each $\theta \in \mathbb{R}_+$.

As a conclusion of this condition, we can assume without loss of generality that $\mathcal{P}$ is finite in case we analyze particular flows over time until some given point in time. Further, this condition prevents some pathological examples like the following.

Example 3.35. Consider the network shown in Figure 3.9 where each edge has a constant transit time of 0 and a constant capacity of 1. We consider the classical flow over time model (see Section 2.5) with the additional assumption that flow must not wait at the intermediate nodes v and w. In fact, we have a path-based flow model where flow is only allowed to enter a path if all edges of this path have free capacities.

For each $i \in \mathbb{N}$ let P_i be the s-t-path which contains $i-1$ times the cycle C induced by v and w. So we have $\mathcal{P} := \{P_i \mid i \in \mathbb{N}\}$. Then a path P_i is able to carry a flow rate of $\frac{1}{i}$ implying that the edge vw is used up to its capacity. Hence, if we wait long enough, we can send an arbitrary amount of flow along each path P_i separately. But in case i goes to infinity, we obtain only a circulation sending flow at a rate of 1 along C. Thus, in the limit no flow is sent to t which is counterintuitive. This can be avoided by allowing, for all $\theta \in \mathbb{R}_+$, only a finite number of paths along which flow can be sent until time θ. Nevertheless, as θ goes to ∞, all paths of $\mathcal{P}$ could be used for sending flow.

In addition, this scenario contradicts the derivation of the Nash flow definition as explained in Section 4.2.

The contradictory behavior of the scenario in Example 3.35 is caused by the following. Although the sequence of flows over time in this scenario converges to zero, flow still remains in the network as a nonzero circulation. If we require that $\mathcal{P}$ must be locally finite for a given flow model, this behavior is avoided.

Summarizing all mentioned aspects of this section (including this and the previous subsections), we define consistent flow models as follows.

Definition 3.36 (Consistent Flow over Time Model)**.** A flow over time model (either edge- or path-based) is called *consistent* if it is $\mathcal{F}$-continuous, satisfies FiFo, and obeys past-orientation. Further, for a consistent path-based flow model the set $\mathcal{P}$ must be locally finite.

3.4 Foresighted Arrival and Transit Times

In this subsection we identify another family of arrival times suitable for an arbitrary consistent flow model. This is necessary from the following point of view. As in static Nash equilibria also in Nash equilibria over time, no flow particle should have an incentive to switch to another path. Therefore, it is quite important to know, how additional flow changes the current transit or arrival time. Note that neither lower nor upper arrival times store such information. More precisely, given a flow over time $\mathcal{F}^+$, then $\check{\ell}_P(\theta)$ equals the time at which the last flow particle entering a path P until a given point in time θ arrives at t. On the other hand, $\hat{\ell}_P(\theta)$ can be interpreted as the arrival time of the first flow particle departing after time θ. In particular, if no flow particle enters P after time θ, the current upper arrival time is equal to ∞ and has no meaning. But especially in this case, we want to know at which time further flow is able to arrive at the sink t via P. This information is important in Section 4.2 where we compute Nash equilibria iteratively. Here, the General Iterative algorithm has to deal with such scenarios in each iteration.

For establishing arrival times storing information about additional flow, we first define

$$\begin{aligned}\mathcal{G}_\theta := \{G : \mathbb{R}_+ \to \mathbb{R}_+ \mid\ & G \text{ is nondecreasing},\\ & G \text{ is absolutely continuous},\\ & G(\vartheta) = 0 \text{ for all } \vartheta \le \theta, \text{ and}\\ & G(\vartheta) > 0 \text{ for all } \vartheta > \theta\}\ .\end{aligned}$$

Hence, $\mathcal{G}_\theta$ contains all functions representing additional (continuous) flow which could be sent from time θ on.

Definition 3.37 (Foresighted Arrival Times)**.** Consider a consistent path-based flow model and let $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$ be a flow over time. Moreover, let $(F_P^-)_{P\in\mathcal{P}}$ be the corresponding family of outflows and $\theta \in \mathbb{R}_+$ be a point in time. For some additional flow $G \in \mathcal{G}_\theta$ we define the *current foresighted arrival time* $\bar{\ell}_P(\mathcal{F}^+)(\theta)$ at time θ of a path P by

$$\bar{\ell}_P(\mathcal{F}^+)(\theta) := \hat{\ell}_P(\mathcal{F}^+ + G)(\theta)\ , \tag{3.32}$$

Here, we define $\mathcal{F}^+ + G := (\tilde{F}_{P'}^+)_{P'\in\mathcal{P}}$, where $\tilde{F}_{P'}^+ := \begin{cases} F_{P'}^+ + G & \text{if } P' = P\\ F_{P'}^+ & \text{otherwise}\end{cases}$.

The definition of foresighted arrival times relies on the additional flow G. For no reason it is desired to get different current arrival times for different G as this makes the definition not only meaningless but also inconsistent. The next example shows that this is the case if the flow model is not consistent, even for a relaxed definition of foresighted arrival times. Nevertheless, we see subsequently that foresighted arrival time functions are well-defined on consistent flow models. Note that FiFo is already required in Definition 3.37 because the upper arrival time $\hat{\ell}_P$ is only defined on such models. Further, for the existence of foresighted arrival time functions, it is quite obvious that $\mathcal{F}$-continuity is essential. Otherwise, for different $G_1, G_2 \in G_\theta$, the flows $\mathcal{F}^+ + G_1$ and $\mathcal{F}^+ + G_2$ can show a completely different behavior after time θ, even if the flow model satisfies FiFo, is past-oriented and G_1 and G_2 are close to each other. For these reasons we concentrate on the past-orientation in the following example.

Example 3.38. In this example we discuss three scenarios illustrated in Figure 3.10 which show that a unique foresighted arrival time function need not

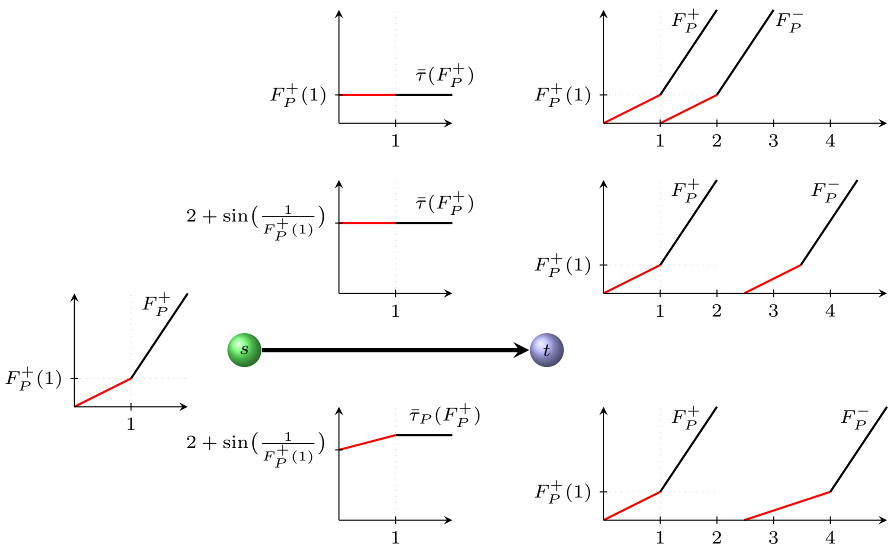


Figure 3.10: The network and the flow behavior of Example 3.38. The red color corresponds to flow entering $P := st$ over the time interval $[0, 1]$ and the black color corresponds to flow entering P afterwards.

exist for arbitrary flow models. The network in all three scenarios consists only of a single edge st and the unique s-t-path is denoted by $P := st$.

The *first scenario* corresponds to the diagrams at the top of Figure 3.10. Here, we consider a flow model where the transit time function of P is set to the amount of flow entering P until time 1, i.e., $\tau_P(\theta) := F_P^+(1)$ for all $\theta \in \mathbb{R}_+$. This flow model is clearly $\mathcal{F}$-continuous. On the other hand, it is not past-oriented because the transit time for a flow particle entering P at a point in time $\theta \in [0, 1)$ depends on the flow entering P over the time interval $[\theta, 1]$. The inflow F_P^+ is given by $f_P^+ := \frac{1}{2}\chi_{[0,1)} + \chi_{[1,\infty)}$. This means, flow enters P at a rate of $\frac{1}{2}$ until time 1 and at a rate of 1 afterwards. To see that the foresighted arrival time function is not well-defined, fix a $\theta \in [0, 1)$. From the definition of $\mathcal{G}_\theta$ we observe that additional flow $G \in \mathcal{G}_\theta$ can amount to any value of $\mathbb{R}_+$ until time 1. Thus, $(F_P^+ + G)(1)$ can take any value from $[F_P^+(1), \infty)$. Since (3.32) shows $\bar{\ell}_P(F_P^+)(\theta) = \theta + (F_P^+ + G)(1)$, this implies that the current foresighted arrival time is not unique.

As already mentioned, one motivation for defining foresighted arrival times arises out of the characterization of Nash equilibria. But instead of an arbitrary amount of additional flow, it is enough to consider how a small portion of additional flow changes the flow behavior. In this sense we relax the definition of the foresighted arrival time function as follows:

$$\bar{\ell}_P(\mathcal{F}^+)(\theta) \;:=\; \lim_{\epsilon \to 0} \hat{\ell}_P(\mathcal{F}^+ + G|_{\leq \theta+\epsilon})(\theta)\ . \qquad (3.33)$$

Here, the role of $G \in \mathcal{G}_\theta$ is changed. Intuitively, it only shows the direction in which the inflow of P is modified by G, whereas G represents a complete additional flow unit in (3.32). Since consistent flow models are past-oriented, both definitions coincide in a consistent environment. But using (3.33) instead of (3.32), the foresighted arrival time function of the above scenario is well defined. In fact, we have $\bar{\ell}_P(\mathcal{F}^+)(\theta) = F_P^+(1)$ for all $\theta \in \mathbb{R}_+$.

However, (3.33) does also not lead to well-defined current foresighted arrival times if the underlying flow model is not consistent. Even for a given $G \in \mathcal{G}_\theta$ the value $\bar{\ell}_P(\mathcal{F}^+)(\theta)$ can be undefined. This is observed in the *second scenario* shown in the middle of Figure 3.10. Here, we consider again a constant transit time function defined as $\tau_P(\mathcal{F}^+)(\theta) := 2 + \sin(\frac{1}{F_P^+(1)})$ if $F_P^+(1) > 0$ holds and $\tau_P(\mathcal{F}^+)(\theta) := 2$ if $F_P^+(1) = 0$ is valid. Consider the zero flow $\mathcal{F}^+ := 0$ and let $G \in \mathcal{G}_\theta$ be some additional flow entering P at a point in time $\theta \in [0, 1)$. Then the value $(\mathcal{F}^+ + G|_{\leq\theta+\epsilon})(1)$ can be any small enough positive real number depending on ϵ. Hence, as ϵ goes to zero, $\hat{\ell}_P(\mathcal{F}^+ + G|_{\leq\theta+\epsilon})(\theta)$ oscillates within the interval $[1, 3]$ implying that $\bar{\ell}_P(\mathcal{F}^+)(\theta)$ is undefined. As before, this flow model is not past-oriented, but, in addition, also not $\mathcal{F}$-continuous at the zero flow. Thus, it is not really a surprise that this flow model does not support foresighted arrival times.

In the *third scenario* of this example illustrated at the bottom of Figure 3.10, we modify the previous flow model such that it becomes $\mathcal{F}$-continuous, but still violates the past-orientation. For all times greater than 1, we set the current transit time to 3, i.e., $\tau_P(\mathcal{F}^+)(\theta) := 3$ for all $\theta \geq 1$. Further, we assume that if flow enters P over the time interval $[0, 1)$, the first flow arrives at t at time $2 + \sin(\frac{1}{F_P^+(1)})$. In order to complete the graph of the outflow F_P^-, we connect the points $(2 + \sin(\frac{1}{F_P^+(1)}), 0)$ and $(4, F_P^+(1))$ with a line segment. As before, the foresighted arrival time function is undefined over $[0, 1)$ for the zero flow. But in contrast, this flow model is $\mathcal{F}$-continuous.

Example 3.38 demonstrates that the foresighted arrival time function may not exist if the underlying flow model is not consistent. The next lemma proves that the foresighted arrival time function is well-defined, i.e., independent on special choices of G, if the underlying flow model is consistent. Further, it shows that the foresighted arrival time function is, in fact, an equivalent arrival time function for the underlying flow over time.

Lemma 3.39. *Consider a consistent flow model and let $\bar{\ell}$ denote the foresighted arrival time function of some given path P. Then $\bar{\ell}$ is well-defined in the sense that it is independent on the particular choice of G. Further, $\bar{\ell}$ represents the flow behavior on P, i.e.,*

$$F^+(\theta) = F^-\big(\bar{\ell}(\theta)\big) \tag{3.34}$$

where F^+ and F^- are the corresponding inflow and outflow function of P, respectively.

Proof. Let $\theta \in \mathbb{R}_+$ be some point in time. Before we prove that $\bar{\ell}$ is independent on the special choice of G, we show that $\bar{\ell}$ satisfies (3.34) for every $G \in \mathcal{G}_\theta$. For this, observe that applying Proposition 3.33 twice implies

$$\bar{\ell}(\theta) = \hat{\ell}\big((\mathcal{F}^+ + G)|_{\leq\vartheta}\big)(\theta) = \hat{\ell}\big(\mathcal{F}^+ + G|_{\leq\vartheta}\big)(\theta) \qquad \forall \vartheta > \theta \ .$$

Further, $\mathcal{F}^+ + G|_{\leq\vartheta}$ converges to $\mathcal{F}^+$ if ϑ goes to θ from right. Defining F_ϑ^- as the outflow function on P with respect to $\mathcal{F}^+ + G|_{\leq\vartheta}$ shows

$$\begin{aligned}
F^-\big(\bar\ell(\theta)\big) &= \lim_{\vartheta\searrow\theta} F^-_\vartheta\big(\bar\ell(\theta)\big) && (\mathcal{F}\text{-continuity})\\
&= \lim_{\vartheta\searrow\theta} F^-_\vartheta\big(\hat\ell(\mathcal{F}^+ + G|_{\leq\vartheta})(\theta)\big) && (\text{definition of } \bar\ell)\\
&= \lim_{\vartheta\searrow\theta}\big(F^+ + G|_{\leq\vartheta}\big)(\theta) && (\text{definition of } F^-_\vartheta)\\
&= F^+(\theta) && (\text{definition of } G)
\end{aligned}$$

and (3.34) is established. For proving that $\bar\ell$ is well-defined we observe that

$$F^-\big(\bar\ell(\theta)\big) \;=\; F^+(\theta) \;=\; \big(F^+ + G|_{\leq\vartheta}\big)(\theta) \;=\; F^-_\vartheta\big(\bar\ell(\theta)\big) \tag{3.35}$$

holds for all $\vartheta > \theta$ because of (3.34) and the definition of F^-_ϑ.

In order to prove that $\bar\ell(\theta)$ is independent on G, consider two functions G_1 and G_2 from $\mathcal{G}_\theta$, let

$$\bar\ell_1(\theta) = \hat\ell(\mathcal{F}^+ + G_1)(\theta) \qquad\text{and}\qquad \bar\ell_2(\theta) = \hat\ell(\mathcal{F}^+ + G_2)(\theta)\ ,$$

and assume without loss of generality $\bar\ell_1(\theta) < \bar\ell_2(\theta)$. For each $\vartheta > \theta$ define

$$g_\vartheta := g_1 + (g_2 - g_1)|_{[0,\vartheta]}\ .$$

Since $g_\vartheta = g_1|_{(\vartheta,\infty)} + g_2|_{[0,\vartheta]}$, we obtain $G_\vartheta \in \mathcal{G}_\theta$ and, hence, $\mathcal{F}^+ + G_\vartheta$ is a valid inflow. Further, since the flow model is past-oriented, we have

$$\hat\ell(\mathcal{F}^+ + G_\vartheta)(\theta) = \hat\ell(\mathcal{F}^+ + G_2)(\theta) = \bar\ell_2(\theta)\ .$$

Further, because of (3.35), this shows

$$F^-\big(\bar\ell_2(\theta)\big) = F^+(\theta) = \tilde F^-_\vartheta\big(\bar\ell_2(\theta)\big)$$

where $\tilde F^-_\vartheta$ is the outflow with respect to $\mathcal{F}^+ + G_\vartheta$. On the other hand, we know that $\mathcal{F}^+ + G_\vartheta$ converges to $\mathcal{F}^+ + G_1$ as ϑ goes from right to θ. So from the $\mathcal{F}$-continuity of the flow model, we obtain

$$F^-\big(\bar\ell_2(\theta)\big) = F^+(\theta) = F^-_1\big(\bar\ell_2(\theta)\big)\ .$$

where F^-_1 is the outflow with respect to $F^+ + G_1$. Applying (3.34) to G_1 yields $F^-_1(\bar\ell_1(\theta)) = F^-_1(\bar\ell_2(\theta))$. Because of the definition of the upper arrival time function in (3.21) applied to the flow $\mathcal{F}^+ + G_1$, this shows $\bar\ell_1(\theta) \geq \bar\ell_2(\theta)$ contradicting our assumption $\bar\ell_1(\theta) < \bar\ell_2(\theta)$. Hence, we have $\bar\ell_1(\theta) = \bar\ell_2(\theta)$ and $\bar\ell$ is well-defined. □

The following lemma shows that foresighted arrival times are upper semi-continuous and past-oriented. In this manner, they combine the upper semi-continuity of upper arrival times with the past-orientation of lower arrival times.

Lemma 3.40. *Consider a consistent flow over time model and let $\bar\ell$ be the foresighted arrival time of some path P. Then $\bar\ell$ is upper semi-continuous and past-oriented, i.e.,*

$$\bar\ell(\mathcal{F}^+)(\theta') = \bar\ell(\mathcal{F}^+|_{\leq\theta})(\theta') \qquad \forall\theta \geq \theta'\ . \tag{3.36}$$

Proof. Let $\mathcal{F}_i^+$ be a sequence of flows over time converging to some $\mathcal{F}^+$. We have to show that for all $\epsilon > 0$ there exists an $N \in \mathbb{N}$ such that

$$\bar{\ell}(\mathcal{F}_i^+)(\theta) \;\leq\; \bar{\ell}(\mathcal{F}^+)(\theta) + \epsilon \;.$$

Let $G \in \mathcal{G}_\theta$ be some additional flow. Then $\mathcal{F}_i + G$ converges to $\mathcal{F} + G$ and the previous inequality is equivalent to

$$\hat{\ell}(\mathcal{F}_i^+ + G)(\theta) \;\leq\; \hat{\ell}(\mathcal{F}^+ + G)(\theta) + \epsilon \;.$$

Since the upper arrival time is upper semi-continuous by Lemma 3.30, this inequality is valid implying that $\bar{\ell}$ is upper semi-continuous.

In order to observe the past-orientation of $\bar{\ell}$, consider an additional flow unit $G \in \mathcal{G}_{\theta'}$ for some point in time $\theta' \in \mathbb{R}_+$. Then by Proposition 3.33 we know for all $\theta \in \mathbb{R}_+$ with $\theta > \theta'$

$$\bar{\ell}(\mathcal{F}^+)(\theta') \;=\; \hat{\ell}(\mathcal{F}^+ + G)(\theta') \;=\; \hat{\ell}(\mathcal{F}^+|_{\leq\theta} + G)(\theta') \;=\; \bar{\ell}(\mathcal{F}^+|_{\leq\theta})(\theta')$$

because $(F^+ + G)(\theta) > (F^+ + G)(\theta')$ holds by the definition of G. It remains to prove equation (3.36) for $\theta' = \theta$. Using the kind of past-orientation which is already shown and the upper semi-continuity of $\bar{\ell}$ we obtain

$$\bar{\ell}(\mathcal{F}^+)(\theta) \;=\; \lim_{\vartheta\searrow\theta} \bar{\ell}(\mathcal{F}^+|_{\leq\vartheta})(\theta) \;\leq\; \bar{\ell}(\mathcal{F}^+|_{\leq\theta})(\theta) \;.$$

For proving the other relation note that $\bar{\ell}(\mathcal{F}^+)(\theta) = \lim_{\vartheta\searrow\theta} \check{\ell}(\mathcal{F}^+ + G)(\vartheta)$ holds for all points in time $\theta \in \mathbb{R}_+$ and for all flows over time $\mathcal{F}^+$ because of Proposition 3.22(iii). This finally shows

$$\begin{aligned}
\bar{\ell}(\mathcal{F}^+|_{\leq\theta})(\theta) &= \lim_{\vartheta\searrow\theta} \check{\ell}(\mathcal{F}^+|_{\leq\theta} + G)(\vartheta) \\
&\leq \lim_{\vartheta\searrow\theta} \big(\liminf_{\vartheta_1\searrow\theta} \check{\ell}(\mathcal{F}^+|_{\leq\vartheta_1} + G)(\vartheta)\big) && \text{(semi-continuity of } \check{\ell}) \\
&\leq \lim_{\vartheta\searrow\theta} \big(\liminf_{\vartheta_1\searrow\theta} \check{\ell}(\mathcal{F}^+|_{\leq\vartheta_1} + G)(\vartheta_1)\big) && \text{(monotonicity of } \check{\ell}) \\
&= \liminf_{\vartheta_1\searrow\theta} \check{\ell}(\mathcal{F}^+|_{\leq\vartheta_1} + G)(\vartheta_1) \\
&= \lim_{\vartheta_1\searrow\theta} \check{\ell}(\mathcal{F}^+ + G)(\vartheta_1) && \text{(past-orientation of } \check{\ell}) \\
&= \bar{\ell}(\mathcal{F}^+)(\theta) \;.
\end{aligned}$$
□

We conclude this section with some basic properties of the foresighted arrival time functions.

Proposition 3.41. *Let $\bar{\ell}$ be a foresighted arrival time function and let F^+ and F^- be corresponding inflow and outflow functions, respectively. Then the following statements hold:*

(i) $F^+(\theta) = F^-(\bar{\ell}(\theta))$ *for all* $\theta \in \mathbb{R}_+$.

(ii) The foresighted arrival time function $\bar{\ell}$ is nondecreasing.

(iii) The foresighted arrival time function is right continuous.

Proof. Statement (i) is proven in Lemma 3.39. To show (ii), let $\theta_2 > \theta_1 \geq 0$. Since F^+ is nondecreasing, statement (i) implies $F^-(\bar{\ell}(\theta_2)) \geq F^-(\bar{\ell}(\theta_1))$. In case of $F^-(\bar{\ell}(\theta_2)) > F^-(\bar{\ell}(\theta_1))$ we directly obtain $\bar{\ell}(\theta_2) > \bar{\ell}(\theta_1)$ because F^- is nondecreasing. Therefore, assume that $F^-(\bar{\ell}(\theta_2)) = F^-(\bar{\ell}(\theta_1))$ and let $G_1 \in \mathcal{G}_{\theta_1}$ and $G_2 \in \mathcal{G}_{\theta_2}$ be additional flow which is sent from time θ_1 and time θ_2 on, respectively. For a point in time $\vartheta > \theta_1$ we define $G_\vartheta := G_1|_{[0,\vartheta]} + G_2$. Then we have $G_\vartheta \in \mathcal{G}_{\theta_1}$ for all $\vartheta > \theta_1$. Now let F_2^- be the outflow of $F^+ + G_2$ and F_ϑ^- be the outflow of $F^+ + G_\vartheta$ for all $\vartheta > \theta_1$. Then we have

$$\begin{aligned}
F_2^-\big(\bar{\ell}(\theta_1)\big) &= \lim_{\vartheta \searrow \theta_1} F_\vartheta^-\big(\bar{\ell}(\theta_1)\big) && (\mathcal{F}\text{-continuity}) \\
&= \lim_{\vartheta \searrow \theta_1} F^+(\theta_1) && (\text{because of (i)}) \\
&= F^+(\theta_1) = F^+(\theta_2)\ .
\end{aligned}$$

But this implies $\bar{\ell}(\theta_1) \leq \bar{\ell}(\theta_2)$ because of (3.32) and the maximum in (3.21) which proves (ii).

For proving (iii), we have to show $\lim_{\theta' \searrow \theta} \bar{\ell}(\theta') = \bar{\ell}(\theta)$ for all $\theta \in \mathbb{R}_+$. From (ii) we get that $\lim_{\theta' \searrow \theta} \bar{\ell}(\theta')$ exists and that $\lim_{\theta' \searrow \theta} \bar{\ell}(\theta') \geq \bar{\ell}(\theta)$ holds. In order to show $\lim_{\theta' \searrow \theta} \bar{\ell}(\theta') \leq \bar{\ell}(\theta)$, let $G \in \mathcal{G}_\theta$. Then $G_{\theta'} := G + (\theta - \theta') \in \mathcal{G}_{\theta'}$ holds for all $\theta' \geq \theta$. Further, let F_ϑ^- be the outflow with respect to $F^+ + G_\vartheta$ for all $\vartheta > \theta$. Now it follows from (ii) and the definition of $\bar{\ell}$ in (3.32) that for all $\vartheta > \theta$

$$F_\vartheta^-\big(\lim_{\theta' \searrow \theta} \bar{\ell}(\theta')\big) \;\leq\; F_\vartheta^-\big(\bar{\ell}(\vartheta)\big) \;=\; (F^+ + G_\vartheta)(\vartheta) \;=\; F^+(\vartheta)\ .$$

Now let ϑ tend to θ from right. Hence, the $\mathcal{F}$-continuity of the flow model and the continuity of F^+ imply

$$F_\theta^-\big(\lim_{\theta' \searrow \theta} \bar{\ell}(\theta')\big) \;\leq\; F^+(\theta) \;=\; (F^+ + G)(\theta)\ .$$

Finally, recalling the definition of $\bar{\ell}(\theta)$, this together with the maximum in (3.21) shows

$$\lim_{\theta' \searrow \theta} \bar{\ell}(\theta') \;\leq\; \bar{\ell}(\theta)\ .$$

Thus, $\bar{\ell}$ is right continuous. $\square$

3.5 Dynamic Network Loading and Path Decomposition

In this subsection, we briefly discuss the dynamic network loading problem and the flow over time decomposition problem. Since these two questions only touch the main focus of this thesis – routing games over times –, we avoid a deep presentation and omit some technical details. Rather and in contrast to all other parts of this thesis, we only make use of intuition whenever possible.

The starting point of this subsection is an *edge-based* flow over time model. On the other hand, as we see in Chapter 4, path-based flows over time are the fundamental ingredient when analyzing Nash equilibria over time. Therefore, we have to answer the following two questions which are inverse to each other. First, given a path-based flow, how does this flow evolve over the edges using the edge-based flow model. Second, given an edge-based flow over time, along which s-t-paths flow is sent. The first question results in the dynamic network loading problem which, in fact, builds a path-based flow model upon a given edge-based one. The second question results in the flow decomposition problem which enables us to follow a particular flow particle through the network.

For formalizing this, we consider a subproblem which appears in both problems. So let $(F_P)_{P\in\mathcal{P}}$ and $(F_e^+, F_e^-)_{e\in E}$ be a path-based and an edge-based flow over time, respectively. Given some consistent edge-based flow model, we consider the question whether or not $(F_P)_{P\in\mathcal{P}}$ and $(F_e^+, F_e^-)_{e\in E}$ represent the same flow over time $\mathcal{F}$. For answering this question, let τ_e and ℓ_e be the foresighted transit and arrival time function of an edge $e \in E$ with respect to $(F_e^+, F_e^-)_{e\in E}$. Next, we consider a path $P := (e_1, \dots, e_{|P|})$ and let $F_{P,i}^+$ and $F_{P,i}^-$ be the inflow and outflow function on the i-th edge of P caused by the flow F_P on P, respectively. Following the flow along P, the functions $F_{P,i}^+$ and $F_{P,i}^-$ are recursively computable by

$$F_{P,1}^+ := F_P, \quad F_{P,i}^-(\theta) := \int_{\ell_{e_i}^{-1}([0,\theta])} f_{P,i}^+(\vartheta)\, d\vartheta, \quad \text{and} \quad F_{P,i+1}^+ := F_{P,i}^- \tag{3.37}$$

having in mind (3.8). On the other hand, we know that starting from s along P at time θ we arrive at time $\ell_{P,i}(\theta) := \theta + \sum_{j=1}^{i} \tau_{e_j}(\ell_{P,j-1}(\theta))$ at the head of e_i (which equals the tail of e_{i+1} by definition). Thus, setting $\ell_{P,0}(\theta) := \theta$, the functions $F_{P,i}^+$ and $F_{P,i}^-$ are also directly computable by

$$F_{P,i+1}^+(\theta) = F_{P,i}^-(\theta) = \int_{\ell_{P,i}^{-1}([0,\theta])} f_P^+(\vartheta)\, d\vartheta\ . \tag{3.38}$$

Of course, if $(F_P)_{P\in\mathcal{P}}$ and $(F_e^+, F_e^-)_{e\in E}$ represent the same flow over time $\mathcal{F}$, the inflow functions $F_{P,i}^+$, where the i-th edge e_i^P of P equals some given edge e, have to sum up to the edge inflow function F_e^+, i.e.,

$$F_e^+ = \sum_{P,i|e_i^P=e} F_{P,i}^+ \qquad \forall e \in E\ . \tag{3.39}$$

This means that the routing behavior of $(F_P)_{P\in\mathcal{P}}$ over the edges of the network is given by $(F_e^+, F_e^-)_{e\in E}$. In addition, (3.39) ensures that (3.37) is well-defined as the foresighted arrival time function of an edge e depends on the (total) inflow function F_e^+. Further, comparing (3.37) with the definition of the edge outflow function in (3.8) we obtain $F_e^- = \sum_{P,i|e_i^P=e} F_{P,i}^-$ because the integral operator works linear for nondecreasing sequences as in (3.39). The discussion above motivates the following definition.

Dynamic Network Loading Problem

Input:	A consistent edge-based flow over time model with foresighted arrival times $(\ell_e)_{e\in E}$ and a path-based flow over time $(F_P)_{P\in\mathcal{P}}$ with $\sum_{P\in\mathcal{P}} |P| \cdot F_P(\theta) < \infty$ for all $\theta \in \mathbb{R}_+$.
Task:	An edge-based flow $(F_e^+, F_e^-)_{e\in E}$ representing $(F_P)_{P\in\mathcal{P}}$.

Note that the input only contains the edge arrival times as maps and not as functions. In fact, the dynamic network loading problem has to construct the corresponding edge arrival time functions simultaneously with an edge-based flow. Further, (3.39) only makes sense if we have $\sum_{P \in \mathcal{P}} |P| \cdot F_P(\theta) < \infty$ for all $\theta \in \mathbb{R}$ as required for the input of the dynamic network loading problem. Otherwise at least one edge $e \in E$ has to carry an infinite amount of flow or an infinite current foresighted arrival time in finite time contradicting the definition of an edge-based flow over time.

Lemma 3.42. *Consider an edge-based flow over time model. Then the dynamic network loading problem always admits a feasible solution.*

Proof. We have to show that the flow behavior of a given path-based flow over time $(F_P)_{P \in \mathcal{P}}$ can be represented by an edge-based flow over time. For this let $\mathcal{P}^+$ be the set of s-t-paths which carry flow, i.e., $\mathcal{P}^+ := \{P \in \mathcal{P} \mid F_P \neq 0\}$. It is rather obvious that there exists a corresponding edge-based flow over time if the paths in $\mathcal{P}^+$ form an acyclic network. In this case consider the edges in some topological order and route the entire inflow F_e^+ along the current edge e as it is required by the map ℓ_e. Due to our assumption, the inflow F_e^+ given by (3.39) is already fixed when considering e. Hence, the solution of the dynamic network loading problem is also unique for this scenario.

The dynamic network loading problem is also feasible if there exists an $\epsilon > 0$ such that the underlying edge-based flow model does not support arrival time functions leading to cycles with a transit time smaller than ϵ. That is, for every edge-based flow over time $(\mathcal{F}^+, \mathcal{F}^-)$, we have $\tau_C(\mathcal{F}^+, \mathcal{F}^-) \geq \epsilon$ for all cycles C. To see this, fix an arbitrary order of the edges. Consider one edge e after another and route the current inflow F_e^+ over e respecting the map ℓ_e. The nature of the Bellman Ford algorithm for computing shortest paths results in the following observation. Applying this procedure kn times, the inflow of each edge e restricted to the time interval $[0, k\epsilon]$ is unique and remains constant over all subsequent calls of this procedure. In this manner, we are able to construct the unique solution of the dynamic network loading problem.

If the underlying edge-based flow model does not fit the previous assumptions, we make a small detour in order to find a feasible solution of the dynamic network loading problem. Instead of considering the given edge-based flow model, we consider the model using a family $\tilde{\mathcal{T}}$ of arrival times defined by $\tilde{\tau}_e(\mathcal{F}^+, \mathcal{F}^-)(\theta) = \max\{\epsilon, \tau_e(\mathcal{F}^+, \mathcal{F}^-)(\theta)\}$ for some given $\epsilon > 0$. Clearly, these flow models do not support cycles with a transit time smaller than ϵ. So the approach in the previous paragraph shows that, for all $i \in \mathbb{N}$, there exists an edge-based flow over time $(\mathcal{F}_i^+, \mathcal{F}_i^-)$ which is a solution of the dynamic shortest path problem for the modified edge-based flow model with $\epsilon := \frac{1}{i}$. Now Helly's Selection Theorem and, in particular, Lemma 2.36 show that $(\mathcal{F}_i^+, \mathcal{F}_i^-)_{i \in \mathbb{N}}$ contains a convergent subsequence because the number of edges is finite. For this note that the inflow functions of each edge e restricted to some interval $[0, \theta]$ are uniformly bounded by $\sum_{P \in \mathcal{P}} |P| \cdot F_P(\theta) < \infty$. The limit of that subsequence is a solution of the dynamic network loading problem on the original instance as the modified arrival time functions converge uniformly to the given initial arrival times. $\square$

Lemma 3.42 shows that the dynamic network loading problem is always feasible in the sense that there always exists a feasible solution. In addition,

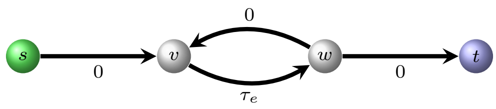


Figure 3.11: A network where each edge except $e = vw$ has a constant transit time of 0.

if the underlying edge-based flow model does not support cycles of zero transit times, the solution is also unique. The following example shows that this is no longer the case if the edge-based flow model allows cycles of zero transit times.

Example 3.43. During this example let $\mathrm{id} : \mathbb{R}_+ \to \mathbb{R}_+$ be the identity function, i.e., $\mathrm{id}(\theta) = \theta$ for all $\theta \in \mathbb{R}_+$. We consider the network depicted in Figure 3.11. The transit time of each edge except of the crucial edge $e = vw$ is constant and set to 0. Further, the transit time of edge e is given by

$$\tau_e(F_e^+) \ := \ \max\bigl\{0,\ 6 \cdot (\mathrm{id} - F_e^+)\bigr\} \ .$$

This means that the transit time of e equals 0 if the inflow function F_e^+ is greater than the identity function. In contrast, for the points in time at which F_e^+ is not greater than the identity function the current transit time is equal to six times the vertical distance between F_e^+ and the identity function.

In the following we consider the dynamic network loading problem which sends flow at a rate of $\frac{1}{2}$ into the s-t-path P which uses the cycle $\{e, wv\}$ exactly once, i.e., $P := \{sv, e, wv, e, wt\}$, $F_P := \frac{1}{2}\mathrm{id}$, and $F_{P'} := 0$ for all other s-t-paths P'. We show that this dynamic network loading problem has at least two solutions.

We obtain the first solution by assuming $\tau_e = 0$. Using (3.37) this implies $F_{P,i}^+ = F_P$ for all $i = 1, 2, 3, 4, 5$. Because of (3.39) this results in the edge-based flow over time given by

$$F_{sv}^+ \ = \ F_{wv}^+ \ = \ F_{wt}^+ \ = \ F_P \ = \ \frac{1}{2}\mathrm{id} \qquad \text{and} \qquad F_e^+ \ = \ 2 \cdot F_P \ = \ \mathrm{id} \ .$$

For checking that this edge-based flow over time is a feasible solution, we have to verify that the transit time function of e is always 0. Since the inflow F_e^+ is equal to identity function, this is the case and we have found a feasible solution for the dynamic network loading problem.

For the second solution we assume that the transit time function τ_e is equal to the identity function, i.e., $\tau_e = \mathrm{id}$. Using (3.37) this implies

$$F_{P,1}^+ \ = \ F_{P,2}^+ \ = \ \frac{1}{2}\mathrm{id}, \qquad F_{P,3}^+ \ = \ F_{P,4}^+ \ = \ \frac{1}{3}\mathrm{id}, \qquad \text{and} \qquad F_{P,5}^+ \ = \ \frac{1}{4}\mathrm{id} \ .$$

Because of (3.39) this results in the edge-based flow over time given by

$$F_{sv}^+ \ = \ \frac{1}{2}\mathrm{id}, \qquad F_e^+ \ = \ \frac{5}{6}\mathrm{id}, \qquad F_{wv}^+ \ = \ \frac{1}{3}\mathrm{id}, \qquad \text{and} \qquad F_{wt}^+ \ = \ \frac{1}{4}\mathrm{id} \ .$$

For checking that this edge-based flow over time is a feasible solution, we have to verify that the transit time function of e is equal to id. In fact, this is the case as $\tau_e(F^+) = \max\bigl\{0,\ 6 \cdot (\mathrm{id} - \frac{5}{6}\mathrm{id})\bigr\} = \mathrm{id}$. Thus, we have found a second feasible solution for the dynamic network loading problem.

The previous approach shows that there exist at least two different solutions to this dynamic network loading problem. Unfortunately, there exist many more. On closer inspection, we see that for each $\theta \in \mathbb{R}_+$ a solution results out of setting $\tau_e|_{[0,\theta]} := 0$ and $\tau_e|_{[\theta,\infty]} := \mathrm{id} - \theta$. That is, until time θ we route the flow on P as in the first solution and after that we use the second solution in order to evolve this flow over the edges. Note that using the second solution before the first solution results not in a feasible edge-based flow over time.

As already mentioned, the dynamic network loading problem works as a definition of a path-based flow model which is built upon an edge-based flow model. In the following, we consider which consistency properties carry over to the path-based flow model in case the edge-based model is consistent. Clearly, the FiFo principle remains true for the path-based flow over time model as the sum of nondecreasing functions remains nondecreasing. In addition, the following Lemma shows that one can assume that the dynamic network loading problem preserves $\mathcal{F}$-continuity.

Lemma 3.44. *Consider a consistent edge-based flow over time model. The dynamic network loading problem behaves continuously. This means a path-based flow model which is built upon a consistent edge-based flow model can be assumed to be $\mathcal{F}$-continuous.*

In fact, if there exists an $\epsilon > 0$ such the underlying edge-based model does not support cycles with a transit time smaller than ϵ, the underlying path-based flow over time model is $\mathcal{F}$-continous. Otherwise every convergent sequence of path-based flows over time with limit point $\mathcal{F}_*$ contains a subsequence such that the corresponding families of outflow functions converge to a family of outflow functions of $\mathcal{F}_*$.

Proof. As in the proof of Lemma 3.42, we first assume that the underlying edge-based flow model does not support cycles with a transit time less than a given number $\epsilon > 0$. By induction over a natural $k \in \mathbb{N}$ we observe the following. Since the edge-based flow model is $\mathcal{F}$-continuous, each edge inflow restricted to the time interval $[0, k\epsilon]$ behaves continuously in $(F_P)_{P\in\mathcal{P}}$. This follows directly from our assumption which implies that the directed graph containing an edge vw for $v, w \in V$ if and only if $\ell_{vw}((k-1)\epsilon) < k\epsilon$ is acyclic. So using a topological order of this graph shows, again inductively, the induction step of the major induction.

For the general case, we consider a sequence $(\mathcal{F}_i)_{i\in\mathbb{N}}$ of path-based flows over time which converges to some $\mathcal{F}_*$. Lemma 3.42 shows that, for each $\mathcal{F}_i$, there exists a corresponding edge-based flow over time $(\mathcal{F}_i^+, \mathcal{F}_i^-)$. Since the number of edges is finite, there exist a convergent subsequence of $(\mathcal{F}_i^+, \mathcal{F}_i^-)_{i\in\mathbb{N}}$ by Lemma 2.36. Let $(\mathcal{F}_*^+, \mathcal{F}_*^-)$ be the corresponding limit point. Then the $\mathcal{F}$-continuity of the edge-based flow model directly implies that $(\mathcal{F}_*^+, \mathcal{F}_*^-)$ represents the path-based flow $\mathcal{F}_*$. □

Lemma 3.44 shows that a path-based flow model which is based on a consistent edge-based flow model is somehow $\mathcal{F}$-continuous. Unfortunately, $\mathcal{F}$-continuity in terms of Definition 3.24 is not obtained in general. This is mainly due to the fact that a solution to the dynamic network loading problem need not to be unique. Nevertheless, it turns out that the continuity property of Lemma 3.44 is enough for the purposes of this thesis.

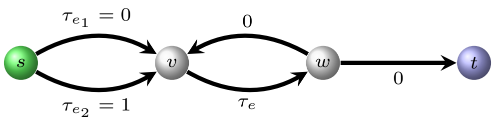


Figure 3.12: A network illustrating pathological behavior of the dynamic network loading problem explained in Example 3.45. Constant transit times are shown on the edges.

The following example shows that this is not the whole truth. Here, we consider a scenario, which contradicts our intuition behind continuity. The main reason for that is related to the question whether or not the set $\mathcal{P}$ of s-t-paths is locally finite with respect to the dynamic network loading problem. We give a particular answer to this question directly after the example. In addition, this example shows that a path-based flow model based on the dynamic network loading problem is not past-oriented in general.

Example 3.45. In this example we consider the network shown in Figure 3.12. The transit of each edge except e is constant. For a point in time $\theta \in \mathbb{R}_+$ the current arrival time on e is defined as $\ell_e(F_e^+)(\theta) = \max\{\theta, F_e^+(\theta)\}$. So the outflow function F_e^- of e is equal to the identity function if the inflow F_e^+ is not less than the identity function.

For each $i \in \mathbb{N}$ let P_i be the s-t-path which traverses e_1 and i times the cycle along the edges e and wv. Now we consider the following network loading problems. Each of these sends flow at a rate of 1 over the time interval $[0, 1]$ along the simple s-t-path $P := (e_2, e, wt)$. In addition, the i-th problem sends flow at a rate of $\frac{1}{i}$ over the time interval $[0, 1]$ along the path P_{i^2}. So the i-th network loading problem is given by the path-based flow over time $\mathcal{F}_i := ((P_{i^2}, \frac{1}{i}\mathrm{id}|_{[0,1]}), (P, \mathrm{id}|_{[0,1]}))$. It is not hard to see that this sequence converges to the flow over time $\mathcal{F}_* := (P, \mathrm{id}|_{[0,1]})$ if i goes to ∞. This leads to an outflow function on P which is equal to $\mathrm{id}|_{[0,1]} - 1$. That is, flow arrives at t over the time interval $[1, 2]$ at a rate of 1.

But what is the outflow function which we obtain as a limit of the network loading problems. So fix an $i \in \mathbb{N}$. First observe that the path P_{i^2} sends in total a flow value of i over the edge e. Further, assuming that the transit time of e equals 0 this happens at a rate of i over the time interval $[0, 1]$. However, this causes intuitively that the outflow of e equals the identity function. So flow on P_{i^2} leaves e at a rate of 1. Now, as i goes to ∞, nearly all of this flow uses the cycle again. Hence, in the limit the path P_{i^2} forces that flow enters e at a rate of 1. Together with the flow on P this leads to an inflow on e of $F_e^+ := \mathrm{id} + \mathrm{id}|_{[1,2]}$. Thus, flow entering e at time 1 leaves e at time 2 and flow entering e at time 2 leaves e at time 3. This shows that the outflow of P equals $\frac{1}{2}\mathrm{id}|_{[0,2]} - 1$. That is, flow arrives at t along P over the time interval $[1, 3]$ at a rate of $\frac{1}{2}$.

Beside this, each dynamic network loading problem shows the following behavior which is not compatible with the definition of past-orientation. The flow behavior on P depends on the entire flow on P_{i^2}. In particular, the behavior of the flow entering P over the time interval $[0, \frac{1}{2}]$ depends on the behavior of the flow entering P_{i^2} over the time interval $[\frac{1}{2}, 1]$. So flow originating later influences flow originating earlier which is actually forbidden for past-oriented path-based flow models.

The main reason for the contradictorily flow behavior is that flow originating until time 1 can be send along paths with an arbitrary number of edges. To overcome this problem we assume the following. The path-based flow defining a dynamic network loading problem must send the flow originating until a given point in time θ over s-t-paths with a bounded number of edges. That is, for each θ there exist an M_θ such that $F_P|_{[0,\theta]} = 0$ for all $P \in \mathcal{P}$ with $|P| > M_\theta$. With this assumption the proof of Lemma 3.44 works. In addition, this also ensures that $\mathcal{P}$ is locally finite with respect to the dynamic network loading problem.

In the following, we turn our attention to the flow decomposition problem. As already mentioned, this problem can be seen as the inverse of the dynamic network loading problem. Unfortunately, as in static flow theory, a decomposition of a flow over time contains flows on cycles in general.

Definition 3.46 (Flow Decomposition). Let $(\mathcal{F}^+, \mathcal{F}^-)$ be an edge-based flow over time. A family $\mathcal{F} := (F_P)_{P \in \mathcal{P} \cup \mathcal{C}}$ of path and cycle flows is called *flow decomposition* if and only if condition (3.39) is satisfied for all edges $e \in E$ using (3.37). Furthermore, the flow on every cycle $C \in \mathcal{C}$ must be a circulation, i.e., $F^-_{C,|C|} = F^+_{C,1}$.

The edge-based flow over time $(\mathcal{F}^+, \mathcal{F}^-)$ is called *path decomposable* if and only if there exists a flow decomposition $\mathcal{F} := (F_P)_{P \in \mathcal{P} \cup \mathcal{C}}$ with $F_P = 0$ for all cycles $P \in \mathcal{C}$. In this case $\mathcal{F}$ is called *path decomposition.*

Path decompositions are quite important when analyzing Nash flows over time arising out of egoistic routing behavior of the flow particle. The reason for that is that in routing games over time every flow particle has to start at the given source s. This implies that flow on any edge must be ascribable to flow originating at s using paths starting at s. The following lemma characterizes such path decomposable flows.

Lemma 3.47. *Let $(\mathcal{F}^+, \mathcal{F}^-)$ be an edge-based, flow decomposable flow over time and let $\mathcal{F} := (F_P)_{P \in \mathcal{P} \cup \mathcal{C}}$ be a corresponding flow decomposition. Then $(\mathcal{F}^+, \mathcal{F}^-)$ is path decomposable if and only if the flow on every cycle $C \in \mathcal{C}$ is covered by path flows, i.e.,*

$$\operatorname{supp}(F_C) \subseteq \bigcup_{P,j \mid v_j^P \in V(C)} \operatorname{supp}(F_{P,j}) \ . \tag{3.40}$$

Proof. Consider a cycle C and an s-t-path P, such that there exists a common node $v \in V(C) \cap V(P)$. Further, assume that v is the j-th node of P and let $S := \operatorname{supp}(F_{P,j})$ be the set of points in time at which flow is send through v along P. By routing the flow arriving at v via P first along C before this flow enters the next edge of P, we are able to reduce the flow on C to zero. For this we are allowed to travel along the cycle C several times. In fact, we have use the cycle C even an infinite but uncountable number of times, in general.

Intuitively, this works as follows. Let θ be some point in time and assume that flow enters or leaves v along C or P at a rate of $f_C(\theta)$ and $f_{P,j}(\theta)$. Define $k := \lceil \frac{f_C(\theta)}{f_{P,j}(\theta)} \rceil$ and $f := \frac{f_C(\theta)}{k} \leq f_{P,j}(\theta)$. Now we reduce the flow on C at time θ to 0 by sending flow on P at a rate of f around C for k times. That is, we reduce the flow on P by f and increase the flow on the s-t-path consisting of P and k times the cycle C by f. $\square$

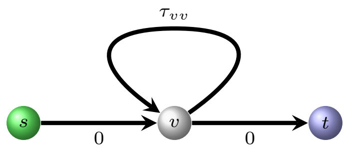


Figure 3.13: Network of Example 3.48 and 3.49. In Example 3.48 the current transit time of the loop is always equal to 1. In the second example this transit time is nonincreasing.

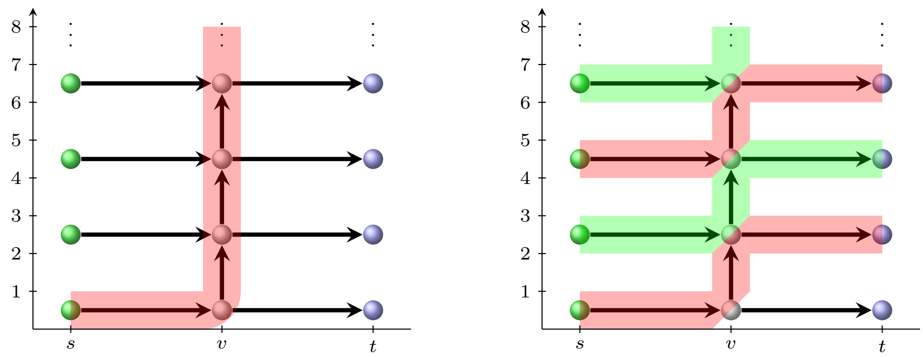


Figure 3.14: Flows considered in Example 3.48.

An essential difference to the dynamic loading problem is that the edge arrival time function are known because the edge-based flow over time is given. From this point of view the flow decomposition problem may be considered as easier than the dynamic network loading problem.

Flow Decomposition Problem	
Input:	An edge-based flow over time $(F_e^+, F_e^-)_{e\in E}$
Task:	A flow decomposition $(F_P)_{P\in\mathcal{P}\cup\mathcal{C}}$ of $(F_e^+, F_e^-)_{e\in E}$.

The following examples illustrate problems which can occur when computing a flow decomposition. Intuitively, a flow decomposition should always exist. However, Example 3.48 shows that a flow decomposition does not exist in general using the current definition of edge-based flows over time. Even if a flow decomposition exists, sending a wrong amount of flow along a wrong path can lead to a dead end. This means that the remaining flow has no flow decomposition. Therefore, an algorithm which searches for a flow decomposition must be carefully designed in order to output a flow decomposition whenever one exist. The second example shows that a flow decomposition can consist of an infinite number of path, even if flow is only sent over a finite time interval. In this sense, an algorithm can only find an approximation in finite time. However, the goal is to close the gap the longer the algorithm runs. The underlying network of both examples is shown in Figure 3.13.

Example 3.48. Considering the network in Figure 3.13, we assume that the current transit time of the loop is always equal to 1, i.e., $\tau_{vv}(\theta) = 1$ for all $\theta \in \mathbb{R}_+$. In the following, we consider two edge-based flows over time. The first scenario sends flow at a rate of 1 over the time interval $[0, 1)$ into the edge sv which,

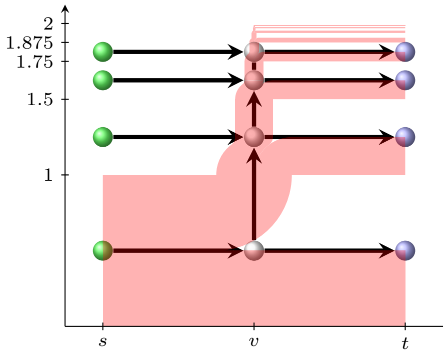


Figure 3.15: Flow over considered in Example 3.49.

subsequently, travels along the loop vv for an infinite number of times (see the left part of Figure 3.14). That is, the edge-based flow is given by

$$f^+_{sv} = f^-_{sv} = \chi_{[0,1)}, \quad f^+_{vv} = \chi_{[0,\infty)}, \quad f^-_{vv} = \chi_{[1,\infty)}, \quad \text{and} \quad f^+_{vt} = f^-_{vt} = 0 \; .$$

Although this is a feasible edge-based flow over time in terms of Definition 3.9, there exist, of course, no corresponding flow decomposition.

For the second scenario we send, in addition, flow over the simple path svt at a rate of 1 over every time interval $[2i, 2i+1)$ with $i \in \mathbb{N}_{>0}$. This results in the following edge-based flow over time,

$$f^+_{sv} = f^-_{sv} = \sum_{i=0}^{\infty} \chi_{[2i,2i+1)}, \qquad f^+_{vv} = \chi_{[0,\infty)}, \qquad f^-_{vv} = \chi_{[1,\infty)},$$

$$\text{and} \qquad f^+_{vt} = f^-_{vt} = \sum_{i=0}^{\infty} \chi_{[2i+1,2i+2)} \; .$$

A closer look at this scenario shows that this edge-based flow has a flow decomposition. In fact, we can send the entire flow along the s-t-path $P := (sv, vv, vt)$ which uses the loop exactly once. For this we have to set the inflow function of P to $f_P = \sum_{i=0}^{\infty} \chi_{[2i,2i+1)}$ (see the right part of Figure 3.14).

In addition, the second scenario shows that we have to be careful when computing a flow decomposition. If we send as much flow as possible along the simple path (sv, vt) – as a first step of some flow decomposition algorithm – then we end up in the first scenario. This shows that we do not obtain a flow decomposition in this manner even though one exists. Note that *static* flows do not show such a controversial behavior.

Example 3.48 shows that iteratively routing as much flow as possible along an s-t-path with the minimum number of edges may not result in a feasible flow decomposition. The next example shows that also choosing an s-t-path with respect to the currently smallest transit or arrival time is not a good idea in general.

Example 3.49. The underlying network shown in Figure 3.13 is the same as in Example 3.48. The idea behind the edge-based flow over time illustrated in Figure 3.15 is the following. The flow originates at a rate of 1 over the time interval $[0, 1)$ and enters the edge sv. Further, this flow travels along the loop vv, but each time the node v is visited, the first half of the current flow leaves the cycle and travels along vt to t. In addition, the current transit time of the loop vv is halved for the second half of the current flow which enters the loop again. For this we initialize the transit time of vv with $\frac{1}{2}$.

So the routing behaves as follows. After the flow enters the edge sv, it arrives at v for the first time at a rate of 1 over the time interval $[0, 1)$. The first half of this flow arriving at v over the time interval $[0, \frac{1}{2})$ travels directly to t along vt. The second half enters v over the time interval $[\frac{1}{2}, 1)$ and proceeds via the loop. Since the loop has currently a transit time of $\frac{1}{2}$, this flow revisits v over the time interval $[1, 1 + \frac{1}{2})$. Again, the first half of this flow, which enters v over $[1, 1 + \frac{1}{4})$, is send directly to t via vt. For the second half, which arrives at v over $[1 + \frac{1}{4}, 1 + \frac{1}{2})$, we halve the transit time of the loop to $\frac{1}{4}$. Hence, this flow arrives at v for the third time over $[1 + \frac{1}{2}, 1 + \frac{1}{2} + \frac{1}{4})$.

Iterating the procedure above over and over again results edge-based flow over time representable by

$$
\begin{aligned}
& f_{sv}^{+} := f_{sv}^{-} := \chi_{[0,1)} \ , \\
& f_{vv}^{+} := \sum_{i=0}^{\infty} \chi_{\left[2-\left(\frac{1}{2}\right)^{k}-\left(\frac{1}{2}\right)^{k+1},\, 2-\left(\frac{1}{2}\right)^{k}\right)} \ , \qquad f_{vv}^{-} := \chi_{[1,2)} \ , \\
\text{and} \qquad & f_{vt}^{+} := f_{vt}^{-} := \sum_{i=0}^{\infty} \chi_{\left[2-\left(\frac{1}{2}\right)^{k-1},\, 2-\left(\frac{1}{2}\right)^{k}-\left(\frac{1}{2}\right)^{k+1}\right)} \ .
\end{aligned}
$$

Further, the transit time function of the loop vv is given by

$$
\tau_{vv}\big|_{\left[2-\left(\frac{1}{2}\right)^{k}-\left(\frac{1}{2}\right)^{k+1},\, 2-\left(\frac{1}{2}\right)^{k}\right)} := \left(\tfrac{1}{2}\right)^{k+1} \qquad \forall k \in \mathbb{N} \ .
$$

At times where τ_{vv} is not set it can be defined arbitrarily. It is not hard to verify that this defines an edge-based flow over time in terms of Definition 3.9.

For obtaining a flow decomposition for this example, let P_k be the s-t-path containing the loop k times for all $k \in \mathbb{N}_0$. Then the first half of the entire flow is sent along P_0, the first half of the remaining flow is sent along P_1, and so on. This shows that a flow decomposition is defined by

$$
f_{P_i} = \chi_{\left[1-\left(\frac{1}{2}\right)^{k},\, 1-\left(\frac{1}{2}\right)^{k+1}\right)} \qquad \forall i \in \mathbb{N} \ .
$$

Now assume that, in addition, some new flow is sent through the network after time 2. Further, let the transit time of each edge be greater than 2 after time 2. Clearly, if an algorithm for finding a flow decomposition always choose a path with the currently minimal transit or arrival time, it never decompose the new flow.

The following algorithm resolves the problems explained in Example 3.48 and 3.49. First it eliminates all flow carrying cycles. Note that we only have to consider simple cycles and the number of simple cycles is finite. Subsequently, it decomposes nearly all flow originating until a given point in time θ. That is, it routes flow originating until time θ along paths until the the remaining flow originating until time θ is smaller than some given number $\epsilon > 0$. This can be done using a finite number of path only. Then the algorithm increases θ, decreases ϵ, and starts decomposing the remaining flow originating until time θ

again. In this manner, we iteratively construct a sequence flow decompositions which converges to a flow decomposition of a given edge-based flow over time provided that the number of iterations goes to infinity.

Flow Decomposition Algorithm	
Input:	An edge-based flow over time $(\mathcal{F}^+, \mathcal{F}^-)$
Output:	A path-based flow over time $\mathcal{F}$ including flow carrying cycles.

(1) Let $C_1, \ldots, C_k$ with $k \in \mathbb{N}$ be an order of the simple cycles.

(2) For each $i := 1$ to k do:

  (a) Let F_{C_i} be the maximum flow originating until time θ which can be sent along C_i.

  (b) Delete the cycle flow from $(\mathcal{F}^+, \mathcal{F}^-)$, i.e., set $F^+_{e_j} := F^+_{e_j} - F_{C_i}$ and $F^-_{e_j} := F^-_{e_j} - F_{C_i}$ for all $j = 1, \ldots, |C_i|$.

(3) Let $(P_i)_{i\in\mathbb{N}}$ be an order of the s-t-paths such that $|P_i| < |P_j|$ holds whenever $i < j$. Set $\theta := 1$ and $\epsilon := 1$.

(4) Set $i := 0$.

(5) Until $\displaystyle\sum_{e\in\delta^+(s)} F^+_e(\theta) \geq \epsilon$ do

  (a) Let F_{P_i} be the maximum flow originating until time θ which can be sent along P_i.

  (b) Delete the path flow from $(\mathcal{F}^+, \mathcal{F}^-)$, i.e., set $f^+_{e_j} := f^+_{e_j} - f_{P_i,j-1}$ and $f^-_{e_j} := f^-_{e_j} - f_{P_i,j}$ for all $j = 1, \ldots, |P_i|$.

  (c) Set $i := i + 1$.

(6) Set $\theta := 2\theta$ and $\epsilon := \frac{\epsilon}{2}$ and go to (4).

Before we analyze this algorithm we have to remark the following. Flow which is send back to s or which leaves t causes problems, in general. For resolving this, we add an artificial source s_0 and an artificial sink t_0 which we connect with s and t, respectively. The flow on the new edges is set to the net outflow of s and the net inflow of t, respectively. Clearly, on the new instance no flow is send back to s_0 and no flow leaves t_0.

As already mentioned, a flow decomposition of an edge-based flow over time may not exist. However, the Flow Decomposition algorithm converges to a flow decomposition if the underlying edge-based flow over time is decomposable. For observing this, we only give a proof idea which is strongly based on intuition. Firstly, we divide a run of this algorithm into *phases*. A new phase starts whenever Step (6) is called. Hence, θ is doubled and ϵ is halved from one phase to the next. Of course, a phase consists of several *iterations*. An iteration is determined by the treatment of one path in the loop in Step (5), i.e., every time the condition in Step (5) is checked, a new iteration starts.

For proving the correctness of the Flow Decomposition algorithm, we need the following observations provided that the initial edge-based flow over time has a flow decomposition:

(i) The remaining flow before each iteration is path decomposable.

(ii) Every phase consists of a finite number of iterations.

(iii) If the number of phases goes to infinity, the remaining flow on each edge decreases locally to 0.

It is somehow obvious that the Flow Decomposition algorithm works as desired if these three statements are valid. The first statement ensures that we do not run into a point of no return. Based on this, statement (ii) ensures that we reach any phase if the algorithm runs long enough. Therefore, the path flows computed during an application of the algorithm converge to a flow decomposition because of statement (iii).

In the following, we intuitively discuss why each of the statements above holds. To see that statement (i) is valid, let $(\mathcal{F}^+, \mathcal{F}^-)$ be a path decomposable edge-based flow over time and $P^* \in \mathcal{P}$ be an s-t-path with the smallest number of edges which is able to carry a positive amount of flow $\tilde{F}_{P^*}|_{\leq\theta}$ originating until a given point in time $\theta \in \mathbb{R}_+$. Given a path decomposition $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ we modify $\mathcal{F}$ by iterating over the edges of P^*. In each iteration i, we route as much flow as possible to the next edge $e_{i+1}^{P^*}$ of P^*. This is done by interchanging flow between s-t-paths containing the first i edges of P^* but not $e_{i+1}^{P^*}$ and s-t-paths containing $e_{i+1}^{P^*}$ but not the predecessor edge $e_i^{P^*}$. In the end, this shows that there exists a path decomposition $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ with $F_{P^*}|_{\leq\theta} = \tilde{F}_{P^*}|_{\leq\theta}$. In this manner, statement (i) is inductively proven.

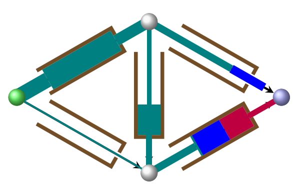
In order to show that statement (ii) is valid, we use the fact that the Flow Decomposition algorithm sends flow along currently shortest path. So in each phase the algorithm behaves like the well known Successive Shortest Path algorithm for computing minimum cost flows. In fact, if we introduce a cost of 1 on each edge, the cost of the current path decomposition decreases from one iteration to the next. Another problem which can be solved by the Successive Shortest Path algorithm is the computation of an earliest arrival flow. So if we interpret the new costs as transit times, we increase the arrival pattern of the current path decomposition from one iteration to the next. That is, the value $\sum_{P|k\geq|P|} F_P(\theta)\cdot(k-|P|)$ increases for all points in time k. This shows that each phase consists of a finite number of iterations.

To see that statement (iii) holds, we observe that, as the number of phases goes to infinity, the flow leaving s until a given time θ decreases to 0. So statement (iii) follows directly from the flow conservation constraint (3.7) in the definition of an edge-based flow over time (see Definition 3.9) as flow on cycles is deleted during the initial phase of the Flow Decomposition algorithm.

3.6 Shortest Paths for Increasing Arrival Times

It is a crucial feature of Nash flows over time that flow particles choose paths along which they arrive at the sink t as early as possible. For defining *currently shortest s-t-paths* with respect to a given flow over time $\mathcal{F}$, we consider the problem of sending an additional flow particle at time $\theta \in \mathbb{R}_+$ from the source s to the sink t as quickly as possible. Let $\ell_v(\theta)$ be the earliest point in time at which this flow particle can arrive at a node $v \in V$. Then it holds

$$\ell_e\big(\ell_v(\theta)\big) \;\geq\; \ell_w(\theta) \qquad\qquad \forall e = vw \in \delta^-(w)\ . \tag{3.41}$$



On the other hand, for each node $w \in V \setminus \{s\}$, there must exist at least one incoming edge $e = vw \in \delta^-(w)$ such that equality holds in (3.41). That is, the flow particle can use edge e in order to arrive at node w as early as possible, i.e., at time $\ell_w(\theta)$. Moreover, we have $\ell_s(\theta) = \theta$ for all $\theta \in \mathbb{R}_+$. Therefore, we define the *node arrival time functions* $\ell_w : \mathbb{R}_+ \to \mathbb{R}_+^\infty$ by

$$\ell_w(\theta) \ := \begin{cases} \theta & \text{if } w = s \ , \\ \min\limits_{e=vw} \ell_e\big(\ell_v(\theta)\big) & \text{if } w \in V \setminus \{s\} \ . \end{cases} \tag{3.42}$$

The node arrival time functions can be computed simultaneously for all times θ by adapting the shortest path algorithm of Bellman and Ford[2] as follows.

Dynamic Bellmann-Ford Algorithm	
Input:	A network $(G, \mathcal{U}, \mathcal{L}, s, t)$.
Output:	The lengths $\ell_v(\theta)$ of a shortest s-v-paths for all nodes v and for all starting times $\theta \geq 0$.

(1) Set $\ell_v(.) := \infty$ for all $v \in V \setminus \{s\}$ and $\ell_s := \text{id}$.

(2) For $i = 1$ to $|V| - 1$ do:
  For all edges $e = vw \in E$ do:
    Set $\ell_w := \min\{\ell_w, \ell_e \circ \ell_v\}$.

Note that this algorithm matches exactly the algorithm of Bellmann and Ford for computing shortest path on a static network except that the update procedure (2) is applied to functions and not to constant real numbers. Thus, the update procedure is applied to $\ell_w(\theta)$ for all times $\theta \in \mathbb{R}_+$ simultaneously.

As we see below, the correctness of the Classical Bellmann-Ford algorithm ensures the existence of the node arrival time functions $\ell_v : \mathbb{R}_+ \to \mathbb{R}_+^\infty$ according to (3.42).

Lemma 3.50. *If all arrival time functions are nondecreasing, a currently shortest path can be assumed to be is simple.*

Proof. Let $P := (e_1, \ldots, e_{|P|})$ be a nonsimple path and let $C := (e_{i+1}, \ldots, e_j)$ be a cycle in P for some $i+1, j \in \{1, \ldots, |P|\}$ with $i < j$. We show that the current transit times $\tau_{P'}(\theta)$ of the path $P' := (e_1, \ldots, e_i, e_{j+1}, \ldots, e_{|P|})$ where C is deleted from P is not greater than the current transit times $\tau_P(\theta)$ of P, i.e., $\tau_{P'}(\theta) \leq \tau_P(\theta)$ holds for all times $\theta \in \mathbb{R}_+$. But first note that P' is a path because $\text{head}(e_i) = \text{tail}(e_{i+1}) = \text{head}(e_j) = \text{tail}(e_{j+1})$ follows from the definition of paths and cycles.

Let $\ell_{P,i}(\theta)$ be the time at which flow is able to leave the i-th edge of P. Since current transit times are nonnegative by definition, we have $\tau_C(\ell_{P,i}(\theta)) \geq 0$. Here, $\tau_C(\ell_{P,i}(\theta))$ is the time which flow assigned to P at time θ has to take in order to traverse the cycle C while traversing P. Hence, this flow leaves the j-th edge of P at time $\ell_{P,j}(\theta) = \ell_{P,i}(\theta) + \tau_C(\ell_{P,i}(\theta)) \geq \ell_{P,i}(\theta)$.

[2]The update procedure of Bellman-Ford for a certain label (arrival time) $\ell_w(\theta)$ is applied for all times θ simultaneously and, hence, is seen as an operation on functions. If we use Dijkstra's algorithm instead, we have to maintain the set of already finalized nodes separately for each time θ. Thus, we also have to apply the update procedure of Dijkstra separately for each θ. Therefore, we adapt the Bellmann-Ford algorithm.

On the other hand, flow assigned to P' at time θ enters the head of e_j, which equals the head of the i-th edge of P', at time $\ell_{P',i}(\theta) = \ell_{P,i}(\theta) \leq \ell_{P,j}(\theta)$. Since the arrival time function are nondecreasing, we inductively obtain

$$\ell_{P',i'}(\theta) \;=\; \ell_{e^{P'}_{i'}}\big(\ell_{P',i'-1}(\theta)\big) \;\leq\; \ell_{e_{i'+j-i}}\big(\ell_{P,i'+j-i-1}(\theta)\big) \;=\; \ell_{P,i'+j-i}(\theta)$$

for $i' \in \{i+1, \ldots, |P'|\}$ because $e^{P'}_{i'} = e_{i'+j-i}$. Thus, for $i' = |P'|$ we get

$$\begin{aligned}\tau_{P'}(\theta) \;&=\; \ell_{P',|P'|}(\theta) - \theta \\ &\leq\; \ell_{P,|P'|+j-i}(\theta) - \theta \;=\; \ell_{P,|P|}(\theta) - \theta \;=\; \tau_P(\theta)\end{aligned}$$

as desired. □

Theorem 3.51 (Orda and Rom, [65]). *The* Dynamic Bellmann-Ford *algorithm works correctly if all arrival time functions are nondecreasing. Moreover, if the node arrival time functions ℓ_v are nondecreasing and if the transit time functions are (left/right) continuous then the node arrival time functions too.*

Proof. Firstly, we show that the Dynamic Bellmann-Ford algorithm computes currently minimal arrival times. That is, at the end of the algorithm, $\ell_v(\theta)$ equals the minimal current arrival time over all s-v-paths for all nodes $v \in V$ at all times $\theta \in \mathbb{R}_+$. For this, first observe inductively that there exists an s-v-path P with $\ell_P(\theta) = \ell_v(\theta)$ whenever $\ell_v(\theta)$ is finite at some point during a run of the algorithm. Moreover, we can assume without loss of generality that P contains at most k edges, where k equals the number of calls of Step (2). Hence, $\ell_v(\theta)$ is an upper bound on the minimal current arrival time. Next, we show that $\ell_v(\theta)$ is also a lower bound.

In the following, we prove by induction that after the k-th call of Step (2) the values $\ell_v(\theta)$ are not greater than the minimal current arrival time over all s-v-paths containing at most k edges. Note that this is true before the first call of Step (2), i.e., for $k = 0$, because of the initialization in Step (1). So assume that the induction assumption is established for some $k-1$ with $k \geq 1$ and let P be an s-w-path which uses at most k edges. Let v be the predecessor node of w on P, i.e., $v := \text{head}(e^P_{|P|})$. Since the subpath P' using the first $|P|-1$ edges from P is an s-v-path with at most $k-1$ edges, we know that $\ell_v(\theta) \leq \ell_{P'}(\theta)$ holds after the $(k-1)$-st call of Step (2). Hence, if edge $e^P_{|P|}$ is considered during the subsequent call of Step (2), we derive

$$\ell_w(\theta) \;\leq\; \ell_v(\theta) + \tau_{e^P_{|P|}}(\ell_v(\theta)) \;\leq\; \ell_{P'}(\theta) + \tau_{e^P_{|P|}}(\ell_{P'}(\theta)) \;=\; \ell_P(\theta)$$

because the edge arrival time functions are nondecreasing. Since P is chosen as an arbitrary s-w-path containing at most k edges, the induction assumption is proven.

During a run of the Dynamic Bellmann-Ford algorithm, Step (2) is called exactly $|V|-1$ times. Since a simple path contains at most $|V|-1$ edges, we know that, at the end of the Dynamic Bellmann-Ford algorithm, $\ell_v(\theta)$ is not greater than the currently minimal arrival time over all simple s-v-paths. So Lemma 3.50 implies that $\ell_v(\theta)$ is a lower bound on and, hence, equals the currently minimal arrival time over all s-v-paths.

To see that the second statement of this lemma is valid, observe the following when considering Step (2). The composition just as the minimum of nondecreasing functions is nondecreasing. Further, the composition just as the minimum of (left/right) continuous functions is (left/right) continuous. Hence, the second statement follows inductively from the fact that the update step in (2) is applied only a finite number times $|E|(|V|-1)$ times. Note that the initialization in Step (1) ensures the basis of the induction. □

It is worth to mention that both the nonnegativity of the transit time functions and the FiFo principle are essential for the validity of Theorem 3.51. So far, we have seen that node arrival time functions exist. Clearly, the concrete functions depend on the underlying flow over time. In the following, we discuss some properties for node arrival times arising out of foresighted arrival times.

As verified by Example 3.45, the path-based flow model which is build upon a consistent edge-based one using the dynamic network loading problem need not to be past-oriented. Nevertheless, corresponding foresighted node arrival times are past-oriented which is established in the following Lemma. Further, we show that foresighted node arrival times are upper semi-continuous.

Lemma 3.52. *Consider a consistent flow over time model. Then the foresighted node arrival time $\bar{\ell}_t$ is upper semi-continuous and past-oriented. If the flow model is edge-based, this holds for all node arrival times $\bar{\ell}_v$ independently on the particular flow decomposition.*

Proof. For consistent path-based flow over time models, the semi-continuity follows directly from the fact that the pointwise minimum (infimum) of upper semi-continuous maps is again upper semi-continuous. For consistent edge-based flow over time models, we additionally need that the finite sum of upper semi-continuous maps remains upper semi-continuous.

It remains to prove the past-orientation. For path-based flow models, this follows directly from the definition of $\bar{\ell}_t$ and the past-orientation of $\bar{\ell}_P$ for each $P \in \mathcal{P}$ (see Lemma 3.40) which imply

$$\bar{\ell}_t(\mathcal{F}^+)(\theta') \;=\; \min_{P\in\mathcal{P}} \bar{\ell}_P(\mathcal{F}^+)(\theta') \;=\; \min_{P\in\mathcal{P}} \bar{\ell}_P(\mathcal{F}^+|_{\leq\theta})(\theta') \;=\; \bar{\ell}_P(\mathcal{F}^+|_{\leq\theta})(\theta')$$

for all points in time $\theta', \theta \in \mathbb{R}_+$ with $\theta' \leq \theta \in \mathbb{R}_+$.

In order to prove the past-orientation statement for an edge-based flow over time $\mathcal{F}$, let $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}\cup\mathcal{C}}$ be a flow decomposition of $\mathcal{F}$. Consider a point in time $\theta \in \mathbb{R}_+$ and recall the definition of a restricted path-based flow over time which states $\mathcal{F}^+|_{\leq\theta} = (F_P^+|_{\leq\theta})_{P\in\mathcal{P}\cup\mathcal{C}}$. Since the flow on each edge satisfies FiFo, we get inductively from equation (3.23) in Lemma 3.21

$$F_{P,i}^+(\mathcal{F}^+)(\vartheta) \;=\; F_{P,i}^+(\mathcal{F}^+|_{\leq\theta})(\vartheta) \qquad \forall \vartheta \leq \bar{\ell}_{P,i-1}(\mathcal{F}^+)(\theta)\;. \tag{3.43}$$

Note that the expression $F_{P,i}^+(\mathcal{F}^+|_{\leq\theta})(\vartheta)$ is evaluated using the arrival time functions arising out of $\mathcal{F}^+$ and *not* the those arising out of $\mathcal{F}^+|_{\leq\theta}$. Next, consider an edge $e = vw \in E$. Since $\bar{\ell}_v(\mathcal{F}^+)(\theta) \leq \bar{\ell}_{P,i-1}(\mathcal{F}^+)(\theta)$ holds for paths P with $e_i^P = e$, equation (3.43) shows

$$\begin{aligned} F_e^+(\mathcal{F}^+)(\vartheta) &= \sum_{P | e_i^P = e} F_{P,i}^+(\mathcal{F}^+)(\vartheta) \\ &= \sum_{P | e_i^P = e} F_{P,i}^+(\mathcal{F}^+|_{\leq\theta})(\vartheta) \\ &= F_e^+(\mathcal{F}^+|_{\leq\theta})(\vartheta) \qquad\qquad \forall \vartheta \leq \bar{\ell}_v(\mathcal{F}^+)(\theta)\;. \end{aligned}$$

Since the flow on e is past-oriented, this shows

$$\begin{aligned}\bar{\ell}_e\big(F_e^+(\mathcal{F}^+)\big)(\vartheta) &= \bar{\ell}_e\big(F_e^+|_{\leq\bar{\ell}_v(\theta)}(\mathcal{F}^+)\big)(\vartheta)\\ &= \bar{\ell}_e\big(F_e^+|_{\leq\bar{\ell}_v(\theta)}(\mathcal{F}^+|_{\leq\theta})\big)(\vartheta)\\ &= \bar{\ell}_e\big(F_e^+(\mathcal{F}^+|_{\leq\theta})\big)(\vartheta) \qquad \forall\vartheta\leq\bar{\ell}_v(\theta) \qquad (3.44)\end{aligned}$$

where past-orientation is applied to the first and to the third "="-sign. Moreover, the flow on e satisfies FiFo implying

$$\begin{aligned}\bar{\ell}_e\big(F_e^+(\mathcal{F}^+)\big)(\bar{\ell}_v(\theta)) &= \bar{\ell}_e\big(F_e^+(\mathcal{F}^+|_{\leq\theta})\big)(\bar{\ell}_v(\theta))\\ &\leq \bar{\ell}_e\big(F_e^+(\mathcal{F}^+|_{\leq\theta})\big)(\vartheta) \qquad \forall\vartheta>\bar{\ell}_v(\theta)\ . \qquad (3.45)\end{aligned}$$

Using the definition of the node arrival time functions in (3.42) with respect to $\mathcal{F}^+$ and $\mathcal{F}^+|_{\leq\theta}$, we obtain from (3.44) and (3.45)

$$\begin{aligned}\bar{\ell}_{P,i-1}(\mathcal{F}^+)(\theta') &\geq \bar{\ell}_v(\mathcal{F}^+|_{\leq\theta})(\theta') \qquad \forall\theta'\leq\theta\\ \bar{\ell}_{P,i-1}(\mathcal{F}^+)(\theta') &\geq \bar{\ell}_v(\mathcal{F}^+|_{\leq\theta})(\theta) \qquad \forall\theta'\geq\theta\end{aligned}$$

by induction over i. Here, v is the tail of the i-th edge in P. Moreover, the first inequality is satisfied with equality if P visits each node on its way to v as early as possible. Finally, for all nodes $v\in V$, this shows

$$\bar{\ell}_v(\mathcal{F}^+)(\theta') = \bar{\ell}_v(\mathcal{F}^+|_{\leq\theta})(\theta') \qquad \forall\theta'\leq\theta\ . \qquad \square$$

We conclude this section by proving some continuity property which we need for establishing the existence of Nash flows over time in Section 4.2.

Lemma 3.53. *Let $(\mathcal{F}_k^+,\mathcal{F}_k^-)_{k\in\mathbb{N}}$ be a convergent sequence of flows over times with limit point $(\mathcal{F}^+,\mathcal{F}^-)$. Consider an s-t-path P and let F_k^- and F^- be the outflow functions and $\bar{\ell}^k$ and $\bar{\ell}$ be the foresighted arrival time functions of P with respect to $(\mathcal{F}_k^+,\mathcal{F}_k^-)$, for all $k\in\mathbb{N}$, and $(\mathcal{F}^+,\mathcal{F}^-)$, respectively. Moreover, assume that for each $\epsilon>0$ there exists a $K\in\mathbb{N}$ such that*

$$F_k^-\big(\bar{\ell}^k(\theta)\big)-\epsilon \leq F_k^-\big(\bar{\ell}_t^k(\theta)\big) \leq F_k^-\big(\bar{\ell}^k(\theta)\big)\ . \qquad (3.46)$$

holds for all $k\geq K$ and all points in time $\theta\in\mathbb{R}_+$. Then we have

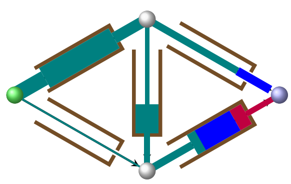
$$\lim_{k\to\infty} F_k^-\big(\bar{\ell}_t^k(\theta)\big) = F^-\big(\bar{\ell}_t(\theta)\big)\ .$$

for all $\theta\in\mathbb{R}_+$.

Proof. Let $\theta\in\mathbb{R}_+$ be a point in time. As foresighted node arrival times are upper semi-continuous by Lemma 3.52, we know $\lim_{k\to\infty}F_k^-(\bar{\ell}_t^k(\theta))\leq F^-(\bar{\ell}_t(\theta))$. The other direction follows from equation (3.34) proven in Lemma 3.39 which shows $\lim_{k\to\infty}F_k^-(\bar{\ell}^k(\theta))=\lim_{k\to\infty}F_k^+(\theta)=F^+(\theta)=F^-(\bar{\ell}(\theta))$. Thus, from equation (3.46) we get

$$\lim_{k\to\infty} F_k^-\big(\bar{\ell}_t^k(\theta)\big) = F^-\big(\bar{\ell}(\theta)\big) \geq F^-\big(\bar{\ell}_t(\theta)\big)\ .$$

as $\bar{\ell}(\theta)\geq\bar{\ell}_t(\theta)$ holds by definition and F^- is nondecreasing. $\square$



3.7 Discrete Flows over Time

In this section we briefly discuss discrete flows over time. We focus on the differences to continuous flows over time and show how the notion of consistency works for discrete flows over time.

To get an idea of flows over time, we assume that we stand at the side of a street in a traffic network because we want to record the flow situation at this place. We notice for each vehicle, which passes by, the current time and the type resulting in a sequence of pairs. We only distinguish between passenger cars, pickups, and trucks. So the type of a vehicle is, in fact, a synonym for its size which is represented by an amount of flow in a discrete flow. Thus, the main difference to continuous flows over time is that, instead of infinitesimal flow particles, units carrying a real flow value are routed through a network. This does not mean that a flow unit is entirely concrete. In fact, we assume that a flow unit has no physical dimension. In our traffic scenario, one can think of a flow unit as a vehicle which carries all of its weight in its *front* bumper.

Formalizing this idea, we are given a set of flow units, either finite or countable and possibly empty. So we use $\mathbb{N}$ to denote the set of flow units. The natural linear order on $\mathbb{N}$ turns out to be useful as we see below. Each flow unit $i \in \mathbb{N}$ is given by a real number $f_i \in \mathbb{R}_+$ determining the amount of flow which is concentrated in this unit. We also allow flow units of size 0 for modeling a finite or maybe also empty set of flow units. The flow situation at a particular place of a network is represented by a *discrete flow pattern* $F := (i_j, \theta_j)_{j\in[N]_0}$ where $N \subseteq \mathbb{N}$ and $N = \{i_0, \ldots, i_{|N|-1}\}$. Here, the pair (i_j, θ_j) means that flow unit i_j passes by at time $\theta_j \in \mathbb{R}_+$. In addition, we require $\theta_{j-1} \leq \theta_j$ for all $j \in [N]$ which means that flow unit i_j travels in front of flow unit $i_{j'}$ whenever $j < j'$. Further, if $\theta_{j-1} = \theta_j$ holds for $j \in [N]$, we assume $i_{j-1} < i_j$. This ensures that if two or more flow units pass by at the same point in time, there is a global priority rule deciding which flow unit has to be considered first. Note that this *order condition* is not a limitation in case a flow pattern is given via a finite number of flow units because such families can be reordered for satisfying the order condition. However, when dealing with an infinite but countable number of flow units not every family of pairs from $\mathbb{N} \times \mathbb{R}_+$ can be considered as a flow pattern. For example a family containing a pair (i_j, q) for all $q \in \mathbb{Q}_+$ cannot be reordered such that the order condition is met.

As in the continuous case, a discrete flow over time is given by the points in time at which a flow unit enters a particular edge or path. Therefore, the state of the flow on an edge or a path can be described via discrete flow patterns.

Definition 3.54 (Discrete Flow over Time)**.** Let $\mathbb{N}$ be the set of flow units and let $f_i \in \mathbb{R}_+$ be the size of flow unit i for each $i \in \mathbb{N}$. A family $\mathcal{F} := (F_P^+)_{P\in\mathcal{P}}$ is called *discrete path-based flow over time* if and only if $F_P^+ := (i_j, \theta_j)_{j\in[N_P]_0}$ with $N_P \subseteq \mathbb{N}$ is a feasible discrete flow pattern for all $P \in \mathcal{P}$.

A family of pairs $\mathcal{F} := (F_e^+, F_e^-)_{e\in E}$ is called a *discrete edge-based flow over time* if and only if $F_e^+ := (i_j, \theta_j)_{j\in[N_e^+]_0}$ and $F_e^- := (i_j, \theta_j)_{j\in[N_e^-]_0}$ are feasible flow patterns for all edges $e \in E$. Since no flow unit should be lost or originate on an edge, we require $N_e^+ = N_e^- =: N_e \subseteq \mathbb{N}$ for all $e \in E$. The same should hold for the nodes which is ensured by:

$$\bigcup_{e\in\delta^-(v)} N_e = \bigcup_{e\in\delta^+(v)} N_e \qquad\qquad \forall v \in V \setminus \{s, t\} \ .$$

In addition, we should ensures that one flow unit is not routed several times in discrete flows over time. Therefore, for path-based flows we require that the sets N_P are pairwise disjoint. For edge-based flows we assume that the set $\{e \mid i \in N_e\}$ of edges used by flow unit i is empty or defines either one s-t-path or one cycle for all $i \in \mathbb{N}$. But nevertheless, if this constraint is violated we could also think of multiple copies of the same flow unit. Having in mind Hilbert's Hotel, these copies can be clearly added to the set $\mathbb{N}$ of flow units.

Recalling the continuous case the inflow of an edge or a path is given via a nondecreasing, (absolutely) continuous function. Also for the continuous case it seems to be appropriate to encode a flow pattern $(i_j, \theta_j)_{j \in [N']_0}$ via a nondecreasing function $F : \mathbb{R}_+ \to \mathbb{R}_+$ given by

$$\theta \;\mapsto\; F(\theta) \;:=\; \sum_{j \mid \theta_j \leq \theta} f_{i_j} \;.$$

So F is a step function and each step defines a flow unit where the height of the jump equals the amount of flow in the corresponding unit. In addition, the time where a flow unit passes by can also be read off F because it equals the time at which the step occurs. Hence, we could define discrete flow models in the same manner as the continuous counterpart – with time-dependent transit time function. Unfortunately, this is not the entire story. What happens if two flow units enter the same edge or the same path at the same point in time? Then these two flow units would be merged into one flow unit and are no longer recognizable as two flow units. This observation also shows that time-dependent transit time function are not the right way for defining discrete flow models. In this case flow units entering the same edge at the same point in time would always leave this edge at the same point in time. But what would be the output if such an edge models a street segment where the number of lanes decreases from 2 to 1? For this reason we define discrete flow over time models via flow unit-dependent transit time functions. That is, a discrete flow model provides a real transit time for each flow unit and not for each point in time. And as in the continuous case, these real transit times depend still on the entire flow situation.

Definition 3.55 (Discrete Flow over Time Model)**.** A *discrete* path-based flow over time model defines a family $\mathcal{T} := (\tau_P)_{P \in \mathcal{P}}$ of maps τ_P, $P \in \mathcal{P}$ assigning a function $\tau_P(\mathcal{F}) : N_P \to \mathbb{R}_+$ to each flow over time $\mathcal{F}$. The map τ_P is called *transit time* of $P \in \mathcal{P}$ and $\tau_P(\mathcal{F})$ is called *transit time function* with respect to $\mathcal{F}$. The value $\tau_P(\mathcal{F})(i_j)$ is called *real transit time* of flow unit i_j.

A corresponding family $\mathcal{L} := (\ell_P)_{P \in \mathcal{P}}$ of *arrival times* is given by an *arrival time function* $\ell_P(\mathcal{F}) : N_P \to \mathbb{R}_+$ defined as $\ell_P(\mathcal{F})(i_j) := \theta_j + \tau_P(\mathcal{F})(i_j)$ for each path $P \in \mathcal{P}$. The value $\ell_P(\mathcal{F})(i_j)$ is called *real arrival time*.

Similarly, a *discrete* edge-based flow over time model is defined by a family $\mathcal{T} := (\tau_e)_{e \in E}$ of maps τ_e, $e \in E$ assigning a function $\tau_e(F_e^+) : N_e \to \mathbb{R}_+$ to each possible inflow F_e^+. The map τ_e is called *transit time* of $e \in E$ and $\tau_e(F_e^+)$ is called *transit time function* with respect to F_e^+. The value $\tau_e(F_e^+)(i_j)$ is called *real transit time* of flow unit i_j.

A corresponding family $\mathcal{L} := (\ell_e)_{e \in E}$ of *arrival times* is given by an *arrival time function* $\ell_e(F_e^+) : N_e \to \mathbb{R}_+$ defined as $\ell_e(F_e^+)(i_j) := \theta_j + \tau_e(F_e^+)(\theta_j)$ for each path $e \in E$. The value $\ell_e(F_e^+)(i_j)$ is called *real arrival time*.

As in the continuous setting, a discrete path transit time τ_P depends on the entire flow $\mathcal{F}$ whereas a discrete edge transit time τ_e depends only on the inflow F_e^+ on the edge e. So from an abstract point of view, the flow on an edge can be seen as a special case of a path flow.

Consider a discrete flow over time $\mathcal{F}$ either path- or edge-based. Then path or edge transit times are used to represent the flow behavior on a path or an edge. In fact, if flow unit i enters a path P at time θ, it arrives at t at time $\ell_P(\mathcal{F})(i) = \theta + \tau_P(\mathcal{F})(i)$. Similarly, if flow unit i enters an edge e at time θ, it arrives at the head of e at time $\ell_e(F_e^+)(i) = \theta + \tau_e(F_e^+)(i)$. Hence, the discrete outflow patterns F_P^- and F_e^- of a path P and an e are representable as

$$\begin{aligned} & F_P^- = \big\{\big(i,\ \ \ell(\mathcal{F})(i)\big) \mid (i,\theta) \in F_P^+\big\} \\ \text{and} \quad & F_e^- = \big\{\big(i, \ell(F_e^+)(i)\big) \mid (i,\theta) \in F_e^+\big\}\ , \end{aligned} \tag{3.47}$$

respectively. In contrast to the path-based setting, (3.47) is an explicit constraint which a *feasible* (with respect to a discrete edge-based flow model) discrete edge-based flow over time has to satisfy. Also the arrival time of a flow unit on one edge must be equal to the departure time of that flow unit on its next edge. Further, the outflow of an edge or a path should be a flow pattern. Of course, this is always satisfied if the corresponding inflow pattern consists of a finite number of flow units. Unfortunately, if the inflow pattern consists of an infinite but countable number of flow units, we have to check the order condition. This imposes additional assumptions on feasible transit times, especially if all flow units enter the underlying path or edge in finite time. Nevertheless, it should be clear how discrete flow over time models work. Therefore, we omit further technical details.

In the following we extend the notion of consistent flow over time models to the discrete case. For this we observe that the continuity of a discrete flow model is not that important for the following reasons. It turns out in Chapter 4 that the continuity of a flow model is needed in order to describe how the flow behavior is changed if some flow switches to another path. However, in discrete flow models always an entire flow unit is able to change its path. Therefore, no limit approaches are needed in this case and we only have to address the FiFo principle and the past-orientation.

One main difference between continuous flows and discrete flows over time is that flow behavior is defined with respect to time in the continuous case and with respect to the flow units in the discrete case. This differentiation remains also when defining consistency. Of course, the discrete FiFo principle should also ensure that a flow unit which enters a path or an edge first also leaves this path or edge first. But, in addition, if two flow units i and j enter a path or an edge at the same point in time, the one with the lower global priority should arrive first – that is flow unit $\min\{i,j\}$. So the *FiFo principle* is satisfied on a path P or an edge e if and only if

$$\begin{aligned} & \ell_P(\mathcal{F})(i_{j-1}) \ \le\ \ell_P(\mathcal{F})(i_j) && \forall j \in [N_P]_0 \\ \text{and} \qquad & \ell_e(F_e^+)(i_{j-1}) \ \le\ \ell_e(F_e^+)(i_j) && \forall j \in [N_e]_0 \end{aligned}$$

holds, respectively. Note that the priority given by the subscript j ensures that the FiFo principle works as desired.

As in the continuous time setting, discrete past-orientation should ensure that the transit time of a flow unit does not depend on flow units entering the same edge or the same path later. In case two or more flow units enter a path or an edge at the same point in time, past-orientation should also ensure that the transit time of a flow unit only depends on the flow units having a lower priority. In particular, the transit time of a flow unit does *not* depend on the flow unit itself. That is why we should think of a vehicle as a flow unit carrying all of its weight in the *front* and not in its rear bumper. This motivates that a discrete flow model is *past-oriented* if and only if, for every path P or every edge e, is holds

$$\begin{aligned} & \tau_P(\mathcal{F})(i_j) &=& \ \tau_P(\mathcal{F}|_{<i_j})(i_j) && \forall j \in [N_P]_0 \\ \text{and} \quad & \tau_P(F_e^+)(i_j) &=& \ \tau_P(F_e^+|_{<i_j})(i_j) && \forall j \in [N_e]_0 \end{aligned}$$

holds, respectively, where $\mathcal{F}|_{<i_j}$ and $F_e^+|_{<i_j}$ denote the restricted flows of the flow units departing from s and entering e before and in front of flow unit i_j, respectively.

Next, we ask for the discrete analogon of the foresighted arrival time function. For this we use the following interpretation which extends the meaning of discrete transit times. In addition to the real transit times of flow units which are routed over the corresponding path P or edge e, transit times also determine the *potential* transit time of an additional flow unit which enters P or e at some given point in time θ. Of course, these transit times also depend on the global priority of that additional flow unit.

In the last part of this section we shortly discuss the discrete network loading and flow decomposition problem. The easier discrete flow decomposition problem asks for the paths and cycles which are taken by the flow units in an edge-based flow over time. It is rather clear that the path or the cycle of a flow unit can be determined by simply following this flow unit through the network. Note that we assume that a particular flow unit is only allowed to use either one path or one cycle. Therefore, the discrete flow decomposition problem is always solvable.

The same holds for the discrete dynamic network loading problem in case of a consistent flow model. Here, we iteratively route an uniquely determined flow unit i through the next edge e of its path P. In fact, we send the flow unit i with the currently smallest entering time θ and if there are several, we choose the flow unit with the smallest global priority i. Since all transit times are nonnegative by definition, this ensures that the flow behavior of i on e is not changed any more as we deal with consistent flow models. In this manner we find the corresponding edge-based flow of a given path-based one.

3.8 Capacitated Networks

In this section we generalize the notion of classical flows over time (see Section 2.5) using arbitrary time varying capacity functions. For the purpose of this thesis we restrict to the case where storage at nodes is forbidden. For this setting we define the residual graph with respect to a given flow over time. Using the residual graph we generalize path-based representations of flows over time. Finally, we prove some technically lemmas which are needed in Section 5.2 where we extend the Classical Successive Shortest path algorithm to this scenario.

Similarly to Section 2.5, we deal with a network $(G,\mathcal{U},\mathcal{T},s,t)$ consisting of directed graph G, a source s, and a sink t. Further, a family $\mathcal{U} := (u_e)_{e\in E}$ of edge capacities u_e is given. In contrast to Section 2.5, these edge capacities are not constant but given via Lebesgue integrable edge capacity functions $u_e : \mathbb{R}_+ \to \mathbb{R}_+$. For some point in time $\theta \in \mathbb{R}_+$, the value $u_e(\theta)$ bounds the rate of flow leaving an edge $e \in E$. Moreover, the flow behavior is determined by a family $\mathcal{T} := (\tau_e)_{e\in E}$ of constant transit times $\tau_e \in \mathbb{R}_+$ determining the amount of time needed for traversing an edge. More precisely, if a flow particle enters an edge at some point in time $\theta \in \mathbb{R}_+$, it leaves e at time $\theta + \tau_e$.

As in Section 2.5 a flow over time is given by a family $\mathcal{F} := (f_e)_{e\in E}$ of Lebesgue integrable functions $f_e : \mathbb{R}_+ \to \mathbb{R}_+$. The value $f_e(\theta)$ stands for the rate at which flow leaves the corresponding edge e at a certain time $\theta \in \mathbb{R}_+$. Since transit times are constant, $f_e(\theta + \tau_e) = (f_e + \tau_e)(\theta)$ equals the rate at which flow enters e at time θ. A feasible flow over time $\mathcal{F} := (f_e)_{e\in E}$ has to obey the edge capacities, i.e., we require

$$f_e \;\leq\; u_e \qquad\qquad \forall e \in E\ .$$

Since we forbid storage at nodes, $\mathcal{F}$ has to satisfy strict flow conservation, i.e,

$$\sum_{e\in\delta^-(v)} f_e \;=\; \sum_{e\in\delta^+(v)} f_e + \tau_e \qquad\qquad \forall v \in V \setminus \{s,t\}\ .$$

Further, we assume that flow only originates at s and only vanishes at t. In particular, this ensures that this definition of an edge-based flow over time fits the general notion of edge-based flows over time stated in Definition 3.9.

Adapting the static notion of the residual networks, we define residual networks for flows over time as follows.

Definition 3.56 (Residual Network)**.** Let $\mathcal{N} := (G,\mathcal{U},\mathcal{T},s,t)$ be a network and $\mathcal{F} := (f_e)_{e\in E}$ be a flow over time. For an edge $e = (v,w)$ we denote the corresponding *backward edge* by $\overleftarrow{e} := (w,v)$. The residual transit time of a backward edge $\overleftarrow{e}$ with $e \in E$ is defined by $\tau^r_{\overleftarrow{e}} := -\tau_e$ whereas the residual transit times of original edges remain unchanged, i.e., $\tau^r_e := \tau_e$ for all $e \in E$. Notice that the transit time of a backward edge is negative in general. We denote the set of all backward edges by $\overleftarrow{E}$ and set $E^r := E \cup \overleftarrow{E}$. For each edge $e \in E$ we define the *residual capacity* of e and the corresponding backward edge $\overleftarrow{e}$ as $u^r_e := u_e - f_e$ and $u^r_{\overleftarrow{e}} := f_e + \tau_e$, respectively.

The network $\mathcal{N}^r := (G^r,\mathcal{U}^r,\mathcal{T}^r,s,t)$ consisting of the graph $G^r := (V,E^r)$, the capacities $\mathcal{U}^r := (u^r_e)_{e\in E^r}$, the transit times $\mathcal{T}^r := (\tau^r_e)_{e\in E^r}$, the source s, and the sink t is called the *residual network* of $\mathcal{N}$ with respect to $\mathcal{F}$.

Note that residual capacities are the only part of the residual network which depends on $\mathcal{F}$. For an edge $e \in E$ the current residual capacities $u^r_e(\theta)$ and $u^r_{\overleftarrow{e}}(\theta)$ are interpretable as the maximum amount by which the current flow can be increased and reduced, respectively, without violating the capacity constraints. Further, the current definition of flows over time is easily extendable to the case of negative transit times. Thus, we are able to consider feasible flows over time on the residual network leading to the notion of augmentation.

Definition 3.57 (Augmentation)**.** Let $\mathcal{N}^r$ be the residual network with respect to some flow over time $\mathcal{F} := (f_e)_{e\in E}$. A feasible, nonzero flow $\mathcal{F}^r := (f^r_e)_{e\in E^r}$

on $\mathcal{N}^r$ is called an *augmenting s-t-flow*. Further, the corresponding *augmented flow* $\mathcal{F}^a := (f_e^a)_{e\in E}$ is defined by

$$f_e^a \;:=\; f_e + f_e^r - (f^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \qquad\qquad \forall e \in E \tag{3.48}$$

Like in static flow theory, augmented flows over time remain feasible on the original network. This is verified by the following lemma.

Lemma 3.58. *Let $\mathcal{F}$ be a flow on a network $\mathcal{N} := (G,\mathcal{U},\mathcal{T},s,t)$. Further, let $\mathcal{F}^r$ be a flow over time on the residual network $\mathcal{N}^r$ of $\mathcal{F}$. Then the augmented flow $\mathcal{F}^a$ defined by*

$$F_e^a := F_e + F_e^r - (F^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \qquad\qquad \forall e \in E$$

is feasible on $\mathcal{N}$ with value $\mathrm{val}(\mathcal{F}^a) = \mathrm{val}(\mathcal{F}) + \mathrm{val}(\mathcal{F}^r)$.

Proof. First we show that $0 \le f_e^a \le u_e$ for all edges $e \in E$. Because of the definition of residual capacities, for each edge $e \in E$, we get essentially

$$\begin{aligned} f_e^a &= f_e + f_e^r - (f^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \;\le\; f_e + f_e^r \;\le\; f_e + (u_e - f_e) = u_e \qquad \text{and}\\ f_e^a &= f_e + f_e^r - (f^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \;\ge\; f_e - (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \;\ge\; f_e - \big((f_e + \tau_e) - \tau_e\big) = 0\;. \end{aligned}$$

Next we show $\mathcal{F}^a$ satisfies the strict flow conservation constraints. For this let $v \in V \setminus \{s,t\}$ be some node disjoint from s and t. By the definition of $\mathcal{F}^a$ we obtain

$$\begin{aligned} &\sum_{e\in\delta^-(v)} F_e^a - \sum_{e\in\delta^+(v)} (F_e^a + \tau_e)\\ =&\sum_{e\in\delta^-(v)} \big(F_e + F_e^r - (F^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}})\big) - \sum_{e\in\delta^+(v)} \Big(\big(F_e + F_e^r - (F^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}})\big) + \tau_e\Big)\;. \end{aligned}$$

Note that $(-(f_{\overleftarrow{e}} - \tau_e)) + \tau_e$ is equal to $-f_{\overleftarrow{e}}$ and *not* equal to $-f_{\overleftarrow{e}} + 2\tau_e$ because addition and subtraction of a real number work as (horizontal) shiftings of functions. Thus, since $\mathcal{F}$ satisfies strict flow conservation on $\mathcal{N}$, expanding the shifting in the second sum leads to

$$\begin{aligned} &\sum_{e\in\delta^-(v)} F_e^a - \sum_{e\in\delta^+(v)} (F_e^a + \tau_e)\\ =&\sum_{e\in\delta^-(v)} F_e^r + \sum_{e\in\delta^+(v)} F^r_{\overleftarrow{e}} - \sum_{e\in\delta^-(v)} (F^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) - \sum_{e\in\delta^+(v)} (F_e^r + \tau_e)\;. \end{aligned}$$

If an arc e is contained in $\delta^+(v)$ and $\delta^-(v)$ then the backward arc $\overleftarrow{e}$ is contained in $\delta^-_{G^r}(v)$ and $\delta^+_{G^r}(v)$, respectively. Hence, we obtain

$$\begin{aligned} &\sum_{e\in\delta^-(v)} F_e^a - \sum_{e\in\delta^+(v)} (F_e^a + \tau_e)\\ =&\sum_{e\in\delta^-_{G^r}(v)} F_e^r - \sum_{e\in\delta^+_{G^r}(v)} (F_e^r + \tau_e^r)\;. \end{aligned}$$

Since $\mathcal{F}^r$ satisfies the strict flow conservation on $\mathcal{N}^r$, the right hand side of the last equation is equal to 0 implying that $\mathcal{F}^a$ satisfies the strict flow conservation constraints.

It remains to prove $\mathrm{val}(\mathcal{F}^a) = \mathrm{val}(\mathcal{F}) + \mathrm{val}(\mathcal{F}^r)$. Since we have

$$\mathrm{val}(\mathcal{F}^a) = -\sum_{e\in\delta^+(t)} (F_e^a + \tau_e) + \sum_{e\in\delta^+(t)} F_e^a \ ,$$

using the same argumentation as above shows:

$$\begin{aligned}\mathrm{val}(\mathcal{F}^a) &= -\sum_{e\in\delta^+(t)} (F_e^a + \tau_e) + \sum_{e\in\delta^-(t)} F_e^a \\ &= -\sum_{e\in\delta^+(t)} (F_e + \tau_e) + \sum_{e\in\delta^-(t)} F_e - \sum_{e\in\delta^+_{G^r}(t)} (F_e^r + \tau_e^r) + \sum_{e\in\delta^-_{G^r}(t)} F_e^r \\ &= \mathrm{val}(\mathcal{F}) + \mathrm{val}(\mathcal{F}^r) \ .\end{aligned}$$

This completes the proof. □

Using Lemma 3.58 we are able to establish a third representation of flows over time – in addition, to path- and edge-based flows over time. This representation relies on the notion of generalized temporally repeated classical flows over time (see Section 2.5). For this let $\mathcal{P}^r$ be the set of s-t-path in G^r.

Definition 3.59 (General Path-based Flow over time)**.** Let $\mathcal{F} := (F_P)_{P\in\mathcal{P}^r}$ be a family of nondecreasing, absolutely continuous functions $F_P : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $F_P(0) = 0$. Then $\mathcal{F}$ is called *general path-based flow over time.*

Given a network $(G, \mathcal{U}, \mathcal{T}, s, t)$ the flow $\mathcal{F}$ is called *feasible* if and only if

$$f_e := \sum_{P,i|e_i^P=e} (f_P + \tau_{P,i}) - \sum_{P,i|e_i^P=\overleftarrow{e}} (f^r_{\overleftarrow{e}} + \tau^r_{P,i-1}) \ \leq\ u_e \qquad \forall e\in E$$

holds for the corresponding edge-based representation $(f_e)_{e\in E}$ of $\mathcal{F}$.

Assume that that there exist an order of $\mathcal{P}^r = \{P_1, P_2, \ldots\}$ such that F_{P_k} is a feasible flow over time on the residual network $\mathcal{N}^r$ with respect to $(F_{P_i})_{i\in[k]}$. Then it is the nature of augmentation that, especially in this case, the generalized path-based flow over time is feasible.

In the following we establish some technical results which are primary used in Section 5.2. Further, they provide a detailed view into the mechanism of iterative augmentation. The next lemma shows a possibility to define the difference of two feasible flows over time.

Lemma 3.60. *Let $\mathcal{F}$ be a flow over time on some network $\mathcal{N} := (G, \mathcal{U}, \mathcal{T}, s, t)$ and $\mathcal{N}^r$ be the corresponding residual network. For another feasible flow over time $\tilde{\mathcal{F}}$ define $\mathcal{F}^r := (F_e^r)_{e\in E^r}$ by*

$$f_e^r := \max\{0, \tilde{f}_e - f_e\} \qquad \textit{and} \qquad f^r_{\overleftarrow{e}} := \max\{0, f_e - \tilde{f}_e\} + \tau_e \qquad \forall e\in E$$

Then $\mathcal{F}^r$ is a feasible flow on $\mathcal{N}^r$ with value $\mathrm{val}(\mathcal{F}^r) = \mathrm{val}(\tilde{\mathcal{F}}) - \mathrm{val}(\mathcal{F})$.

Proof. That $\mathcal{F}^r$ obeys the capacity constraints follows directly from the definitions of the residual capacities which, for all $e \in E$, imply

$$\begin{aligned} & f_e^r &= \max\{0, \tilde{f}_e - f_e\} \ \leq\ u_e - f_e \ =\ u_e^r \\ \text{and}\quad & f^r_{\overleftarrow{e}} &= \max\{0, f_e - \tilde{f}_e\} + \tau_e \ \leq\ f_e + \tau_e \ =\ u^r_{\overleftarrow{e}} \ .\end{aligned}$$

Next we prove that $\mathcal{F}^r$ satisfies the flow conservation constraints. For this consider a node $v \in V$ and let $e \in E$ be an incident edge. If $e \in \delta^+(v)$ we know $\overleftarrow{e} \in \delta^-(v)$ and

$$(f^r_e + \tau^r_e) - f^r_{\overleftarrow{e}} = (\tilde{f}_e - f_e) + \tau_e \ .$$

If $e \in \delta^-(v)$ we know $\overleftarrow{e} \in \delta^+(v)$ and

$$(f^r_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) - f^r_e = \tilde{f}_e - f_e \ .$$

Summing up the corresponding equations over all edges $e \in E$ which are incident to v we obtain

$$\begin{aligned}
&\sum_{e \in \delta^+_{G^r}(v)} (f^r_e + \tau^r_e) - \sum_{e \in \delta^-_{G^r}(v)} f^r_e \\
&= \sum_{e \in \delta^+_{G^r}(v)} (\tilde{f}_e + \tau^r_e) - \sum_{e \in \delta^-_{G^r}(v)} \tilde{f}_e - \Big(\sum_{e \in \delta^+_{G^r}(v)} (f_e + \tau^r_e) - \sum_{e \in \delta^-_{G^r}(v)} f_e \Big) \ .
\end{aligned}$$

Hence, for all $v \in V \setminus \{s,t\}$ we know that the right hand side is 0 because of the flow conservation constraints for $\mathcal{F}$ and $\tilde{\mathcal{F}}$. This shows that $\mathcal{F}^r$ satisfies flow conservation. Furthermore, for $v = t$ we get:

$$\mathrm{val}(\mathcal{F}^r) = \mathrm{val}(\tilde{\mathcal{F}}) - \mathrm{val}(\mathcal{F}) \qquad \square$$

In the following we use the symbol $\overleftarrow{e}$ not only for original edges but for all edges $e \in E^r$. In particular, if $e \in E^r$ is some backward edge then $\overleftarrow{e} \in E$ refers to the corresponding original edge. The next Lemma shows how the residual capacities are changed after an augmentation.

Lemma 3.61. *Let $\mathcal{N}^1 := (G^r, \mathcal{U}^1, \mathcal{T}^r, s, t)$ be the residual network of some network $\mathcal{N}$ with respect to some flow over time. Consider a feasible flow over time $\mathcal{F} := (f_e)_{e \in E^r}$ on $\mathcal{N}^1$. Further, let $\mathcal{N}^2$ be the residual network of $\mathcal{N}$ with respect to the augmented flow over time. Then we have for all $e \in E^r$*

$$u^2_e \ = \ u^1_e - f_e + (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \ .$$

Proof. Let $\mathcal{F}^1 := (f^1_e)_{e \in E}$ be the flow over time on $\mathcal{N}$ which defines the residual network $\mathcal{N}^1$. Because of the definition of the residual capacities we have for each original edge $e \in E$

$$u^1_e \ = \ u_e - f^1_e \qquad \text{and} \qquad u^1_{\overleftarrow{e}} \ = \ f^1_e + \tau_e \ .$$

Augment $\mathcal{F}^1$ with $\mathcal{F}$ and let $\mathcal{F}^2 := (f^2_e)_{e \in E}$ be the edge-based representation of the augmented flow. Then from (3.48) we know

$$f^2_e \ = \ f^1_e + f_e - (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \qquad \forall e \in E \ .$$

Since $\mathcal{N}^2$ is the residual network with respect to $\mathcal{F}^2$, we obtain for each original edge $e \in E$

$$u^2_e \ = \ u_e - \big(f^1_e + f_e - (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}})\big) \quad \text{and} \quad u^2_{\overleftarrow{e}} \ = \ \big(f^1_e + f_e - (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}})\big) + \tau_e.$$

Since $+\tau_e$ and $+\tau_{\overleftarrow{e}}$ work as shifts, this is equivalent to:

$$\begin{aligned} u_e^2 &= u_e^1 - f_e + (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}) \\ \text{and} \qquad u_{\overleftarrow{e}}^2 &= u_{\overleftarrow{e}}^1 + (f_e + \tau_e) + \big((-f_{\overleftarrow{e}}) + \tau^r_{\overleftarrow{e}} + \tau_e)\big) \\ &= u_{\overleftarrow{e}}^1 - f_{\overleftarrow{e}} + (f_e + \tau^r_e) \ . \end{aligned}$$

This shows $u_e^2 = u_e^1 - f_e + (f_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}})$ for all residual edges $e \in E^r$. □

The following Lemma provides some intuition for the composition of two augmentations.

Lemma 3.62. *Let $\mathcal{N}^1 := (G^r, \mathcal{U}^1, \mathcal{T}^r, s, t)$ be the residual network of some network $\mathcal{N}$ with respect to some flow over time. Consider a feasible flow over time $\mathcal{F}^1 := (f_e^1)_{e\in E^r}$ on $\mathcal{N}^1$ and let $\mathcal{N}^2$ be the residual network with respect to the corresponding augmented flow. Further, let $\mathcal{F}^2 := (f_e^2)_{e\in E^r}$ be a feasible flow on $\mathcal{N}^2$ and $\tilde{\mathcal{F}}_1 := (\tilde{f}_e^1)_{e\in E^r}$ be a flow satisfying $\tilde{f}_e^1 \le f_e^1$ for all residual edges $e \in E^r$ where equality holds for all points in time $\theta \in \mathbb{R}_+$ with $f_e^2(\theta) > 0$ or $f_{\overleftarrow{e}}^2(\theta + \tau^r_{\overleftarrow{e}}) > 0$. Then the family $\mathcal{F} := (f_e)_{e\in E^r}$ defined by*

$$f_e \ := \ \max\Big\{0,\ \tilde{f}_e^1 + f_e^2 - \big((\tilde{f}_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau^r_{\overleftarrow{e}}\big)\Big\} \qquad \forall e \in E^r$$

obeys the capacity constraints on $\mathcal{N}^1$.

Proof. Consider a residual edge $e \in E^r$ and let $\theta \in \mathbb{R}_+$ be some point in time. If $f_e^2(\theta) = 0$, we obtain

$$\Big(\tilde{f}_e^1 + f_e^2 - \big((\tilde{f}_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau^r_{\overleftarrow{e}}\big)\Big)(\theta) \ \le \ \tilde{f}_e^1(\theta) \ \le \ f_e^1(\theta) \ \le \ u_e^1(\theta) \ .$$

On the other hand, if $f_e^2(\theta) > 0$ holds implying that $\tilde{f}_e^1(\theta) = f_e^1(\theta)$ is valid just as $\tilde{f}_{\overleftarrow{e}}^1(\theta + \tau^r_{\overleftarrow{e}}) = f_{\overleftarrow{e}}^1(\theta + \tau^r_{\overleftarrow{e}})$, we know

$$\begin{aligned} \Big(\tilde{f}_e^1 + f_e^2 - \big((\tilde{f}_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau^r_{\overleftarrow{e}}\big)\Big)(\theta) &= \Big(f_e^1 + f_e^2 - \big((f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau^r_{\overleftarrow{e}}\big)\Big)(\theta) \\ &\le \big(f_e^1 + u_e^2 - (f_{\overleftarrow{e}}^1 + \tau^r_{\overleftarrow{e}})\big)(\theta) \\ &= u_e^1(\theta) \end{aligned}$$

because of Lemma 3.61. Since residual capacities are always nonnegative, this implies that $\mathcal{F}$ satisfies the capacities constraints on $\mathcal{N}^1$. □

For the final lemma of this section we have to decompose a flow over time which is feasible on some residual network. Unfortunately, the flow decomposition problem discussed in Section 3.5 deals with nonnegative transit times. Since flow decomposition is not the main subject of this thesis, we simply assume that flow decomposition works also on the residual networks used in this section. Nevertheless, we note the following without going into detail. Since the transit times in this section are constant, this ensures the existence of a flow decomposition especially for the case of rational transit times.

Lemma 3.63. *Let $\mathcal{F}^1 := (f_e^1)_{e\in E^r}$ and $\mathcal{F}^2 := (f_e^2)_{e\in E^r}$ be two flows over time on some residual graph G^r meaning, in particular, that $\mathcal{F}^1$ and $\mathcal{F}^2$ need not to satisfy capacity constraints. Then the family $\mathcal{F} := (f_e)_{e\in E^r}$ defined by*

$$f_e \ := \ \max\Big\{0,\ f_e^1 + f_e^2 - \big((f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau^r_{\overleftarrow{e}}\big)\Big\} \qquad \forall e \in E^r$$

is a flow over time, i.e., satisfies flow conservation.

Further, let $(F^1_P)_{P\in\mathcal{P}\cup\mathcal{C}}$, $(F^2_P)_{P\in\mathcal{P}\cup\mathcal{C}}$, *and* $(F_P)_{P\in\mathcal{P}\cup\mathcal{C}}$ *be a flow decomposition of* $\mathcal{F}^1$, $\mathcal{F}^2$, *and* $\mathcal{F}$, *respectively. Then we know*

$$\mathrm{val}(\mathcal{F}^1)+\mathrm{val}(\mathcal{F}^2) \;=\; \mathrm{val}(\mathcal{F})\ , \tag{3.49}$$

$$\sum_{P\in\mathcal{P}\cup\mathcal{C}} \tau^r_P\cdot\Big(\big\|F^1_P\big\|_\infty+\big\|F^2_P\big\|_\infty\Big) \;=\; \sum_{P\in\mathcal{P}\cup\mathcal{C}} \tau^r_P\cdot\big\|F_P\big\|_\infty\ , \tag{3.50}$$

$$\sum_{P\in\mathcal{P}\cup\mathcal{C}} |P|\cdot\Big(\big\|F^1_P\big\|_\infty+\big\|F^2_P\big\|_\infty\Big) \;\geq\; \sum_{P\in\mathcal{P}\cup\mathcal{C}} |P|\cdot\big\|F_P\big\|_\infty \tag{3.51}$$

where equality holds in the third equation if and only if no flow is decreased to zero, i.e., $f_e(\theta)=0$ *if and only if* $f^1_e(\theta)+f^2_e(\theta)=0$ *holds essentially for all* $e\in E^r$.

Proof. First we verify that $\mathcal{F}$ satisfies flow conservation. For this consider a node $v\in V$ and let $e\in E^r$ be a residual edge which is incident to v in G^r. If $e\in\delta^+_{G^r}(v)$ we know $\overleftarrow{e}\in\delta^-_{G^r}(v)$ and

$$\begin{aligned}
(f_e+\tau^r_e)-f_{\overleftarrow{e}} ={}& \max\Big\{0,\ f^1_e+f^2_e-\big((f^1_{\overleftarrow{e}}+f^2_{\overleftarrow{e}})+\tau^r_{\overleftarrow{e}}\big)\Big\}+\tau^r_e\\
&-\max\Big\{0,\ f^1_{\overleftarrow{e}}+f^2_{\overleftarrow{e}}-\big((f^1_e+f^2_e)+\tau^r_e\big)\Big\}\\
={}& \max\Big\{0,\ \big((f^1_e+f^2_e)+\tau^r_e\big)-(f^1_{\overleftarrow{e}}+f^2_{\overleftarrow{e}})\Big\}\\
&-\max\Big\{0,\ f^1_{\overleftarrow{e}}+f^2_{\overleftarrow{e}}-\big((f^1_e+f^2_e)+\tau^r_e\big)\Big\}\\
={}& \big((f^1_e+f^2_e)+\tau^r_e\big)-(f^1_{\overleftarrow{e}}+f^2_{\overleftarrow{e}})\ .
\end{aligned} \tag{3.52}$$

Summing up these equations over all edges $e\in\delta^+_{G^r}(v)$ we obtain

$$\begin{aligned}
&\sum_{e\in\delta^+_{G^r}(v)}(f_e+\tau^r_e)-\sum_{e\in\delta^-_{G^r}(v)}f_e\\
={}&\sum_{e\in\delta^+_{G^r}(v)}(f^1_e+\tau^r_e)-\sum_{e\in\delta^-_{G^r}(v)}f^1_e+\sum_{e\in\delta^+_{G^r}(v)}(f^2_e+\tau^r_e)-\sum_{e\in\delta^-_{G^r}(v)}f^2_e\ .
\end{aligned}$$

Hence, for all $v\in V\setminus\{s,t\}$ we know that the right hand side is essentially 0 because of the flow conservation constraints for $\mathcal{F}^1$ and $\mathcal{F}^2$. This shows that $\mathcal{F}$ satisfies flow conservation. In addition, for the case of $v=t$ we obtain

$$\mathrm{val}(\mathcal{F}^1)+\mathrm{val}(\mathcal{F}^2) \;=\; \mathrm{val}(\mathcal{F})$$

which proves (3.49).

In order to prove (3.50), let $v\in V$ be some node and $e\in\delta^+_{G^r}(v)$ be some outgoing residual edge. From (3.52) we get

$$\begin{aligned}
\tau_e\cdot(F_e+\tau^r_e)+\tau_{\overleftarrow{e}}\cdot F_{\overleftarrow{e}} &= \tau_e\cdot\big((F_e+\tau^r_e)-F_{\overleftarrow{e}}\big)\\
&= \tau_e\cdot\Big(\big((F^1_e+F^2_e)+\tau^r_e\big)-(F^1_{\overleftarrow{e}}+F^2_{\overleftarrow{e}})\Big)\\
&= \tau_e\cdot\big((F^1_e+F^2_e)+\tau^r_e\big)+\tau_{\overleftarrow{e}}\cdot(F^1_{\overleftarrow{e}}+F^2_{\overleftarrow{e}})\ .
\end{aligned}$$

Since each F is nonnegative, nondecreasing and shifting by a constant does not change the limit to ∞, this shows

$$\tau_e \|F_e\|_\infty + \tau_{\overleftarrow{e}} \|F_{\overleftarrow{e}}\|_\infty = \tau_e \Big(\|F_e^1\|_\infty + \|F_e^2\|_\infty) \Big) + \tau_{\overleftarrow{e}} \Big(\|F_{\overleftarrow{e}}^1\|_\infty + \|F_{\overleftarrow{e}}^2\|_\infty \Big) \,.$$

Summing up these equations over all v and all $e \in \delta^+_{G^r}(v)$ we obtain

$$\sum_{e \in E^r} \tau_e \|F_e\|_\infty = \sum_{e \in E^r} \tau_e \Big(\|F_e^1\|_\infty + \|F_e^2\|_\infty) \Big) \,.$$

Finally, observe that the right hand side of this equation equals the left hand side of (3.50) and vice versa. Hence, (3.50) is established.

For proving (3.51) consider a node $v \in V$ and let $e \in \delta^+_{G^r}(v)$ be some outgoing residual edge of v. Because of the definition of $\mathcal{F}$, we know $(f_e + \tau_e^r)(\theta) = 0$ or $f_{\overleftarrow{e}}(\theta) = 0$ for each point in time $\theta \in \mathbb{R}_+$. Because of (3.52) this shows

$$\begin{aligned} (f_e + \tau_e^r) + f_{\overleftarrow{e}} \;=\; \big|(f_e + \tau_e^r) - f_{\overleftarrow{e}}\big| \;&=\; \big|\big((f_e^1 + f_e^2) + \tau_e^r\big) - (f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2)\big| \\ &\leq\; \big((f_e^1 + f_e^2) + \tau_e^r\big) + (f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) \end{aligned}$$

implying

$$\|F_e\|_\infty + \|F_{\overleftarrow{e}}\|_\infty \;\leq\; \|F_e^1\|_\infty + \|F_e^2\|_\infty + \|F_{\overleftarrow{e}}^1\|_\infty + \|F_{\overleftarrow{e}}^2\|_\infty$$

where equality holds if and only if $(f_e^1 + f_e^2) + \tau_e^r$ and $f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2$ are mutually singular. Summing up these inequalities over all v and all $e \in \delta^+_{G^r}(v)$, we obtain

$$\sum_{e \in E^r} \|F_e\|_\infty \;\leq\; \sum_{e \in E^r} \|F_e^1\|_\infty + \|F_e^2\|_\infty$$

where equality holds if and only if $(f_e^1 + f_e^2) + \tau_e^r$ and $f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2$ are mutually singular for all $e \in E$. Now observe that the right hand side of this equation equals the left hand side of (3.50) and vice versa. For establishing (3.50) note the following. The functions $(f_e^1 + f_e^2) + \tau_e^r$ and $f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2$ are mutually singular if and only if $\min\{f_e^1 + f_e^2, (f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau_{\overleftarrow{e}}^r\}$ is essentially 0. Further, $f_e(\theta) = 0$ equals zero if and only if $(f_e^1 + f_e^2)(\theta) \leq ((f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2) + \tau_{\overleftarrow{e}}^r)(\theta)$. As all functions are nonnegative, the mutually singularity of $(f_e^1 + f_e^2) + \tau_e^r$ and $f_{\overleftarrow{e}}^1 + f_{\overleftarrow{e}}^2$ is equivalent to the statement that $f_e(\theta) = 0$ holds if and only if $f_e^1(\theta) + f_e^2(\theta) = 0$ holds essentially. $\square$

3.9 Your Comments

This chapter introduces the notion of flows over time used in this thesis. Based on a quite general definition of flows over time, natural assumption arising from potential real world applications are imposed and analyzed. Besides, several aspects with respect to flow behavior are discussed. Finally, classical flows over time are generalized to networks with time-varying capacity functions.

In literature, one possible criterion for classifying flows over time is the representation of time. *Discrete time models* allow the variation of flow only at a finite or, at most, a countable number of points in time leading to the notion of *discrete* flows over time. As done in Section 3.7, discrete flows over time are directly characterizable via their corresponding flow units. In contrast, in *continuous time models*, flow can vary at every point in time. For that reason, it is standard to define so-called *continuous* flows over time primary via Lebesgue integrable flow rate functions determining the rate at which flow travels through the network. Nevertheless, in Section 3.1 and 3.2, flows over time are primary defined via cumulative flow functions which are the nondecreasing and locally Lebesgue integrable. This is mainly motivated by simplifying reasons as this avoids some integrals. But you find another advantage. In [43] Koch, Nasrabadi, and Skutella define flow via finite Borel measures to combine the discrete and the continuous classical flow over time model in only one model. A finite Borel measure is uniquely determined by its distribution function which is nondecreasing and right continuous. Hence, all results which remain true when replacing the locally absolutely continuity with right-continuity are still valid for such a measure-based formulation of flows over time. One advantage of such a measure-based approach is that there is no need for a discrete model when discretizing a continuous flow over time. One can still work with the same measure-based formulation.

The presented flow over time model is based on flow-dependent transit times. Although measure theory provides a tool for extending this model, it is still quite general mainly due to two aspects. Firstly, a transit time functions is given by a Lebesgue measurable function and, secondly, may depend on past, present, and future flow behavior. In fact, for path-based models, a path transit time function depends on the entire flow over time and, for edge-based models, an edge transit time function depend on the entire flow on that edge. Further, using this general flow over time model, it shown how flow behavior is represented via outflow functions. Interestingly, instead of one outflow function, one needs an uncountable family of so called restricted outflow functions for modeling the entire flow behavior on an edge or a path. Just as inflow function also outflow function are nondecreasing and absolutely continuous because of the compatible-requirement for feasible transit time functions. In this sense, the presented flow over time model is closed.

This flow over time model generalizes known models which are based on transit times (like, e.g., [29, 45, 92]). Further, it shows a way to extend the notion of outflow models which are discussed in, e.g, [15, 48, 55]. So far, the flow behavior in outflow models is defined by exactly one outflow function assigned to an edge or path. As verified by Example 3.5, such outflow models are not completely well-defined. For that reason, it is mainly implicitly assumed that outflow models satisfy the FiFo assumption.

The FiFo principle is one of the most popular assumption used in dynamic network flow theory and related fields. With respect to arrival time functions you find mainly two formulations representing FiFo. Either arrival time functions have to be nondecreasing (see, e.g., [61]) or even *strictly* nondecreasing (see, e.g., [14, 92]). If arrival time functions are allowed to be constant as in the first case, outflow functions may be undefined in continuous flow over time models. This happens, e.g., if a positive amount of flow enters an edge over an interval where the corresponding arrival time functions is constant. This causes that a positive amount of flow leaves that edge at a single point in time which cannot be represented via *continuous* outflow functions. Such a flow behavior is avoided if FiFo is modeled by *strictly* nondecreasing arrival time functions. But this would imply that a probably most popular example of a flow model satisfying FiFo does not fit the definition of FiFo. This addresses the deterministic queuing, which intuitively satisfies FiFo, but for which partially constant arrival time functions are essential. In Subsection 3.3.1 a mix of these two definitions is used which says that whenever a positive amount of flow is sent over some time interval, the arrival time function must be strictly monotonically for at least one point out of this interval. Be that as it may, besides, you observe that for a measure-theoretic approach of flows over time the simpler non-strict version already works fine.

As you see, FiFo is only one aspect of consistent flow over time models, which are discussed in Section 3.3. The $\mathcal{F}$-continuity discussed in subsection 3.3.2 requires that the flow behavior, which is represented by outflow functions, behaves continuously depending on the inflow. This property seems to be not considered in literature explicitly. Nevertheless, it implicitly occurs in nearly all contributions dealing with *continuous* flows over times where the flow behavior is uniquely determined by the inflow. For example, load-dependent transit times just as load-dependent outflow rates are assumed to be continuous depending on the load (see, e.g., [29, 30, 92]). It seems and you strongly believe that this load-continuity implies $\mathcal{F}$-continuity. Nevertheless, a formal proof of that would be nice and, as far as you know, is missing in literature.

Another assumption, which is valid in many real routing scenarios, is given by the past-orientation of a flow model. In Subsection 3.3.3, it states that current flow behavior depends not on future scenarios. More precisely, the routing behavior of a flow particle entering an edge or a path at a certain point in time θ does not depend on flow particles entering this edge or path after time θ. Past-orientation is also known as *causality* in literature (see, e.g., [14]) and ensures together with the FiFo condition that transit times and outflows are somehow separable. Flow models based on load-dependent transit times are past-oriented by definition. In contrast, outflow models defined by load-dependent outflow rates are not past-oriented.

The last assumption which is required for consistent flow models is that the set of s-t-paths must be locally finite. This assumption is only imposed for path-based flow models and does not occur in literature. This assumption is needed in the next chapter for proving the existence of Nash equilibria. Therefore, you discuss this condition in Section 4.4.

The four assumptions – FiFo, $\mathcal{F}$-continuity, past-orientation, and locally finiteness of $\mathcal{P}$ – define consistent flow models. Such consistent flow over time models are representable by a nice arrival time function which is introduced in Section 3.4. The promising feature of this so-called foresighted arrival time functions becomes apparent in Section 4.2 of the next chapter where Nash equilibria for routing games over time are defined. In fact, the properties of foresighted arrival time function by themselves ensures that Nash equilibria over time are well-defined. You find it worth to note that foresighted arrival time functions may have discontinuities with respect to time. In this manner, flow models defined via load-dependent transit times can be consistent even if transit time functions are not continuous with respect to the load.

In Section 3.5 dynamic network loading and dynamic flow decomposition is considered. Although it is stated that this section is mainly based on intuition, you find it sometimes quite technical. So you search for points which are not discussed down to the smallest detail. Beside the obvious part about the correctness of the Flow Decomposition algorithm, you identify three other parts. Firstly, it is not shown that the sequence of flows over time considered at the end of the proof of Lemma 3.42 converges. However, this would be directly implied by the reverse direction of Lemma 3.30, which most likely holds. Secondly, Example 3.43 does not satisfy the FiFo-assumption. However, this could be enforced by considering τ_e as an auxiliary function which defines the actual arrival time function by

$$\begin{aligned} & \ell_e(F_e^+)(\theta) &=& \max_{\theta' \le \theta}\bigl(\theta' + \tau_e(F_e^+)(\theta')\bigr) \\ \text{or} \qquad & \ell_e(F_e^+)(\theta) &=& \min_{\theta' \ge \theta}\bigl(\theta' + \tau_e(F_e^+)(\theta')\bigr) \ . \end{aligned}$$

In both cases FiFo holds and the scenarios discussed in Example 3.43 are not changed.

Thirdly, equation (3.37) and (3.38) describe two ways for computing the inflow $F_{P,i}^+$ contributed by an s-t-path P to its i-th edge e_i. Surprisingly, you does not find a proof in literature showing the correspondence equivalence. Maybe, this seems to be too obvious. Nevertheless, this describes a fundamental property of flows over time. So you try to prove this inductively having the recursion $\ell_{P,i}^{-1}([0,\theta]) = \ell_{P,i-1}^{-1}(\ell_{e_i}^{-1}([0,\theta]))$. Using integration by substitution and the fundamental theorem of calculus, the induction step could look like (please do not follow the details):

$$\begin{aligned} F_{P,i+1}^+ &= \int_{\ell_{P,i}^{-1}([0,\theta])} f_P^+(\vartheta)\, d\vartheta = \int_{\ell_{P,i-1}^{-1}(\ell_{e_i}^{-1}([0,\theta]))} f_P^+(\vartheta)\, d\vartheta \\ &= \int_{\ell_{e_i}^{-1}([0,\theta])} (\ell_{P,i-1}^{-1})'(\vartheta) \cdot f_P^+(\ell_{P,i-1}^{-1}(\vartheta))\, d\vartheta \\ &= \int_{\ell_{e_i}^{-1}([0,\theta])} \frac{d}{d\vartheta}\Bigl(\int_{[0,\vartheta]} (\ell_{P,i-1}^{-1})'(\vartheta') \cdot f_P^+(\ell_{P,i-1}^{-1}(\vartheta'))d\vartheta'\Bigr)\, d\vartheta \\ &= \int_{\ell_{P,i}^{-1}([0,\theta])} \frac{d}{d\vartheta}\Bigl(\int_{\ell_{P,i-1}^{-1}([0,\vartheta])} f_P^+(\vartheta')\ \vartheta'\Bigr) d\vartheta \\ &= \int_{\ell_{e_i}^{-1}([0,\theta])} f_{P,i-1}^-(\vartheta)\, d\vartheta = \int_{\ell_{e_i}^{-1}([0,\theta])} f_{P,i}^+(\vartheta)\, d\vartheta \ . \end{aligned}$$

Luckily, you know a little bit about measure theory which shortens this to one line. To see this, let μ_{i+1} be the measure denoting the flow $F_{P,i+1}^+$. That is, for a measurable set B, the value $\mu_{i+1}(B)$ equals the amount of flow entering the i-th edge e_{i+1} over the times in B. Then it holds for all measurable sets B:

$$\mu_{i+1}(B) = \mu_P\bigl(\ell_{P,i}^{-1}(B)\bigr) = \mu_P\bigl(\ell_{P,i-1}^{-1}(\ell_{e_i}^{-1}(B)\bigr) = \mu_i\bigl(\ell_{e_i}^{-1}(B)\bigr) \ .$$

This shows that using measure theory would simplify this particular proof. However, it should be noted that, inserting $[0,\theta]$ for B, the well-known relation $\mu_{i+1}([0,\theta]) = F_{P,i+1}^+(\theta)$ establishes this fundamental property for flow models satisfying FiFo also in one line.

Further, you are interested in literature addressing dynamic network loading and dynamic flow decomposition. The dynamic network loading problem is analyzed by, e.g., Xu et. al [92]. They consider continuous load-dependent transit times under the strict FiFo-assumption and show existence and uniqueness of a solution to the dynamic network loading problem. In this sense, the approach in Section 3.5 generalizes there results. Unfortunately, it seems that there is no contribution addressing dynamic flow decomposition so far. But if you find one you will directly inform the author of this thesis.

According to the shortest path problem, considered in Section 3.6, you find the article of Orda and Rom [65] from 1990 which you have already inserted at the right place in this section. Orda and Rom solve the problem of finding all shortest s-t-path simultaneously for all points in time in the same manner. They make the following really interesting remark. The DYNAMIC BELLMANN-FORD algorithm can be seen as a polynomial algorithm in the space of functions, i.e., the number of simple operation between functions is polynomial in the input size. You are wondering, if there is any other algorithm working with functions which is polynomial in this sense.

Finally, you discuss Section 3.8 which generalizes classical flows over time to arbitrary time-varying capacities. This generalized model is studied first by Philpott [71] in 1990. The definition of the residual network together with the proof of Lemma 3.58 is done by Koch, Nasrabadi, and Skutella [43] even for the more general case of measure-based flows over time. Besides, waiting is allowed in both cited articles.

Since you do not find a complete, mathematically exact formulation of consistent flow over time models, you collect all ingredients into the following definition. You concentrate on the path-based definition because the edge-based notion can easily be read off this characterization. Besides, it is worth to mention that *no* arrival times are used.

Definition 3.64 (Consistent Flow over Time Model)**.** For all paths P and all times $\theta, \theta' \in \mathbb{R}_+$, a *consistent* path-based flow over time model has to

- be $\mathcal{F}$-continuous, i.e., each map $\mathcal{F}^+ \mapsto F_P^-|_{\leq\theta}$ has to be continuous as a map from $(L_\infty^{\mathrm{loc}})_1^{\mathcal{P}}$ to L_∞^{loc} ,
- satisfy FiFo, i.e., $F_P^-|_{\leq\theta}(\theta') \;=\; \min\{F_P^-(\theta'), F_P^+(\theta)\}$,
- obey past-orientation, i.e., $F_P^-|_{\leq\theta}(\mathcal{F}^+) = F_P^-(\mathcal{F}^+|_{\leq\theta})$.

Further, the set $\mathcal{P}$ of s-t-paths must be locally finite. That is, for each $\theta \in \mathbb{R}_+$ the set $\mathcal{P}_{\leq\theta} \;:=\; \{P \in \mathcal{P} \mid \exists \mathcal{F}^+ : F_P^+(\theta) > 0\}$ is finite.

Chapter 4

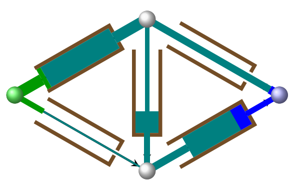
Routing Games over Time

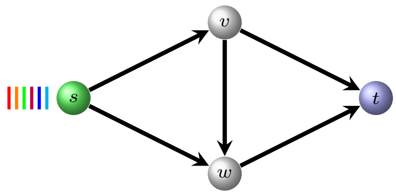
As it is the case in static routing games, also in routing games over time, each player controls some flow which has to be send through a given network. For this each player chooses a unique path from some source to some sink along which all of his flow is sent. But in contrast to the static setting, the time dimension is a key feature of routing games over time. In addition to the flow which is controlled by the players, also the points in time at which the flow originates are given in advance. Moreover, we assume that flow enters the chosen paths immediately after origination. In particular, the players do not have to chose departure times for their flow implying that we only consider so-called route choice models.

The routing decisions of the players lead to a path-based flow over time and the flow behavior is given by some flow over time model. Hence, we are able to determine the arrival times of the flows of all players. In addition, we can check whether or not it is worth for a player to switch unilaterally to another quicker path. This allows us to characterize Nash equilibria for routing games over times based on the assumption that each player wants to send his flow as fast as possible through the network.

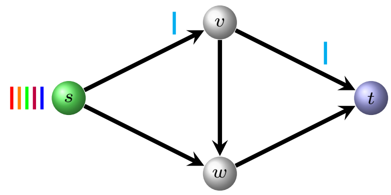
For the purpose of this thesis we restrict to consistent flow over time models in this chapter. Further, we only consider the single commodity case where the entire flow of each player has to be sent from a unique source s to a unique sink t.

In Section 4.1 we consider atomic routing games over time which are determined by a finite or, at most, by a countable number of players each of which controls a positive amount of flow. For this setting we observe that Nash equilibria always exist. The nonatomic case is handled in Section 4.2. It turns out that Nash Equilibria are completely characterized by a family of *static* flows. Based on this, we show that Nash equilibria always exist. This existence is established for path-based just as edge-based flow models. The corresponding proof is constructive and implies on the side that the definition of a nonatomic equilibrium is well-motivated as it occurs as the limit of atomic Nash equilibria. Note that nonatomic games are used for approximating atomic games if the number and the density of players is by far too large for obtaining computational results. Finally, in Section 4.3 we discuss several adequate objectives for defining prices of anarchy and stability which are used to measure the performance of Nash equilibria.

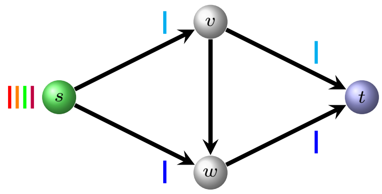




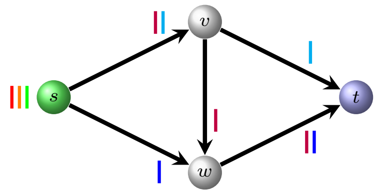
(a) The basic network of this example.



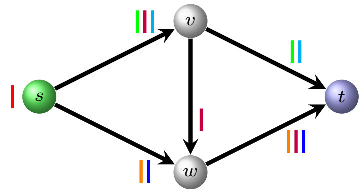
(b) Routing of the first flow unit.



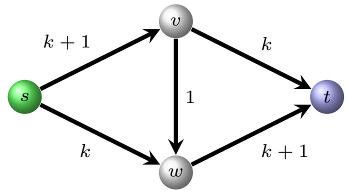
(c) Routing of the first two flow units.



(d) Routing of the first three flow units.



(e) Routing of the first five flow units.



(f) Routing of the first $2k+1$ flow units.

Figure 4.1: Routing behavior of Example 4.1. All transit times are 0 and all capacities are 1 and the flow behavior is given by the discrete deterministic queuing model.

4.1 Atomic Case

In this section we define atomic routing games over time by discussing all necessary components. This enables us to characterize atomic Nash equilibria over time. It turns out that such Nash equilibria always exist provided the routing occurs on a consistent discrete flow model. For explaining the subsequent definitions we start with an example which uses the discrete version of the deterministic queuing model. As already mentioned, in an atomic routing game over time the number of players is at most countable. Further, each player controls a nonnegative amount of flow which originates as a whole. So we can identify each player with his flow unit.

Example 4.1. To follow this example, take a look at Figure 4.1. We assume that all flow units have a size of 1 which are located in an initial waiting queue in front of the unique source s. In that sense all flow units originate at time 0. Further, every flow unit knows the routing of its predecessors in the initial waiting queue, and if two flow units arrive at one edge at the same time, the flow unit with the smaller position in the initial waiting queue traverses the edge first. That is, flow units which are located nearer to the source have priority.

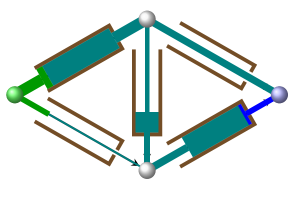
The routing behavior of the flow units on an edge is determined as follows. Each edge has a free flow transit time of 0. This means that if an edge is unused, a flow unit entering an edge e at a point in time $\theta \in \mathbb{R}_+$ arrives at the head of e instantaneously, i.e., at time θ. In addition, there is, in general, a waiting queue at the end of each edge because we require that only one flow unit can leave an edge over one time unit. That is, each edge has a capacity of 1 which remains constant over time. Thus, if a flow unit enters an edge e at time θ it arrives at the head of e at time $\theta + q_e(\theta)$ where $q_e(\theta)$ is the number of flow units waiting at the end of e at time θ. Within the waiting queue overtaking is not allowed implying that FiFo holds on an edge. Further, as already mentioned, if two flow units enter an edge at the same point in time, the one with the lower position in the initial waiting queue has priority.

Since the flow units choose their paths through the network by themselves, we assume that every flow unit chooses an s-t-path which will lead them as quickly as possible to the unique sink t. In the following we consider such a routing on the network in Figure 4.1a.

We refer to the s-t-path svt as the upper path, to swt as the lower path, and to $svwt$ as the zigzag path. Of course, the first cyan flow unit can choose every path in order to arrive at t at time 0. We assume that the first cyan flow unit takes the upper path (see Figure 4.1b). If the second blue flow unit would also use the edge sv, it must wait one time unit at the end of this edge. Therefore, if the second blue flow unit would choose the upper or the zigzag path, it would be at t at time 1. Of course, if the second blue flow unit uses the lower path, it arrives at t at time 0. Thus, it chooses the lower path (see Figure 4.1c). Because the third purple flow unit must wait at the end of sv and sw one time unit, it arrives at node v or w at time 1. But at time 1 flow unit 1 and 2 are already at t such that the third purple flow unit has no further waiting time after reaching v or w and arrives at t at time 1 whatever which path would be chosen. To get a more interesting situation we assume that flow unit 3 traverses the network through the zigzag path to be at t at time 1 (see Figure 4.1d).

If the fourth green flow unit traverses the edge sv, it must wait two time units and reaches v at time 2. Since the first three flow units are already at t at this time, it would arrive at t at time 2 using the upper or the zigzag path. If the fourth green flow unit would use the lower path, it must wait one time unit at the end of the edge sw in order to arrive at w at time 1. But at time 1 the third purple flow unit enters the edge sw such that the fourth green flow unit must again wait one time unit at the end of this edge and would be again at t at time 2. Thus, we may conclude that the fourth green flow unit can take every path. But if the fourth green flow unit takes the zigzag path, it would arrive at w at time 2, and therefore, the next orange flow unit would choose the lower path in order to be at w at time 1. This means that the fifth orange flow unit would overtake the fourth green flow unit which cannot be tolerated by the fourth green flow unit. Therefore, flow unit 4 chooses either the upper or the lower path. If the fourth green flow unit takes the upper path, the fifth orange flow unit uses the lower path and vice versa (see Figure 4.1e).

The situation for the next red flow unit is the same as for the fourth green flow unit such that one of the next two flow units must choose the lower and the other the upper path in order to avoid being overtaken. This leads again to the same situation for the next two flow units, and from now on the routing of further flow units is always done in this manner. Thus, k of the first $2k+1$



flow units take the upper path, one the zigzag, and the other k take the lower path (see Figure 4.1f).

We see that the flow units are routed such that they are at every node of the chosen path as early as possible. This ensures that, in the resulting routing, every flow unit arrives the sink t as early as possible. Thus, we have a Nash equilibrium. On the other hand, simply choosing a currently quickest path (without being as early as possible at every intermediate node) does not lead to an equilibrium as explained for the fourth green flow unit.

In the following we formalize the routing behavior in Example 4.1 for an arbitrary consistent discrete flow over time model. We start with a brief description of the scenario leading to the notion of atomic routing games over time.

Scenario. As already mentioned, we have players who want to send their flow units through a network from a given source s to a given sink t in a selfish manner. The interesting aspect is that the routing occurs over time. In particular, this means that flow units have predetermined departure times which can also be interpreted as origination times. Further, each player acts free in order to choose an s-t-path along which his positive amount of flow is sent. The underlying discrete flow over time model defines how the flow units behave while traveling. In particular, every flow unit travels as a whole, i.e., has no physical dimension ensuring, e.g., that departure and arrival times are unique. Using this knowledge, players try to send their flow units to t as early as possible. If no player can improve the arrival time of his flow unit, the current routing is stable and, in fact, a Nash equilibrium. In the following we discuss all ingredients of an atomic routing game more in detail.

Players. In an atomic routing game over time we are given an at most countable set of players which is represented by $\mathbb{N}$. Each *player* $i \in \mathbb{N}$ is determined by an *origination time* $\theta_i \in \mathbb{R}_+$ and by a *flow value* $d_i \in \mathbb{R}_+$. That is, the flow of player i originates at s at time θ_i and controls flow of value d_i which should be send from s to t along one s-t-path. In particular, we allow that a flow unit carries no flow in order to model a finite set of players. We assume that the players are ordered with respect to the origination time, i.e., $\theta_i \geq \theta_j$ for all $i > j$. Note that this assumption is, in fact, a requirement in case of countable many players. Intuitively, player i originates at the i-th position in front of the source s.

When resolving conflicts in case two players want to enter the same edge or the same path at the same time, player i has priority against another player j if $i < j$. Hence, the origination time or, in fact, the position of a player serves as a priority.

Finally, we assume that the flow controlled by each player has no physical dimension like a length or a width. So it originates as a whole, travels through the network as a whole, and arrives at t as a whole. In the following, we denote by $D := (d_i, \theta_i)_{i \in \mathbb{N}}$ the origination pattern of the flow units which are controlled by the players.

Strategies. A player wants to send his entire flow from the source s to the sink t along one s-t-path. So the strategy set of each player equals the set $\mathcal{P}$ of all s-t-paths. Let P_i be the chosen path of player i. Then the routing decisions of all players yield a discrete path-based flow over time $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$. Vice versa, every discrete flow over time can be interpreted as some strategy profile if all flow units of the players are sent through the network.

Definition 4.2 (Strategy Profile)**.** Let $\mathcal{F} := (F_P^+)_{P\in\mathcal{P}}$ be a discrete path-based flow over time. Then $\mathcal{F}$ is called *strategy profile* of the players $D := (\theta_i, d_i)_{i\in\mathbb{N}}$ if and only if

$$\bigcup_{P\in\mathcal{P}} N_P \subseteq \{i \in \mathbb{N} \mid d_i > 0\} .$$

Recall that N_P is the set of flow units traversing s-t-path P (see Definition 3.54).

Costs. The costs occurring in an atomic routing game over time are based on some discrete consistent flow model $\mathcal{T}$. Assume that player i has chosen s-t-path P_i. We define the cost of player i in some strategy profile $\mathcal{F}$ as the real arrival time $\ell_{P_i}(\mathcal{F})(\theta_i) := \theta_i + \tau_{P_i}(\mathcal{F})(\theta_i)$ of the corresponding flow unit. Note that in case of an edge-based flow model we first have to solve the discrete dynamic network loading problem before we are able to compute the costs.

Definition 4.3 (Atomic Routing Game over Time)**.** An *atomic routing game over time* is a tuple $(G, \mathcal{T}, s, t, D)$ consisting of a directed graph $G := (V, E)$, a source $s \in V$, a sink $t \in V$, and a discrete flow pattern D denoting the set of players together with a discrete flow over time model $\mathcal{T}$.

Like in static atomic routing games, a Nash equilibrium is characterized by a flow over time $\mathcal{F}$ where no player has an incentive to change the chosen path in order to reduce the cost.

Definition 4.4 (Nash Flow over Time)**.** Let $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ be a discrete flow over time determining the routing decisions of the players in an atomic routing game over time. Then $\mathcal{F}$ is a *discrete Nash equilibrium* (*discrete Nash flow over time*) if and only if, for all $i \in \mathbb{N}$, it holds that

$$\ell_{P_i}(\mathcal{F})(\theta_i) \leq \ell_P(\mathcal{F}^{i\to P'})(\theta_i) \qquad \forall P' \in \mathcal{P} .$$

Here, $\mathcal{F}^{i\to P'} := (F_P^{i\to P'})_{P\in\mathcal{P}}$ denotes the flow which is obtained in case that player i switches from path P_i to path P', i.e.,

$$F_P^{i\to P'} := \begin{cases} F_P \setminus (d_i, \theta_i) & \text{if } P = P_i \\ F_P \cup (d_i, \theta_i) & \text{if } P = P' \\ F_P & \text{otherwise} \end{cases} .$$

In Example 4.1 a Nash equilibrium arises if each flow unit is sent over currently shortest path such that it is as early as possible at each intermediate node. The following theorems shows that this works for all consistent atomic routing games over time. In particular, this ensures the existence of Nash equilibria.

Theorem 4.5. *For an atomic routing game over time $(G, \mathcal{T}, s, t, D)$ a Nash equilibrium always exists if the underlying discrete flow model is consistent.*

Proof. Consecutively, route the flow of every player along a currently shortest path. That is, we treat the players represented by $\mathbb{N}$ in increasing order. In particular, this ensures that the departure times of the considered the flow units are nondecreasing. For the current flow unit i we compute a shortest path P_i from s to t. If the underlying flow model is edge-based i must also arrive as early as possible at each intermediate node of P. This can be achieved by using Djikstra because the edge arrival times are nondecreasing with time as the flow model satisfies FiFo. We show that the resulting routing $\mathcal{F}$ is a Nash equilibrium. This is rather clear if the consistent flow model is path-based. So we concentrate on edge-based models in the following.

Let $\ell_v(i)$ be the earliest points in time at which flow unit i could arrive at node $v \in V$ if player i is handled by the algorithm. We first observe by induction over the sequence of players that $\ell_v : \mathbb{N} \to \mathbb{R}_+$ is nondecreasing for each $v \in V$. For this assume that everything works fine until flow unit i, and let $j := i + 1$ be the next flow unit. Further, assume that our induction assumption fails for j, and let w be a node such that $\ell_w(j) < \ell_w(i)$ and consider some edge $e = vw \in \delta^-(w)$ along which j is able to arrive at w at time $\ell_w(j)$. If $\ell_v(j) \geq \ell_v(i)$, we would obtain

$$\begin{aligned} \ell_w(j) \;&=\; \ell_e\big(F_e^+|_{<j}\big)\big(\ell_v(j)\big) \\ &\geq\; \ell_e\big(F_e^+|_{<j}\big)\big(\ell_v(i)\big) \;\geq\; \ell_e\big(F_e^+|_{<i}\big)\big(\ell_v(i)\big) \;\geq\; \ell_w(i) \end{aligned}$$

because the flow model satisfies the FiFo principle and is past-oriented. Thus we have $\ell_v(j) < \ell_v(i)$. Proceeding in this manner, we end up with $\ell_s(j) < \ell_s(i)$ which is a contradiction as $i < j$ implies that j departures not before i, i.e., as we know $\ell_s(j) \geq \ell_s(i)$. Hence, the induction assumption holds and ℓ_v is nondecreasing in the order of N for each $v \in V$.

So far we have shown that no flow unit j overtakes a flow unit $i < j$. But this immediately implies that no flow unit has an incentive to switch to another path. To see this, observe that the restricted entering flows until some flow unit i remains constant on each edge after flow unit i is handled by the algorithm. So the Lemma follows from the past-orientation of the flow model. □

4.2 Nonatomic Case

In this section we consider nonatomic routing games over time. The main difference to the atomic counterpart is that the number of players is no longer finite or countable. In fact, we deal with an uncountable number of players. Furthermore, each player only controls only a negligible fraction of flow instead of a concrete flow unit. Like in atomic routing games all players want to send their flow particles selfishly through a network leading to the notion of nonatomic Nash equilibria over time. Characterizing these Nash equilibria is the main subject of this section.

This section is organized as follows. First, we explain all ingredients of a nonatomic routing game starting with an informal description. This part ends with a definition of nonatomic Nash equilibria over time. In what follows, we establish several equivalent formulations of the underlying Nash flows over time. Finally, we show that a Nash equilibrium always exists. This existence proof is based on a discretization of the nonatomic routing game in atomic counterparts and, therefore, constructive. As a side product, we see that a nonatomic Nash equilibrium over time can be approximated by atomic Nash flows. In this sense, the given definition of Nash equilibria is well-motivated.

The behavior of the controlled flow particles in a nonatomic routing game is determined by a flow model. In this section we only deal with consistent flow models over time. On that account, the symbol ℓ always refers to *foresighted* arrival times.

It lies in the nature of routing games that the behavior of the players results in path-based flows over time. Even if the flow model is *edge-based*, we have to consider *path-based* flows over time using the dynamic network loading problem. Nevertheless, edge-based flows are much more interesting especially with respect to computational aspects. For this reason, it is one main objective of this section to characterize edge-based flows over time which imply a path-based Nash flow over time via path decomposition. Therefore, we only consider *path-decomposable* edge-based flows over time in this section.

Scenario. Like atomic routing games over time also nonatomic routing games over time occur on a given network. Each player controls an infinitesimal flow unit, a so-called flow particle, which should be send from a unique source s to a unique sink t. Since the departure time of each flow particle is given in advance, each player has to select an s-t-path. In this connection, all players act against each other having in mind to minimize the arrival time of their particles. A Nash equilibrium arises out of the selfish routing decisions if all players cannot decrease the arrival time of their flow particle by switching to another s-t-path. In this sense, a Nash equilibrium is stable with respect to the egoistic behavior of the players. In the following we discuss this scenario more formally.

Players. As already mentioned, each player corresponds to a flow particle. How these flow particles originate at s is represented by a *(cumulative) supply function* $D : \mathbb{R}_+ \to \mathbb{R}_+$. The supply function D is assumed to be absolutely continuous and nondecreasing with $D(0) = 0$. Here, $D(\theta)$ equals the total amount of flow which originates at s until some point in time $\theta \in \mathbb{R}_+$. Or equivalently, the departure time of the flow particle at position $D(\theta)$ is θ. Since D is continuous, each player can be identified by the origination time of the corresponding flow particle in contrast to the atomic counterpart. In addition, the assumptions on D imply the existence of a *supply rate function* $d : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\frac{dD}{d\theta}(\theta) = d(\theta)$ for almost all $\theta \in \mathbb{R}_+$ and $D(\theta) := \int_0^\theta d(\vartheta)\, d\vartheta$ for all $\theta \in \mathbb{R}_+$. For some point in time $\theta \in \mathbb{R}_+$, the value $d(\theta)$ equals the rate at which flow controlled by the players originates at s.

Strategies. Since all players want to send their flow particles from the source s to the sink t, the strategy set of each player equals the set $\mathcal{P}$ of all s-t-paths. Hence, the routing decisions of all players yield a continuous path-based flow over time $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ where the total net outflow of s equals the demand. Vice versa, every such continuous flow over time corresponds to some strategy profile.

Definition 4.6 (Strategy Profile)**.** Let $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ be a continuous path-based flow over time. Then $\mathcal{F}$ is called strategy profile of the nonatomic players D if and only if

$$\sum_{P \in \mathcal{P}} F_P \;=\; D \;.$$

Costs. The costs occurring in nonatomic routing games over time are based on consistent flow models. Let $\mathcal{L}$ be some family of foresighted arrival times and assume that a flow particle originating at time θ is sent along the s-t-path P. We define the cost of the corresponding player in some strategy profile $\mathcal{F}$ as the arrival time $\ell_P(\mathcal{F})(\theta)$ of this flow particle. Herewith, the discussion of the ingredients of a nonatomic routing game over time ends.

Definition 4.7 (Nonatomic Routing Game over Time)**.** A *nonatomic routing game over time* $(G, \mathcal{L}, s, t, D)$ consists of a directed graph $G := (V, E)$, a source $s \in V$, a sink $t \in V$, and a supply function $D : \mathbb{R}_+ \to \mathbb{R}_+$ encoding the set of players together with a consistent flow model given by the family $\mathcal{L}$ of arrival times.

Like in static nonatomic routing games, a Nash equilibrium over time is characterized by a flow over time $\mathcal{F}$ where no player has an incentive to change the chosen path in order to reduce the cost.

Definition 4.8 (Nonatomic Nash Flow over Time). Let $\mathcal{F} := (F_P)_{P \in \mathcal{P}}$ be a continuous flow over time determining the routing decisions of the players in a nonatomic routing game over time $(G, \mathcal{L}, s, t, D)$. Then $\mathcal{F}$ is a *nonatomic Nash equilibrium* (*nonatomic Nash flow over time*) if and only if

$$f_P(\theta) \;>\; 0 \qquad \Rightarrow \qquad \ell_P(\mathcal{F})(\theta) \;=\; \min\bigl\{\ell_{P'}(\mathcal{F})(\theta) \;\big|\; P' \in \mathcal{P}\bigr\} \;.$$

holds for almost all times $\theta \in \mathbb{R}_+$.

As we see at the end of this section, Definition 4.8 is well-motivated. But before we proceed with the main discussion of this section, we address the following question. Why we use foresighted arrival times instead of lower or upper arrival times? The answer is given by Example 5.14 dealing with the directed flow model. It shows that neither lower nor upper arrival times reflect our intuition about Nash equilibria in general. Also note that the considered scenario shows exactly the same behavior in case we are given the very popular deterministic queuing model presented in Chapter 7. The following theorem gives an alternative characterization of Nash flows over time which holds for path-based flows over time.

Theorem 4.9. *Let $\mathcal{F} := (F^+_P)_{P \in \mathcal{P}}$ be path-based flow over time on a consistent flow model. Then $\mathcal{F}$ is a Nash flow over time if and only if*

$$F^-_P\bigl(\ell_P(\theta)\bigr) \;=\; F^-_P\bigl(\ell_t(\theta)\bigr) \qquad\qquad \forall \theta \in \mathbb{R}_+, P \in \mathcal{P} \;. \tag{4.1}$$

Proof. First assume that $\mathcal{F}$ is a Nash flow over time and let P be an s-t-path and $\theta \in \mathbb{R}_+$ be a point in time. If $\ell_P(\theta) = \ell_t(\theta)$ holds, (4.1) is obviously satisfied by P at time θ. Hence, we only have to consider the case $\ell_P(\theta) > \ell_t(\theta)$.

Let $\theta_1 := \inf\{\theta' \geq 0 \mid \ell_P(\theta') > \ell_t(\theta)\}$ be the earliest point in time at which a flow particle would arrive later than $\ell_t(\theta)$ along P. Since $\theta_1 \leq \theta$ holds, we know

$$\ell_P(\theta') \;>\; \ell_t(\theta) \;\geq\; \ell_t(\theta')$$

for all $\theta' \in (\theta_1, \theta]$ because node arrival time functions are nondecreasing. This shows that P is not a currently shortest path for all departure times $\theta' \in (\theta_1, \theta]$ implying that F_P^+ is constant over $(\theta_1, \theta]$ by Definition 4.8. This shows

$$F_P^+(\theta) \;=\; F_P^+(\theta_1) \;=\; F_P^-\big(\ell_P(\theta_1)\big) \;\le\; F_P^-\big(\ell_t(\theta)\big) \;\le\; F_P^-\big(\ell_P(\theta)\big) \;=\; F_P^+(\theta)\ .$$

implying that (4.1) holds for P at time θ.

For proving the other direction let P be an s-t-path and assume that (4.1) holds for P. We have to show that the condition in Definition 4.8 is valid for P. If $\ell_P(\theta) = \ell_t(\theta)$ holds for some point in time θ nothing has be shown. Therefore, let $\theta \in \mathbb{R}_+$ be a point in time with $\ell_P(\theta) > \ell_t(\theta)$. Since the foresighted node arrival time function ℓ_t is right continuous, there exists an $\epsilon > 0$ such that $\ell_P(\theta) > \ell_t(\theta + \epsilon)$ implying

$$F_P^-\big(\ell_P(\theta)\big) \;\ge\; F_P^-\big(\ell_t(\theta+\epsilon)\big) = F_P^-\big(\ell_t(\theta)\big).$$

because F_P^- is nondecreasing. Because of (4.1) this shows $F^+(\theta) = F^+(\theta+\epsilon)$ implying that f_P^+ is essentially zero on $(\theta, \theta+\epsilon)$. Thus, also the condition in Definition 4.8 is essentially valid for P over $(\theta, \theta+\epsilon)$. Hence, using Corollary 2.8 this proof is done. □

Intuitively, Definition 4.8 means that flow should only be send over currently shortest paths in a Nash flow. On the other hand, equation (4.1) implies that all flow particles entering an s-t-path P before time θ are already at t at time $\ell_t(\theta)$. Since no particle departing at time θ is able to arrive at t before time $\ell_t(\theta)$, this means that no particle has the possibility to overtake any other flow particle. In this sense, Theorem 4.9 shows that routing along currently shortest paths is equivalent to this nonovertaking condition.

In the following we show that this observation also holds for edge-based flows over time. As already seen, putting this aspect into a formal framework is straight forward for path-based flow models. However, if the underlying flow model is edge-based, this requires a little bit more technical work. In particular, we have to consider the question, whether or not an *edge-based* flow over time $\mathcal{F}$ has a path decomposition satisfying Definition 4.8. In this case we also call $\mathcal{F}$ a *Nash equilibrium*. In fact, it turns out that, for an edge-based Nash flow over time, every path decomposition satisfies Definition 4.8.

In the next part of this section, $\mathcal{F}$ takes the role of an edge-based flow over time which is path decomposable. That is, $\mathcal{F}$ contains no isolated flow carrying cycles because flow on such cycles cannot be supplied from the source s. If $\mathcal{F}$ is exceptionally path-based, it is mentioned explicitly. For the sake of clarity, as $\mathcal{F}$ is given, we use the symbol ℓ for denoting foresighted arrival time *functions* $\ell(\mathcal{F})$. The following intuitive real-world example illustrate the subsequent definitions and results.

Example 4.10. Suppose you are at the airport and, since you are already late, you want to get to your departure gate as quickly as possible. But first you have to check-in. Afterwards, you head for the security check in order to finally get to your gate and board the aircraft. But there is a waiting queue in front of the check-in counter. The question is how quickly you should approach the end of the waiting queue at the check-in counter. Of course, as long as the last person in line remains the same, i. e., no one else enters the line, it does not matter at what time you line up – you always leave the check-in counter at the same time. However, if there are people behind you who want to check in at the same counter, they could overtake you if you do not line up immediately.

In a Nash equilibrium, as already mentioned, flow should always be sent over currently shortest s-t-paths. Formalizing this, let ℓ_v be the foresighted node arrival time function of a node $v \in V$. Recall that $\ell_v(\theta)$ equals the minimum time at which a flow particle departing at s at some point in time $\theta \in \mathbb{R}_+$ is able to arrive at v. We say that an edge $e = vw \in E$ *is contained in a shortest path at time* θ0 if and only if $\ell_w(\theta) = \ell_e(\ell_v(\theta))$. Of course, if an edge $e = vw \in E$ does not lie on a shortest s-t-path at a certain time $\theta \in \mathbb{R}_+$, no flow should be assigned to that edge at time $\ell_v(\theta)$ in a Nash flow.

Definition 4.11. We say that *flow is only sent along currently shortest paths* if and only if, for each edge $e = vw \in E$, the following condition holds for almost all times $\theta \in \mathbb{R}_+$

$$\ell_w(\theta) \;<\; \ell_e\big(\ell_v(\theta)\big) \qquad\qquad \Rightarrow \qquad\qquad f_e^+\big(\ell_v(\theta)\big) \;=\; 0\;.$$

We emphasize the following aspect of a routing satisfying Definition 4.11: In a *static* shortest path all subpaths are also shortest paths. But this is no longer true if we consider the dynamic case as illustrated in Example 4.10. Here, as long as the last person in line remains the same you always leave the check-in counter at the same time. So in principle, you could decide to make a detour – maybe for buying a small present for your family – and you can still leave the check-in counter as early as possible. However, in this case you will use at least one edge which does not lie on a currently shortest path. Since Definition 4.11 forbids entering that edge, you have to line up at the check-in counter as early as possible.

As we see below, the condition in Definition 4.11 is equivalent to the condition that every particle tries to overtake as much other flow as possible while not being overtaken. The latter condition is, in fact, a *universal FIFO condition.* That is, it is equivalent to the statement that no flow particle can possibly overtake any other flow particle. As already proven in Theorem 4.9, this aspect holds for path-based flow models.

In order to model the universal FIFO condition more formally, we consider an additional flow particle originating at s at time $\theta \in \mathbb{R}_+$. Of course, in order to ensure that no flow particle has the possibility to overtake this particle, it is necessary to take a shortest s-t-path. Therefore, for each edge $e = vw \in E$, we define the amount of flow $x_e^+(\theta)$ assigned to e before this particle can reach v and the amount of flow $x_e^-(\theta)$ leaving e before this particle can reach w as follows:

$$x_e^+(\theta) \;:=\; F_e^+\big(\ell_v(\theta)\big)\;, \qquad x_e^-(\theta) \;:=\; F_e^-\big(\ell_w(\theta)\big) \qquad \forall \theta \in \mathbb{R}_+\;. \tag{4.2}$$

Thus, the amount of flow $b_s(\theta) := D(\theta)$ that has originated at s before our flow particle occurs at s and the amount of flow $-b_t(\theta) := \mathrm{val}(\mathcal{F})(\ell_t(\theta))$ arriving at t before our flow particle can reach t satisfy

$$\begin{aligned} b_s(\theta) \;&=\; \sum_{e\in\delta^+(s)} x_e^+(\theta) - \sum_{e\in\delta^-(s)} x_e^-(\theta) \\ \text{and} \quad b_t(\theta) \;&=\; \sum_{e\in\delta^+(t)} x_e^+(\theta) - \sum_{e\in\delta^-(t)} x_e^-(\theta)\;. \end{aligned} \tag{4.3}$$

By definition, $b_s(\theta)$ is always nonnegative and $b_t(\theta)$ is always nonpositive. Intuitively, if $b_s(\theta) > -b_t(\theta)$, the considered flow particle overtakes other flow particles. And if $b_s(\theta) < -b_t(\theta)$, the flow particle is overtaken by other flow particles. This motivates the following definition which is simultaneously applicable if the underlying flow over time is path-based.

Definition 4.12. We say that *no flow overtakes any other flow* if and only if, for each point in time $\theta \geq 0$, it holds that $b_s(\theta) = -b_t(\theta)$.

Intuitively, this definition must be satisfied by a Nash flow over time: Assume that a flow particle p_2 originating at the source at time θ_2 overtakes an earlier flow particle p_1 originating at the source at time $\theta_1 < \theta_2$. That is, p_2 arrives at the sink before p_1. Because the foresighted arrival time functions ℓ_e are nondecreasing for all edges e (see Proposition 3.41(ii)), flow particle p_1 can avoid being overtaken by p_2 and improve its cost (arrival time at the sink) by choosing the same path as p_2.

Now we are able to prove the equivalence of the universal FIFO condition and the condition that flow only uses currently shortest paths. Further, both conditions characterize edge-based Nash flows over time. In addition, a further equivalent characterization is given. With respect to Example 4.10, Theorem 4.13 also tells you why you should line up immediately. It says that if you do not reach the end of the waiting queue as early as possible, other persons may overtake you.

Theorem 4.13. *For a given path decomposable edge-based flow over time $\mathcal{F}$, the following statements are equivalent:*

(i) Flow is only sent along currently shortest paths.

(ii) For each edge $e \in E$ and at all times $\theta \in \mathbb{R}_+$, it holds that $x_e^+(\theta) = x_e^-(\theta)$.

(iii) No flow overtakes any other flow.

(iv) The flow $\mathcal{F}$ is a Nash flow over time.

In the proof of Theorem 4.13, the following lemma plays an important role. It gives a more global characterization of when flow is being sent only along currently shortest paths (Definition 4.11 gives only a pointwise characterization). Formally, it can also be interpreted as an edge-based analogon of Theorem 4.9.

Lemma 4.14. *For a given path decomposable flow over time $\mathcal{F}$, the following statements are equivalent:*

(i) Flow is only being sent along currently shortest paths.

(ii) For each edge $e = vw \in E$ and for all $\theta \in \mathbb{R}_+$, it holds that

$$F_e^-\big(\ell_e(\ell_v(\theta))\big) \;=\; F_e^-\big(\ell_w(\theta)\big) \;. \tag{4.4}$$

Proof. Equation (4.4) is obviously fulfilled if edge e is contained in a shortest path at time θ. Therefore, we only consider edges e and times θ such that e does not lie on a shortest path at time θ.

(i)$\Rightarrow$(ii): Let $\theta \in \mathbb{R}_+$ and $e = vw \in E$ be an edge which is not contained in a shortest path at time θ, i. e., $\ell_w(\theta) < \ell_e(\ell_v(\theta))$. Further, let $(F_P^+)_{P\in\mathcal{P}}$ be a path decomposition of $\mathcal{F}$ implying $F_e^- = \sum_{P,i|e_i^P=e} F_{P,i}^-$ in particular. Hence, we have to show

$$F_{P,i}^-\big(\ell_e(\ell_v(\theta))\big) \;=\; F_{P,i}^-\big(\ell_w(\theta)\big) \tag{4.5}$$

for each pair (i,P) with $e_i^P = e$. Consider a pair (P,i) with $e_i^P = e$ and define θ_1, θ_2 such that flow entering P over the interval (θ_1, θ_2) arrive at w after time $\ell_w(\theta)$ and before time $\ell_e(\ell_v(\theta))$. Further, (θ_1, θ_2) should be inclusionwise maximal with this property, i.e,

$$\begin{aligned} &\theta_1 \;:=\; \max\Big\{0,\ \sup\big\{\theta' \geq 0 \mid \ell_{P,i+1}(\theta') \leq \ell_w(\theta)\big\}\Big\} \\ \text{and}\qquad &\theta_2 \;:=\; \inf\big\{\theta' \geq 0 \mid \ell_e\big(\ell_v(\theta)\big) \leq \ell_{P,i+1}(\theta')\big\}\ . \end{aligned}$$

Next consider a time $\theta' \in (\theta_1, \theta_2)$. We observe that a flow particle traversing P with departure time θ' visits an edge e_j^P with $j \leq i$ which does not lie on a shortest path at time θ'. To see this, assume the opposite which would result in $\ell_w(\theta') = \ell_{P,i+1}(\theta') = \ell_w(\theta')$. As $\ell_w(\theta) < \ell_{P,i+1}(\theta')$ holds, this shows $\theta' > \theta$. On the other hand, we know $\ell_{P,i}(\theta') < \ell_v(\theta)$ as $\ell_{P,i+1}(\theta') < \ell_e(\ell_v(\theta))$ implying $\theta' < \theta$. Thus, we have a contradiction and the flow particle visits an edge e_j^P with $j \leq i$ which does not lie on a shortest path at time θ'.

Since flow is only being sent along currently shortest path, essentially no flow enters P over the time interval (θ_1, θ_2). This shows

$$\begin{aligned} F_P^+(\theta_2) = F_P^+(\theta_1) = F_{P,i}^-\big(\ell_{P,i+1}(\theta_1)\big) &= F_{P,i}^-\big(\ell_w(\theta)\big) \\ &\leq F_{P,i}^-\big(\ell_e\big(\ell_v(\theta)\big)\big) = F_{P,i}^-\big(\ell_{P,i+1}(\theta_2)\big) = F_P^+(\theta_2)\ . \end{aligned}$$

Hence, (4.5) and, thus, (4.4) holds and this part of the proof is done.

(ii)$\Rightarrow$(i): Let $\theta \in \mathbb{R}_+$ be some point in time and $e = vw \in E$ be an edge that is not contained in a shortest path at time θ, i.e., $\ell_w(\theta) < \ell_e(\ell_v(\theta))$. Since ℓ_w is right continuous by Proposition 3.41(iii) and 3.51, there exists an $\epsilon > 0$ such that $\ell_w(\theta + \epsilon) < \ell_e(\ell_v(\theta))$. So the nonnegativity of the flow rate functions yields

$$\begin{aligned} 0 \;\leq\; \int_{\ell_v(\theta)}^{\ell_v(\theta+\epsilon)} f_e^+(\vartheta)\, d\vartheta \;&=\; \int_{\ell_e(\ell_v(\theta))}^{\ell_e(\ell_v(\theta+\epsilon))} f_e^-(\vartheta)\, d\vartheta \\ &\leq\; \int_{\ell_w(\theta+\epsilon)}^{\ell_e(\ell_v(\theta+\epsilon))} f_e^-(\vartheta)\, d\vartheta \;=\; 0\ . \end{aligned}$$

This shows that the statement (i) holds for almost all $\theta' \in (\theta, \theta + \epsilon)$. Thus, recalling Corollary 2.8, this proof is finished. ☐

Proof of Theorem 4.13. The main observation we need is the following equation which we get from Proposition 3.41(i) and the definitions of x_e^+, x_e^- in (4.2):

$$\begin{aligned} x_e^+(\theta) - x_e^-(\theta) \;&=\; F_e^+\big(\ell_v(\theta)\big) - F_e^-\big(\ell_w(\theta)\big) \\ &=\; F_e^-\big(\ell_e(\ell_v(\theta))\big) - F_e^-\big(\ell_w(\theta)\big)\ . \end{aligned} \tag{4.6}$$

Because of Lemma 4.14, this equation implies the equivalence of (i) and (ii).

For proving the equivalence of (ii) and (iii), we construct a static b-flow instance. We replace each edge $e = vw \in E$ by a new node v_e and two edges vv_e

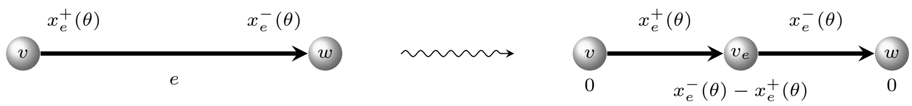


Figure 4.2: Construction of the b-flow instance used in the proof of Theorem 4.13.

and $v_e w$ (see Figure 4.2). The supply-demand vector of the corresponding b-flow instance is defined as follows. For each node $v \in V \setminus \{s,t\}$, we set $b_v(\theta) := 0$, and for each new node v_e with $e \in E$, we define $b_{v_e}(\theta) := x_e^-(\theta) - x_e^+(\theta)$. An intuitive explanation of $b_{v_e}(\theta)$ is as follows. Recall that, at the time when a flow particle originating at the source at time θ gets to node v along a shortest path, the amount of flow having previously entered edge e is $x_e^+(\theta)$. Similarly, when the same flow particle travels along a shortest path to w, the amount of flow that has previously arrived at w via edge e is $x_e^-(\theta)$. Hence, a flow particle which arrives at w via e can increase the amount of overtaken flow by $-b_{v_e}(\theta)$ if it goes directly along a currently shortest path to w. However, the condition in (ii) states that this flow particle cannot improve its situation, i. e., $b_{v_e}(\theta) = 0$.

Note that we have defined $b_s(\theta)$ and $b_t(\theta)$ in (4.3). It follows from (4.6) and the nonnegativity of the outflow rate functions that only node s has a supply, i. e., a positive b-value.

Consider the following *static* flow. For each edge $e = vw \in E$, set the flow value on edge vv_e to $x_e^+(\theta)$ and the flow value on edge $v_e w$ to $x_e^-(\theta)$. We claim that this static flow is a feasible b-flow. To prove this we need to check the flow conservation constraints. By construction and (4.3), flow conservation is fulfilled at nodes s, t, and also at the new nodes v_e, $e \in E$. It remains to verify flow conservation at nodes $v \in V \setminus \{s,t\}$. The following equation follows from linearity of the integral operator and condition (3.7).

$$\sum_{e\in\delta^-(v)} x_e^-(\theta) \;=\; \sum_{e\in\delta^-(v)} F_e^-(\ell_v(\theta)) \;=\; \sum_{e\in\delta^+(v)} F_e^+(\ell_v(\theta)) \;=\; \sum_{e\in\delta^+(v)} x_e^+(\theta) \;.$$

Thus, we have a feasible b-flow on the constructed instance. In particular, the sum over all supplies and demands is equal to zero. That is,

$$\sum_{v\in V} b_v(\theta) + \sum_{e\in E} b_{v_e}(\theta) \;=\; 0 \;.$$

Note that this shows $b_s(\theta) \geq -b_t(\theta)$ for all $\theta \geq 0$ which we use later in this proof. Because the source s is the only node with a positive b-value, the supply of s is equal to the demand of t if and only if the b-values of all other nodes are 0. This proves the equivalence of (ii) and (iii).

In order to finalize the proof also for edge-based flows over time, we show the equivalence of (iii) and (iv). For this let $(F_P^+)_{P\in\mathcal{P}}$ be an arbitrary path-decomposition of $\mathcal{F}$. Since $\min_{P\in\mathcal{P}} \ell_P(\theta) = \ell_t(\theta)$ holds for all $\theta \in \mathbb{R}_+$ by definition, we know

$$b_s(\theta) \;=\; \sum_{P\in\mathcal{P}} F_P^+(\theta) \;=\; \sum_{P\in\mathcal{P}} F_P^-\big(\ell_P(\theta)\big) \;\geq\; \sum_{P\in\mathcal{P}} F_P^-\big(\ell_t(\theta)\big) \;=\; -b_t(\theta)$$

by the definition of a path decomposition (see Definition 3.46). Thus, (iii) holds if and only if

$$F_P^-\big(\ell_P(\theta)\big) \;=\; F_P^-\big(\ell_t(\theta)\big) \qquad\qquad \forall \theta \in \mathbb{R}_+, P \in \mathcal{P} \;.$$

Hence, it remains to show that these equations hold if and only if Definition 4.8 is valid. But this is exactly the statement of Theorem 4.9. □

Note that x^+ and x^- coincide, whenever one of the four statements in Theorem 4.13 holds. In this case $x_e^+(\theta) = x_e^-(\theta)$ is a static s-t-flow of value $b_s(\theta)$ for all $\theta \in \mathbb{R}_+$. This static flow plays an important role when analyzing Nash flows for the deterministic queuing model in Chapter 7 motivating the following definition.

Definition 4.15 (Underlying Static Flow.)**.** Let $\mathcal{F}$ be an edge-based Nash flow over time. For all $\theta \in \mathbb{R}_+$ the underlying static flow $x(\theta) := (x_e(\theta))_{e\in E}$ at time θ is defined by $x_e(\theta) := x_e^+(\theta) = x^-(\theta)$ for all $e \in E$.

The last part of this section is devoted to the existence of Nash flows over time. We show that consistent flow models, either edge- or path-based, permit a Nash flow over time on any network. The idea works as follows. Iteratively, we compute a currently shortest s-t-path P and send all newly originating flow particles along P until their cumulative amount equals some given small $\epsilon > 0$. It turns out that, in this manner, a particular flow particle can overtake only a small amount of flow. In fact, the amount of overtaken flow is bounded linearly in ϵ. This shows, if ϵ tends to 0, no flow particle overtakes any other flow particle. Thus, recalling Theorem 4.13, if the corresponding sequence of flows over time converges, the limit is a Nash equilibrium. The following algorithm describes the routing strategy explained above for a given $\epsilon > 0$. Note that we need to solve a dynamic network loading problem in Step (3) of the algorithm, if the underlying flow model is edge-based.

General Iterative Algorithm

Input:	A nonatomic routing game over time $(G, \mathcal{T}, s, t, D)$ based on a consistent flow model and $\epsilon > 0$.
Output:	A path-based flow over time $\mathcal{F}$.

(1) Compute times $(\theta_i)_{i\in\mathbb{N}}$ such that $\theta_0 = 0$ and $D(\theta_{i+1}) - D(\theta_i) = \epsilon$.

(2) Set $\mathcal{F} := 0$, and $i := 0$.

(3) Compute a shortest s-t-path P at time θ_i according to $\mathcal{F}$ and $\mathcal{T}$.

(4) Route the flow originating within the time interval $[\theta_i, \theta_{i+1})$ along P, i.e., set $f_P^+ := f_P^+ + d|_{[\theta_i,\theta_{i+1})}$.

(5) Set $i := i + 1$ and go to (3).

This algorithm never terminates, but it serves as a definition for the constructed flow over time $\mathcal{F}$. Nevertheless, since we work with consistent flow models, it can be interrupted to obtain restricted flows of $\mathcal{F}$. But note that the current $\mathcal{F}$, updated in Step (4), is not equal to the restricted flow until time θ_i if the flow model is edge-based. In this case we have to run some more iterations. To be more precise, let $\mathcal{F}_i$ be the current $\mathcal{F}$ before the flow originating over the time interval $[\theta_i, \theta_{i+1})$ is assigned to some path. Further, consider a point in time $\theta \in \mathbb{R}_+$ and let $k \in \mathbb{N}$ be such that $\theta_k \leq \theta < \theta_{k+1}$. Let P_k and P_{k+1} be the s-t-paths found by the algorithm in iterations k and $k + 1$, respectively, and assume that they share an edge e. Especially, if P_1 and P_2 are disjoint on their way from s to e, flow on both paths is able to arrive simultaneously at the head of e. Hence, the flow behavior on P_k depends on the flow of P_{k+1} and, in particular, this implies $\mathcal{F}_{k+1}|_{\leq\theta} \neq \mathcal{F}_{k+2}|_{\leq\theta}$. Note that in this scenario, the transit time function of P_{k+1} must increases slower over $[\theta_k, \theta_{k+1})$ than the one of P_k with respect to $\mathcal{F}_k$.

Although we do not obtain the restricted flow until time $\theta \in [\theta_k, \theta_{k+1})$ after iteration k, we are nevertheless in a situation to extract important information. Since a consistent flow model is past-oriented, $\ell_t(\mathcal{F}_i)(\theta')$ remains constant for all $i \geq k$ and all $\theta' \leq \theta_k$. Therefore, after iteration k, we know the node arrival time function $\ell_t(\mathcal{F})$ until time θ_k. And, in fact, we know the node arrival time functions $\ell_v(\mathcal{F})$ until time θ_k for all nodes $v \in V$ (see Lemma 3.52). Further, since FiFo is satisfied and, hence, the node arrival time functions are nondecreasing, no flow originating after time θ_k is able to arrive at t before time $\ell_t(\mathcal{F})(\theta_k)$. Therefore, after iteration k, we have a complete picture of the behavior of flow particles arriving at t before time $\ell_t(\mathcal{F}_k)(\theta_k)$. Note the difference to Example 3.45 which shows that a path-based flow model of an edge-based past-oriented flow model is not past-oriented in general.

This has two interesting consequences. In order to compute $(\mathcal{F}^+|_{\leq\theta}, \mathcal{F}^-|_{\leq\theta})$ for some $\theta \in [\theta_k, \theta_{k+1})$, we have to wait until $\ell_{P_{k'}}(\mathcal{F}_i)(\theta) \leq \ell_t(\mathcal{F}_i)(\theta_i)$ holds for all $k' \leq k$. But even more important, after iteration $k-1$, we know the restricted path flow $(F^+_{P_k}|_{\leq\theta_k}, F^-_{P_k}|_{\leq\theta_k})$ as $\ell_{P_k}(\mathcal{F}_k)(\theta_k) = \ell_t(\mathcal{F}_k)(\theta_k)$.

As we have a first intuition about the General Iterative algorithm, we direct the focus on the existence of continuous Nash equilibria over time. For this, we show that a flow over time constructed by the algorithm violates the nonovertaking condition only slightly. In Step (1) of the algorithm neighboring particles are collected into a small portion of flow which is sent to the sink along one path. These small portion of flow can be interpreted as *flow units* which distribute its flow over time. In the following we explain intuitively that a particle of one flow unit can only overtake particles from at most finitely many other flow units.

The first thing we need is the observation that whenever two flow units use the same s-t-path P, the unit which departs later at s cannot overtake flow of the other, i.e., they traverse P one after another. For path-based flow models this is directly implied by the FiFo-principle ensuring that ℓ_P strictly increases if flow is send along P. For edge-based flow models this follows from induction over the number of edges of the considered path using again the past-orientation and the FiFo-principle.

The second thing we need is the fact that a particle of one flow unit can overtake at most one complete flow unit on every other path. Assume that there is one flow unit using s-t-path P_1 which overtakes more than one flow unit of an s-t-path P_2. Therefore, the last flow unit using P_2 reaches t strictly later than the flow unit on P_1. But because the node labels do not decrease, the length of P_1 is shorter than the length of P_2 at the time where the last flow unit on P_2 starts traversing the network. This is a contradiction because the General Iterative algorithm sends the flow units always along a currently shortest s-t-path. Since every used path is simple in case of an edge-based flow model, one flow unit can overtake flow of at most finitely many other flow units. Note that past-orientation is used implicitly during this argumentation.

In the following we formalize the discussion above. Firstly, we show that one flow particle can overtake at most flow of value $\epsilon > 0$ on every path. But before note that the outflow functions of $\mathcal{F}$ exists and are well-defined. For path-based flow models this follows directly from the past-orientation. For edge-based flow models this is directly implied by the properties of the network loading problem (see Lemma 3.44). Note that the inflow functions converge obviously in $(L_\infty^{\mathrm{loc}})^{\mathcal{P}}$ if the number of iterations of the General Iterative algorithm goes to infinity. In fact, they remain constant on any compact subset of $\mathbb{R}_+$ after a while.

Lemma 4.16. *Let $\mathcal{F}$ be the output of the* General Iterative *algorithm for some given consistent routing game over time and some given $\epsilon > 0$. Consider an s-t-path P and let F^- be the outflow function and ℓ be the foresighted arrival time function of P. Then for all points in time $\theta \in \mathbb{R}_+$ we have*

$$F^-\big(\ell(\theta)\big) - \epsilon \;\le\; F^-\big(\ell_t(\theta)\big) \;\le\; F^-\big(\ell(\theta)\big)\ .$$

Proof. Fix some point in time $\theta \in \mathbb{R}_+$. Since $\ell_t \le \ell$ holds by definition and F^- is nondecreasing, we obtain directly $F^-(\ell_t(\theta)) \le F^-(\ell(\theta))$. To see that also the first inequality holds, consider an $i \in \mathbb{N}$ such that $\ell(\theta_i) \le \ell_t(\theta) < \ell(\theta_{i+1})$ holds implying $\theta_i \le \theta$ in particular. On the one hand, this shows

$$F^-\big(\ell(\theta_i)\big) \;\le\; F^-\big(\ell_t(\theta)\big) \;\le\; F^-\big(\ell(\theta_{i+1})\big) \tag{4.7}$$

because F^- is nondecreasing. On the other hand, since ℓ_t just as ℓ are nondecreasing, we know $\ell_t(\theta_j) \le \ell_t(\theta) < \ell(\theta_{i+1}) \le \ell(\theta_j)$ for all $j > i$ with $\theta_j \le \theta$. This shows that the General Iterative algorithm assigns no flow to P over the time interval $[\theta_{i+1}, \theta]$ if $\theta_{i+1} \le \theta$. Hence, from $F^+ = F^- \circ \ell$ (see Lemma 3.41(i)) we obtain

$$F^-\big(\ell(\theta)\big) \;=\; F^+(\theta) \;=\; F^+(\theta_{i+1}) \;=\; F^-\big(\ell(\theta_{i+1})\big)$$

if $\theta_{i+1} \le \theta$. Further, by definition we know $\ell(\theta_i) \le \ell(\theta)$. Since F^- is nondecreasing, this leads to

$$F^-\big(\ell(\theta_i)\big) \;\le\; F^-\big(\ell(\theta)\big) \;\le\; F^-\big(\ell(\theta_{i+1})\big)\ . \tag{4.8}$$

Note that in case of $\theta_{i+1} > \theta$ the inequality $F^-(\ell(\theta)) \le F^-(\ell(\theta_{i+1}))$ obviously holds as F^- is nondecreasing. Since the General Iterative algorithm assigns flow of value ϵ over the time interval $[\theta_i, \theta_{i+1}]$ to one s-t-path, we know $F^-(\ell(\theta_{i+1})) - F^-(\ell(\theta_i)) \le \epsilon$. Thus, $F^-(\ell(\theta)) - \epsilon \le F^-(\ell_t(\theta))$ follows directly from (4.7) and (4.8). □

Lemma 4.16 shows that if ϵ goes to 0, the limit flow would satisfy the assumptions of Theorem 4.9. Thus, if the limit flow exists, it is a Nash flow over time. In addition, Lemma 4.16 implies a nice corollary for edge-based flows over time, which shows that the output of the General Iterative algorithm violates the non-overtaking condition in terms of Definition 4.12 only slightly.

Corollary 4.17. *Let $\mathcal{F}$ be the output of the* General Iterative *algorithm for some given consistent routing game over time and some given $\epsilon > 0$. If the underlying flow model is edge-based, we have for all $\theta \in \mathbb{R}_+$*

$$|b_s(\theta) + b_t(\theta)| \le M\epsilon$$

where M equals the number of all simple s-t-path.

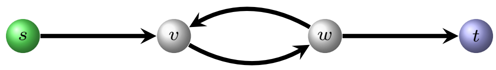


Figure 4.3: A network of Example 4.18 where each edge has a constant capacity of 1 and a constant transit time of 0.

Proof. Because $b_s(\theta) + b_t(\theta) \geq 0$ holds, it is enough to show $b_s(\theta) + b_t(\theta) \leq M\epsilon$. Since a currently shortest path which is simple exists (see Lemma 3.50), we are able to assume that the General Iterative algorithm routes flow only along simple s-t-paths. Let $\mathcal{P}' \subseteq \mathcal{P}$ be the set of simple s-t-paths and $M := |\mathcal{P}'|$. Then we obtain directly from Lemma 4.16

$$\begin{aligned} b_s(\theta) \;=\; \sum_{P \in \mathcal{P}'} F_P^+(\theta) \;&=\; \sum_{P \in \mathcal{P}'} F_P^-\big(\ell_P(\theta)\big) \\ &\leq\; M\epsilon + \sum_{P \in \mathcal{P}'} F_P^-\big(\ell_t(\theta)\big) \;=\; M\epsilon - b_t(\theta) \end{aligned}$$

for all points in time $\theta \in \mathbb{R}_+$ which completes the proof. □

By construction the output flow of the General Iterative algorithm is a feasible strategy profile of the routing game (G, τ, s, t, D). Lemma 4.16 and Corollary 4.17 show that when ϵ decreases then the non-overtaking condition becomes less violated. Thus, if we let ϵ tend to 0 we could expect to obtain a Nash flow over time in the limit. But before we work out the details of this approach we reconsider Example 3.35 which can lead to an unexpected behavior of the General Iterative algorithm.

Example 4.18. We redraw the underlying network in Figure 4.3. As in Example 3.35, we consider the classical flow over time model (see Section 2.5) with the additional assumption that flow must not wait at the intermediate nodes v and w. So in fact, we have a path-based flow model where flow is only allowed to enter a path if all edges of this path have free capacities. Hence, if flow wants to traverse a certain path P, it has to wait until all edges of P has some capacity left before it enters P. In fact, this equals the direct flow model, which we present in Chapter 5. In this scenario, each edge has a constant capacity of 1 and a constant transit time of 0. Further, we assume that the flow controlled by the players arrive at s at a constant rate of 1, i.e., $D := \mathrm{id}$.

For each $i \in \mathbb{N}$ let P_i be the s-t-path which contains $i-1$ times the cycle C induced by v and w. So we have $\mathcal{P} := \{\mathcal{P}_i \mid i \in \mathbb{N}\}$. It turns out that in each iteration of the General Iterative algorithm, the edge vw is the bottleneck whatever which path is chosen. That means, in every iteration every path $P \in \mathcal{P}$ is a shortest one and the currently minimal arrival time equals the earliest point in time at which edge vw becomes free.

For each $k \in \mathbb{N}$ we apply the General Iterative algorithm with $\epsilon := \frac{1}{k}$ to this instance and assume that in iteration i path P_i is chosen in Step (3). So if k goes to ∞ implying that ϵ tends to 0, we obtain the zero flow in the limit. Of course, the zero flow is a Nash equilibrium, but not for the players given by $D = \mathrm{id}$ as the zero flow is not a feasible strategy profile. But note, that the output of the General Iterative algorithm is already a Nash flow for each $k \in \mathbb{N}$.

Example 4.18 shows that one critical point of our approach is the feasibility of the strategy profile in the limit. Ensuring this feasibility is the main reason for introducing the locally finiteness of $\mathcal{P}$ which is not satisfied by the routing in Example 4.18. However, as explained in Example 3.35, the flow model in Example 4.18 can be easily modified in order to ensure that $\mathcal{P}$ is locally finite. Having in mind that $\mathcal{P}$ is locally finite for consistent flow models, the following theorem shows that a nonatomic routing game over time always admits a Nash equilibrium over time.

Theorem 4.19. *Let $(G, \mathcal{T}, s, t, D)$ be a nonatomic routing game over time which is based on a consistent flow over time model. Then a Nash flow over time exists.*

Proof. For all $k \in \mathbb{N}$ let $\mathcal{F}_k$ be the output of the GENERAL ITERATIVE algorithm with $\epsilon := \frac{1}{k}$. We prove that for every time horizon $\theta \in \mathbb{R}_+$ the sequence $(\mathcal{F}_k|_{\leq\theta})_{k\in\mathbb{N}}$ of the restricted flows has an accumulation point. Subsequently, based on Lemma 4.16, we show that every accumulation point is a Nash flow over time. Depending on whether the underlying flow model is edge- or path-based let $\mathcal{P}$ be the set of simple s-t-paths in G or the set of s-t-paths which are able to carry flow until time θ, respectively. Note that $\mathcal{P}$ is finite in both cases. Further, if flow is send along some path P in $\mathcal{F}_k$ then P is contained in $\mathcal{P}$. To see this note that the GENERIC ITERATIVE algorithm sends flow only along s-t-paths which are currently shortest at some point in time. Hence, for path-based flow models this follows as consistency implies that the set of s-t-paths must be locally finite. For edge-based flow models this follows from Lemma 3.50 verifying that there always exists a simple shortest s-t-path.

Let $\mathcal{F}_k|_{\leq\theta} := (F_P^k)_{P\in\mathcal{P}}$ be the path-based representation of the restricted flow $\mathcal{F}_k|_{\leq\theta}$. Since $\mathcal{F}_k$ is a feasible strategy profile, we know, for all functions F_P^k and all $\theta_1, \theta_2 \in \mathbb{R}_+$ with $\theta_1 \leq \theta_2$, that $F_P^k(\theta_2) - F_P^k(\theta_1) \leq D(\theta_2) - D(\theta_1)$ is valid. This shows that $D - F_P^k$ is nondecreasing ensuring that Lemma 2.36 is applicable.

In the following we show that the sequence of the restricted path-based flows over time $\mathcal{F}_k|_{\leq\theta} := (F_P^k)_{p\in\mathcal{P}}$ has an accumulation point which is a feasible strategy profile of $(G, \mathcal{T}, s, t, D)$. For this let $P_1, \dots P_{|\mathcal{P}|}$ be an arbitrary order of the paths in $\mathcal{P}$. Then by Lemma 2.36 there exists a countable subset $J_1 \subseteq \mathbb{N}$ such that $(F_{P_1}^k)_{i\in J_1}$ converges uniformly to some $F_{P_1}^*$. Next we apply Lemma 2.36 to the new subsequence $(\mathcal{F}_k)_{k\in J_1}$ and obtain a countable subset $J_2 \subseteq J_1$ such that $(F_{P_2}^k)_{k\in J_2}$ converges uniformly to some $F_{P_2}^*$. Note that $(F_{P_1}^k)_{k\in J_2}$ still converges uniform to $F_{P_1}^*$. Proceeding in this manner we get after $|\mathcal{P}| < \infty$ iterations a subsequence $(\mathcal{F}_k)_{k\in J_{|\mathcal{P}|}}$ such that $(F_{P_j}^k)_{k\in J_{|\mathcal{P}|}}$ converges uniformly to some $F_{P_j}^*$ for all $j \in \{1, \dots, |\mathcal{P}|\}$. Further, Lemma 2.36 ensures that each F_P^* is nondecreasing and absolutely continuous which makes $\mathcal{F}^* := (F_P^*)_{P\in\mathcal{P}}$ to a continuous path-based flow over time. Now let $J_{|\mathcal{P}|} = \{k_i | i \in \mathbb{N}\}$ be such that $k_i < k_{i+1}$ for all $i \in \mathbb{N}$. Since $\mathcal{P}$ is finite, we obtain for each $\theta' \in [0, \theta]$

$$D(\theta') = \lim_{i\to\infty} \sum_{P\in\mathcal{P}} F_P^{k_i}(\theta') = \sum_{P\in\mathcal{P}} \lim_{i\to\infty} F_P^{k_i}(\theta') = \sum_{P\in\mathcal{P}} F_P^*(\theta')$$

implying that $(F_P^*)_{P\in\mathcal{P}}$ is a feasible strategy profile.

In order to prove that $(F_P^*)_{P\in\mathcal{P}}$ is a Nash equilibrium we apply Lemma 4.16

to $\mathcal{F}^{k_i}$ for all $i \geq \mathbb{N}$ implying

$$F_P^{k_i-}\big(\ell_P(\theta')\big) - \frac{1}{k_i} \;\leq\; F_P^{k_i-}\big(\ell_t(\theta')\big) \;\leq\; F_P^{k_i-}\big(\ell_P(\theta')\big) \qquad \forall \theta' \in \mathbb{R}_+ \ . \tag{4.9}$$

In particular, using Lemma 3.53 this shows $\lim_{i\to\infty} F_P^{k_i-}(\ell_t(\theta')) = F_P^{*-}(\ell_t(\theta'))$. Further, because of Proposition 3.41(i) we know

$$\lim_{i\to\infty} F_P^{k_i-}(\ell_P(\theta')) \;=\; \lim_{i\to\infty} F_P^{k_i}(\theta') \;=\; F_P^{*}(\theta') \;=\; F_P^{*-}(\ell_P(\theta')) \quad \forall \theta' \in \mathbb{R}_+ \ .$$

Hence, if i goes to infinity, we obtain from (4.9) that

$$F_P^{*-}\big(\ell_P(\theta')\big) \;=\; F_P^{*-}\big(\ell_t(\theta')\big) \qquad\qquad \forall \theta' \in \mathbb{R}_+, P \in \mathcal{P} \ .$$

Thus, by Theorem 4.9 the flow over time $\mathcal{F}^*$ is a Nash flow over time. □

We conclude this with the following remark addressing the question how well-defined a *nonatomic* Nash equilibrium in terms of Definition 4.8 is.

Remark 4.20. The General Iterative algorithm is also interesting from another point of view. The initialization in Step (1) discretizes the supply into flow units (each of size ϵ). Hence, interpreting the θ_i as origination times of these flow units we obtain an atomic routing game over time out of the nonatomic description. Recalling the proof of Theorem 4.5, the General Iterative algorithm computes a Nash equilibrium for this atomic routing game over time. In this context, the proof of Theorem 4.19 shows that there exists a sequence of atomic Nash equilibria which converges to a nonatomic Nash equilibria in terms of Definition 4.8. Thus, as nonatomic routing games usually arise out of limit approaches of atomic routing games, this shows that Definition 4.8 is well-motivated.

4.3 Price of Anarchy and Price of Stability

Theorem 4.19 shows that a Nash equilibrium always exists. In this section we provide the basis for another aspect of routing games over time. The main aspect of selfish routing is that all players try to minimize their own costs regardless of potential consequences for the efficiency of the routing. Hence, the routing of a Nash flow over time reflects somehow anarchy. In contrast, a system optimum for a routing game is a flow which optimizes some objective function, maybe for measuring social welfare. How bad or good a Nash flow acts with respects to such an objective function is reflected by the price of anarchy. Given a particular routing game, the price of anarchy measures the worst-case behavior of a Nash flow with respect to a system optimum.

In addition to the price of anarchy, there exists also another popular way of measuring the efficiency of Nash equilibria. Here, instead of total anarchy, a central authority has the possibility to guide the players into a Nash equilibrium. The idea behind is that as long as a player sees no possibility to decrease his cost, he follows the advice of the supervisor. In this sense, Nash equilibria are stable because nobody should have an obvious incentive to act against the central authority. Of course, the goal is to guide the players into the best Nash equilibrium with respect to some given objective function. The ratio between the system optimum and the best Nash equilibrium is called price of stability.

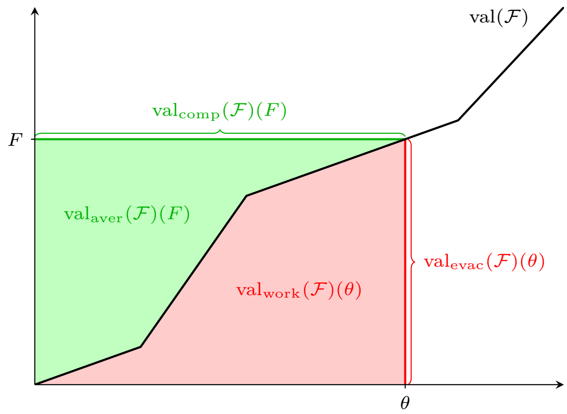


Figure 4.4: Illustration of four objective functions which are based on the arrival pattern $\mathrm{val}(\mathcal{F})$ of a flow over time $\mathcal{F}$.

In this section we present four different objective functions for routing games over time. Each objective function is based on the arrival pattern of a flow over time. But note that the flow pattern is interpreted as a value *over time*. In this sense, each objective value results out of a point in time or a flow value. Each objective function is illustrated in Figure 4.4.

Evacuation price of anarchy. Consider an evacuation scenario. Assuming that the latest point in time of escape is θ, a system optimum tries to save as many people as possible until time θ. Thus, a natural objective function for evacuation scenarios equals the flow pattern

$$\mathrm{val}_{\mathrm{evac}}(\mathcal{F})(\theta) := \mathrm{val}(\mathcal{F})(\theta) \ . \tag{4.10}$$

Definition 4.21 (Evacuation Price of Anarchy)**.** Let $\mathcal{I}$ be an instance of a routing game over time, $\mathcal{F}_{\mathrm{O}}$ be a system optimum with respect to (4.10), and $\mathcal{F}_{\mathrm{N}}$ be a Nash equilibrium. Given a point in time $\theta \in \mathbb{R}_+$, the *evacuation price of anarchy of* $\mathcal{F}_{\mathrm{N}}$ *at* θ is defined by

$$\rho_{\mathrm{evac}}(\mathcal{F}_{\mathrm{N}})(\theta) := \frac{\mathrm{val}_{\mathrm{evac}}(\mathcal{F}_{\mathrm{O}})(\theta)}{\mathrm{val}_{\mathrm{evac}}(\mathcal{F}_{\mathrm{N}})(\theta)} \ .$$

The *evacuation price of anarchy of* $\mathcal{F}_{\mathrm{N}}$ equals the worst-case ratio over all points in time, i.e.,

$$\rho_{\mathrm{evac}}(\mathcal{F}_{\mathrm{N}}) := \sup_{\theta \in \mathbb{R}_+} \rho_{\mathrm{evac}}(\mathcal{F}_{\mathrm{N}})(\theta) \ .$$

Further, the *evacuation price of anarchy of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho^{\mathrm{PoA}}_{\mathrm{evac}}(\mathcal{I}) := \sup_{\mathcal{F}_{\mathrm{N}}} \rho_{\mathrm{evac}}(\mathcal{F}_{\mathrm{N}}) \qquad \text{and} \qquad \rho^{\mathrm{PoA}}_{\mathrm{evac}}(\mathcal{F}_{\mathrm{N}}) := \sup_{\mathcal{I}} \rho^{\mathrm{PoA}}_{\mathrm{evac}}(\mathcal{I}) \ ,$$

respectively.

Definition 4.22 (Evacuation Price of Stability)**.** Let $\mathcal{I}$ be an instance of a routing game over time. The *evacuation price of stability of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho_{\text{evac}}^{\text{PoS}}(\mathcal{I}) := \inf_{\mathcal{F}_{\text{N}}} \rho_{\text{evac}}(\mathcal{F}_{\text{N}}) \qquad \text{and} \qquad \rho_{\text{evac}}^{\text{PoS}} := \sup_{\mathcal{I}} \rho_{\text{evac}}^{\text{PoS}}(\mathcal{I}) \ ,$$

respectively.

Working price of Anarchy. Assume, you are the boss of some enterprise which revenues mainly depends on manpower. Hence, you are interested in the number of working hours of your employees. Of course, you expect that everyone starts working directly after they arrive at their corresponding workplaces. Thus, you would maximize the area under a flow pattern, especially, if there exists a project which should be finished yesterday. So your objective function is

$$\text{val}_{\text{work}}(\mathcal{F})(\theta) = \int_0^\theta \text{val}(\mathcal{F})(\vartheta) \; d\vartheta \ . \tag{4.11}$$

Definition 4.23 (Working Price of Anarchy)**.** Let $\mathcal{I}$ be an instance of a routing game over time, $\mathcal{F}_{\text{O}}$ be a system optimum with respect to (4.11), and $\mathcal{F}_{\text{N}}$ be a Nash equilibrium. Given a point in time $\theta \in \mathbb{R}_+$, the *working price of anarchy of* $\mathcal{F}_{\text{N}}$ *at* θ is defined by

$$\rho_{\text{work}}(\mathcal{F}_{\text{N}})(\theta) := \frac{\text{val}_{\text{work}}(\mathcal{F}_{\text{O}})(\theta)}{\text{val}_{\text{work}}(\mathcal{F}_{\text{N}})(\theta)} \ .$$

The *working price of anarchy of* $\mathcal{F}_{\text{N}}$ equals the worst-case ratio over all points in time, i.e.,

$$\rho_{\text{work}}(\mathcal{F}_{\text{N}}) := \sup_{\theta \in \mathbb{R}_+} \rho_{\text{work}}(\mathcal{F}_{\text{N}})(\theta) \ .$$

Further, the *working price of anarchy of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho_{\text{work}}^{\text{PoA}}(\mathcal{I}) := \sup_{\mathcal{F}_{\text{N}}} \rho_{\text{work}}(\mathcal{F}_{\text{N}}) \qquad \text{and} \qquad \rho_{\text{work}}^{\text{PoA}}(\mathcal{F}_{\text{N}}) := \sup_{\mathcal{I}} \rho_{\text{work}}^{\text{PoA}}(\mathcal{I}) \ ,$$

respectively.

Definition 4.24 (Working Price of Stability)**.** Let $\mathcal{I}$ be an instance of a routing game over time. The *working price of stability of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho_{\text{work}}^{\text{PoS}}(\mathcal{I}) := \inf_{\mathcal{F}_{\text{N}}} \rho_{\text{work}}(\mathcal{F}_{\text{N}}) \qquad \text{and} \qquad \rho_{\text{work}}^{\text{PoS}} := \sup_{\mathcal{I}} \rho_{\text{work}}^{\text{PoS}}(\mathcal{I}) \ ,$$

respectively.

Completion Time Price of Anarchy. Consider a manufacturer of a route guidance system of the next generation which takes also into account social aspects. Of course, the primary feature of such a route guidance system must be that every user reaches his destination as fast as possible. Otherwise, if some user identifies another faster path, this would reduce the acceptance of the route guidance system. However, a unique selling point of the manufacturer could be the following. Given a set of users, the route guidance system serves the corresponding requests such that the arrival time of the last user is minimized. Or in flow terminology, given a flow value $F \in \mathbb{R}_+$, minimize the completion time, i.e., the time at which last particle of this flow arrives at the sink[1]. This leads to the following objective function:

$$\mathrm{val}_{\mathrm{comp}}(\mathcal{F})(F) = \min\{\theta \mid \mathrm{val}(\mathcal{F})(\theta) = F\} \ . \tag{4.12}$$

In contrast to the two objective functions presented above, the goal is to minimize $\mathrm{val}_{\mathrm{comp}}(\mathcal{F})(F)$. Therefore, we have to adjust the corresponding definitions of the price of anarchy and stability.

Definition 4.25 (Completion Time Price of Anarchy)**.** Let $\mathcal{I}$ be an instance of a routing game over time, $\mathcal{F}_{\mathrm{O}}$ be a system optimum with respect to (4.12), and $\mathcal{F}_{\mathrm{N}}$ be a Nash equilibrium. Given a point in time $F \in \mathbb{R}_+$, the *completion time price of anarchy of* $\mathcal{F}_{\mathrm{N}}$ *at* F is defined by

$$\rho_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}})(F) := \frac{\mathrm{val}_{\mathrm{comp}}(\mathcal{F}_{\mathrm{O}})(F)}{\mathrm{val}_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}})(F)} \ .$$

The *completion time price of anarchy of* $\mathcal{F}_{\mathrm{N}}$ equals the worst-case ratio over all points in time, i.e.,

$$\rho_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}}) := \inf_{F \in \mathbb{R}_+} \rho_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}})(F) \ .$$

Further, the *completion time price of anarchy of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho^{\mathrm{PoA}}_{\mathrm{comp}}(\mathcal{I}) := \inf_{\mathcal{F}_{\mathrm{N}}} \rho_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}}) \qquad \text{and} \qquad \rho^{\mathrm{PoA}}_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}}) := \inf_{\mathcal{I}} \rho^{\mathrm{PoA}}_{\mathrm{comp}}(\mathcal{I}) \ ,$$

respectively.

Definition 4.26 (Completion Time Price of Stability)**.** Let $\mathcal{I}$ be an instance of a routing game over time. The *completion time price of stability of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho^{\mathrm{PoS}}_{\mathrm{comp}}(\mathcal{I}) := \sup_{\mathcal{F}_{\mathrm{N}}} \rho_{\mathrm{comp}}(\mathcal{F}_{\mathrm{N}}) \qquad \text{and} \qquad \rho^{\mathrm{PoS}}_{\mathrm{comp}} := \inf_{\mathcal{I}} \rho^{\mathrm{PoS}}_{\mathrm{comp}}(\mathcal{I}) \ ,$$

respectively.

Average Arrival Time Price of Anarchy. Consider again the manufacturer of a route guidance system. Another interesting feature could be to minimize the average arrival time of the users. So if all users departure around time zero, this minimizes the average travel time. Note that minimizing the average travel time is equivalent to minimize the total travel time of all users, which is of particular importance when considering the network load or economic aspects. So in this scenario we have the following objective function:

$$\mathrm{val}_{\mathrm{aver}}(\mathcal{F})(F) = \int_0^F \mathrm{val}_{\mathrm{comp}}(\mathcal{F})(F')\, dF' \ . \tag{4.13}$$

[1]Usual the symbol F determines a function interpreted as a flow distribution over time. But here, we use the symbol F for a real number representing a flow *value*.

Definition 4.27 (Average Arrival Time Price of Anarchy)**.** Let $\mathcal{I}$ be an instance of a routing game over time, $\mathcal{F}_{\mathrm{O}}$ be a system optimum with respect to (4.13), and $\mathcal{F}_{\mathrm{N}}$ be a Nash equilibrium. Given a point in time $F \in \mathbb{R}_+$, the *average arrival time price of anarchy of* $\mathcal{F}_{\mathrm{N}}$ *at* F is defined by

$$\rho_{\text{aver}}(\mathcal{F}_{\mathrm{N}})(F) := \frac{\text{val}_{\text{aver}}(\mathcal{F}_{\mathrm{O}})(F)}{\text{val}_{\text{aver}}(\mathcal{F}_{\mathrm{N}})(F)} \ .$$

The *average arrival time price of anarchy of* $\mathcal{F}_{\mathrm{N}}$ equals the worst-case ratio over all points in time, i.e.,

$$\rho_{\text{aver}}(\mathcal{F}_{\mathrm{N}}) := \inf_{F \in \mathbb{R}_+} \rho_{\text{aver}}(\mathcal{F}_{\mathrm{N}})(F) \ .$$

Further, the *average arrival time price of anarchy of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho^{\text{PoA}}_{\text{aver}}(\mathcal{I}) := \inf_{\mathcal{F}_{\mathrm{N}}} \rho_{\text{aver}}(\mathcal{F}_{\mathrm{N}}) \qquad \text{and} \qquad \rho^{\text{PoA}}_{\text{aver}}(\mathcal{F}_{\mathrm{N}}) := \inf_{\mathcal{I}} \rho^{\text{PoA}}_{\text{aver}}(\mathcal{I}) \ ,$$

respectively.

Definition 4.28 (Average Arrival Time Price of Stability)**.** Let $\mathcal{I}$ be an instance of a routing game over time. The *average arrival time price of stability of* $\mathcal{I}$ and *of the routing game over time* are defined by

$$\rho^{\text{PoS}}_{\text{aver}}(\mathcal{I}) := \sup_{\mathcal{F}_{\mathrm{N}}} \rho_{\text{aver}}(\mathcal{F}_{\mathrm{N}}) \qquad \text{and} \qquad \rho^{\text{PoS}}_{\text{aver}} := \inf_{\mathcal{I}} \rho^{\text{PoS}}_{\text{aver}}(\mathcal{I}) \ ,$$

respectively.

4.4 Your Comments

So far, routing games over time are mostly studied within the traffic community. After Pigou [73] introduces his example in 1920, Vickrey [89] and Yagar [93] are the first who discuss the topic of Nash equilibria in 1969 and 1971, respectively. Up to the middle of the 1980's, nearly all contributions consider Nash equilibria over time on given small instances (see, e.g. [38, 53, 67, 89]). Since then, the number of publications in this area has increased rapidly and Nash equilibria where modeled mathematically. For a survey see, e.g., [69]. The considered routing models can be grouped into for categories: mathematical programming (see, e.g., [35, 52]), optimal control (see, e.g., [30, 74]), variational inequalities (see, e.g., [16, 29, 75, 84, 87]), and simulation-based approaches (see, e.g., [7, 10, 54, 88, 93]). Mathematical programming is mainly applied to atomic routing games over time. With respect to nonatomic routing games, optimal control is used to model social optima and variational inequalities are the most common formulation for analyzing Nash equilibria. Many models mentioned above use a path-based formulation of flows over time. Therefore, they are often computationally intractable. Edge-based formulations are, for example, considered in [2, 16, 75].

Recently, Anshelevich and Ukkusuri [4] analyze discrete routing game models for Nash equilibria in the context of flows over time. They consider how a single splittable flow unit present at the source s at time 0 would traverse a network assuming every flow particle is controlled by a different player. The underlying flow model allows to send a positive amount of flow over an edge at each integral points in time. Moreover, the transit times are assumed to be constant. Hoefer et al. [39] also consider a discrete routing game. They study existence and complexity properties of pure Nash equilibria and best-response strategies.

According to the contribution of this chapter, routing games over time and, in particular, Nash equilibria over time are defined for the very general class of consistent flow models. It is shown that Nash equilibria are characterized by flows over time where no flow particle is able to overtake any other flow particle. As a consequence of the universal FiFo condition, Nash equilibria are completely defined via a family of static flows which are called underlying static flows. You think that studying these static flows (as done in Chapter 7) is a promising research direction for understanding Nash flows over time.

Further, the universal FiFo condition is essentially for proving existence of Nash equilibria for all consistent flow models. So far, the most general result in this direction is the work of Zhu and Marcotte [94]. They prove the existence of Nash equilibria for flow models based on load dependent transit times. In addition, they require that the transit time functions increase at least linearly depending on time, which is even stronger than requiring strict FiFo. But, beside existence, this FiFo condition enables them to establish also the uniqueness of Nash equilibria.

Theorem 4.19 extends the existence result of Zhu and Marcotte to arbitrary consistent flow models. As you already comment in Section 3.9, the locally finiteness of $\mathcal{P}$ is essential for proving existence in path-based flow over time models. In fact, this property on $\mathcal{P}$ ensures that the considered limit is a feasible strategy profile. But you have the feeling that the definition of locally finiteness is somewhat to restrictive. You believe that it is enough to require the following. For any current arrival time $\ell \in \mathbb{R}_+$ the set of path along which flow is able to arrive at the sink before time ℓ is finite. For consistent path-based models, using the family $\mathcal{L} := (\ell_P)_{P\in\mathcal{P}}$ of foresighted arrival times, this can be equivalently restated quite simple as $|\{P \in \mathcal{P} \mid \ell_P(0) \leq \ell\}| < \infty$ for all $\ell \in \mathbb{R}_+$.

Reconsidering the existence proof of Theorem 4.19 your idea is the following. As in the proof of Theorem 4.19, let $\mathcal{F}^* := (F^*_P)_{P\in\mathcal{P}}$ be the limit point of a sequence $(\mathcal{F}^k)_{k\in\mathbb{N}}$ of path-based flows over time $\mathcal{F}^k := (F^k_P)_{P\in\mathcal{P}}$. Since the foresighted arrival times are upper semi-continuous by Lemma 3.40, the foresighted arrival time functions of the flows $(\mathcal{F}^k)$ become bounded from above in terms of the foresighted arrival time functions of $\mathcal{F}^*$. Hence, by the new definition of the locally finiteness of $\mathcal{P}$ the set of paths at which flow is sent in $(\mathcal{F}^k)$ is finite for large enough k. Now you follow the same line of arguments as in the proof of Theorem 4.19 to show that $\mathcal{F}^*$ is a feasible strategy profile and a Nash equilibrium.

So far so good, but where are difficulties. Constructing the sequence $(\mathcal{F}^k)_{k\in\mathcal{P}}$ happens no longer under the assumption that $\mathcal{P}$ can be assumed to be finite. But using a diagonalization argument this can be unproblematicly resolved. The other aspect, you have to be aware of, is that $(\mathcal{F}^k)_{k\in\mathbb{N}}$ converges to $\mathcal{F}^*$ not in the sense as required for the definition of $\mathcal{F}$-continuity (see Definition 3.24). In fact, each $(F^k_P)_{k\in\mathbb{N}}$ converges to F^*_P separately for each $P \in \mathcal{P}$ and not uniformly for all $P \in \mathcal{P}$ as desired. Thus, based on this new kind of convergence you would have to install a another kind of $\mathcal{F}$-continuity and, in particular, you would have to reprove Lemma 3.40 for this new definition of $\mathcal{F}$-continuity. Nevertheless, the question arises whether or not, for consistent path-based flow models, the new definitions of $\mathcal{F}$-continuity and locally finiteness of $\mathcal{P}$ imply the known $\mathcal{F}$-continuity in terms of Definition 3.24. Proving this, your idea immediately establishes the existence of Nash equilibria under the weaker definition of locally finiteness of $\mathcal{P}$.

Although Definition 4.8 is an immediate generalization of static Nash flows, it is still a highly nontrivial problem to come up with an appropriate flow over time model, i.e., a particular flow over time model which allows an exact computation of Nash flows over time in finite time and an analysis of the prices of anarchy defined in Section 4.3. Especially for edge-based flow over time models, the main issue are the arrival time functions of a path which also serve as cost functions for the corresponding routing game. For static routing games, these cost functions are not given explicitly, but implicitly via edge arrival times. The arrival time of an edge e is a function of the flow x_e of that edge which can easily be computed as $x_e := \sum_{P,i:e=e_i^P} x_P$.

The situation is considerably more complicated for flows over time. Here, it is usually a highly nontrivial problem to compute the flow behavior on an edge e given a path-based flow over time. Consider a flow particle that enters a certain path $P \in \mathcal{P}$ containing e at a certain time θ. Notice that the time at which this particle arrives at an edge e depends on the current arrival times on the predecessor edges on path P which in turn depend on the flow behavior on these predecessor edges. But the flow behavior on the predecessor edges depends at least on the flow of the paths using these edges. This shows that, in contrast to the static case, the flow behavior on e also depends on flow of paths which do not traverse e in general. This fact induces involved dependencies among the functions representing the flow behavior on the edges. As a consequence, given a path-based flow over time, determining the edge arrival time functions is a highly nontrivial task in general. Clearly, this is exactly the task of the dynamic network loading problem discussed in Section 3.5.

Finally, you think, it is worth to mention that there are at least to other interesting objection function which can be used to define the price of anarchy. Due to personal communication with Elliot Anshelevich, these objective functions are maximum and average current *transit* times of flow particles of a given value. But note that an arbitrary earliest arrival flow is no longer a social optimum in this case. In fact, you need an earliest arrival which is additionally a latest departure flow, i.e., a flow attaining the earliest arrival flow value where each flow particles depart as lately as possible. Luckily, for classical flows over time the SUCCESSIVE SHORTEST PATH algorithm already outputs such a latest departure flow (see, e.g., [23]).

Because of your affinity to measure theory, you conclude these comments with the remark that it would be interesting to extend the notion of routing games over time to measure based flows over time.

Chapter 5

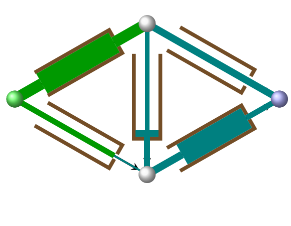
Direct Nash Flows over Time

In this chapter we analyze a routing game over time which is of particular interest if waiting is not allowed or not possible. For example, this feature occurs naturally in communication networks. Here, it is by no means acceptable that a phone call is interrupted or the voice is distorted due to time varying transit times. Also, with respect to the modeling of air lanes, it is not a good idea to stop an airplane during a flight. Other applications could arise out of security aspect in case waiting means increasing the risk.

Direct flows over time are build upon the classical flow over time model (see Section 2.5). That is, we are given a network and each edge has a fixed transit time. In addition, we assign a capacity to each edge bounding the rate, at which flow is allowed to cross an edge. Originating flow has to chose immediately a path for traversing the network. But however the case may be, flow is only allowed to enter a path if each edge of this path has free capacity ensuring that the flow is able to traverse the path without waiting at intermediates nodes or on edges. Conflicts between different flow particles are resolved by giving priority to the particle originating earlier. This means, if a flow particle has chosen some path, it has to wait in front of the source until the path is no longer blocked by particles with a smaller origination time. Hence, the flow-dependent path transit time consists of the (flow-dependent) waiting time plus the (constant) transit time.

In Section 5.1 we give a formal definition of the direct flow over time model with a strong mathematical background. In particular, we develop foresighted arrival times for this flow model. In Section 5.2 we show that direct earliest arrival flows exist and how a general path-based formulation is computable. This approach generalizes the Successive Shortest Path algorithm (see Section 2.5) to arbitrary time varying capacities. However, we have to make the assumption that all transit times are rational. Based on these general path-based representation, we give a very easy characterization of specific Nash equilibria in Section 5.3. In particular, we show that good Nash equilibria are computable by considering networks consisting only of parallel edges connecting the source with the sink.

In Section 5.4 we show that these Nash equilibria lead to nice bounds on the prices of stability. In terms of tightness, this analysis is nearly complete. In particular, in Subsection 5.4.2 we observe that the evacuation price of stability is bounded by 2. This bound directly carries over to the working price of anarchy.



However, if we restrict to constant capacities, the bound for the working price of anarchy decreases to 1.439 as we see in Subsection 5.4.3. Further, we observe that for the case of arbitrary capacities the (average) completion time price of stability is unbounded. In contrast, as observed in Subsection 5.4.4 and 5.4.5, we obtain a bound of $\frac{3}{4}$ for the case of constant capacities. We strongly conjecture that this bound can be further improved for the average completion time setting. Surprisingly, if the the supply tends to infinity, the nice direct Nash flows become earliest arrival flows implying that all prices of stability tend to 1 in such scenarios.

5.1 Direct Flow Model

Direct routing is based on three essential assumptions. The main feature is, that no waiting is supported. That is, after entering a path a flow particle has to cross this path immediately without waiting. Or in other words, the time needed for traversing a path after a flow particle has entered this path is flow-independent. The decision whether or not a path can be entered relies on time-dependent edge capacities. This means, if a particle would have to enter an edge which is currently used up to its capacity, this particle must not enter the corresponding path. Whether or not an edge is currently fully loaded only depends on flow which has already originated. Thus, we give priority to flow particles originating earlier which is the second assumption. So the routing occurs as follows. After origination a flow particle chooses some path P. Then it waits in front of the source (or in front of P) until this path becomes free. After P becomes free the particle enters P and travels directly to the sink. The third assumption simply requires that if there is a waiting queue in front of a path, this path is used up to its capacity, i.e., no flow particle waits longer than it has to. As we will see, this ensures FiFo of the underlying path-based flow model.

In this section we define the direct flow model and give a precise mathematical statement of the assumptions explained above. For this we consider edge-based flows over time. However, these edge-based flows are only used for defining the path-based direct flow model. Besides, this definition is not based on the dynamic network loading problem. In fact, as we will see, the direct flow model itself has no obvious consistent edge-based formulation.

We work on a directed graph $G := (V, E)$ with (finite) node set V and (finite) edge set E. For each edge $e \in E$ a *free-flow* transit time $\tau_e \in \mathbb{R}_+$ is given. If a flow particle enters e at a particular point in time $\theta \in \mathbb{R}_+$, it arrives at the head of e at time $\theta + \tau_e$. In addition, we assign to each edge e a capacity function $u_e : \mathbb{R}_+ \to \mathbb{R}_+$ bounding the outflow rate of e. That is, for a time $\theta \in \mathbb{R}_+$ and an edge e, the value $u_e(\theta)$ bounds the rate at which flow leaves e at time θ, and since waiting on an edge is not allowed, $u_e(\theta)$ also bounds the rate at which flow is able to enter e at time $\theta - \tau_e$. Especially for the latter interpretation, constant edge transit times are essential. So we deal with a network $(G, \mathcal{U}, \mathcal{T}, s, t)$ containing a source $s \in V$ and sink $t \in V$ where $\mathcal{U} := (u_e)_{e \in E}$ and $\mathcal{T} := (\tau_e)_{e \in E}$. Note that, in contrast to the previous chapters 3 and 4, the symbol τ_e stands for a nonnegative real number.

As already mentioned, the direct flow model is path-based and should be consistent. Thus, we have to ensure that the set $\mathcal{P}$ of all s-t-paths is locally finite in terms of Definition 3.34. So we require that an s-t-path $P := (e_1, \dots, e_{|P|})$ is contained in $\mathcal{P}$ if and only if P contains no cycle with a transit time of 0. That is, an edge of P can be only revisited at different points in time. More formally, let $\tau_{P,i} := \sum_{j=1}^{i} \tau_{e_j}$ be the time needed for arriving at the head of e_i from s along P. Then we require $\tau_{P,i} < \tau_{P,j}$ in case $e_i = e_j$ for some $i < j$. Note that, in general, nodes can be revisited at the same point in time.

In the following we determine the flow which is caused by the particles that have entered their chosen paths. For this consider a flow particle which has chosen a certain path $P := (e_1, \dots, e_{|P|})$. We say that this flow particle is *assigned* to P. If this flow particle starts traversing P, we say that this particle *enters* P. As already mentioned, a flow particle, which is assigned to P, has to wait until P is free in order to enter P. Assume that this flow particle enters P at some point in time $\theta \in \mathbb{R}_+$. Then it arrives at the head of e_i at time $\theta + \tau_{P,i}$. Thus, it arrives at t at time $\theta + \tau_P$ where we call $\tau_P := \sum_{j=1}^{i} \tau_{e_j}$ the *free flow transit time* of P. Or in other words, a flow particle which arrives at t at time θ enters P at time $\theta - \tau_P$ and arrives the head of edge e_i at time $\theta - \tau_P + \tau_{P,i}$. Recalling the shifting of a functions in Definition 2.10, this motivates the following definition. Note that e_i^P denotes the i-edge of P.

Definition 5.1 (Entering Flow)**.** Let $(F_P^-)_{P\in\mathcal{P}}$ be the outflow of some path-based flow over time. Then the *entering flow over time* $(F_P)_{P\in\mathcal{P}}$ is defined as $F_P := F_P^- + \tau_P$. Further, an edge-based representation $(F_e)_{e\in E}$ is given by

$$F_e := \sum_{P,i|e_i^P=e} F_P^- + \tau_P - \tau_{P,i} = \sum_{P,i|e_i^P=e} F_P - \tau_{P,i} \qquad \forall e \in E\ . \tag{5.1}$$

Note that F_e is, in fact, the outflow of edge e as $\tau_P - \tau_{P,i}$ equals the time needed for traveling from the head of the i-th edge $e_i^P = e$ to t along P.

Although Definition 5.1 is just a definition, it is worth to discuss consistency because of the following. It is quite natural to interpret the entering flow $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ as an inflow of some flow over time. Then $\mathcal{F}$ should lead to the family $\mathcal{F}^- := (F_P^-)_{P\in\mathcal{P}}$ of outflow functions in terms of (3.1). Further, in this scenario, the equations (5.1) represent a solution to the complex dynamic network loading problem (see Section 3.5) which have to be verified. However, as we deal with constant transit times, it is not that hard to observe that Definition 5.1 is consistent with respect to both questions. In this connection direct flows over time admit a very simple solution to the dynamic network loading problem. This becomes even clearer if we assume that each free flow transit time is equal to zero. Then the notion of (5.1) coincide with the notion of the *static* network loading problem – interpret the flow functions as flow values. Hence, direct flows over time seems to be very similar to static flows and as static routing games are well-understood, it is promising to start analyzing dynamic Nash equilibria under this flow over time model.

As already mentioned, in a feasible direct flow over time a particle is only allowed to enter the network if there is free capacity on every edge of the corresponding path. Hence, the entering flow has to obey the following *capacity constraints*:

$$f_e \le u_e \qquad \forall e \in E\ . \tag{5.2}$$

In the following we give a precise definition of a direct flow model. In fact, we define the model implicitly by stating conditions relating inflow, outflow, and arrival time functions. First, we define a candidate for the foresighted arrival time function. For this consider a path-based flow over time $(F_P^+)_{P\in\mathcal{P}}$ with corresponding restricted outflow functions $F_P^-|_{\leq\theta}$ for all $P \in \mathcal{P}$ and $\theta \in \mathbb{R}_+$. Recall that a path-based flow over time is completely defined via these functions (see Section 3.1). Bearing in mind Definition 5.1, the restricted entering flow is defined via $F_P|_{\leq\theta} := F_P^-|_{\leq\theta}+\tau_P$ with the corresponding edge-based formulation

$$F_e|_{\leq\theta} := \sum_{P,i|e_i^P=e} F_P|_{\leq\theta} - \tau_{P,i} \qquad \forall e \in E, \theta \in \mathbb{R}_+ \ . \tag{5.3}$$

Further, the remaining capacity $\mathcal{U}|_{\leq\theta} := (u_e|_{\leq\theta})_{e\in E}$ is given by:

$$u_e|_{\leq\theta} := u_e - f_e|_{\leq\theta} \qquad \forall e \in E, \theta \in \mathbb{R}_+ \ . \tag{5.4}$$

With these definition we are able to define the times at which a flow is allowed to enter a certain path $P = (e_1, \ldots, e_{|P|})$. For this consider flow which is assigned to P directly after time $\theta \in \mathbb{R}_+$. Since all flow particles originating until time θ have priority against this further flow, we have to consider the restricted flow until time θ. Hence, the remaining capacity $u|_{\leq\theta}$ bounds the rate at which further flow is able to traverse a particular edge. Since flow could traverse an edge if the corresponding current capacity is positive, further flow is able to enter P at time θ' if and only if $u_{e_i}|_{\leq\theta}(\theta' + \tau_{P,i}) > 0$ holds for all $i = 1, \ldots, |P|$. Defining the *remaining path capacity* as

$$u_P|_{\leq\theta} \ := \min_{i=1,\ldots,|P|} u_{e_i}|_{\leq\theta} + \tau_{P,i} \tag{5.5}$$

this can be restated as $u_P|_{\leq\theta}(\theta') > 0$. However, we require, in addition, that no flow waits longer than it has to. Therefore,

$$\bar{q}_P(\theta) = \max\left\{\Delta \geq 0 \;\middle|\; \int_\theta^{\theta+\Delta} u_P|_{\leq\theta}(\vartheta)\, d\vartheta \leq 0\right\} \tag{5.6}$$

seems to be a good choice for defining this minimal waiting time because: If the entering flow and, hence, also every restricted entering flow satisfies the capacity constraint (5.2), we are allowed to replace the “$\leq$”-sign by an “$=$”-sign. This means that during the interval $[\theta, \theta + \bar{q}(\theta)]$ no further flow can enter P as $u_P|_{\leq\theta}$ is essentially 0 on this interval. Secondly, from time $\theta + \bar{q}(\theta)$ further flow is able to enter P as

$$\int_{\theta+\bar{q}_P(\theta)}^{\theta+\bar{q}_P(\theta)+\epsilon} u_P|_{\leq\theta}(\vartheta)\, d\vartheta > 0$$

has to hold for all $\epsilon > 0$. However, if the underlying network only consists of one s-t-path P and if all capacity functions are strictly positive then setting all current waiting times to 0 results in a feasible $\bar{q}_P$. In this case the entering flow equals the inflow. Therefore, the capacity constraint is violated by the entering flow if the inflow is larger than the capacity of P. This shows that the capacity condition (5.2) must be an explicit constraint for defining feasible direct flows over time.

Summarizing the above discussion, the routing occurs as follows. After a flow particle originates at some point in time θ, it is assigned to some path P. Then it has to wait for $\bar{q}(\theta)$ time units before it can enter P at time $\theta + \bar{q}_P(\theta)$. Since traveling along P takes τ_P time units, it arrives at t at time

$$\bar{\ell}_P(\theta) := \theta + \bar{q}_P(\theta) + \tau_P \ . \tag{5.7}$$

This motivates the following definition. Note that due to the previous definitions a flow over time is completely defined in terms of the families of inflow and entering flow functions.

Definition 5.2 (Direct Flows over Time)**.** Let $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$ be a path-based flow over time and $\mathcal{F}^- := (F_P^-|_{\leq\theta})_{P\in\mathcal{P},\theta\in\mathbb{R}_+}$ be an entire family of restricted outflow functions. For an s-t-path P and a point time $\theta \in \mathbb{R}_+$, let $F_P|_{\leq\theta}$ be the restricted entering flow defined in (5.3). Then $(\mathcal{F}^+, \mathcal{F}^-)$ denotes a *feasible direct flow over time* if and only if $\bar{\ell}_P$ defined by (5.6) and (5.7) satisfies (3.2), i.e,

$$F_P^-|_{\leq\theta}(\theta') = \int_{\bar{\ell}^{-1}([0,\theta'])} f_P^+|_{[0,\theta]}(\vartheta)\ d\vartheta \ . \tag{5.8}$$

for all times $\theta \in \mathbb{R}_+$ and paths $P \in \mathcal{P}$. Moreover, for all $\theta \in \mathbb{R}_+$, the restricted entering flow $\mathcal{F}|_{\leq\theta} := (F_P|_{\leq\theta})_{P\in\mathcal{P}}$ hat to satisfy the capacity constraint (5.2).

In the following we observe that Definition 5.2 results in a consistent path-based flow over time model where foresighted arrival times are defined by (5.7).

Lemma 5.3. *The direct flow over time model is consistent. Furthermore, the family $\bar{\mathcal{L}} := (\bar{\ell}_P)_{P\in\mathcal{P}}$ of foresighted arrival times is given by* (5.7)*.*

Proof. Firstly, observe that the direct flow over time model is past-oriented as, for all s-t-paths P and all points in time $\theta \in \mathbb{R}_+$, the definition of $\bar{\ell}_P(\theta)$ only depends on $\mathcal{F}^+|_{\leq\theta}$.

In order to prove that the direct flow over time model satisfies FiFo, we apply Lemma 3.19. For this we show that all arrival time functions defined by (5.7) are nondecreasing. Let P be some s-t-path and θ_1, $\theta_2 \in \mathbb{R}_+$ be two points in time with $\theta_1 < \theta_2$. In case of $\theta_1 + \bar{q}_P(\theta_1) < \theta_2$ we obtain directly $\bar{\ell}_P(\theta_1) < \bar{\ell}_P(\theta_2)$ as $\bar{q}_P$ is nonnegative by definition. So assume without loss of generality that $\theta_2 \leq \theta_1 + \bar{q}_P(\theta_1)$. Since the corresponding restricted outflow functions satisfy $F_P^-|_{\leq\theta_1} \leq F_P^-|_{\leq\theta_2}$, we know $F_P|_{\leq\theta_1} \leq F_P|_{\leq\theta_2}$ by the definitions in (5.3) implying $u_P|_{\leq\theta_1} \geq u_P|_{\leq\theta_2} \geq 0$. Thus, we get because of the definition of $\bar{q}_P$ in (5.6):

$$0 \leq \int_{\theta_2}^{\theta_2+(\theta_1-\theta_2+\bar{q}_P(\theta_1))} u_P|_{\leq\theta_2}(\vartheta)\ d\vartheta \ \leq \ \int_{\theta_1}^{\theta_1+\bar{q}_P(\theta_1)} u_P|_{\leq\theta_1}(\vartheta)\ d\vartheta \ = \ 0 \ .$$

Because of the maximum in (5.6) this leads to $\theta_1 - \theta_2 + \bar{q}_P(\theta_1) \leq \bar{q}_P(\theta_2)$ implying $\theta_1 + \bar{q}_P(\theta_1) \leq \theta_2 + \bar{q}_P(\theta_2)$. Thus, $\bar{\ell}_P$ is nondecreasing. Since $\bar{\ell}$ is nonnegative by definition, we have $\bar{\ell}^{-1}([0,\bar{\ell}(\theta))) \subseteq [0,\theta] \subseteq \bar{\ell}^{-1}([0,\bar{\ell}(\theta)])$ for all $\theta \in \mathbb{R}_+$. As F^- is in particular continuous and nondecreasing, we conclude

$$\begin{aligned} F^-(\bar{\ell}(\theta)) \ &= \ \int_{\bar{\ell}^{-1}([0,\bar{\ell}(\theta)])} f^+(\vartheta)\ d\vartheta \ \geq \ \int_{[0,\theta]} f^+(\vartheta)\ d\vartheta = F^+(\theta) \\ \text{and} \qquad F^-(\bar{\ell}(\theta)) \ &= \ \int_{\bar{\ell}^{-1}([0,\bar{\ell}(\theta)))} f^+(\vartheta)\ d\vartheta \ \leq \ \int_{[0,\theta]} f^+(\vartheta)\ d\vartheta = F^+(\theta) \ . \end{aligned}$$

This shows that FiFo is satisfied by Lemma 3.19.

Instead of proving that Definition 3.24 holds for the direct flow over time model, we show a somewhat weaker kind of $\mathcal{F}$-continuity which is still strong enough for ensuring existence of Nash equilibria using Theorem 4.19. We show that for each sequence $(\mathcal{F}_k^+)_{k\in\mathbb{N}}$ converging to some path-based flow over time $\mathcal{F}^+$ there exists a subsequence $(\mathcal{F}_{k_i}^+)_{i\in\mathbb{N}}$ such that, for all s-t-path P, the corresponding subsequence $(F_{k_i}^-)_{i\in\mathbb{N}}$ of outflow functions converges to the outflow function F_P^- of P with respect to $\mathcal{F}^+$.

Consider some s-t-path P. Since the entering flow of P is always bounded by $u_P := \min_{i=1,\dots,|P|}\ u_{e_i} + \tau_{P,i}$, we are able to apply Lemma 2.36 to the sequence $(F_k^-)_{k\in\mathbb{N}}$ of outflow functions of P where the upper bound is set to U_P. This shows that there exists a subsequence $(F_{k_{i^P}}^-)_{i^P\in\mathbb{N}}$ converging to some F^-. This procedure we apply iteratively to each s-t-path P while remaining current subsequences. Using a diagonalization technique, we construct the subsequence $(\mathcal{F}_{k_i}^+)_{i\in\mathbb{N}}$ of $(\mathcal{F}_k^+)_{k\in\mathbb{N}}$ as follows. The index k_1 is set to the first index of the subsequence constructed with respect to the first considered s-t-path. The index k_2 is set to the second index of the subsequence constructed with respect to the second considered s-t-path and so on. In this manner we construct a sequence $(\mathcal{F}_{k_i}^+)_{i\in\mathbb{N}}$ such that, for all s-t-path P, the corresponding subsequence $(F_{k_i}^-)_{i\in\mathbb{N}}$ of outflow functions converges to the outflow function F_P^- of P with respect to $\mathcal{F}^+$.

It remains to observe that the arrival time functions defined by (5.7) are, in fact, foresighted arrival time functions. To see this, note that the definition of the waiting time in 5.6 ensures that whenever additional flow is sent at a certain point in time $\theta \in \mathbb{R}_+$ along some s-t-path P, this flow arrives at time $\bar{\ell}_P(\theta)$ at the sink t. □

5.2 Direct Earliest Arrival Flows

In order to measure the performance of direct Nash flows which are characterized subsequently in Section 5.3, we need a notion of optimal direct flows. This would enable us to define the price of anarchy just as the price of stability for direct flow over time models. If all edge capacities are constant the direct flow model is nothing more than the classical flow over time model where waiting is not allowed. As explained in Section 2.5, earliest arrival flows serve as optimal flows in this classical setting. Recall that an earliest arrival flow maximizes the amount of flow which has been arriving at the sink for all points in time simultaneously. It is well-known that, in classical flow over time scenarios, there always exist an earliest arrival flow where flow does not wait at intermediate nodes. So it seems to be promising to extend the notion of earliest arrival flows to the direct flow over time model.

In this section we show that direct flow over time models always admit an earliest arrival flow. We do this in a constructive manner. It turns out that direct earliest arrival flows are computable via a Successive Shortest Path algorithm. In particular, this Successive Shortest Path algorithm generalizes the corresponding algorithm of the classical setting to arbitrary time-varying capacities. Nevertheless, this algorithm is stated as a generalization of the Static Successive Shortest Path algorithm for computing *static*

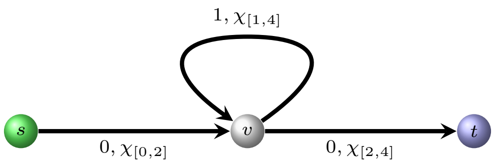


Figure 5.1: Network related to Example 5.5 showing that the path capacity does not result in a feasible flow in general. On the edges the transit times and capacities are shown in this order.

minimum cost flows. Also we prove the correctness of this algorithm along the same lines as it works for the static case. In particular, this avoids the extension of the classical dynamic MaxFlow-MinCut Theorem to the direct setting with arbitrary time-varying capacities. Further, for the termination of the Direct Successive Shortest Path algorithm presented in this section the following assumption is essential.

Asumption 5.4. We assume that $\tau_e \in \mathbb{Q}_+$ is rational for each edge $e \in E$.

As mentioned in Section 5.1, a direct flow over time is completely characterized by a family of inflow functions and by a family of entering flow functions. However, how the flow arrives at the sink depends primary on the entering flow over time. Therefore, we only consider entering flows in this section. That is, every flow occurring during this section is interpreted as an entering flow over time.

In the Static Successive Shortest Path algorithm flow is always send along minimum cost paths in the residual network. In order to transfer this algorithm to direct flow models the role of the costs is engaged by the transit times. Further, it is crucial feature of the Static Successive Shortest Path algorithm that flow is always send along simple paths implying that the minimum edge capacity can be send feasible along this path. Unfortunately, the Direct Successive Shortest Path algorithm has to send flow along paths which contain cycles in general. Therefore, we Must be careful with the amount of flow which shall be send along a path. This is explained in the following example which shows that the flow on some s-t-path $P \in \mathcal{P}$ resulting out of the path capacity u_P can violate the capacity condition.

Example 5.5. Consider the network shown in Figure 5.1 consisting of three nodes s, v, and t which are connected with two edges sv and vt of zero transit time. Further, there is a loop vv with transit time 1. The capacities bounding the flow rate at the end of the edges are

$$u_{sv} := \chi_{[0,2]}\ , \qquad u_{vv} := \chi_{[1,4]}, \qquad \text{and} \qquad u_{vt} := \chi_{[2,4]}\ .$$

Next we consider the s-t-path $P := (sv, vv, vv, vt)$ which traverses the loop vv twice. Since we have

$$u_{sv} + 0 = \chi_{[0,2]}, \quad u_{vv} + 1 = \chi_{[0,3]}, \quad u_{vv} + 2 = \chi_{[-1,2]}, \quad \text{and} \quad u_{vt} + 2 = \chi_{[0,2]},$$

the path capacity u_P is equal to $u_P = \chi_{[0,2]}$. Hence, it seems feasible to send flow at a rate of 1 into the s-t-path P over the time interval $[0, 2]$, i.e., $f_P := \chi_{[0,2]}$. This flow enters the loop vv over the time intervals $[0, 2]$ and $[1, 2]$ implying that flow enters vv at a rate of 2 from time 1 to time 2. But this shows that f_P is infeasible because flow is only allowed to enter vv at a maximum rate of 1. However, P is not a shortest s-t-path on this network which is able to carry flow. The shortest s-t-path is $P' := (sv, vv, vt)$ which traverses the loop vv only once. Furthermore, we have $u_{P'} = \chi_{[1,2]}$ which results in a feasible direct flow over time.

In the following we formulate the Successive Shortest Path algorithm for direct flows over time. As its static counterpart, it sends flow successively along currently shortest paths in the corresponding residual network. Since transit times take the role of costs in the algorithm, the length of a path is measured with respect to transit times and *not* with respect to arrival times. We mentioned this in order to avoid confusion in advance – beside the fact that considering arrival times makes no sense during the algorithm as there are no departure times.

Further, a path in a static residual network is always able to carry flow. Since this is not the case for the dynamic residual network defined in Section 3.8, we explicitly have to require that only paths are considered which are able to carry flow. As the static algorithm, we also send the minimal edge capacity through the computed paths. Unfortunately, Example 5.5 shows that this may not be feasible implying that we also have to concentrate our attention to this point when analyzing this algorithm. Nevertheless, Example 5.5 already shows that this could work as the "bad" path is not a shortest one. Finally, note that the Successive Shortest Path algorithm returns a flow over time which is given via a general path-based representation (see Section 3.8).

Successive Shortest Path Algorithm (SSP)

Input: A network $\mathcal{N} := (G, \mathcal{U}, \mathcal{T}, s, t)$ consisting of a directed graph G, a family $\mathcal{U}$ of time-dependent capacity functions, a family $\mathcal{T}$ of real transit times, a source s, and a sink t.

Output: A direct earliest arrival flow over time given by a generalized path-based flow over time $\mathcal{F} := (f_P)_{P \in \mathcal{P}^r}$.

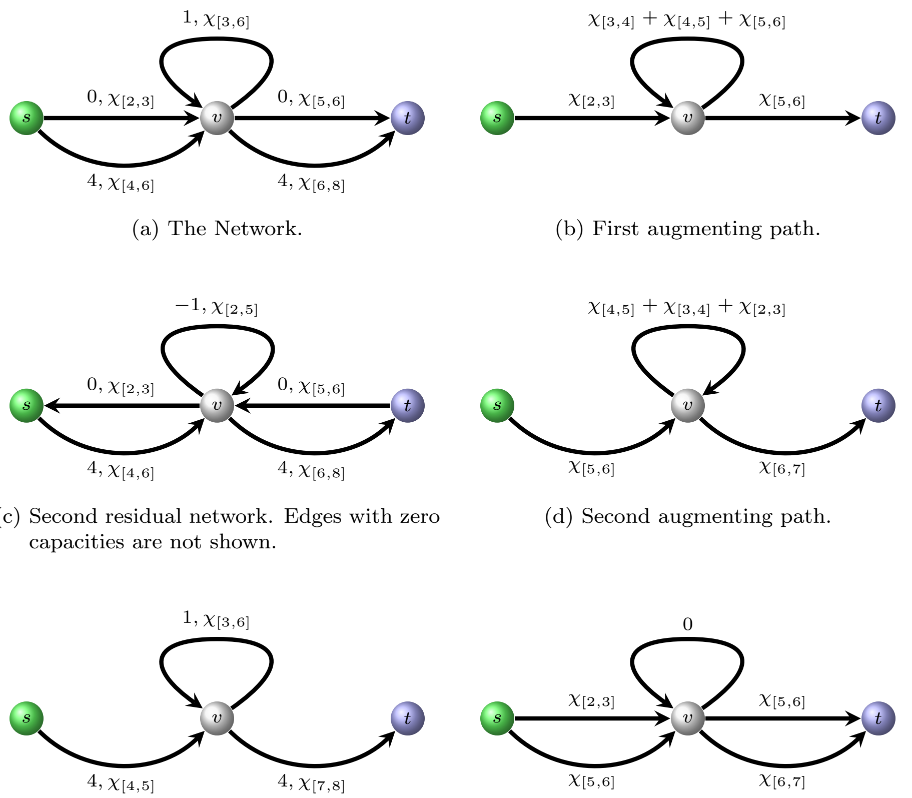
(1) Let $\mathcal{F} := (f_P)_{P \in \mathcal{P}^r}$ be the zero flow, i.e., set $f_P := 0$ for all $P \in \mathcal{P}^r$.

(2) Compute the residual network $(G^r, \mathcal{U}^r, \mathcal{T}^r, s, t)$ with respect to $\mathcal{F}$.

(3) Let $P \in \mathcal{P}^r$ be a shortest path which is able to carry flow, i.e., $U^r_P \gneq 0$.
If there are several take one with the smallest number of edges.
If no such path exists return $\mathcal{F}$.

(4) Set $f_P := u^r_P$ and go to (2).

Although the SSP algorithm has a termination condition in Step (3), it is a priori not clear whether or not this condition is met in finite time. In fact, the SSP algorithm does not terminate in general. However, the main goal of this section is to show that the SSP algorithm can be interrupted after a finite number of iterations in order to compute a direct earliest arrival flow for a given time horizon $T \in \mathbb{R}_+$, i.e., a flow which sends the maximal amount of flow to t simultaneously until all times $\theta \leq T$. This is done along the same lines as proving the correctness of the *static* SSP algorithm which minimizes the cost of



(a) The Network.

(b) First augmenting path.

(c) Second residual network. Edges with zero capacities are not shown.

(d) Second augmenting path.

(e) Third residual network. Backward edges with head s or tail t are not shown.

(f) The edge-based direct earliest arrival flow.

Figure 5.2: Shows a run of the SSP algorithm explained in Example 5.6. Left the current residual networks are drawn (transit times and capacities are shown on the edges in this order). Right the shortest augmenting paths and the computed earliest arrival flow are drawn (the flow on the edges is shown).

a *static* flows. But before, we apply the SSP algorithm to a particular instance which is illustrated in Figure 5.2.

Example 5.6. Consider the network in Figure 5.2a consisting of three nodes s, v, and t. The source s is connected with v via two edges e_1 and e_2 with transit times 0 and 4 and capacities $\chi_{[2,3]}$ and $\chi_{[4,6]}$, respectively. Over the node v there is a loop e_3 with transit time 1 and capacity $\chi_{[3,6]}$. Completing the constructing, v is connected to the sink t via two edges e_4 and e_5 with transit times 0 and 4 and capacities $\chi_{[4,6]}$ and $\chi_{[6,8]}$, respectively.

After initializing $\mathcal{F}$ with the zero flow in Step (1), the residual network computed during the subsequent call of Step (2) equals the original network except that all backward arc are added with zero capacity. Hence, in Step (3) the SSP algorithm has to find a shortest path which is able to carry flow on the original network. The shortest augmenting path is the path $P_1 := (e_1, e_3, e_3, e_3, e_4)$ which goes along the short sv and vt edges and uses the loop e_3 three times (see Figure 5.2b). For verifying that this is the shortest augmenting path, note the following two facts. Firstly,every path which is able to carry flow and which uses the edge e_1 and e_4 has to traverse the loop e_3 exactly three times. Secondly, since the transit time of P is 3, a shortest augmenting path never uses the edges e_2 or e_5 because they both have a transit time of 4. Hence, P_1 is the s-t-path constructed during the first call of Step (3). Moreover, we have $u^r_{P_1} = \chi_{[2,3]}$ which can be send feasible along P_1 (see Figure 5.2b for the corresponding edge-based flow). So at the end of the first iteration, $\mathcal{F}$ is set to $(P_1, \chi_{[2,3]})$ in Step (4).

Now we construct the residual network shown in Figure 5.2c which is computed during the second call of Step (2). Since the path flow $(P_1, \chi_{[2,3]})$ uses the edges e_1, e_3, and e_4 up to its capacity, the residual capacities of these edges vanish. Further, the residual capacities of the corresponding backward edges are computed by

$$u_{\overleftarrow{e}_1} = \chi_{[2,3]} \ , \qquad u_{\overleftarrow{e_3}} = \chi_{[3,6]} + 1 = \chi_{[2,5]} \ , \qquad \text{and} \qquad \chi_{\overleftarrow{e}_4} = \chi_{[5,6]} \ .$$

Since P_1 does not use other edges, the residual capacities of the remaining edges are unchanged. The shortest augmenting path on this residual network shown in Figure 5.2d equals the path $P_2 := (e_2, \overleftarrow{e}_3, \overleftarrow{e}_3, \overleftarrow{e}_3, e_5)$ with a capacity of $u^r_{P_2} = \chi_{[1,2]}$. So after the second execution of Step (4), the flow $\mathcal{F}$ equals the general path-based flow $((P_1, \chi_{[2,3]}), (P_2, \chi_{[1,2]}))$ (see Figure 5.2f for an edge-based representation).

The residual network computed at the beginning of the third iteration is shown in Figure 5.2e. It turns out that there exists no s-t-path which is able to carry flow on this network. Hence, $((P_1, \chi_{[2,3]}), (P_2, \chi_{[1,2]}))$ is the output of the SSP algorithm, and it is not hard to see that it is a direct earliest arrival flow on this network. For this take a look at the original instance shown in Figure 5.2a and observe that it is impossible to send flow which arrives at t over a path containing both long edges e_2 and e_5.

In the following we prove the correctness of the SSP algorithm. As already mentioned, the main line of arguments is borrowed from the correctness proof of the static SSP algorithm. The following lemma shows that the flow $\mathcal{F}$ which is maintained by the algorithm always obeys the capacity constraints of the original network. Note that this lemma has no static counterpart as it is not needed, i.e., it follows directly from the definition of the static residual network and the simplicity of shortest path. The proof of this lemma is based on Lemma 3.58 which shows that a feasible flow which is augmented by a feasible flow in the corresponding residual network remains feasible. Therefore, establishing the following lemma reduces to proving the feasibility of the path flows in the residual network. For this it worth to recall Example 5.5 which also illustrates the main idea of the proof.

Lemma 5.7. *At every call of Step (2) the flow over time $\mathcal{F}$ is feasible on $\mathcal{N}$.*

Proof. We show that the flow (f_P, P) computed during an iteration of the SSP algorithm is feasible on the current residual network $\mathcal{N}^r := (G^r, \mathcal{U}^r, \mathcal{T}^r, s, t)$. For this we use that P is a currently shortest s-t-path which is able to carry flow. Then this lemma is inductively established by Lemma 3.58 as the zero flow is obviously feasible on the original network.

Firstly, we observe that if an edge e is revisited by P, the old and the new flow leave e at different points in time. Formally, let i_1 and i_2 be two positions of P with $1 \leq i_1 < i_2 \leq |P|$ and $e^P_{i_1} = e = e^P_{i_2}$. Then we claim that the flow $f_{i_1} := f_P - \tau^r_{P,i_1}$ and the flow $f_{i_2} := f_P - \tau^r_{P,i_2}$ are mutually singular, i.e., $\min\{f_{i_1}, f_{i_2}\}$ is essentially equal to 0.

Assuming the opposite we show that P cannot be a shortest path which is able to carry flow. For this we consider the cycle $C := (e^P_{i_1}, \ldots, e^P_{i_2-1})$ resulting out of the subpath of P from the tail of $e^P_{i_1}$ to the tail of $e^P_{i_2}$. The transit time τ^r_C of C equals $\tau^r_{P,i_2} - \tau^r_{P,i_1}$. Hence, since we assume that $\min\{f_{i_1},\ f_{i_2}\}$ is not essentially 0, shifting both f_{i_1} and f_{i_2} by τ_{P,i_1} units to the left shows that

$$f := \min\{f_P,\ f_P - \tau^r_C\}$$

is not essentially 0. Next we distinguish two cases depending on whether τ^r_C is positive or negative. The case $\tau^r_C > 0$ corresponds to Example 5.5. Note that $\tau^r_C \neq 0$ because P contains no cycle of zero transit time by definition.

If $\tau^r_C > 0$ holds, we consider the path $Q := (e^P_1, \ldots, e^P_{i_1-1}, e^P_{i_2}, \ldots, e^P_{|P|})$ where the cycle C is removed from P_k. Since $\tau^r_C > 0$, we have $\tau^r_Q < \tau^r_P$ on the one hand. On the other hand, we have

$$\begin{aligned} u^r_{e^Q_i} + \tau^r_{Q,i} &= u^r_{e^P_i} + \tau^r_{P,i} \\ &\geq f_P \geq f && \forall 1 \leq i < i_1 \ , \\ u^r_{e^Q_i} + \tau^r_{Q,i} &= u^r_{e^P_{i-i_1+i_2}} + (\tau^r_{P,i-i_1+i_2} - \tau^r_C) \\ &\geq f_P - \tau^r_C \geq f && \forall i_1 \leq i \leq |Q| \ . \end{aligned}$$

Hence, Q is an s-t-path with a transit time smaller than τ^r_P which is able to carry flow in the residual network $\mathcal{N}^r$ contradicting the minimality of P.

If $\tau^r_C < 0$ holds, let $Q := (e^P_1, \ldots, e^P_{i_2-1}, e^P_{i_1}, \ldots, e^P_{|P|})$ where the cycle C is traversed once more. Since $\tau^r_C < 0$, we have $\tau^r_Q < \tau^r_P$ on the one hand. On the other hand, we have

$$\begin{aligned} u^r_{e^Q_i} + \tau^r_{Q,i} &= u^r_{e^P_i} + \tau^r_{P,i} \\ &\geq f_P \geq f + \tau^r_C && \forall 1 \leq i < i_2 \ , \\ u^r_{e^Q_i} + \tau^r_{Q,i} &= u^r_{e^P_{i-i_2+i_1}} + (\tau^r_{P,i-i_2+i_1} + \tau^r_C) \\ &\geq f_P + \tau^r_C \geq f + \tau^r_C && \forall i_2 \leq i \leq |Q| \ . \end{aligned}$$

Since $f + \tau^r_C := \min\{f_P + \tau^r_C,\ f_P\} \gneq 0$ is valid, Q is again an s-t-path with a transit time smaller than τ_P which is able to carry flow in the residual network $\mathcal{N}^r$. Thus, if an edge e is revisited by P then the old and the new flow enter e at different points in time.

In order to show that (f_P, P) is feasible with respect to u^r, let e be an edge of P and let $J := \{i \mid e^P_i = e\}$ be the set of positions at which P traverses e. Since we have shown so far that the elements of $(f_i)_{i \in J}$ are pairwise mutually singular, there exists a set S_i for each $i \in J$ such that $S_i \cap S_j = \emptyset$ for all $i, j \in J$ with $i \neq j$ and f_i is essentially equal to $f_i|_{S_i}$ for all $i \in J$. Hence, the following chain holds essentially

$$\begin{aligned} f_e = \sum_{i \mid e^P_i = e} f_i = \sum_{i \mid e^P_i = e} f_i|_{S_i} &= \sum_{i \mid e^P_i = e} (u^r_P - \tau^r_{P,i})|_{S_i} \\ &\leq \sum_{i \mid e^P_i = e} (u^r_e)|_{S_i} \leq u^r_e \ . \end{aligned}$$

□

As already mentioned, we do not know a priory whether or not the SSP algorithm terminates. Let $K \in \mathbb{N}^\infty$ be the number of iteration, either finite or infinite. For the rest of this section, we introduce the following notation. For each $k \in [K]$, we denote by P_k the s-t-path which is constructed during the k-th call of Step (3) and by $\mathcal{F}^k := (f_e^k)_{e \in E}$ the corresponding *edge-based* flow which is sent along P_k during the subsequent call of Step (4). Note that (f_{P_k}, P_k) is the path-based flow of $\mathcal{F}^k$. Further, for all $k \in [K+1]$ we use $\mathcal{U}^k := (u_e^k)_{e \in E^r}$ for denoting the residual capacities with respect to $(f_{P_i})_{i \in [k]}$. That is, $\mathcal{U}^k$ is equal to $\mathcal{U}^r$ computed during the k-th call of Step (2) implying that $\mathcal{F}^k$ is a feasible flow over time on the residual network $\mathcal{N}^k := (G^r, \mathcal{U}^k, \mathcal{T}^r, s, t)$. The following recursion is a direct consequence of Lemma 3.61:

$$u_e^{k+1} = u_e^k - f_e^k + (f_{\overleftarrow{e}}^k + \tau_{\overleftarrow{e}}^r) \qquad \forall k \in [K] \; . \tag{5.9}$$

A cycle C of some residual network is called *negative cycle* if the transit time τ_C^r of C is negative. Further, for a cycle or a path $P := \{e_1, \ldots, e_{|P|}\}$, we denote with $\tau_P^+ := \sum_{i | \tau_{e_i}^r > 0} \tau_{e_i}^r$ and with $\tau_P^- := \sum_{i | \tau_{e_i}^r < 0} \tau_{e_i}^r$ the sum of the positive and negative residual transit times of the edges traversed by P, respectively. In this sense, a negative cycle C satisfies $\tau_C^- > \tau_C^+$.

The following lemma shows that, in each residual network $\mathcal{N}^k$, no negative cycle exists around which flow could be sent. In the static setting it is well-known that this negative cycle condition is an optimality criterion. This optimality criterion is met by the static SSP, ensuring that it outputs a minimum cost flow. Along these lines, Lemma 5.8 is subsequently used for proving that the output of the direct SSP algorithm is an earliest arrival flow if the algorithm terminates.

Before we prove the negative cycle condition, note the following aspect. Since flow is routed directly, every flow carrying cycle C must have a transit time of 0 unless C contains s or t. Thus, every negative cycle around which flow could be sent must contain either s, t, or both.

Lemma 5.8. *For all $k \in [K+1]$ the residual network $\mathcal{N}^k$ contains no negative cycle which is able to carry flow.*

Proof. We prove this by induction over k. For $k = 1$ nothing has to be shown because the original transit times $(\tau_e)_{e \in E}$ are nonnegative and residual capacities of the backward edges are 0. So assume that this lemma is established for some $k \in [K]$.

Let C be a cycle in G^r and let $\mathcal{F}^C := (f_e^C)_{e \in E^r}$ be a nonzero direct flow over time which can be feasible sent around C in the residual network $\mathcal{N}^{k+1}$. Further, assume without loss of generality that there exists a $\theta_2 \in \mathbb{R}_+$ such that $f_e^C(\theta) = 0$ holds for all $e \in E^r$ and almost all $\theta > \theta_2$. Next, we construct a flow $\bar{\mathcal{F}}$. For this let $\mathcal{F}^k$ be the flow which is send along P_k by the SSP algorithm and let $\theta_1 := \theta_2 + \tau_{P_k}^-$. Send the restricted flow $\mathcal{F}^k|_{\leq \theta_1}$ along P_k and $\mathcal{F}^C$ along C. Further, delete flow whenever it is sent simultaneously along e and $\overleftarrow{e}$ for some $e \in E$, i.e., for all times reduce the flow on both edges until the flow on one edge becomes 0. The resulting flow is $\bar{\mathcal{F}} := (\bar{f}_e)_{e \in E}$. Since $\mathcal{F}^k|_{\leq \theta_1} \leq \mathcal{F}^k$ and $\mathcal{F}^C$ satisfy the assumptions of Lemma 3.62 and 3.63, the flow over time $\bar{\mathcal{F}}$ is feasible on the residual network $\mathcal{N}^k$.

Next, decompose $\bar{\mathcal{F}} := (\bar{F}_P)_{P \in \bar{\mathcal{P}} \cup \bar{\mathcal{C}}}$ into a set $\bar{\mathcal{P}}$ of flow carrying s-t-paths and a set $\bar{\mathcal{C}}$ of flow carrying cycles. Because P_k is a shortest flow carrying s-t-path in $\mathcal{N}^k$, we know $\tau^r_{P_k} \leq \tau^r_{P'}$ for each $P' \in \bar{\mathcal{P}}$, and because of the induction assumption, we know $\tau^r_{C'} \geq 0$ for each $C' \in \bar{\mathcal{C}}$. Note that, in fact, $\tau^r_{C'} = 0$ holds for all $C' \in \bar{\mathcal{C}}$ which contain neither s nor t because waiting is not allowed at intermediate nodes. Thus, equation (3.50) in Lemma 3.63 shows

$$\begin{aligned} \sum_{P' \in \bar{\mathcal{P}}} \left\|\bar{F}_{P'}\right\|_\infty \cdot \tau^r_{P_k} &\leq \sum_{P' \in \bar{\mathcal{P}}} \left\|\bar{F}_{P'}\right\|_\infty \cdot \tau^r_{P'} + \sum_{C' \in \bar{\mathcal{C}}} \left\|\bar{F}_{C'}\right\|_\infty \cdot \tau^r_{C'} \\ &= F_{P_k}(\theta_1) \cdot \tau^r_{P_k} + F_C(\theta_2) \cdot \tau^r_C \ . \end{aligned}$$

Hence, because of equation (3.49) in Lemma 3.63, we obtain $0 \leq F_C(\theta_2) \cdot \tau^r_C$ implying $\tau^r_C \geq 0$ because $F_C(\theta_2) > 0$ holds by definition. □

The following lemma shows that the transit times of the paths found in Step (3) tend to infinity in case the SSP algorithm never terminates. As we will see, this ensures that we can interrupt the algorithm if we want to compute a direct earliest arrival flow for a given time horizon $T \in \mathbb{R}_+$. Besides, the lemma provides also basic insights in the functionality of the algorithm.

Lemma 5.9. *The following holds:*

(i) The sequence $(\tau^r_{P_k})_{k \in [K]}$ is nondecreasing.

(ii) If $\tau^r_{P_k} = \tau^r_{P_{k+1}}$ for some $k \in [K]$ then $|P_k| \leq |P_{k+1}|$.

(iii) If $\tau^r_{P_{k_1}} = \tau^r_{P_{k_2}}$ for some $k_1, k_2 \in [K]$ with $k_1 < k_2$ such that $f^{k_1}_{\overleftarrow{e}} + \tau^r_{\overleftarrow{e}}$ and $f^{k_2}_e$ are not mutually singular for some $e \in E^r$ then $|P_k| < |P_{k+1}|$.

(iv) The sequence $(\tau^r_{P_k})_{k \in [K]}$ is unbounded if $K = \infty$.

Proof. For proving (i) and (ii), let θ_2 such that $F^{k+1}_P(\theta_2) > 0$, i.e., a positive amount of flow departing before time θ_2 is sent through P_{k+1} by the SSP algorithm. Further, let $\theta_1 := \theta_2 + \tau^-_{P_k} + \tau^+_{P_{k+1}}$. In order to construct the flow $\bar{\mathcal{F}}$ send $\mathcal{F}^k|_{\theta_1}$ through P_k and $\mathcal{F}^{k+1}|_{\theta_2}$ through P_{k+1} and delete as much flow as possible on cycles consisting of exactly one pair of antiparallel edges. Then $\mathcal{F}^k|_{\theta_1} \leq \mathcal{F}^k$ and $\mathcal{F}^{k+1}|_{\leq\theta_2}$ satisfy the assumptions of Lemma 3.62 and 3.63 implying that $\bar{\mathcal{F}}$ is feasible a feasible flow over time on the residual network $\mathcal{N}^k$.

Next, we decompose $\bar{\mathcal{F}} := (\bar{F}_P)_{P \in \bar{\mathcal{P}} \cup \bar{\mathcal{C}}}$ into a set $\bar{\mathcal{P}}$ of flow carrying s-t-paths and a set $\bar{\mathcal{C}}$ of flow carrying cycles. Then Lemma 3.63 shows

$$F_{P_k}(\theta_1) + F_{P_{k+1}}(\theta_2) = \sum_{P' \in \bar{\mathcal{P}}} \left\|F_{P'}\right\|_\infty \qquad \text{and} \tag{5.10}$$

$$\tau^r_{P_k} F_{P_k}(\theta_1) + \tau^r_{P_{k+1}} F_{P_{k+1}}(\theta_2) = \sum_{P' \in \bar{\mathcal{P}}} \tau^r_{P'} \left\|F_{P'}\right\|_\infty + \sum_{C' \in \bar{\mathcal{C}}} \tau^r_{C'} \left\|F_{C'}\right\|_\infty . \tag{5.11}$$

Since $\bar{\mathcal{F}}$ is feasible on $\mathcal{N}^k$, we know from the construction of P_k in Step (3) that $\tau^r_{P_k} \leq \tau^r_{P'}$ for all $P' \in \bar{\mathcal{P}}$. Furthermore, Lemma 5.8 tells us $0 \leq \tau^r_{C'}$ for all $C' \in \bar{\mathcal{C}}$. Hence, we obtain from (5.11)

$$\tau^r_{P_k} \cdot F_{P_k}(\theta_1) + \tau^r_{P_{k+1}} \cdot F_{P_{k+1}}(\theta_2) \geq \tau^r_{P_k} \cdot \sum_{P' \in \bar{\mathcal{P}}} \left\|F_{P'}\right\|_\infty .$$

Together with (5.10) this leads to $\tau^r_{P_{k+1}} \geq \tau^r_{P_k}$ and (i) is proven.

Assume $\tau^r_{P_{k+1}} = \tau^r_{P_k}$ in order to prove (ii). Using $0 \leq \tau^r_{C'}$ for all $C' \in \bar{\mathcal{C}}$, we obtain from (5.11)

$$\tau^r_{P_k} \;\geq\; \sum_{P' \in \bar{\mathcal{P}}} \frac{\left\|F_{P'}\right\|_\infty}{F_{P_k}(\theta_1) + F_{P_{k+1}}(\theta_2)} \cdot \tau^r_{P'} \; .$$

Because of (5.10) the left hand side is a convex combination of the values $\tau^r_{P'}$. On the other hand, we know $\tau^r_{P_k} \leq \tau^r_{P'}$ for all $P' \in \bar{\mathcal{P}}$ implying $\tau^r_{P_k} = \tau^r_{P'}$ for all $P' \in \bar{\mathcal{P}}$. Thus, from the second criterion defining P_k in Step (3) we get $|P_k| \leq |P'|$ for all $P' \in \bar{\mathcal{P}}$. So because of inequality (3.51) in Lemma 3.63, we know

$$\begin{aligned} F_{P_k}(\theta_1) \cdot |P_k| + F_{P_{k+1}}(\theta_2) \cdot |P_{k+1}| \;&\geq\; \sum_{P' \in \bar{\mathcal{P}}} \left\|F_{P'}\right\|_\infty \cdot |P'| + \sum_{C' \in \bar{\mathcal{C}}} \left\|F_{C'}\right\|_\infty \cdot |C'| \\ &\geq\; |P_k| \cdot \sum_{P' \in \bar{\mathcal{P}}} \left\|F_{P'}\right\|_\infty \; . \end{aligned}$$

Together with (5.10) this leads to $|P_{k+1}| \geq |P_k|$ and (ii) is proven.

For proving (iii) we assume that $f^k_{\overleftarrow{e}} - \tau^r_{\overleftarrow{e}}$ and $f^{k_2}_e$ are mutually singular for all $k \in \{k_1+1, \ldots, k_2-1\}$ and $e \in E^r$. This causes no loss of generality because (ii) holds. Further, this means that, for all $k \in \{k_1+1, \ldots, k_2-1\}$, the path P_k sends essentially no flow along an edge $\overleftarrow{e}$ at some point in time $\theta - \tau^r_{\overleftarrow{e}}$ if the corresponding forward edge e is used by P_{k_2} at time θ. Hence, the sequence $(u^k_e(\theta))_{k_1 < k \leq k_2}$ is nonincreasing for almost all points in time θ at which P_{k_2} sends flow over this edge. But this shows that $\mathcal{F}^{k_2}$ is feasible with respect to $\mathcal{U}^{k_1+1}$.

Now we proceed similarly to the proof of (ii). Let $\bar{\mathcal{F}} = (\bar{f}_e)_{e \in E^r}$ be the flow arising out of the path-based flow $((P_{k_1}, F_{P_{k_1}}|_{\leq\theta_1}), (P_{k_2}, F_{P_{k_2}}|_{\leq\theta_2}))$ after deleting flow on antiparallel edges. Here, $\theta_2 \in \mathbb{R}_+$ is chosen such that $F_{P_{k_2}}(\theta_2) > 0$ and θ_1 is set to $\theta_2 + \tau^-_{P_k} + \tau^+_{P_{k+1}}$. This shows that $F_{P_{k_1}}|_{\leq\theta_1} \leq F_{P_{k_1}}$ and $F_{P_{k_2}}|_{\leq\theta_2}$ satisfy the assumptions of Lemma 3.62 and 3.63. In particular, this ensures that $\bar{\mathcal{F}}$ is a feasible flow over time on $\mathcal{N}^{k_1}$.

Next, we decompose $\bar{\mathcal{F}} := (\bar{F}_P)_{P \in \bar{\mathcal{P}} \cup \bar{\mathcal{C}}}$ into a set $\bar{\mathcal{P}}$ of flow carrying s-t-paths and a set $\bar{\mathcal{C}}$ of flow carrying cycles. Now following the same lines of arguments as above, we observe that $\tau^r_{P_{k_1}} = \tau^r_{P'}$ holds for all $P' \in \bar{\mathcal{P}}$. Thus, from the second criterion defining P_{k_1} in Step (3), we get $|P_{k_1}| \leq |P'|$ for all $P' \in \bar{\mathcal{P}}$. Since flow is simultaneously sent along antiparallel edges, we know that the inequality (3.51) in Lemma 3.63 is satisfied strictly implying

$$F_{P_{k_1}}(\theta_1) \cdot |P_{k_1}| + F_{P_{k_2}}(\theta_2) \cdot |P_{k_2}| > \sum_{P' \in \bar{\mathcal{P}}} \left\|F_{P'}\right\|_\infty \cdot |P'| \geq |P_{k_1}| \cdot \sum_{P' \in \bar{\mathcal{P}}} \left\|F_{P'}\right\|_\infty \; .$$

Since we have $F_{P_{k_1}}(\theta_1) + F_{P_{k_2}}(\theta_2) = \sum_{P' \in \bar{\mathcal{P}}} \|F_{P'}\|_\infty$, this implies $|P_{k_2}| > |P_{k_1}|$ and (iii) is proven.

For establishing (iv), we first observe that each path $P \in \mathcal{P}^r$ occurs at most once in the family $(P_k)_{k \in K}$. For this assume the opposite and let $k_1, k_2 \in K$ be two indices with $k_1 < k_2$ and $P_{k_1} = P_{k_2}$. Then (i) and (ii) imply $\tau^r_{P_k} = \tau^r_{P_{k_1}}$ and $|P_k| = |P_{k_1}|$ for all $k \in \{k_1+1, \ldots, k_2-1\}$. Hence, (iii) shows that $f^k_{\overleftarrow{e}} - \tau^r_{\overleftarrow{e}}$ and $f^{k_2}_e$ are mutually singular for all $k \in \{k_1, \ldots, k_2\}$ and all edges $e \in E^r$.

Further, recalling Step (3) and (4) we know that $u^{k_1}_{P_{k_1}} = 0$. This shows

$$0 \;=\; u^{k_1+1}_{P_{k_1}} \;\geq\; u^{k_2}_{P_{k_1}} \;\geq\; 0$$

meaning that P_{k_2} is not able to carry flow which is a contradiction. Hence, each path $P \in \mathcal{P}^r$ occurs at most once in the family $(P_k)_{k \in K}$.

Since the number of paths in $\mathcal{P}^r$ having the same given transit time are finite, after a finite number of iterations the transit time must strictly increase. Further, because all transit times are rational by Assumption 5.4, it increases by some constant which is known a priori. Hence, (iv) holds. □

The following lemma is the reason why we are able to identify nice Nash equilibria in Section 5.3. Beside this, it is used in the subsequent lemma.

Lemma 5.10. *For all nodes $v \in V$, let $\tau^k_{v\to t} : \mathbb{R}_+ \to \mathbb{R}_+$ be a function defined as $\tau^k_{v\to t}(\theta) := \min\{\tau_P \mid P \in \mathcal{P}^r_{v\to t}, \theta \in \operatorname{supp}(u^k_P)\}$. That is, $\tau^k_{v\to t}(\theta)$ equals the minimum transit time over all v-t-paths P which are able to carry flow at time $\theta \in \mathbb{R}_+$ on the residual network $\mathcal{N}^k$. Then the sequence $(\tau^k_{v\to t})_{k\in[K+1]}$ is essentially nondecreasing for all $v \in V$, i.e., $\tau^k_{v\to t} \leq \tau^{k+1}_{v\to t}$ for all $k \in [K]$.*

Proof. Consider an iteration $k+1 \in [K+1]$ and let $v \in V$ be a node and $\theta \in \mathbb{R}_+$ be a point in time. Further, let P be a v-t-path defining $\tau^{k+1}_{v\to t}(\theta)$ for a set $\Theta \subseteq \mathbb{R}_+$ of points in time $\theta \in \mathbb{R}_+$ which is not a null set, i.e., $\lambda(\Theta) > 0$ and $\tau_P = \tau^{k+1}_{v\to t}(\theta)$ and $\theta \in \operatorname{supp}(u^{k+1}_P)$ for all $\theta \in \Theta$. In the following we assume without loss of generality that $\tau^k_{v\to t}(\theta)$ is constant over Θ. Otherwise we partition Θ in sets over which $\tau^k_{v\to t}(\theta)$ is constant. As the set of v-t-paths is countable, there is only a countable number of different values for $\tau^k_{v\to t}(\theta)$. Therefore, the partition of Θ is countable. Hence, proving the following for each subset of the partition results in a proof for the entire set Θ.

Defining $f_P := u^{k+1}_P|_\Theta$ and $\mathcal{F} := (P, F_P)$, we proceed similarly to the proof of 5.9(i). Choose $\theta_2 \in \mathbb{R}_+$ such that $F_P(\theta_2) > 0$ and set $\theta_1 := \theta_2 + \tau^-_{P_k} + \tau^+_P$. Let $\bar{\mathcal{F}} = (\bar{f}_e)_{e\in E^r}$ be the flow arising out of $((P_k, F_{P_k}|_{\leq\theta_1}), (P, F_P|_{\leq\theta_2}))$ after deleting flow on antiparallel edges. Since $F_{P_k}|_{\leq\theta_1} \leq F_{P_k}$ and $F_P|_{\leq\theta_2}$ satisfy the assumptions of Lemma 3.62 and 3.63, we know that $\bar{\mathcal{F}}$ is a feasible flow over time on $\mathcal{N}^k$.

Next we decompose $\bar{\mathcal{F}} := (\bar{F}_P)_{P\in\bar{\mathcal{P}}_{s\to t}\cup\bar{\mathcal{P}}_{v\to t}\cup\bar{\mathcal{C}}}$ into a set $\bar{\mathcal{P}}_{s\to t}$ of flow carrying s-t-paths, a set $\bar{\mathcal{P}}_{v\to t}$ of flow carrying v-t-paths, and a set $\bar{\mathcal{C}}$ of flow carrying cycles. Then Lemma 3.63 shows

$$F_{P_k}(\theta_1) = \sum_{P'\in\bar{\mathcal{P}}_{s\to t}} \left\|F_{P'}\right\|_\infty , \tag{5.12}$$

$$F_P(\theta_2) = \sum_{P'\in\bar{\mathcal{P}}_{v\to t}} \left\|F_{P'}\right\|_\infty , \qquad \text{and} \tag{5.13}$$

$$\begin{aligned} &F_{P_k}(\theta_1)\cdot\tau^r_{P_k} + F_P(\theta_2)\cdot\tau^r_P \\ &\quad = \sum_{P'\in\bar{\mathcal{P}}_{s\to t}} \left\|F_{P'}\right\|_\infty \cdot \tau^r_{P'} + \sum_{P'\in\bar{\mathcal{P}}_{v\to t}} \left\|F_{P'}\right\|_\infty \cdot \tau^r_{P'} + \sum_{C'\in\bar{\mathcal{C}}} \left\|F_{C'}\right\|_\infty \cdot \tau^r_{C'}. \end{aligned} \tag{5.14}$$

Since $\bar{\mathcal{F}}$ is feasible with respect to $\mathcal{U}^k$, we know from the construction of P_k in Step (3) that $\tau^r_{P_k} \leq \tau^r_{P'}$ for all $P' \in \bar{\mathcal{P}}_{s\to t}$. Furthermore, Lemma 5.8 tells us $0 \leq \tau^r_{C'}$ for all $C' \in \bar{\mathcal{C}}$. Hence, we obtain from (5.14) using (5.12)

$$F_P(\theta_2) \cdot \tau^r_P \;\geq\; \sum_{P' \in \bar{\mathcal{P}}_{v \to t}} \tau^r_{P'} \cdot \left\| F_{P'} \right\|_\infty . \tag{5.15}$$

Further, if $v \neq s$ we also know $F_P = \sum_{P' \in \bar{\mathcal{P}}_{v\to t}} F_{P'}$ as incoming flow is directly assigned to some outgoing edge of v. This shows that $\tau^k_{v\to t}(\theta) \leq \tau^r_{P'}$ holds for all $P' \in \bar{\mathcal{P}}_{v\to t}$ and $\theta \in \Theta$. Together with (5.13) and (5.15) this finally leads to $\tau^r_P \geq \tau^k_{v\to t}(\theta)$. ☐

In the following section, we also need the following Lemma which is quite similar to the previous one.

Lemma 5.11. *For all nodes $v \in V$ let $\tau^k_{s\to v} : \mathbb{R}_+ \to \mathbb{R}_+$ be a function defined as $\tau^k_{s\to v}(\theta) := \min\{\tau_P \mid P \in \mathcal{P}^r_{s\to v}, \theta \in \operatorname{supp}(u^k_P - \tau^r_P)\}$. That is, $\tau^k_{v\to t}(\theta)$ equals the minimum transit time over all v-t-paths P along which flow is able to arrive at v at time $\theta \in \mathbb{R}_+$ on the residual network $\mathcal{N}^k$. Then the sequence $(\tau^k_{s\to v})_{k\in[K+1]}$ is nondecreasing for all $v \in V$, i.e., $\tau^k_{s\to v} \leq \tau^{k+1}_{s\to v}$ for all $k \in [K]$.*

Proof. This is proven along the same lines as Lemma 5.10 and requires the same technique which is already used in the three previous lemmas. Therefore, we omit further details here. ☐

As already mentioned, our goal is to prove that the SSP algorithm can be interrupted in order to compute a direct earliest arrival flow for some given time horizon $T \in \mathbb{R}_+$. In particular, if $\mathcal{F} := (F_{P_k})_{k\in[K]}$ is the output of the SSP algorithm, $\tilde{\mathcal{F}} := (F_{P_k}|_{\leq T-\tau^r_{P_k}})_{k\in[K]}$ represents the flow behavior of the particles arriving at t not later than T. Unfortunately, the results so far are proven only for $\mathcal{F}$ instead of $\tilde{\mathcal{F}}$. The following lemma shows that all results remain true for $\tilde{\mathcal{F}}$. More precisely, we show that the sequence $\tilde{F}$ could be constructed by the SSP algorithm on the original network when restricting the capacity u_e of each edge $e \in \delta^-(t)$ to $[0, T]$.

Lemma 5.12. *Consider the residual network $\mathcal{N}^k := (G^r, \mathcal{U}^k, \mathcal{T}^r, s, t)$ with respect to the direct flow over time $\mathcal{F}^{<k} := (F_{P_{k'}})_{k'\in[k]}$. Let $T \in \mathbb{R}_+$ be a time horizon, and for each $j \in [k]$, let $\tilde{F}_{P_j} := F_{P_j}|_{[0,T-\tau^r_{P_j}]}$ be the flow on P_j which arrives at t before time T. Further, let $\tilde{\mathcal{N}}^k(G^r, \tilde{\mathcal{U}}^k, \mathcal{T}^r, s, t)$ be the residual network with respect to the flow $\tilde{\mathcal{F}}^{<k} := (\tilde{F}_{P_i})_{k'\in[k]}$. Then it holds for each s-t-path $P \in \mathcal{P}^r$*

$$u^k_P|_{[0,T-\tau^r_P]} = \tilde{u}^k_P|_{[0,T-\tau^r_P]} . \tag{5.16}$$

Proof. Let $P \in \mathcal{P}^r$ be an s-t-path of the residual graph G^r and $I := [0, T - \tau^r_P]$ be the time window over which flow must enter P in order to arrive at t before time T. Then such flow leaves the i-th edge of P during the time window $I_i := I + \tau^r_{P,i}$. For proving (5.16) we partition I depending on whether flow is able to enter P or not. That is, we have $I = S \cup \bar{S}$ where $S := \operatorname{supp}(u^k_P) \cap I$ and $\bar{S} := I \setminus S$. In the following we show $u^k_P|_S = \tilde{u}^k_P|_S$ and $u^k_P|_{\bar{S}} = \tilde{u}^k_P|_{\bar{S}}$.

But first we derive an equation which we need for both cases. For this let $e := e^P_i$ be the i-th edge of P for some $i \in \{1, \ldots |P|\}$ and v be the head of e.

Further, let S_i be the set of points in time at which flow is able to travel from v to t along the subpath $Q := \{e^P_{i+1}, \dots, e^P_{|P|}\}$ of P, i.e., $S_i := \operatorname{supp}(u^k_Q) \cap I_i$. Then as we subsequently see, the equation

$$u^k_e|_{S_i} = \tilde{u}^k_e|_{S_i} \tag{5.17}$$

is essentially valid. In order to prove (5.17) we have to consider all s-t-paths $P_{k'}$ with $k' \in [k]$ containing e or $\overleftarrow{e}$. First let $(P_{k'}, j)$ be such that the j-th edge of the s-t-path $P_{k'}$ equals e, i.e., $e^{P_{k'}}_j = e$. If no flow enters v along $P_{k'}$ over S_i, we obtain

$$\begin{aligned} 0 \;=\; f^{k'}_j|_{S_i} \;\geq\; \tilde{f}^{k'}_j|_{S_i} \geq 0 \\ \Rightarrow \qquad f^{k'}_j|_{S_i} \;=\; \tilde{f}^{k'}_j|_{S_i} \end{aligned} \tag{5.18}$$

Next we show that (5.18) also holds if flow enters v along $P_{k'}$ over S_i. In this case we know that there exists a θ in $S_i \cap (\operatorname{supp}(f_{P_{k'}}) + \tau^r_{P_{k'},j})$. Then from the definition of S_i, we obtain $\tau^r_P - \tau^r_{P,i} \;\geq\; \tau^k_{v\to t}(\theta)$, and from the definition of $P_{k'}$, we know $\tau^{k'}_{v\to t}(\theta) \;=\; \tau^r_{P_{k'}} - \tau^r_{P_{k'},j}$. Hence, Lemma 5.10 shows:

$$\tau^r_P - \tau^r_{P,i} \;\geq\; \tau^r_{P_{k'}} - \tau^r_{P_{k'},j}$$

Defining $I^{k'} := [0, T - \tau^r_{P_{k'}}]$ as the time window over which flow must enter $P_{k'}$ in order to arrive at t before time T this shows:

$$\begin{aligned} (I_i - \tau^r_{P_{k'},j}) \backslash I^{k'} &= [\tau^r_{P,i} - \tau^r_{P_{k'},j}, T - \tau^r_P + \tau^r_{P,i} - \tau^r_{P_{k'},j}] \backslash [0, T - \tau^r_{P_{k'}}] \\ &\subseteq [\tau^r_{P,i} - \tau^r_{P_{k'},j}, T - \tau^r_{P_{k'}}] \backslash [0, T - \tau^r_{P_{k'}}] \\ &\subseteq \mathbb{R}_- \ . \end{aligned}$$

Hence, we obtain:

$$\begin{aligned} 0 \;\geq\; f^{k'}_j|_{I_i} - \tilde{f}^{k'}_j|_{I_i} &= ((f_{P_{k'}} - f_{P_{k'}}|_{I^{k'}}) - \tau^r_{P_{k'},j})|_{I_i} \\ &= (f_{P_{k'}}|_{I_i - \tau^r_{P_{k'},j}} - f_{P_{k'}}|_{I^{k'} \cap (I_i - \tau^r_{P_{k'},j})}) - \tau^r_{P_{k'},j} \;=\; 0 \end{aligned}$$

Since $S_i - \tau^r_e \subseteq I_i$, equation (5.18) holds also in this case.

Next we consider a pair (P'_k, j) such that the j-th edge of $P_{k'}$ equals $\overleftarrow{e}$, i.e., $e^{P_{k'}}_j = \overleftarrow{e}$. Similarly to the proof of (5.18), we show

$$f^{k'}_j|_{S_i + \tau^r_{\overleftarrow{e}}} \;=\; \tilde{f}^{k'}_j|_{S_i + \tau^r_{\overleftarrow{e}}} \ . \tag{5.19}$$

For this, first observe that (5.19) holds in case no flow enters v along $P_{k'}$ over S_i because this implies $0 = f^{k'}_j|_{S_i+\tau^r_{\overleftarrow{e}}} \;\geq\; \tilde{f}^k_j|_{S_i+\tau^r_{\overleftarrow{e}}} \;\geq 0$. Hence, assume that there exists a θ in $S_i \cap \operatorname{supp}(f_{P_{k'}}) + \tau^r_{P_{k'},j-1}$. Then from the definition of S_i, we obtain $\tau^r_P - \tau^r_{P,i} \;\geq\; \tau^k_{v\to t}(\theta)$, and from the definition of $P_{k'}$ we know $\tau^{k'}_{v\to t}(\theta) \;=\; \tau^r_{P_{k'}} - \tau^r_{P_{k'},j-1}$. By Lemma 5.10 this implies:

$$\tau^r_P - \tau^r_{P,i} \;\geq\; \tau^r_{P_{k'}} - \tau^r_{P_{k'},j-1} \qquad \Rightarrow \qquad \tau^r_P - \tau^r_{P,i-1} \;\geq\; \tau^r_{P_{k'}} - \tau^r_{P_{k'},j}$$

Defining $I^{k'} := [0, T - \tau^r_{P_{k'}}]$ as the time window over which flow must enter $P_{k'}$ in order to arrive at t before time T this shows:

$$\begin{aligned}
&(I_{i-1} - \tau^r_{P_{k'},j}) \setminus I^{k'} \\
&\quad = [\tau^r_{P,i-1} - \tau^r_{P_{k'},j}, T - \tau^r_P + \tau^r_{P,i-1} - \tau^r_{P_{k'},j}] \setminus [0, T - \tau^r_{P_{k'}}] \\
&\quad \subseteq [\tau^r_{P,i-1} - \tau^r_{P_{k'},j}, T - \tau^r_{P_{k'}}] \setminus [0, T - \tau^r_{P_{k'}}] \\
&\quad \subseteq \mathbb{R}_- \ .
\end{aligned}$$

Hence, we obtain:

$$\begin{aligned}
0 \geq f^{k'}_j|_{I_{i-1}} - \tilde{f}^{k'}_j|_{I_{i-1}} &= ((f_{P_{k'}} - f_{P_{k'}}|_{I^{k'}}) - \tau^r_{P_{k'},j})|_{I_{i-1}} \\
&= (f_{P_{k'}}|_{I_{i-1} - \tau^r_{P_{k'},j}} - f_{P_{k'}}|_{I^{k'} \cap (I_{i-1} - \tau^r_{P_{k'},j})}) - \tau^r_{P_{k'},j} = 0
\end{aligned}$$

Since $S_i + \tau^r_{\overleftarrow{e}} \subseteq I_{i-1}$, equation (5.19) is proven.

With (5.18) and (5.19) we are prepared for proving (5.17). First observe that for all $k' \in [k]$ we have

$$f^{k'}_e|_{S_i} = \sum_{j | e^{P_{k'}}_j = e} f^{k'}_j|_{S_i} = \sum_{j | e^{P_{k'}}_j = e} \tilde{f}^{k'}_j|_{S_i} = \tilde{f}^{k'}_e|_{S_i}$$

$$\text{and} \quad f^{k'}_{\overleftarrow{e}}|_{S_i + \tau^r_{\overleftarrow{e}}} = \sum_{j | e^{P_{k'}}_j = \overleftarrow{e}} f^{k'}_j|_{S_i + \tau^r_{\overleftarrow{e}}} = \sum_{j | e^{P_{k'}}_j = \overleftarrow{e}} \tilde{f}^{k'}_j|_{S_i + \tau^r_{\overleftarrow{e}}} = \tilde{f}^{k'}_{\overleftarrow{e}}|_{S_i + \tau^r_{\overleftarrow{e}}}$$

because of (5.18) and (5.19), respectively. By iteratively applying (5.9) this shows (5.17) because

$$\begin{aligned}
u^k_e|_{S_i} &= u^1_e|_{S_i} - \sum_{k'=1}^{k} \left(f^{k'}_e|_{S_i} - \left(f^{k'}_{\overleftarrow{e}}|_{S_i + \tau^r_{\overleftarrow{e}}} + \tau^r_{\overleftarrow{e}} \right) \right) \\
&= u^1_e|_{S_i} - \sum_{k'=1}^{k} \left(\tilde{f}^{k'}_e|_{S_i} + \left(\tilde{f}^{k'}_{\overleftarrow{e}}|_{S_i + \tau^r_{\overleftarrow{e}}} + \tau^r_{\overleftarrow{e}} \right) \right) = \tilde{u}^k_e|_{S_i} \ .
\end{aligned}$$

For proving (5.16), we observe that $S + \tau^r_{P,i} \subseteq S_i$ holds for all $i \in \{1, \ldots, |P|\}$. Hence, because of (5.17) we obtain

$$u^k_P|_S = \tilde{u}^k_P|_S \ .$$

Finally, we show $u^k_P|_{\bar{S}} = 0 = \tilde{u}^k_P|_{\bar{S}}$. For this we partition $\bar{S} := \bigcup_{i=1}^{|P|} \bar{S}_i$ into sets $\bar{S}_i$ with $i \in \{1, \ldots, |P|\}$ ensuring $(u^k_{e^P_i} + \tau^r_{P,i})|_{\bar{S}_i} = 0$ and $(u^k_{e^P_{i'}} + \tau^r_{P,i'})|_{\bar{S}_i} > 0$ for all $i' > i$ essentially. That is, e^P_i is the last edge at which u^k_P attains its minimum over almost all points in time in $\bar{S}_i$. Hence, we know $\bar{S}_i + \tau^r_{P,i} \subseteq S_i$ for all $i \in \{1, \ldots, |P|\}$. Thus, (5.17) shows

$$\begin{aligned}
0 = u^k_P|_{\bar{S}} = \sum_{i=1}^{|P|} u^k_P|_{\bar{S}_i} &= \sum_{i=1}^{|P|} (u^k_{e^P_I}|_{\bar{S}_i + \tau^r_{P,i}} + \tau^r_{P,i}) \\
&= \sum_{i=1}^{|P|} (\tilde{u}^k_{e^P_I}|_{\bar{S}_i + \tau^r_{P,i}} + \tau^r_{P,i}) \geq \sum_{i=1}^{|P|} \tilde{u}^k_P|_{\bar{S}_i} = \tilde{u}^k_P|_{\bar{S}} \ .
\end{aligned}$$

Therefore, we are done if $\tilde{u}^k_P|_{\bar{S}} \geq 0$ which would be implied by the feasibility of $\tilde{\mathcal{F}}^{<k}$. By induction this follows from the validity of (5.16) for $k-1$ which shows the feasibility of $\tilde{F}_{P_{k-1}}$ on the residual network $\tilde{\mathcal{N}}^{k-1}$. $\square$

Lemma 5.12 is the last result which is needed for proving the correctness of the SSP algorithm.

Theorem 5.13. *The SSP algorithm works correctly. That is, for each time horizon $T \in \mathbb{R}_+$ it computes a direct earliest arrival flow on a given network $(G, \mathcal{U}, \mathcal{T}, s, t)$ in finite time. Furthermore, this earliest arrival flow is given by the general path-based flow over time $(f_{P_k}|_{[0,T-\tau^r_{P_k}]})_{k\in[K]}$.*

Proof. Let $K \in \mathbb{N}$ be the smallest iteration for which $\tau^r_{P_K} > T$ holds or at which the SSP algorithm terminates. Clearly, K exists and is finite by Lemma 5.9(iv). Next we show that the flow $\mathcal{F}^* := (f^*_{P_k})_{k\in[K]}$ given by $f^*_{P_k} := f_{P_k}|_{[0,T-\tau_P]}$ for all $k \in [K]$ is an earliest arrival flow for the time horizon T.

By Lemma 5.12 we know that the sequence $(f^*_{P_k})_{k\in[K]}$ is an output of the SSP algorithm on the network $\tilde{\mathcal{N}} := (G, \tilde{\mathcal{U}}, \mathcal{T}, s, t)$ where $\tilde{\mathcal{U}} := (\tilde{u}_e)_{e\in E}$ is given by $\tilde{u}_e := u_e|_{[0,T]}$ for all $e \in E$.

Consider an arbitrary feasible flow over time on $\mathcal{F} := (f_e)_{P\in\mathcal{P}}$ on $\tilde{\mathcal{N}}$. In the following we show $\mathrm{val}(\mathcal{F}^*) > \mathrm{val}(\mathcal{F})$ which establishes this theorem. For this let $(f^*_e)_{e\in E}$ be the edge-based representation of $\mathcal{F}^*$. Because of Lemma 3.60 the flow $\mathcal{F}^r := (f_e)_{e\in E^r}$ defined by

$$f^r_e := \max\{0, \tilde{f}_e - f_e\} \qquad \text{and} \qquad f^r_{\overleftarrow{e}} := \max\{0, f_e - \tilde{f}_e\} + \tau_e \qquad \forall e \in E$$

is feasible on the residual network $\tilde{\mathcal{N}}^K$ with value $\mathrm{val}(\mathcal{F}^r) = \mathrm{val}(\mathcal{F}) - \mathrm{val}(\mathcal{F}^*)$. Further, by Lemma 5.12 each flow carrying s-t-path in $\tilde{\mathcal{N}}^K$ must have a transit time strictly greater than T. Hence, there exists no flow carrying s-t-path by the definition of $\tilde{\mathcal{N}}$ implying $\mathrm{val}(\mathcal{F}^r) \leq 0$. This shows $\mathrm{val}(\mathcal{F}^*) > \mathrm{val}(\mathcal{F})$ proving that $\mathcal{F}^*$ is an earliest arrival flow for the time horizon T. $\square$

5.3 Characterization of Nash Equilibria

In this section we discuss Nash equilibria for direct flows over time where waiting is not allowed at intermediate nodes. We start with an example verifying that neither lower nor upper arrival times show the expected selfish routing behavior of flow particles and, therefore, motivating the usage of foresighted arrival times. Subsequently, we present a characterization of Nash equilibria for the direct flow model. We see that there may exists several Nash equilibria for a given direct routing games over time. Finally, based on an earliest arrival flow, we identify Nash equilibria showing a good performance with respect to the price of anarchy which are further analyzed in Section 5.4. In particular, these Nash equilibria allow to compute the prices of stability by only considering networks consisting of parallel s-t-edges.

The following simple example provides a first intuition behind direct Nash flows over time. In particular, it motivates the consideration of foresighted arrival times. Further, note that this example is one-to-one transferable to the quite popular deterministic queuing model considered in Section 7.

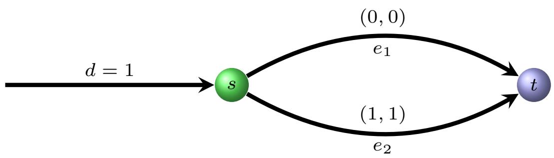


Figure 5.3: Network of Example 5.14 showing the necessity of foresighted arrival time functions.

Example 5.14. Consider the network in Figure 5.3 consisting only of two parallel s-t-edges e_1 and e_2. The transit time of e_1 is 0 and its capacity vanishes for all points in time. The transit time of e_2 equals 1 and its capacity function is constantly equal to 1. Further, the demand rate is also constantly equal to 1. Thus, the only feasible direct flow over time $\mathcal{F}$ sends the entire flow over e_2, i.e., $\mathcal{F} := ((0, e_1), (\mathrm{id}, e_2))$.

Intuitively, $\mathcal{F}$ is of course a Nash equilibrium as no flow particle would switch to e_1. Computing the lower arrival time functions of the two s-t-paths, we observe

$$\check{\ell}_{(e_1)} \;=\; \mathrm{id} \qquad\qquad \text{and} \qquad\qquad \check{\ell}_{(e_1)} \;=\; \mathrm{id} + 1\ .$$

Thus, for all departure times, the current lower arrival time of e_1 is strictly smaller than the one of e_2. This would imply that $\mathcal{F}$ is no Nash equilibrium in terms of Definition 4.8. On the other hand, the foresighted arrival time functions are given by

$$\bar{\ell}_{(e_1)} \;=\; \infty \qquad\qquad \text{and} \qquad\qquad \bar{\ell}_{(e_1)} \;=\; \mathrm{id} + 1$$

implying that $\mathcal{F}$ is Nash equilibrium in terms of Definition 4.8.

In order to see that also upper arrival times contradict our intuition behind Nash equilibria increase the capacity of e_1 to 1. In this scenario, a Nash flow sends all flow particles along e_1 implying that $\mathcal{F}$ is no longer a Nash equilibrium. But considering the upper arrival time functions of $\mathcal{F}$ which are given by

$$\hat{\ell}_{(e_1)} \;=\; \infty \qquad\qquad \text{and} \qquad\qquad \hat{\ell}_{(e_1)} \;=\; \mathrm{id} + 1,$$

the flow over time $\mathcal{F}$ satisfies Definition 4.8 and would become a Nash flow. On the other hand, the corresponding foresighted arrival time functions coincide with our intuition. They evaluate to

$$\bar{\ell}_{(e_1)} \;=\; \mathrm{id} \qquad\qquad \text{and} \qquad\qquad \bar{\ell}_{(e_1)} \;=\; \mathrm{id} + 1$$

showing that $\mathcal{F}$ is no Nash equilibrium.

In the following we characterize direct Nash flows over time in terms of entering flows. This characterization is based on the following insight. Consider a direct Nash flow and a flow particle departing at a certain time $\theta \geq 0$ which is sent along some s-t-path P. Let v_i be the i-th node of P. Since in a Nash equilibrium this flow particle arrives at t at time $\ell_{\mathrm{N}}(\theta)$, it enters v_i at time $\ell^i_N(\theta) := \ell_N(\theta) - \tau_P + \tau_{P,i}$. Note that $\tau_P - \tau_{P,i}$ is the time which a flow particle has to take in order to travel from v_i to t along P. Then each v_i-t-path P' with a transit time smaller than $\tau_P - \tau_{P,i}$ must be used up to its capacity at time $\ell^i_{\mathrm{N}}(\theta)$. Otherwise this flow particle would switch to Q after entering v_i in order to arrive at t earlier. On the other hand, assume that the remaining transit time $\tau_P - \tau_{P,i}$ of P is maximal among all flow carrying s-t-paths containing v_i. Then each s-t-path P' containing v_i for which the remaining transit time from v_i to t is larger than $\tau_P - \tau_{P,i}$ carries no flow. These observations lead to the following characterization of direct Nash flows over time.

Lemma 5.15. *Let $\mathcal{F} := (F_P)_{P\in\mathcal{P}}$ be a family of absolutely continuous, nondecreasing functions representing some entering flow which satisfy the capacity constraint* (5.2)*, and let D be a supply function with $D \geq \sum_{P\in\mathcal{P}} F_P$. Moreover, for each node $v \in V \setminus \{t\}$ there exists a function $\tau_v : \mathbb{R}_+ \to \mathbb{R}_+$ such that:*

(i) For all points in time $\theta \in \mathbb{R}_+$ it holds $D(\theta) = \mathrm{val}(\mathcal{F})\big(\theta + \tau_s(\theta)\big)$.

(ii) For almost all points in time $\theta \in \mathbb{R}_+$ and all s-t-paths P containing v at some position i, i.e., $v = \mathrm{head}(e^P_i)$ it holds

$$\tau_P - \tau_{P,i} \ > \ \tau_v(\theta) \qquad\qquad \Rightarrow \qquad\qquad f_P(\theta - \tau_{P,i}) \ = \ 0 \ .$$

(iii) For almost all times $\theta \geq 0$ and all v-t-paths P with $\tau_P < \tau_v(\theta)$ it holds

$$\exists\, i : \sum_{P',j | e^{P'}_j = e^P_i} f_{P'}(\theta + \tau_{P,i} - \tau_{P',j}) = u_e(\theta + \tau_{P,i}) \ .$$

Then there exists a family $\mathcal{F}^+ := (F^+_P)_{P\in\mathcal{P}}$ of inflow functions such that $(\mathcal{F}^+, \mathcal{F})$ is a direct Nash flow over time with minimum transit time function $\tau_s = \tau_{\mathrm{N}}$.

Proof. In order to obtain a direct Nash flow over time $\mathcal{F}$, we have to define the family $\mathcal{F}^+ := (F^+_P)_{P\in\mathcal{P}}$ of inflow functions appropriately. For this let P be an s-t-path and $\theta \in \mathbb{R}_+$ be a point in time. Since τ_s should become the minimum transit time function, P could potentially carry flow if $\tau_P \leq \tau_s(\theta)$. In this case the current arrival time should be equal to $\theta + \tau_s(\theta)$ implying a current waiting time of $\tau_s(\theta) - \tau_P$ in front of P. On the other hand, P must carry no flow if $\tau_P > \tau_s(\theta)$ which would also imply that no flow waits in front of P at time θ. Hence, we define the current waiting time $q_P(\theta)$ in front of P at time θ by $q_P(\theta) := \max\{0, \tau_s(\theta) - \tau_P\}$. As $\ell_P(\theta) := \theta + q_P(\theta) + \tau_P$ should be a feasible arrival time function for the resulting direct flow over time, we define the cumulative inflow function of P by

$$F^+_P(\theta) := F_P\big(\theta + q_P(\theta)\big) \tag{5.20}$$

ensuring that FiFo is directly satisfied. Next we show that $\mathcal{F}^+ := (F^+_P)_{P\in\mathcal{P}}$ defined in this manner becomes valid strategy profile. For this let P be an s-t-path with $\tau_P > \tau_s(\theta)$ for some $\theta \in \mathbb{R}_+$. Since τ_s is nondecreasing because of (i), we know that $f_P(\theta') = 0$ for almost all $\theta' \in [0, \theta]$ because of (ii) applied to s. Thus we know $F_P(\theta) = 0$. This shows, on the one hand, $F_P(\theta + \tau_s(\theta) - \tau_P) = 0$ as $\theta + \tau_s(\theta) - \tau_P < \theta$ and, on the other hand, $F^+_P(\theta) = F_P(\theta) = 0$ because of (5.20) implying

$$\begin{aligned} D(\theta) \ &= \ \mathrm{val}(\mathcal{F})(\theta + \tau_s(\theta)) \ = \ \sum_{P\in\mathcal{P}} F_P(\theta + \tau_s(\theta) - \tau_P) \\ &= \sum_{P|\tau_P \leq \tau_s(\theta)} F_P(\theta + \tau_s(\theta) - \tau_P) \ = \sum_{P|\tau_P \leq \tau_s(\theta)} F^+_P(\theta) \ = \ \sum_{P\in\mathcal{P}} F^+_P(\theta) \ . \end{aligned}$$

Hence, the family $\mathcal{F}^+$ of inflow functions is a feasible strategy profile for the direct routing game over time $(G, \mathcal{U}, \mathcal{T}, s, t, D)$.

Next, we prove that $(\mathcal{F}^+, \mathcal{F})$ is a feasible direct flow over time in terms of Definition 5.2. Recalling Definition 5.1 let $\mathcal{F}^- := (F_P - \tau_P)_{P\in\mathcal{P}}$ be the corresponding family outflow functions. Then $(\mathcal{F}^+, \mathcal{F}^-)$ is a feasible flow over time for the family $\mathcal{L} := (\ell_P)_{P\in\mathcal{P}}$ of arrival time functions $\ell_P : \theta \mapsto \theta + q_P(\theta) + \tau_P$ by definition. Therefore, we have to show that the arrival time functions ℓ_P and $\bar{\ell}_P$ are equivalent with respect to F_P^+ for each s-t-path P.

Because of (i) we know that $\mathrm{id}+\tau_s$ and, hence, all ℓ_P's are nondecreasing implying that $(\mathcal{F}^+, \mathcal{F}^-)$ satisfies FiFo. Because of Theorem 3.18 and Lemma 3.19 it is enough to show $F_P^+(\theta) = F_P(\theta+\bar{q}_P(\theta))$ for all s-t-paths $P \in \mathcal{P}$ and all points in time $\theta \in \mathbb{R}_+$. Consider an s-t-path P and a point in time $\theta \in \mathbb{R}_+$. Since $\mathcal{F}$ satisfies the capacity constraints, we know for each $\theta' \in \mathbb{R}_+$ with $F_P(\theta') > F_P^+(\theta)$

$$\int_\theta^{\theta'} u_P|_{\leq\theta}(\vartheta)\, d\vartheta \;\geq\; F_P(\theta') - F_P^+(\theta) \;>\; 0$$

implying $\theta + \bar{q}_P(\theta) \leq \theta'$ by the definition of $\bar{q}_P$ in (5.6). Because F_P is nondecreasing this shows $F_P^+(\theta) \geq F_P(\theta + \bar{q}_P(\theta))$.

Hence, it remains to prove $F_P^+(\theta) \leq F_P(\theta + \bar{q}_P(\theta))$ for all $P \in \mathcal{P}$ and all $\theta \in \mathbb{R}_+$. For proving this we need the following observation. Let P be an arbitrary s-t-path and $\theta \in \mathbb{R}_+$ be a point in time such that $\tau_P < \tau_s(\theta)$. Thus, condition (iii) is applicable for s meaning that for almost all such θ's at least one edge of P is used up to its capacity. Let i_θ be the position of the last saturated edge, i.e.,

$$i_\theta := \max\Big\{ i \;\Big|\; \sum_{P', j | e_j^{P'} = e_i^P} f_{P'}(\theta + \tau_{P,i} - \tau_{P',j}) = u_e(\theta + \tau_{P,e}) \Big\}\ , \tag{5.21}$$

and let v be the head of $e_{i_\theta}^P$. Further, let P_θ be the subpath of P from v to t, i.e., $P_\theta := (e_{i_\theta+1}^P, \ldots, e_{|P|}^P)$. Then no edge of P_θ is used up to its capacity implying $\tau_P - \tau_{P,i_\theta} = \tau_{P_\theta} \geq \tau_v(\theta)$ because of (iii). Next consider a pair (P', j) with $e_j^{P'} = e_{i_\theta}^P$ such that flow leaves $e_{i_\theta}^P$ at time $\theta + \tau_{P,i_\theta}$ via P'. That is, we have $f_{P'}(\theta + \tau_{P,i_\theta} - \tau_{P',j}) > 0$ implying essentially $\tau_v(\theta) \geq \tau_{P'} - \tau_{P,j}$ because of (ii). This shows

$$\tau_P - \tau_{P,i_\theta} \geq \tau_{P'} - \tau_{P,j}\ . \tag{5.22}$$

Now fix an s-t-path P and a point in time $\theta \in \mathbb{R}_+$. Because of (5.20) it is enough to show $\bar{q}_P(\theta) \geq q_P(\theta)$ as F_P is nondecreasing. If $q_P(\theta) = 0$ nothing has to be shown because $\bar{q}_P$ is nonnegative by definition. Hence, assume $q_P(\theta) > 0$. Since the function $\mathrm{id} + q_P$ do not decrease, we obtain $\vartheta < \theta + q_P(\theta) \leq \vartheta + q_P(\vartheta)$ for all $\vartheta \in [\theta, \theta + q_P(\theta)) =: I$. This shows $q_P(\vartheta) > 0$ implying $\tau_P < \tau_s(\vartheta)$ for all $\vartheta \in I$. From the discussion above we know that i_ϑ defined by (5.21) exists for almost all $\vartheta \in I$. Further, let (P', j) be a pair with $e_j^{P'} = e_{i_\vartheta}^P$ such that flow leaves $e_{i_\vartheta}^P$ at time $\theta + \tau_{P,i_\vartheta}$, i.e., $f_{P'}(\vartheta + \tau_{P,i_\vartheta} - \tau_{P',j}) > 0$. Since (5.22) holds, we obtain

$$\begin{aligned} \vartheta + \tau_{P,i_\vartheta} + \tau_{P'} - \tau_{P,j} \;&\leq\; \vartheta + \tau_P \;\leq\; \theta + q_P(\theta) + \tau_P \\ &=\; \theta + \tau_s(\theta) \;=\; \theta + q_{P'}(\theta) + \tau_{P'} \end{aligned}$$

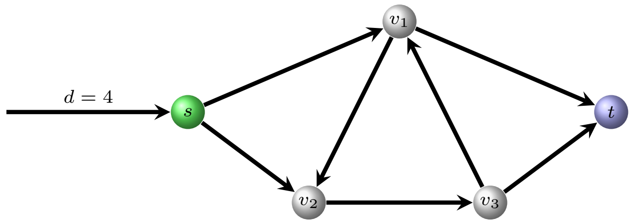


Figure 5.4: Network of Example 5.16 where each edge has a transit time of 0 and a constant capacity rate of 1.

implying $\vartheta + \tau_{P,i_\vartheta} - \tau_{P',j} \leq \theta + q_{P'}(\theta)$ for almost all $\vartheta \in I$. In particular, this shows $F_{P'}|_{\leq\theta}(\vartheta + \tau_{P,i_\vartheta} - \tau_{P',j}) = F_{P'}(\vartheta + \tau_{P,i_\vartheta} - \tau_{P',j})$. Since condition (iii) applied to s and P is essentially satisfied by i_ϑ for all $\vartheta \in I$, we obtain

$$\begin{aligned} 0 \;&\leq\; \int_\theta^{\theta+q_P(\theta)} u_P|_{\leq\theta}(\vartheta)\, d\vartheta \\ &\leq\; \int_\theta^{\theta+q_P(\theta)} u_{e_{i_\vartheta}}|_{\leq\theta}(\vartheta + \tau_{P,i_\vartheta})\, d\vartheta \;=\; \int_\theta^{\theta+q_P(\theta)} u_{e_{i_\vartheta}}(\vartheta + \tau_{P,i_\vartheta})\, d\vartheta \;=\; 0\,. \end{aligned}$$

By the definition of $\bar{q}_P$ in (5.6), this shows $\bar{q}_P(\theta) \geq q_P(\theta)$. Hence, $(\mathcal{F}^+, \mathcal{F}^-)$ is a feasible direct flow over time.

Finally, we observe that $(\mathcal{F}^+, \mathcal{F})$ is a Nash flow over time in terms of Definition 4.8. Since $\mathcal{L}$ and $\bar{\mathcal{L}}$ are equivalent, (i) shows $D(\theta) = \mathrm{val}(\mathcal{F}^+)(\bar{\ell}_{\mathrm{N}}(\theta))$ where we define $\bar{\ell}_{\mathrm{N}} := \min\{\bar{\ell}_P \mid P \in \mathcal{P}\}$. As $\bar{\mathcal{L}}$ is the family of foresighted arrival time functions of $(\mathcal{F}^+, \mathcal{F}^-)$, Theorem 4.13 shows that $(\mathcal{F}^+, \mathcal{F})$ is a direct Nash flow over time as statement (iii) of Theorem 4.13 holds. □

Lemma 5.15 provides a characterization all Nash equilibria of some direct routing game over time has to satisfy. Before we construct a nice Nash equilibrium, we consider to examples explaining the functionality of Lemma 5.15. In addition, the first example shows that Nash equilibria need not be unique and the second shows that the direct price of anarchy is unbounded for each objective function discussed in Section 4.3.

Example 5.16. This example works with the network shown in Figure 5.4 where each edge has a transit time of 0 and the capacity function which is constantly equal to 1. Further, the demand rate of this direct routing game over time is set to the constant function 4. Consider the two s-t-paths $P_1 := sv_1v_2v_3t$ and $P_2 := sv_2v_3v_1t$ and the two flows over time $\mathcal{F}_1^+ := (4\mathrm{id}, P_1)$ and $\mathcal{F}_2^+ := (4\mathrm{id}, P_2)$ which send the entire flow along P_1 and P_2, respectively. Corresponding entering flows are given by $\mathcal{F}_1 := (\mathrm{id}, P_1)$ and $\mathcal{F}_2 := (\mathrm{id}, P_2)$, respectively, which obviously obey the capacity constraints. In fact, both flows are Nash equilibria. We examplarily verify this for $\mathcal{F}_1^+$ using Lemma 5.15.

Since τ_s must be equal to the minimum transit time function, we have to define $\tau_s := \tau_{\mathrm{N}} = 3\mathrm{id}$. Note that $\ell_{\mathrm{N}} = 4\mathrm{id}$ because the cumulative outflow function equals the identity function. For each other node, $\tau_v(\theta)$ equals the maximum remaining transit time $\tau_P - \tau_{P,i}$ of all s-t-paths P traversing v. Hence, for every node $v \in V \setminus \{s,t\}$, we have $\tau = 0$ as each v-t-path has a transit time of 0. Thus, we only have to check condition (iii) of Lemma 5.15 for s which says that there is no free capacity on any s-t-path. Since each s-t-path contains at least

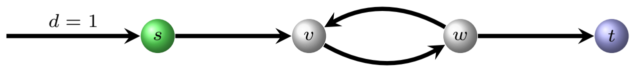


Figure 5.5: Network of Example 5.17 where each edge has a transit time of 0 and a constant capacity rate of 1.

one edge of P_1 and each edge of P_1 is used up to its capacity from time 0 on, statement (iii) holds. This shows that $\mathcal{F}_1$ is a Nash flow. Along the same lines we can use Lemma 5.15 to show that $\mathcal{F}_2$ is a Nash flow as well.

The next example is a continuation of Example 3.35 and shows that each kind of the prices of anarchy is unbounded for the direct flow model.

Example 5.17. Consider the network shown in Figure 5.5 where each edge has a constant transit time of 0 and a constant capacity of 1. Further, let the demand rate function d be constantly equal to 1. For each $i \in \mathbb{N}_0$ let P_i be the s-t path which contains i times the cycle C induced by v and w. Clearly, the social optimum sends the entire flow along the path $P_0 = svwt$. This implies a flow pattern which is equal to the identity function.

Next we consider the sequence $(\mathcal{F}_i^+)_{i\in\mathbb{N}}$ of direct flows over time where the entire flow is sent along path P_i. Paying attention to the capacity constraint on the edge vw, we see that $(\frac{1}{i+1}\mathrm{id}, P_i)$ is the entering flow of $\mathcal{F}_i^+$. As the edge vw is always used up to its capacity, it is not hard to observe that each $\mathcal{F}_i^+$ is a direct Nash flow over time. Further, the corresponding flow pattern is given by $\frac{1}{i+1}\mathrm{id}$. This shows that the evacuation and working price of anarchy equals $i + 1$. In addition, the completion and average arrival time price of anarchy is equal to $\frac{1}{i+1}$. Thus, each direct price of anarchy is unbounded.

Example 5.17 shows that each kind of the price of anarchy of direct routing games over time is unbounded. Further, Example 5.16 shows that the set of Nash equilibria for a given direct routing game over time is not unique. This motivates searching for a direct Nash flow over time, which shows a better performance than worst case Nash equilibria. In the last part of this section, we identify such nice Nash equilibria. Further, the corresponding approach enables us to evaluate each price of stability by considering only networks consisting of parallel s-t-edges.

Nice Nash flows over time are constructed from earliest arrival flows computed via the SSP algorithm. Especially, their general path-based representation is of importance. Using this representation, we construct a network consisting only of parallel s-t-edges – one for each flow carrying path of the residual network. The capacity function and the transit time of an edge is set to the flow and to the transit time of the corresponding s-t-path, respectively. On this new simple network, we compute a Nash flow over time. Finally, considering this Nash flow, we send the flow of an edge along the corresponding path of the residual network. It turns out that this general path-based flow is a Nash flow over time on the original network. More formally, the definition of such so called EA-Nash flows works as follows.

Definition 5.18 (EA-Nash flow)**.** Let $\mathcal{N} := (G,\mathcal{U},\mathcal{T}, s, t, D)$ be a nonatomic direct routing game over time and let $\mathcal{F}_{\mathrm{O}} := (F^{\mathrm{O}}_{P_k})_{k\in[K]}$ be the general path-based representation of an earliest arrival flow computed by the SSP algorithm on $(G,\mathcal{U},\mathcal{T}, s, t)$, where $K \in \mathbb{N}^\infty$ equals the number of needed iterations.

Consider the network $\mathcal{N}_\ominus$ which consists of the source s, the sink t, and, for all $k \in [K]$, an s-t-edge e_k representing s-t-path P_k. Let $\tau_k := \tau^r_{P_k}$ be the transit time and $f^{\mathrm{O}}_{P_k} - \tau_k$ be the capacity of edge e_k. The demand of $\mathcal{N}_\ominus$ is set to the original demand D. Further, let $\mathcal{F}^+_\ominus := (F^+_{(e_k)})_{k \in [K]}$ be the path-based Nash flow on $\mathcal{N}_\ominus$ and $\mathcal{F}_\ominus := (F_{(e_k)})_{k \in [K]}$ be the corresponding path-based entering flow. The term *the* Nash flow refers to the fact that, in case of $\tau^r_{P_{k'}} = \tau^r_{P_k}$, we preferentially send flow into a path $P_{k'}$ if $k' < k$.

Interpret $\mathcal{F}_\ominus$ as a general path-based flow over time $((P_k, F_{e_k}))_{k \in [K]}$ on $\mathcal{N}$ and decompose it into a path-based flow over time $\mathcal{F} := (F_P)_{P \in \mathcal{P}}$. Then the *EA-Nash flow over time* $\mathcal{F}^+ := (F^+_P)$ is given by $\mathcal{F}$ representing the entering flow of $\mathcal{F}^+$. That is, F^+_P is defined by $F^+_P(\theta) := F_P(\max\{\theta, \ell_\ominus(\theta) - \tau_P\})$ where $\ell_\ominus$ is the minimum arrival time function of the Nash flow $\mathcal{F}^+_\ominus$.

The following lemma shows that the EA-Nash flow is really a direct Nash flow over time for a given routing game over time.

Theorem 5.19. *Let $(G, \mathcal{U}, \mathcal{T}, s, t, D)$ be a nonatomic direct routing game over time. Then the EA-Nash flow is a direct Nash flow over time.*

Proof. The outline of this proof is as follows. Firstly, we show that the entering flow $\mathcal{F}$ of the EA-Nash flow over time is feasible, i.e., is nonnegative and obeys the capacity constraints. In the same step, we show that $\mathcal{F}$ also satisfies the assumptions of Lemma 5.15. Thus, the EA-Nash flow is really a Nash flow over time if the corresponding inflow functions represent a feasible strategy profile of the underlying routing game over time. This is shown in the last part of this proof.

For the first part we need the following observation. Assume that $\mathcal{F}_\ominus$ sends flow into an edge e_k at some point in time θ. Then $\mathcal{F}_\ominus$ sends flow into an edge $e_{k'}$ at the maximum possible rate for all departure times $\theta' \geq \theta$ provided that this flow arrives at t not later than the flow on e_k. More precisely, if $\theta' \geq \theta$ and $\theta' + \tau_{k'} \leq \theta + \tau_k$ hold for some $k' < k$ then $\mathcal{F}_\ominus$ sends flow into e'_k at the maximum possible rate at time θ'.

Assume the opposite implying that $e_{k'}$ is not used up to its capacity at time θ' by $\mathcal{F}_\ominus$. Further, we know that $\mathcal{F}_\ominus$ sends now flow into e_k at time θ. Consequently, statement (iii) and (ii) of Lemma 5.15 applied to the Nash flow $\mathcal{F}_\ominus$ for the node s show

$$\tau_\ominus(\theta') \;\leq\; \tau_{k'} \qquad\qquad \text{and} \qquad\qquad \tau_k \;\leq\; \tau_\ominus(\theta)\;,$$

respectively. Note that we use the converses of statement (iii) and (ii). Since we know $\theta' + \tau_{k'} \leq \theta + \tau_k$ by assumption, this shows

$$\ell_\ominus(\theta') \;\leq\; \theta' + \tau_{k'} \;\leq\; \theta + \tau_k \;\leq\; \ell_\ominus(\theta)\;.$$

Because $\ell_\ominus$ is a nondecreasing function and we assume $\theta' \geq \theta$, equality holds at every "$\leq$"-sign. This shows, $\theta' + \tau_{k'} = \theta + \tau_k$ and $\ell_\ominus(\theta') = \ell_\ominus(\theta)$. If $\theta < \theta'$ holds, D must be constant over $[\theta, \theta')$ by Theorem 4.13(iii) contradicting that flow is sent along P_k at time $\theta - \tau_{P_k,i}$. On the other hand, if $\theta = \theta'$ holds, we obtain $\tau_k = \tau'_k$. This contradicts Definition 5.18 which says that flow is sent along $e_{k'}$ at first. Hence, $\mathcal{F}_\ominus$ sends flow into $e_{k'}$ at the maximal possible rate at time θ'.

Next, we use this result to prove that $\mathcal{F}$ is nonnegative, obeys the capacity constraints, and satisfies the assumption of Lemma 5.15. For this we show that $\mathcal{F}$ can be seen as an earliest arrival flow on the following network $\mathcal{N}'$. Add a node s_0 and an edge s_0s to the original network. The transit time function of this new edge is constantly equal to 0 and the capacity is set to the entire inflow of $\mathcal{F}$, i.e., $u_{s_0s} = \sum_{k\in K} f_{e_k}$. Then extending each path flow of $\mathcal{F}$ to the new edge s_0 results in an earliest arrival flow. We prove that $((P_k, F_{e_k}))_{k\in K}$ represents a possible outflow sequence of the SSP algorithm applied to $\mathcal{N}'$. This would imply directly the feasibility of $\mathcal{F}$ with respect to the nonnegativity and the capacity constraints.

In order to see that the assumptions of Lemma 5.15 would be also satisfied, we set $\tau_s := \tau_\ominus$ to the minimum transit time of the Nash flow over time $\mathcal{F}_\ominus$ ensuring the validity of statement (i). Thus, if the assumptions are not satisfied, there exists a node v, a flow carrying v-t-path P_1, and a v-t-path P_2 with free capacity such that $\tau_{P_1} > \tau_{P_2}$. Hence, the cycle which traverses P_1 backwards and P_2 forwards has a positive residual capacity and a negative transit time contradicting the negative cycle condition of an earliest arrival flow. So if $\mathcal{F}$ represents an earliest arrival flow on $\mathcal{N}'$, it satisfies the assumption of Lemma 5.15.

To show that $((P_k, F_{e_k}))_{k\in K}$ represents a possible outflow sequence of the SSP algorithm applied to $\mathcal{N}'$, we consider a $k \in [K]$ and a point in time $\theta \in \mathbb{R}_+$ at which $\mathcal{F}_\ominus$ sends flow into the edge e_k. In the following, we verify that the SSP algorithm may send flow at a rate of $f_{e_k}(\theta)$ into P_k at time θ. First observe that, due to the definition of $\mathcal{N}'$, there is enough capacity on the new edge s_0s. Therefore, it is enough to show that the current residual capacities of the edges of P_k coincide with residual capacities of $(F^{\mathrm{O}}_{P_{k'}})_{k'<k}$ with respect to the original network $\mathcal{N}$.

So consider a node $v \in V$ which is visited by P_k at the i-th position meaning that $\mathcal{F}$ sends flow through v at time $\theta_v := \theta + \tau^r_{P_k,i}$. Further, let $P_{k'}$ with $k' < k$ be another s-t-path along which $\mathcal{F}_{\mathrm{O}}$ sends flow through v at time θ_v. Let j be a position at which $P_{k'}$ traverses node $v = v_j^{P_{k'}}$. We have to show that $\mathcal{F}$ sends flow into P_k at time $\theta - \tau^r_{P_{k'},j}$ at the maximal possible rate. Because of Lemma 5.10 and Lemma 5.11, we know $\tau^r_{P_{k'}} - \tau^r_{P_{k'},j} \le \tau^r_{P_k} - \tau^r_{P_k,i}$ and $\tau^r_{P_{k'},j} \le \tau^r_{P_k,i}$, respectively. This shows that $\mathcal{F}_{\mathrm{O}}$ sends flow into $P_{k'}$ at time $\theta_v - \tau^r_{P_{k'},j} \ge \theta$ which arrives at t at time $\theta_v + \tau^r_{P_{k'}} - \tau^r_{P_{k'},j} \ge \theta + \tau^r_{P_k}$. As shown above, this implies that $\mathcal{F}$ sends flow into P_k at time $\theta_v - \tau^r_{P_{k'},j}$ at the maximum possible rate. Hence, the residual capacity of each edge $e \in \delta^-_{G^r}(v)$ at time θ_v after iteration $k-1$ coincides with respect to $\mathcal{F}$ and $\mathcal{F}_O$. This shows that F_{e_k} can be sent feasible along P_k within the residual network of $\mathcal{N}'$ and $((P_{k'}, F_{e_{k'}}))_{k'<k}$.

It remains to show that P_k is a *shortest* path which is able to carry flow in the residual network of $\mathcal{N}'$ with respect to $((P_{k'}, F_{e_{k'}}))_{k'<k}$. For this we show that each v-t-path starting at the i-th node v of P_k has no free capacity at time $\theta_v := \theta + \tau^r_{P_k,i}$ if $\tau^r_Q < \tau^r_{P_k} - \tau^r_{P_k,i}$. Note that $\mathcal{F}_{\mathrm{O}}$ shows this behavior.

So let v_1v_2 be the first edge having a residual capacity of 0 with respect to $\mathcal{F}_{\mathrm{O}}$. To observe that the capacity of this edge also vanishes with respect to $\mathcal{F}$, we consider the following procedure with respect to $\mathcal{F}_{\mathrm{O}}$. Let $k_1 < k$ be as large as possible such that P_{k_1} sends flow along an intermediate node of the v-v_2-subpath of Q. Further, let i_1 and j_1 be the smallest and the largest position of such an intersection node on Q, respectively. Proceed by defining k_2 as the largest integer smaller than k such that P_{k_2} sends flow along an intermediate node $v_{j_2} \neq \{v_{j_1}, v_2\}$ of the v_{j_1}-v_2-subpath of Q. Again, let i_2 and j_2 be the smallest and the largest position of such an intersection node on Q, respectively. We stop this iteration if v_1 becomes an intersection node.

Defining $P_{k_0} := P_k$ and $i_0 := j_0 := 0$ (implying $v_{i_0} = v$), we observe the following. The sequence $(k_0, k_1, \ldots\})$ is finite, decreases, and ends up in finding the largest $k' < k$ such that $P_{k'}$ traverses node v_1. For each suitable $\alpha \in \mathbb{N}$, the flow on the edges between $v_{j_{\alpha-1}}$ and v_{i_α} is not changed between iteration k_α and k of the SSP algorithm applied to $\mathcal{N}$. On the one hand, for all suitable $\alpha \in \mathbb{N}$ this shows that the path from $v_{j_{\alpha-1}}$ to v_{i_α} proceeding via P_{k_α} to t has free capacity in the residual network if P_{k_α} is computed. On the other hand, we know that the path from s to $v_{j_{\alpha-1}}$ along $P_{k_{\alpha-1}}$ proceeding to v_{i_α} is able to carry flow in the residual network if $P_{k_{\alpha-1}}$ is computed. Hence, Lemma 5.10 and Lemma 5.11 show that flow traveling along $P_{k_{\alpha-1}}$ departs earlier but arrives later than flow traveling along P_{k_α}. By induction, this implies that $\mathcal{F}_{\mathrm{O}}$ sends flow into $P_{k'}$ after time θ which arrives at t before time $\theta + \tau^r_{P_k}$. By the observation above, this shows that $\mathcal{F}$ also sends flow at a maximum possible rate into $P_{k'}$ which arrives at v_1 at the desired time. This ensures that the current capacity of v_1v_2 equals 0 implying that P_k is a shortest path which is able to carry flow in the residual network of $\mathcal{N}'$ with respect to $((P_{k'}, F_{e_{k'}}))_{k'<k}$. Hence, $\mathcal{F}$ represents a possible output of the SSP algorithm applied to $\mathcal{N}'$. Therefore, $\mathcal{F}$ satisfies the assumption of Lemma 5.15.

It remains to show that $\mathcal{F}^+ := (F_P^+)_{P\in\mathcal{P}}$ is a feasible strategy profile. Recalling the first part of the proof of Lemma 5.15 this follows directly from the definition of $\mathcal{F}^+$ as $\mathcal{F}^+_{\ominus}$ is a feasible strategy profile. □

Theorem 5.19 shows that an EA-Nash flow is a Nash flow over time. Section 5.4 verifies the good performance of these EA-Nash flows in order to compute prices of stability for the direct routing game over time. But EA-Nash flows are also interesting from another point of view. Assume that for a given network the demand goes to infinity. In this case the waiting times in front of each path grow rapidly implying that the times at which flow is sent along a particular path decreases to 0. In this manner, the EA-Nash flow becomes an earliest arrival flow. Further, increasing the demand to infinity means that all flow particles already depart at time 0. Hence, for such scenarios the price of stability is equal one, i.e., there exists a Nash flow with an optimal performance. This is formally observed in the following theorem.

Theorem 5.20. *For a direct routing game over time where each flow particle is present at s at time 0, there exists a Nash flow over time which is an earliest arrival flow.*

Proof. Let $\mathrm{val}(\mathcal{F}_{\mathrm{O}})$ be the flow pattern of an earliest arrival flow $(F_P^{\mathrm{O}})_{P\in\mathcal{P}^r}$, and let $D_k(\theta) := \sum_{P\in\mathcal{P}^r} F_P^{\mathrm{O}}(k\cdot\theta)$ for some $k \in \mathbb{N}$ and all $\theta \in \mathbb{R}_+$. In the following, we show that the EA-Nash flow converges to the earliest arrival flow if k goes to infinity.

Let $\mathcal{F}_k^+$ be the EA-Nash flow with respect to the demand function D_k and let ℓ^k_{N} be the corresponding minimum arrival time function. Because of Theorem 4.13 we know $D_k(\theta) = \mathrm{val}(\mathcal{F}_k^+)(\ell^k_{\mathrm{N}}(\theta))$. Further, the flow pattern of a Nash

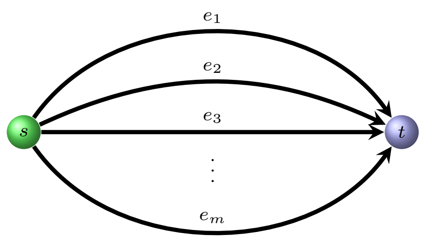


Figure 5.6: Network only consisting of parallel s-t-arcs used for the computation of an EA-Nash flow.

flow is not greater than the flow pattern of an earliest arrival flow which in turn is bounded by the total inflow. This shows

$$\sum_{P \in \mathcal{P}^r} F_P^{\mathrm{O}}(k \cdot \theta) \;\leq\; \mathrm{val}(\mathcal{F}_{\mathrm{O}}) \left(\ell_{\mathrm{N}}^k(\theta)\right) \;\leq\; \sum_{P \in \mathcal{P}^r} F_P^{\mathrm{O}}\left(\ell_{\mathrm{N}}^k(\theta)\right) .$$

Since $\sum_{P \in \mathcal{P}^r} F_P^{\mathrm{O}}$ is nondecreasing, this shows $\ell_{\mathrm{N}}^k > k \cdot \mathrm{id}$. Hence, because of Lemma 5.15 the EA-Nash flow starts sending flow into some s-t-path P at time $\frac{\tau_P^r}{k}$ at the latest. As $\mathcal{F}_{\mathrm{O}}$ already sends flow into P at time 0, the earliest arrival flow and the EA-Nash flow pattern differ by at most

$$\sum_{P \in \mathcal{P}^r} F_P^{\mathrm{O}}\left(\frac{\tau_P^r}{k}\right)$$

As this converges to 0 if k goes to ∞, we are done. □

5.4 Price of Stability

In this section we measure the performance of direct Nash flows over time. Since Example 5.17 shows that the price of anarchy is unbounded in general, we focus on the price of stability. We present bounds on the price of stability for each of the objective functions discussed in Section 4.3. It turns out that Nash equilibria resulting out of direct earliest arrival flows are very appropriate for finding answers in this connection. Recalling Definition 5.18, it is enough to consider only networks consisting of parallel edges connecting the source s to the sink t (see Figure 5.6). Because of Theorem 5.19, bounding a particular price of anarchy on these simple instances directly provides bounds for the corresponding general direct price of stability. To see this, also note that a Nash flows over time on such instances are unique.

Notations and basic valid equations we use throughout this section are explained in Subsection 5.4.1. Subsequently, in Subsection 5.4.2 we show that the direct evacuation and working price of stability are tightly bounded by 2. The corresponding tight instances for the bound on the evacuation price of stability works with constant capacities, whereas the capacities in the tight examples for the working price of anarchy are highly nonconstant. In fact, the working price of stability decreases if we restrict to constant capacity. This is observed in Subsection 5.4.3.

The completion and average arrival time price of stability are analyzed in Subsection 5.4.4 and 5.4.5, respectively. It is not hard to see that both are unbounded for arbitrary time-dependent capacity functions. Therefore, we only consider constant capacities in these subsections. For such scenarios we show that the completion time price of stability is tightly bounded by $\frac{3}{4}$. This bound carries over to the average arrival time price of stability. However, we strongly believe that it is not tight in this case. This is corroborated by empirical results where we compute the average arrival time price of anarchy for concrete instances.

5.4.1 Preliminary Considerations

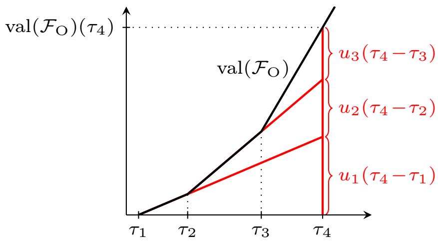
In the following we introduce some notation which we use throughout this section. As already mentioned, we mainly consider networks consisting only of parallel s-t-edges (see Figure 5.6). Let $\mathcal{I}_m$ be denote such an instance which is build on $m \in \mathbb{N}$ parallel s-t-edges $e_1, \ldots, e_m$. The transit time of an edge e_i is stored in $\tau_i := \tau_{e_i} \in \mathbb{R}_+$ and we require that the transit times are ordered increasingly, i.e., $\tau_1 \leq \ldots \leq \tau_m$. Further, let τ_{m+1} be some point in time which is as least as large as τ_m. The capacity of an edge e_i is determined by $u_i := u_{e_i}$. Depending on whether or not we consider scenarios with constant capacities, u_i is either a Lebesgue integrable function $\mathbb{R}_+ \to \mathbb{R}_+$ bounding the rate at which flow is able to traverse e_i or a real number representing the constant value of the capacity function. Further, we assume that the given demand D is large enough to serve all edges up to its capacity, i.e., in particular we require

$$D \;\geq\; \sum_{i=1}^{m} U_i + \tau_i \qquad \text{or} \qquad d \;\geq\; \sum_{i=1}^{m} u_i \tag{5.23}$$

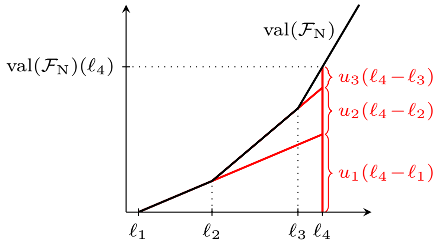
depending on whether we analyze time-varying or constant capacities. In particular, we assume that u_i is essentially zero over the time interval $[0, \tau_i]$ in nonconstant scenarios and that the demand rate is constant in constant scenarios. Further, we also require that the demand is able to use each edge up to its capacity in a Nash flow over time. Unfortunately, this condition is not as easy to state, especially, if we consider time-varying scenarios. For constant capacities this follows already from (5.23).

Since transit times are ordered nondecreasingly, a Nash flow first uses the edge e_1, then e_2, and so on. This shows in particular that a Nash flow on such instances is unique. With θ_i we denote the time when the Nash flow starts sending flow over edge e_i and with $\ell_i := \ell_N(\theta_i)$ be the arrival time of the Nash flow at time θ_i. In particular, we have $\theta_1 = 0$ and $\ell_1 = \tau_1$. Since we assume that the demand is able to support all edges up to its capacity, an edge which is used by the Nash flow at some point in time is used up to its capacity for all future points in time. Further, we use ℓ_{m+1} for a point in time greater than ℓ_m, and let $\theta_{m+1} \in \ell_N^{-1}(\ell_{m+1})$ be a corresponding departure time for which we assume without loss of generality $\theta_{m+1} \geq \theta_m$.

In the following we consider an instance $\mathcal{I}_m$ and evaluate the flow pattern of the earliest arrival and the Nash flow which are denoted by $\mathcal{F}_{\mathrm{O}}$ and $\mathcal{F}_{\mathrm{N}}$, respectively. First, we consider time-varying scenarios. Clearly, because of (5.23) an earliest arrival flow uses all edges up to its capacity. Hence, the flow pat-



(a) Earliest arrival flow pattern



(b) Nash flow pattern

Figure 5.7: Geometric interpretation of the earliest arrival and the Nash flow pattern on $\mathcal{I}_3$.

tern val($\mathcal{F}_{\mathrm{O}}$) of an earliest arrival flow is computable by

$$\mathrm{val}(\mathcal{F}_{\mathrm{O}}) \;=\; \sum_{i=1}^{m} U_i \;. \tag{5.24}$$

Next consider the Nash flow over time $\mathcal{F}_{\mathrm{N}}$. Since we assume that after time ℓ_i the underlying edge e_i is used up to its capacity, we obtain for the Nash flow pattern val($\mathcal{F}_{\mathrm{N}}$)

$$\mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_{k+1}) \;=\; \sum_{i=1}^{k} U_i(\ell_{k+1}) - U_i(\ell_i) \qquad \forall k \in [m+1] \;. \tag{5.25}$$

Now we turn our attention to scenarios where all edge capacities of $\mathcal{I}_m$ are constant. In contrast to time-varying scenarios, (5.23) ensures that an edge which is used at some point in time by $\mathcal{F}_{\mathrm{N}}$ is used up to its capacity for all future points in time. Thus, (5.24) and (5.25) translate to (see Figure 5.7)

$$\mathrm{val}(\mathcal{F}_{\mathrm{O}})(\tau_{k+1}) \;=\; \sum_{i=1}^{k} u_i(\tau_{k+1} - \tau_i) \qquad \forall k \in [m+1]_0 \tag{5.26}$$

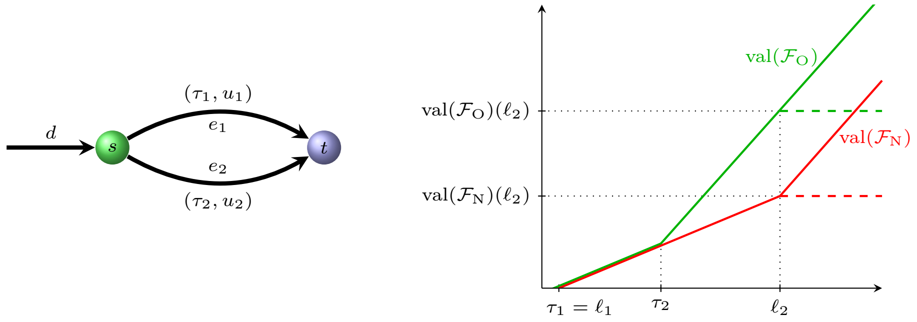
$$\text{and} \qquad \mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_{k+1}) \;=\; \sum_{i=1}^{k} u_i(\ell_{k+1} - \ell_i) \qquad \forall k \in [m+1]_0 \;. \tag{5.27}$$

We conclude this section by establishing equivalences between the τ_i's, θ_i's, and ℓ_i's. The first equation follows directly from the previous definitions which imply that the transit time of the Nash flow at time θ_i is τ_i, i.e.,

$$\ell_i \;=\; \theta_i + \tau_i \qquad\qquad \forall i \in [m+1] \tag{5.28}$$

The next equation is directly implied by Theorem 4.13(iii) which shows that the flow pattern of $\mathcal{F}_{\mathrm{N}}$ satisfies the equation $\mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_i) = D(\theta_i) = d\theta_i$. For this we use $u_{i_1\ldots i_2} := \sum_{j=i_1}^{i_2} u_j$ as a short notation for the sum of the capacities over the edges $e_{i_1}, e_{i_1+1}, \ldots, e_{i_2}$ for some $i_1, i_2 \in [m+1]$ which is set to zero if $i_2 < i_1$. Hence, from (5.27) we get for all $i \in [m+1]$

$$d(\theta_{i+1} - \theta_i) \;=\; \mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_{i+1}) - \mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_i) \;=\; u_{1\ldots i}(\ell_{i+1} - \ell_i) \;.$$



(a) The network of $\mathcal{I}_2$.

(b) The earliest arrival and the Nash flow pattern. The dashed pattern refers to the instance where capacities are set to 0 after time ℓ_2.

Figure 5.8: Illustration of Example 5.21.

Together with (5.28) this shows for all $i \in [m+1]$

$$d(\theta_{i+1} - \theta_i) = u_{1\ldots i}(\ell_{i+1} - \ell_i) , \tag{5.29}$$

$$(d - u_{1\ldots i})(\theta_{i+1} - \theta_i) = u_{1\ldots i}(\tau_{i+1} - \tau_i) , \tag{5.30}$$

and

$$(d - u_{1\ldots i})(\ell_{i+1} - \ell_i) = d\,(\tau_{i+1} - \tau_i) . \tag{5.31}$$

5.4.2 Evacuation and Working Price of Anarchy

In this subsection we analyze the evacuation and working price of stability for the direct flow model. We show that both prices are equal to 2 for time-varying capacities. The following example depicted in Figure 5.8 shows that for each $\epsilon > 0$ there are instances, for which the evacuation just as the working price of stability is larger than $2-\epsilon$. In addition, this example provides a complete analysis of the evacuation price of anarchy for networks consisting of two parallel s-t-edges with constant capacities.

Example 5.21. We consider the instance $\mathcal{I}_2$ drawn in Figure 5.8a consisting of two s-t-edges e_1 and e_2. Let $\mathcal{F}_\mathrm{O}$ and $\mathcal{F}_\mathrm{N}$ be the earliest arrival and the Nash flow on $\mathcal{I}_2$. The corresponding flow pattern are drawn in Figure 5.8b. Recalling the definition of the evacuation price of anarchy we have to estimate the supremum of $\rho_\mathrm{evac}(\theta) := \frac{\mathrm{val}(\mathcal{F}_\mathrm{O})(\theta)}{\mathrm{val}(\mathcal{F}_\mathrm{N})(\theta)}$ over all $\theta \in \mathbb{R}_+$. Using the notation of Subsection 5.4.1 we know $\ell_1 = \tau_1$ and $\ell_2 = \theta_2 + \tau_2 > \tau_2$ implying that the quotient $\rho_\mathrm{evac}(\theta)$ is equal to 1 until time τ_2 and, in general, greater than 1 afterwards. Further, the slope of both, $\mathrm{val}(\mathcal{F}_\mathrm{O})(\theta)$ and $\mathrm{val}(\mathcal{F}_\mathrm{N})(\theta)$, is equal to $u_1 + u_2 > 0$ for all $\theta \geq \ell_2$. This shows that $\rho_\mathrm{evac}(\theta)$ decreases after time ℓ_2. Hence, in order to compute the price of anarchy for this instance it is enough to consider ρ_evac over $[\tau_2, \ell_2]$. Let $\theta \in \mathbb{R}_+$ be some point in time contained in $[\tau_2, \ell_2]$. Because of (5.26) and (5.27) we obtain

$$\begin{aligned}
\mathrm{val}(\mathcal{F}_\mathrm{O})(\theta) &= u_1(\theta - \tau_1) + u_2(\theta - \tau_2) = (u_1 + u_2)(\theta - \tau_1) - u_2(\tau_2 - \tau_1) , \\
\mathrm{val}(\mathcal{F}_\mathrm{N})(\theta) &= u_1(\theta - \ell_1) = u_1\,(\theta - \tau_1) .
\end{aligned}$$

Since this shows

$$\rho_{\text{evac}}(\theta) \;=\; \frac{u_1+u_2}{u_1} - \frac{u_2(\tau_2-\tau_1)}{u_1(\theta-\tau_1)}\;,$$

we observe that $\rho_{\text{evac}}(\theta)$ increases if θ becomes larger. Hence, $\rho_{\text{evac}}(\theta)$ is maximized at $\theta = \ell_2 = \theta_2 + \tau_2$. Using equation (5.31), we obtain for the second term of the difference

$$\frac{u_2(\tau_2-\tau_1)}{u_1(\ell_2-\tau_1)} \;=\; \frac{u_2(\tau_2-\tau_1)}{u_1(\ell-\ell_1)} \;=\; \frac{u_2(d-u_1)}{u_1 d}\;.$$

This leads to

$$\rho_{\text{evac}}(\ell_2) \;=\; \frac{u_1+u_2}{u_1} - \frac{u_2(d-u_1)}{u_1 d} \;=\; 1+\frac{u_2}{d}\;.$$

This shows that if d goes to infinity for fixed capacities, the evacuation price of anarchy on this instance tends to 1. In particular, this verifies Theorem 5.20. Further, the $\rho_{\text{evac}}(\ell_2)$ increases if d decreases. Since we assume $d \le u_1+u_2$, the evacuation price of anarchy becomes maximal if $d = u_1+u_2$ and is bounded from above by 2. Finally, $\rho_{\text{evac}}(\ell_2)$ can be as close as possible to 2 if u_1 is sufficiently small compared to u_2. But note that it jumps to 1 in case u_1 is exactly equal to 0. More precisely, for a given small $\epsilon > 0$ we set $u_1 := \epsilon$, $u_2 := 1-\epsilon$, and $d = 1$ verifying

$$\rho_{\text{evac}} \;\ge\; 2-\epsilon\;.$$

Finally, it is also worth to mention that the evacuation price of anarchy on such instances is independent on the exact values for the transit times provided that $\tau_1 < \tau_2$ holds.

To see that also the working price of anarchy ρ_{work} can be made as close as possible to 2, recall that ρ_{work} refers to the area under the flow pattern graph. Consider an instance $\mathcal{I}$ where the evacuation price of anarchy is nearly equal to 2 at some point in θ. That is, at time θ, the earliest arrival flow pattern is nearly twice as large as the Nash flow pattern. Now reduce all edge capacities to zero after time θ and call the resulting instance $\tilde{\mathcal{I}}$. Then the earliest arrival and Nash flow pattern on $\tilde{\mathcal{I}}$ are given by the corresponding flow patterns on $\mathcal{I}$ until time θ which remain constant for later points in time (see the dashed lines in Figure 5.8b). Increasing the time to infinity shows that the area under the earliest arrival flow pattern become nearly twice as large as the one under the Nash flow pattern.

Example 5.21 shows that the direct evacuation and working price of stability is lower bounded by 2. As already mentioned, this is also the correct value for these prices of stability.

Theorem 5.22. *For the direct flow model the evacuation and working price of stability is equal to* 2.

Proof. Because of Example 5.21 the evacuation and working price of stability are lower bounded by 2 for the direct flow model. In the following we show that this is also an upper bound. As already mentioned and observed in Theorem 5.19, it is enough to consider instances $\mathcal{I}_m$ consisting of m parallel s-t-edges. Let $\mathcal{F}_{\mathrm{O}}$

be an earliest arrival flow on $\mathcal{I}_m$. Because of (5.24) the flow pattern $\mathrm{val}(\mathcal{F}_\mathrm{O})$ of $\mathcal{F}_\mathrm{O}$ is given by

$$\mathrm{val}(\mathcal{F}_\mathrm{O}) \;=\; \sum_{i=1}^{m} U_i \;.$$

Note that capacities bound the flow at the end of the corresponding edge and that we assume $u_e|_{[0,\tau_e]} = 0$.

For the Nash flow $\mathcal{F}_\mathrm{N}$, let ℓ_{m+1} be the time at which the respective price of anarchy is attained and let θ_{m+1} be such that $\theta_{m+1} \in \ell^{-1}(\ell_{m+1})$. Without loss of generality, we assume $\theta_{m+1} \geq \theta_m$. Otherwise we reduce transit times greater than $\ell_{m+1} - \theta_{m+1}$ to this value. On this new instance the Nash flow is not changed until time θ_{m+1} but the earliest arrival flow is improved. Hence, because of (5.25), the flow pattern of $\mathcal{F}_\mathrm{N}$ at time ℓ_{m+1} satisfies

$$\mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_{m+1}) \;=\; D(\theta_{m+1}) \;=\; \sum_{i=1}^{m} \big(U_i(\ell_{m+1}) - U_i(\ell_i)\big)$$

where the first equality sign is implied by Theorem 4.13(iii). This shows

$$\begin{aligned}
\mathrm{val}(\mathcal{F}_\mathrm{O})(\ell_{m+1}) - \mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_{m+1}) \;&=\; \sum_{i=1}^{m} U_i(\ell_i) \;=\; \sum_{i=1}^{m} (U_i + \tau_i)(\theta_i) \\
&\leq\; D(\theta_{m+1}) \;=\; \mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_{m+1})
\end{aligned}$$

proving that the price of anarchy is upper bounded by 2 and, hence, $\rho_{\mathrm{evac}} = 2$.

The upper bound of 2 for the evacuation price of stability carries directly over to the working price of stability. To see this, observe that $\rho_{\mathrm{evac}} = 2$ implies

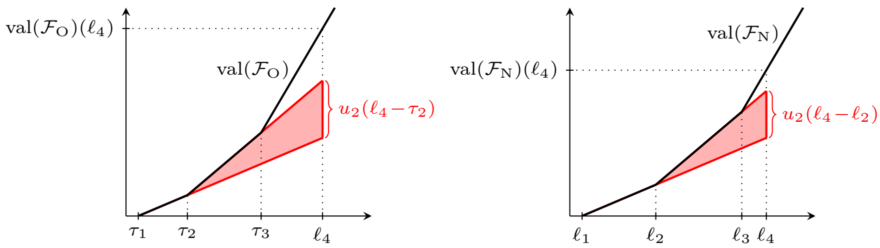
$$\begin{aligned}
\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{O})(\theta) \;&=\; \int_0^\theta \mathrm{val}(\mathcal{F}_\mathrm{O})(\vartheta)\; d\vartheta \\
&\leq\; \int_0^\theta 2 \cdot \mathrm{val}(\mathcal{F}_\mathrm{N})(\vartheta)\; d\vartheta \;=\; 2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{N})(\theta)
\end{aligned}$$

for each point in time $\theta \in \mathbb{R}_+$. $\square$

Summarizing, the direct evacuation and working price of stability is equal to 2. Even if we restrict to constant capacity function, this remains true for the evacuation price of stability as verified by Example 5.21. Unfortunately, this does not hold for the working price of stability. Here, restricting to constant capacities may open the space for further improvements on the price of stability.

5.4.3 Working Price of Anarchy for Constant Flow Rates

In this subsection we analyze the direct working price of stability $\rho^{\mathrm{PoS}}_{\mathrm{work}}$ for constant scenarios. We show that $\rho^{\mathrm{PoS}}_{\mathrm{work}}$ is equal to $\alpha^2 = 1.43923$ where $\alpha = 1.19969$ is the unique positive solution of $2\alpha = \log \frac{\alpha+1}{\alpha-1}$. As already mentioned, for proving this result we can restrict to instances $\mathcal{I}_m$ consisting of $m \in \mathbb{N}$ parallel s-t-edges by Theorem 5.19. Note that for time-varying capacities the working price of stability is equal to 2 as observed in Theorem 5.22.



(a) Working objective function for an arliest arrival flow.

(b) Working objective function for a Nash flow.

Figure 5.9: Computation of the earliest arrival and the Nash flow working objective function on $\mathcal{I}_3$.

So consider an instance $\mathcal{I}_m$. Firstly, we evaluate the working objective functions $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{O})$ and $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{N})$ for the earliest arrival and the Nash flow over time on $\mathcal{I}_m$, respectively. After that we give a brief description of the approach we take in order to compute $\rho^{\mathrm{PoS}}_{\mathrm{work}}$. This approach is worked out subsequently.

For computing $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{O})$ and $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{N})$ we concentrate on the point in time ℓ_{m+1} at which the working price of stability of $\mathcal{I}_m$ is obtained. For the same reason as in Subsection 5.4.2, we assume without loss of generality $\ell_{m+1} \geq \ell_m$. Taking a look at Figure 5.9, we evaluate $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{O})(\ell_{m+1})$ and $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{N})(\ell_{m+1})$ as the sum of the areas of the red triangles which shows

$$2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{O})(\ell_{m+1}) = \sum_{i=1}^{m} u_i(\ell_{m+1} - \tau_i)^2 \tag{5.32}$$

and

$$2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{N})(\ell_{m+1}) = \sum_{i=1}^{m} u_i(\ell_{m+1} - \ell_i)^2 \; . \tag{5.33}$$

For maximizing the quotient $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{O})(\ell_{m+1})$ over $\mathrm{val}_{\mathrm{work}}(\mathcal{F}_\mathrm{N})(\ell_{m+1})$, we take the following approach. Firstly, we observe that this optimization results in maximizing a quadratic form over the unique sphere in $\mathbb{R}^m$. That is, for a suitable symmetric matrix $A \in \mathbb{R}^{m\times m}$, we have to maximize x^TAx over all $x \in \mathbb{R}^m$ with $\sum_{i=1}^m x_i^2 = 1$. It is well known that this maximum equals the maximal eigenvalue of A. Thus, computing $\rho^{\mathrm{PoS}}_{\mathrm{work}}$ reduces to finding the maximal eigenvalue of A.

The eigenvalues of a matrix A are exactly the roots of the characteristic polynom $q(\alpha) := \det(A - \alpha I)$ where $I \in \mathbb{R}^{m\times m}$ is the $m \times m$-identity matrix. In order to estimate the roots of q we transform $A - \alpha I$ into a 3-diagonal matrix, i.e., a matrix where all entries are 0 except those on the main and the two neighboring diagonals. The determinant of a 3-diagonal matrix is recursively computable which results in a linear recursion of order 2 for the characteristic polynom q. Solving this recursion gives an explicit formula for q depending on m. Further, this explicit formula enables us to estimate all roots of q such that we obtain a tight bound on the largest eigenvalue of A. This finishes our approach.

In the following we work out all the details of the approach explained above. Throughout this approach we simplify current expressions at several points.

The first simplification is already made right now. Since we can model arbitrary (rational) edge capacities with unit capacities, we assume without loss of generality that the capacity on each edge is equal to 1. Further, we assume that the demand rate is the constant function m. This imposes no loss of generality because larger demand rates have no impact on the earliest arrival flow but improve the Nash flow. Hence, recalling (5.32) and (5.33), we want to maximize the quotient of the following two values:

$$\begin{aligned} & 2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{O}})(\ell_{m+1}) = \sum_{i=1}^{m} (\ell_{m+1} - \tau_i)^2 = \sum_{i=1}^{m} (\ell_{m+1} - \ell_i + \theta_i)^2 \\ \text{and} \qquad & 2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1}) = \sum_{i=1}^{m} (\ell_{m+1} - \ell_i)^2 \ . \end{aligned}$$

Subtracting both equations leads to

$$2 \cdot \big(\mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{O}})(\ell_{m+1}) - \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1})\big) = \sum_{i=1}^{m} 2\theta_i(\ell_{m+1} - \ell_i) + \theta_i^2. \tag{5.34}$$

Next we show how to express $2 \cdot ((\mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{O}})(\ell_{m+1}) - \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1}))$ in terms of $x_i := \ell_{m+1} - \ell_i$ for all $i \in [m+1]$. This seems to be promising as we have $2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1}) = \sum_{i=1}^{m} x_i^2$. For unique capacities, equation (5.28) translates to $m(\theta_{j+1} - \theta_j) = j(\ell_{j+1} - \ell_j)$ for all $j \in [m+1]$. This shows for all $i \in [m+1]$

$$\begin{aligned} \theta_i &= \sum_{j=1}^{i-1} (\theta_{j+1} - \theta_j) = \sum_{j=1}^{i-1} \frac{j}{m}(\ell_{j+1} - \ell_j) \\ &= \sum_{j=1}^{i-1} \frac{j}{m}(x_j - x_{j+1}) = -\frac{i}{m}x_i + \sum_{j=1}^{i} \frac{1}{m}x_j \ . \end{aligned}$$

Therefore, the summands of the left hand side in (5.34) are expressable as

$$\begin{aligned} \theta_i(2x_i + \theta_i) &= (-\frac{i}{m}x_i + \sum_{j=1}^{i} \frac{1}{m}x_j)(2x_i - \frac{i}{m}x_i + \sum_{j=1}^{i} \frac{1}{m}x_j) \\ &= (-\frac{2i}{m} + \frac{i^2}{m^2})x_i^2 + \frac{1}{m^2}(\sum_{j=1}^{i} x_j)^2 + (\frac{2}{m} - \frac{2i}{m^2}) \sum_{j=1}^{i} x_i x_j \ . \end{aligned}$$

Hence, computing the difference $2 \cdot ((\mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{O}})(\ell_{m+1}) - \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1}))$ we obtain the following coefficient in front of $x_i x_j$ for some $i, j \in [m+1]$

$$\begin{aligned} & \begin{cases} (-\frac{2i}{m} + \frac{i^2}{m^2}) + \frac{m-i+1}{m^2} + (\frac{2}{m} - \frac{2i}{m^2}) & \text{if } i = j \\ \frac{2(m-i+1)}{m^2} + (\frac{2}{m} - \frac{2i}{m^2}) & \text{if } j < i \end{cases} \\ = & \begin{cases} \frac{-2i+3}{m} + \frac{i^2-3i+1}{m^2} & \text{if } i = j \\ \frac{4}{m} + \frac{-4i+2}{m^2} & \text{if } j < i \end{cases} \\ = & \begin{cases} \frac{2}{m} + \frac{-2i+1}{m^2} + \frac{-2i+1}{m} + \frac{i^2-i}{m^2} & \text{if } i = j \\ 2 \cdot (\frac{2}{m} + \frac{-2i+1}{m^2}) & \text{if } j < i \end{cases} \ . \end{aligned}$$

Let $A_m := (a_{ij}^m)_{1 \le i,j \le m}$ be the symmetric matrix defined by

$$a_{ij}^m := \frac{2}{m} + \frac{-2\max\{i,j\}+1}{m^2} + \begin{cases} 0 & \text{if } i \neq j \\ \frac{-2i+1}{m} + \frac{i^2-i}{m^2} & \text{if } i = j \end{cases} . \tag{5.35}$$

With this definition the right hand side of (5.34) is equivalent to $x^T A_m x$ where $x^T := (x_1, \dots, x_m)$. Thus, for computing $\rho_{\text{work}}^{\text{PoS}}$, we maximize the quotient $x^T A_m x / \|x\|_2^2$ over $x \in \mathbb{R}^m$ where $\|.\|_2$ denotes the 2-norm in $\mathbb{R}^m$. Note that $\rho_{\text{work}}^{\text{PoS}}$ is equal to the optimum plus 1.

Since scaling of x does not change the value of the fraction, the problem can be restated as $\max\{x^T A_m x \mid \|x\|_2 = 1\}$. From linear algebra we know that the maximum equals the maximal eigenvalue of A_m. Hence, an upper bound for the eigenvalues of A_m gives an upper bound for the direct working price of stability. We proceed as follows. The eigenvalues of A_m are the roots of the characteristic polynomial $q_m(\beta) := \det(A - \beta I_m)$ where I_m is the identity matrix of dimension $m \times m$. The first step is to find a recursive definition of q_m using the special structure of the matrices A_m. After simplifying the recursion we compute an explicit formula for q_m. From this explicit formula we bound the roots of q_m. The recursion for q_m is derived from the following recursion of the matrices A_m.

Lemma 5.23. *The matrices A_m satisfy the following recursion:*

$$A_m = \begin{pmatrix} a_{11}^m & a_{12}^m & \cdots & a_{1m}^m \\ a_{21}^m & & & \\ \vdots & & \frac{(m-1)^2}{m^2} A_{m-1} - \frac{2m-1}{m^2} I_{m-1} & \\ a_{m1}^m & & & \end{pmatrix} .$$

Proof. Recall the definition of A_m in (5.35). For $1 < j < i \le m$ we have

$$\begin{aligned} m^2 a_{ij}^m &= 2m + (-2i+1) = (2(m-1)+2) + (-2(i-1)-2+1) \\ &= 2(m-1) + (-2(i-1)+1) \\ &= (m-1)^2 a_{i-1j-1}^{m-1} . \end{aligned}$$

With respect to the diagonal elements we obtain for $1 < i \le m$:

$$\begin{aligned} m^2 a_{ii}^m &= (-2i+3)m + i^2 - 3i + 1 \\ &= -2mi + 3m + i^2 - 3i + 1 \\ \text{and} \quad (m-1)^2 a_{i-1i-1}^{m-1} &= (-2(i-1)+3)(m-1) + (i-1)^2 - 3(i-1) + 1 \\ &= (-2i+5)(m-1) + i^2 - 5i - 3i + 5 \\ &= -2mi + 5m + 2i - 5 + i^2 - 5i + 5 \\ &= (-2mi + 3m + i^2 - 3i + 1) + (2m-1) . \end{aligned}$$

This proves the required recursion because A_m and A_{m-1} are symmetric. □

Next we use this recursion in order to find a recursion for the characteristic polynomial q_m of A_m. Consider the matrix $A_m - \beta I_m$ and subtract the second column from the first column and, subsequently, the second row from the first row in order to obtain the matrix $B_m = (b^m_{ij})_{1\le i,j\le m}$. We have

$$\begin{aligned} b^m_{21} = b^m_{12} &= a^m_{21} - (a^m_{22} - \beta) \\ &= \frac{2}{m} + \frac{-2\cdot 2+1}{m^2} - \Big(\frac{2}{m} + \frac{-2\cdot 2+1}{m^2} + \frac{-2\cdot 2+1}{m} + \frac{2^2-2}{m^2} - \beta\Big) \\ &= \frac{3}{m} - \frac{2}{m^2} + \beta \end{aligned}$$

and

$$\begin{aligned} b^m_{11} &= a^m_{11} - \beta - a^m_{21} - b^m_{21} \\ &= \frac{2}{m} + \frac{-2\cdot 1+1}{m^2} + \frac{-2\cdot 1+1}{m} + \frac{1^2-1}{m^2} - \beta \\ &\quad - \Big(\frac{2}{m} + \frac{-2\cdot 2+1}{m^2}\Big) - \Big(\frac{3}{m} - \frac{2}{m^2} + \beta\Big) \\ &= -\frac{4}{m} + \frac{4}{m^2} - 2\beta\ . \end{aligned}$$

Since $b^m_{1i} = b^m_{i1} = 0$ for all $i > 2$, the matrix B_m has the following structure:

$$B_m = \begin{pmatrix} -\frac{4}{m} + \frac{4}{m^2} - 2\beta & \frac{3}{m} - \frac{2}{m^2} + \beta \quad 0 \quad \dots \quad 0 \\ \frac{3}{m} - \frac{2}{m^2} + \beta & \\ 0 & \frac{(m-1)^2}{m^2} A_{m-1} - \big(\frac{2m-1}{m^2} + \beta\big) I_{m-1} \\ \vdots & \\ 0 & \end{pmatrix}.$$

Since column and row transformations do not change the determinant of a matrix, q_m is equal to $\det(B_m)$. We obtain the following recursion formula for q_m.

Lemma 5.24. *The characteristic polynomial q_m of A_m satisfies*

$$\begin{aligned} q_m(\beta) = &\left(-\frac{4}{m} + \frac{4}{m^2} - 2\beta\right)\left(\frac{(m-1)^2}{m^2})\right)^{m-1} \cdot q_{m-1}\left(\frac{m^2\beta}{(m-1)^2} + \frac{2m-1}{(m-1)^2}\right) \\ &- \left(\frac{3}{m} - \frac{2}{m^2} + \beta\right)^2 \left(\frac{(m-2)^2}{m^2}\right)^{m-2} \cdot q_{m-2}\left(\frac{m^2\beta}{(m-2)^2} + \frac{4m-4}{(m-2)^2}\right) \end{aligned}$$

with initial values

$$q_1(\beta) = 1 - \beta \qquad \textit{and} \qquad q_2(\beta) = \beta^2 + \frac{1}{2}\beta - \frac{1}{4}\ .$$

Proof. We use the Laplace formula for verifying the recursion. Firstly, we expand the determinant of B_m along the first column and obtain

$$\det(B_m) = b^m_{11}\det(B^{1,1}_m) - b^m_{21}\det(B^{2,1}_m)$$

where $B^{i,j}_m$ is the matrix that resulting from B_m by removing the i-th row and the j-th column. In order to compute $\det(B^{2,1}_m)$, we use again the Laplace formula but now along the first row leading to

$$\det(B_m) = b^m_{11}\det(B^{1,1}_m) - (b^m_{21})^2\det(\bar{B}_m)$$

where $\bar{B}$ is the matrix that results from B_m by removing the first two rows and columns.

From the formula for B_m and the multi-linearity of the determinant we know

$$\begin{aligned}\det(B_m^{1,1}) &= \det\left(\frac{(m-1)^2}{m^2}A_{m-1} - \Big(\frac{(2m-1)}{m^2}+\beta\Big)I_{m-1}\right)\\ &= \det\left(\frac{(m-1)^2}{m^2}\left(A_{m-1} - \Big(\frac{m^2}{(m-1)^2}\beta+\frac{(2m-1)}{(m-1)^2}\Big)I_{m-1}\right)\right)\\ &= \left(\frac{(m-1)^2}{m^2}\right)^{m-1}\cdot\det\left(A_{m-1} - \Big(\frac{m^2}{(m-1)^2}\beta+\frac{(2m-1)}{(m-1)^2}\Big)I_{m-1}\right)\ .\end{aligned}$$

For computing $\det(\bar{B}_m)$ we apply the recursion formula for A_m twice and obtain

$$\begin{aligned}\bar{B}_m &= \left(\frac{(m-1)^2}{m^2}A_{m-1} - \Big(\frac{(2m-1)}{m^2}+\beta\Big)I_{m-1}\right)^{1,1}\\ &= \frac{(m-1)^2}{m^2}\left(\frac{(m-2)^2}{(m-1)^2}A_{m-2} - \frac{(2m-3)}{(m-1)^2}I_{m-2}\right)\\ &\quad - \Big(\frac{(2m-1)}{m^2}+\beta\Big)I_{m-2}\\ &= \left(\frac{(m-2)^2}{m^2}A_{m-2} - \Big(\frac{(4m-4)}{m^2}+\beta\Big)I_{m-2}\right)\ .\end{aligned}$$

From this we compute $\det(\bar{B}_m)$ similarly to $\det(B_m^{1,1})$ and obtain the desired recursion formula for q_m.

It remains to check the initial values. Since we have

$$A_1 = (0) \qquad \text{and} \qquad A_2 = \begin{pmatrix}\frac{1}{4} & \frac{1}{4}\\ \frac{1}{4} & -\frac{3}{4}\end{pmatrix},$$

we get

$$q_1(\beta) = -\beta \qquad \text{and} \qquad q_2(\beta) = (\frac{1}{4}-\beta)(-\frac{3}{4}-\beta) - \frac{1}{16} = \beta^2 + \frac{1}{2}\beta - \frac{1}{4}$$

and this lemma is established. □

The next step is to simplify the recursion formula for q_m. For eliminating m from the corresponding exponents, we define $\tilde{q}_m(\beta) := m^{2m}\cdot q_m(\frac{\beta}{m^2}-1)$. Using the equations

$$\begin{aligned}\frac{\beta}{(m-1)^2}-1 &= \frac{\frac{m^2}{m^2}\beta - m^2 + m^2 - (m-1)^2}{(m-1)^2} &= \frac{m^2(\frac{\beta}{m^2}-1)}{(m-1)^2} + \frac{2m-1}{(m-1)^2}\ ,\\ \frac{\beta}{(m-2)^2}-1 &= \frac{\frac{m^2}{m^2}\beta - m^2 + m^2 - (m-2)^2}{(m-2)^2} &= \frac{m^2(\frac{\beta}{m^2}-1)}{(m-1)^2} + \frac{4m-4}{(m-1)^2}\ ,\end{aligned}$$

Lemma 5.24 shows

$$\begin{aligned}
\tilde{q}_m(\beta) &= \Big(-\frac{4}{m}+\frac{4}{m^2}-2\Big(\frac{\beta}{m^2}-1\Big)\Big)\, m^2\cdot\tilde{q}_{m-1}(\beta)\\
&\quad -\Big(\frac{3}{m}-\frac{2}{m^2}+\frac{\beta}{m^2}-1\Big)^2 m^4\cdot\tilde{q}_{m-2}(\beta)\\
&= (2m^2-4m+4-2\beta)\cdot\tilde{q}_{m-1}(\beta)-(-m^2+3m-2+\beta)^2\cdot\tilde{q}_{m-2}(\beta)\\
&= 2(m^2-2m+2-\beta)\cdot\tilde{q}_{m-1}(\beta)-(m^2-3m+2-\beta)^2\cdot\tilde{q}_{m-2}(\beta)\ ,\\
\tilde{q}_1(\beta) &= 1-\beta\ ,\\
\tilde{q}_2(\beta) &= 16((\tfrac{\beta}{4}-1)^2+\tfrac{1}{2}(\tfrac{\beta}{4}-1)-\tfrac{1}{4}) = \beta^2-6\beta+4\ .
\end{aligned}$$

Setting $\tilde{q}_0(\beta) := 1$ we obtain $\tilde{q}_2(\beta) = 2(2-\beta)\cdot(1-\beta)-(-\beta)^2\cdot 1 = \beta^2-6\beta+4$ using the recursion. Finally, we substitute $\beta := \alpha^2-\frac{1}{4}$ and set $\bar{q}_m(\alpha) := \tilde{q}_m(\beta)$ which leads to

$$\begin{aligned}
\bar{q}_m(\alpha) &= \quad 2(m^2-2m+2-\alpha^2+\tfrac{1}{4})\cdot\bar{q}_{m-1}(\alpha)\\
&\quad -(m^2-3m+2-\alpha^2+\tfrac{1}{4})^2\cdot\bar{q}_{m-2}(\alpha)\\
&= 2(m^2-2m-\alpha^2+\tfrac{9}{4})\cdot\bar{q}_{m-1}(\alpha)-(m^2-3m-\alpha^2+\tfrac{9}{4})^2\cdot\bar{q}_{m-2}(\alpha)\\
&= 2((-\tfrac{3}{2}+m)^2-\alpha^2+m)\cdot\bar{q}_{m-1}(\alpha)-((-\tfrac{3}{2}+m)^2-\alpha^2)^2\cdot\bar{q}_{m-2}(\alpha)\ ,\\
\bar{q}_0(\alpha) &= 1\ ,\\
\bar{q}_1(\alpha) &= \frac{5}{4}-\alpha^2\ .
\end{aligned}$$

In order to see that this last substitution simplifies the recursion, we have to consider its explicit formula which is verified in the following lemma. In this explicit formula the *gamma function* Γ and the *digamma function* Ψ play an important role. The gamma and the digamma function can be seen as a generalization of the factorial function and the harmonic numbers to real numbers, respectively. That is, they satisfy the functional equations

$$\Gamma(\alpha+1)\ =\ \alpha\cdot\Gamma(\alpha) \qquad\text{and}\qquad \Psi(\alpha+1)\ =\ \Psi(\alpha)+\frac{1}{\alpha}$$

with initial values $\Gamma(1) := 1$ and $\Psi(1) := -\gamma$ where γ is the Euler-Mascheroni constant, i.e., the limiting difference between the harmonic series and the natural logarithm. Subsequently, we use at several points the following well-known *reflection formulas*

$$\Gamma(1-z)\cdot\Gamma(z)\ =\ \frac{\pi}{\sin(\pi z)} \qquad\text{and}\qquad \psi(1-x)-\psi(x)\ =\ \pi\cdot\cot(\pi x)\ .$$

With these definitions we are able to express the explicit formula for $\bar{q}$.

Lemma 5.25. *We have* $\bar{q}_m(\alpha) = \dfrac{q_m^\Gamma(\alpha)\cdot q_m^\Psi(\alpha)}{2\pi\alpha}$ *where*

$$\begin{aligned}
q_m^\Gamma(\alpha) &:= \Gamma\big(\tfrac{1}{2}+m-\alpha\big)\cdot\Gamma\big(\tfrac{1}{2}+m+\alpha\big) \qquad \textit{and}\\
q_m^\Psi(\alpha) &:= \cos(\alpha\pi)\Big(2\alpha+m\Big(\Psi(\tfrac{1}{2}+m-\alpha)-\Psi(\tfrac{1}{2}+m+\alpha)\Big)\Big)+m\pi\sin(\alpha\pi)\ .
\end{aligned}$$

Proof. The strategy is simple: Insert the explicit formula in the recursive definition and try to obtain an equation like "0 = 0". From the functional equation for the gamma function, we obtain

$$q^{\Gamma}_m(\alpha) = \big((-\tfrac{1}{2}+m)^2-\alpha^2\big)\cdot q^{\Gamma}_{m-1}(\alpha)$$

and

$$q^{\Gamma}_{m-1}(\alpha) = \big((-\tfrac{3}{2}+m)^2-\alpha^2\big)\cdot q^{\Gamma}_{m-2}(\alpha)\ .$$

Further, the functional equation of the digamma functions implies

$$\begin{aligned}
\frac{q^{\Psi}_m(\alpha)}{m} &= \cos(\alpha\pi)\bigg(\frac{2\alpha}{m}+\Psi(\tfrac{1}{2}+m-\alpha)-\Psi(\tfrac{1}{2}+m+\alpha)\bigg)+\pi\sin(\alpha\pi)\\
&= \cos(\alpha\pi)\bigg(\frac{2\alpha(m-1)}{m(m-1)}+\frac{1}{-\frac{1}{2}+m-\alpha}+\Psi(-\tfrac{1}{2}+m-\alpha)\\
&\qquad -\frac{1}{-\frac{1}{2}+m+\alpha}-\Psi(-\tfrac{1}{2}+m+\alpha)\bigg)+\pi\sin(\alpha\pi)\\
&= \cos(\alpha\pi)\bigg(\frac{2\alpha}{m-1}+\frac{-2\alpha}{m(m-1)}+\frac{2\alpha}{(-\frac{1}{2}+m)^2-\alpha^2}\\
&\qquad +\Psi(-\tfrac{1}{2}+m-\alpha)-\Psi(-\tfrac{1}{2}+m+\alpha)\bigg)+\pi\sin(\alpha\pi)\\
&= \frac{q^{\Psi}_{m-1}(\alpha)}{m-1}+2\alpha\cos(\alpha\pi)\bigg(-\frac{1}{m(m-1)}+\frac{1}{(-\frac{1}{2}+m)^2-\alpha^2}\bigg)\\
&= \frac{q^{\Psi}_{m-1}(\alpha)}{m-1}+2\alpha\cos(\alpha\pi)\bigg(\frac{\alpha^2-\frac{1}{4}}{m(m-1)\big((-\frac{1}{2}+m)^2-\alpha^2\big)}\bigg)\ .
\end{aligned}$$

Assuming that $\bar{q}_{m-1}$ satisfies the explicit formula, we know that $\bar{q}_m$ also satisfies the explicit formula if and only if we have

$$\begin{aligned}
\bar{q}_m(\alpha) &= \frac{q^{\Gamma}_m(\alpha)\cdot q^{\Psi}_m(\alpha)}{2\pi\alpha}\\
&= \big((-\tfrac{1}{2}+m)^2-\alpha^2\big)\cdot\frac{q^{\Gamma}_{m-1}(\alpha)}{2\pi\alpha}\\
&\quad\cdot\bigg(\frac{m}{m-1}q^{\Psi}_{m-1}(\alpha)+2\alpha\cos(\alpha\pi)\bigg(\frac{\alpha^2-\frac{1}{4}}{(m-1)\big((-\frac{1}{2}+m)^2-\alpha^2\big)}\bigg)\bigg)\\
&= \frac{m\big((-\frac{1}{2}+m)^2-\alpha^2\big)}{m-1}\cdot\bar{q}_{m-1}(\alpha)+\frac{\cos(\alpha\pi)}{\pi}\cdot\frac{\alpha^2-\frac{1}{4}}{m-1}\cdot q^{\Gamma}_{m-1}(\alpha)\ .
\end{aligned}$$

Inserting this order-1-recursion in the left hand side of the order-2-recursion for $\bar{q}_m$ we show that we obtain the same order-1-recursion but now for $\bar{q}_{m-1}$ depending on $\bar{q}_{m-2}$. For this we collect the terms containing $\bar{q}_{m-1}(\alpha)$ on the right hand side and get for the corresponding coefficient the following expression:

$$\begin{aligned}
&2\big((-\tfrac{3}{2}+m)^2-\alpha^2+m\big)-\frac{m\big((-\frac{1}{2}+m)^2-\alpha^2\big)}{m-1}\\
&= \frac{2(m-1)\big((-\frac{3}{2}+m)^2-\alpha^2\big)+2m(m-1)-m\big((-\frac{1}{2}+m)^2-\alpha^2\big)}{m-1}\\
&= \frac{2(m-1)\big((-\frac{3}{2}+m)^2-\alpha^2\big)-m\big((-\frac{3}{2}+m)^2-\alpha^2\big)}{m-1}\\
&= \frac{(m-2)\big((-\frac{3}{2}+m)^2-\alpha^2\big)}{m-1}\ .
\end{aligned}$$

Multiplying both sides with $\dfrac{m-1}{(m-2)\big((-\frac{3}{2}+m)^2-\alpha^2\big)}$ and separating $\bar{q}_{m-1}(\alpha)$ leads to

$$\begin{aligned}\bar{q}_{m-1}(\alpha) &= \frac{(m-1)\big((-\frac{3}{2}+m)^2-\alpha^2\big)}{m-2}\cdot\bar{q}_{m-2}(\alpha)\\ &\quad+\frac{\cos(\alpha\pi)}{\pi}\cdot\frac{q^{\Gamma}_{m-1}(\alpha)}{(-\frac{3}{2}+m)^2-\alpha^2}\cdot\frac{\alpha^2-\frac{1}{4}}{m-2}\\ &= \frac{(m-1)\big((-\frac{3}{2}+m)^2-\alpha^2\big)}{m-2}\cdot\bar{q}_{m-2}(\alpha)\\ &\quad+\frac{\cos(\alpha\pi)}{\pi}\cdot q^{\Gamma}_{m-2}(\alpha)\cdot\frac{\alpha^2-\frac{1}{4}}{m-2}\ .\end{aligned}$$

This establishes the explicit formula for $\bar{q}_{m-1}$ if, in addition to $\bar{q}_{m-1}$, also $\bar{q}_{m-2}$ satisfies the explicit formula. Hence, for completing this induction, it remains to verify the initial values. For $m=0$ we have

$$q^{\Gamma}_0(\alpha) = \Gamma\big(\tfrac{1}{2}-\alpha\big)\cdot\Gamma\big(\tfrac{1}{2}+\alpha\big)$$

and

$$q^{\Psi}_0(\alpha) = 2\alpha\cos(\alpha\pi)\ .$$

From the reflection formula for Γ, we obtain

$$q^{\Gamma}_0(\alpha) = \frac{\pi}{\sin\big(\pi(\frac{1}{2}+\alpha)\big)} = \frac{\pi}{\cos(\pi\alpha)}$$

implying $\bar{q}_0(\alpha)=1$. For $m=1$ we get:

$$q^{\Gamma}_1(\alpha) = \frac{\pi(\frac{1}{4}-\alpha^2)}{\cos(\pi\alpha)} \qquad \text{and}$$

$$\begin{aligned}q^{\Psi}_1(\alpha) &= \cos(\alpha\pi)\Big(2\alpha+\Big(\Psi(\tfrac{3}{2}-\alpha)-\Psi(\tfrac{3}{2}+\alpha)\Big)\Big)+\pi\sin(\alpha\pi)\\ &= \cos(\alpha\pi)\Big(2\alpha+\frac{1}{\frac{1}{2}-\alpha}-\frac{1}{\frac{1}{2}+\alpha}+\Big(\Psi(\tfrac{1}{2}-\alpha)-\Psi(\tfrac{1}{2}+\alpha)\Big)\Big)+\pi\sin(\alpha\pi)\\ &= \cos(\alpha\pi)\Big(2\alpha\Big(1+\frac{1}{\frac{1}{4}-\alpha^2}\Big)+\Big(\Psi(\tfrac{1}{2}-\alpha)-\Psi(\tfrac{1}{2}+\alpha)\Big)\Big)+\pi\sin(\alpha\pi)\ .\end{aligned}$$

From the reflection formula for Ψ, we obtain

$$\Psi(\tfrac{1}{2}-\alpha)-\Psi(\tfrac{1}{2}+\alpha) = \pi\cot\big(\pi(\tfrac{1}{2}+\alpha)\big) = \pi\cdot\frac{\cos\big(\pi(\frac{1}{2}+\alpha)\big)}{\sin\big(\pi(\frac{1}{2}+\alpha)\big)} = \pi\cdot\frac{-\sin(\pi\alpha)}{\cos(\pi\alpha)}\ .$$

Hence, we have

$$q^{\Psi}_1(\alpha) = \cos(\alpha\pi)2\alpha\Big(1+\frac{1}{\frac{1}{4}-\alpha^2}\Big)$$

implying

$$\bar{q}_1(\alpha) = \tfrac{1}{4}-\alpha^2+1 = \tfrac{5}{4}-\alpha^2\ . \qquad \square$$

The next step is to locate the roots of $\bar{q}_m$. The following lemma evaluates the sign of $\bar{q}_m$ at certain values from which we are able to determine bounds for some roots of $\bar{q}_m$.

Lemma 5.26.

(i) For $\alpha = 1, \dots, m-1$, it holds $\operatorname{sign}\big(\bar{q}_m(\alpha + \frac{1}{2})\big) = (-1)^\alpha$.

(ii) For $\alpha = 1, \dots, m-1$, it holds $\operatorname{sign}\big(\bar{q}_m(\alpha)\big) = -(-1)^\alpha$.

Proof. For $0 < \alpha < \frac{1}{2} + m$, we have $q_m^\Gamma(\alpha) > 0$ and $2\pi\alpha > 0$. Hence, the sign of $\bar{q}_m(\alpha)$ is equal to the sign of $q_m^\Psi(\alpha)$. For $\alpha = 1, \dots, m-1$ we have

$$q_m^\Psi\big(\alpha + \tfrac{1}{2}\big) = m\pi \sin\big((\alpha + \tfrac{1}{2})\pi\big) \qquad \Rightarrow \qquad \operatorname{sign}(q_m^\Psi\big(\alpha + \tfrac{1}{2}\big)) = (-1)^\alpha$$

proving (i). In order to prove (ii), we observe that for $\alpha = 1, \dots, m-1$ it holds

$$\begin{aligned} q_m^\Psi(\alpha) &= \cos(\alpha\pi)\bigg(2\alpha + m\Big(\Psi(\tfrac{1}{2} + m - \alpha) - \Psi(\tfrac{1}{2} + m + \alpha)\Big)\bigg) \\ &= \cos(\alpha\pi)\bigg(2\alpha - m\sum_{j=-\alpha}^{\alpha-1} \frac{1}{\frac{1}{2} + m + j}\bigg) . \end{aligned}$$

Using Jensen's inequality, we obtain

$$\frac{1}{2\alpha} \cdot \sum_{j=-\alpha}^{\alpha-1} \frac{1}{\frac{1}{2} + m + j} > \frac{1}{\frac{1}{2\alpha} \cdot \sum\limits_{j=-\alpha}^{\alpha-1} (\frac{1}{2} + m + j)} = \frac{1}{m}$$

$$\Rightarrow \qquad 2\alpha - m\sum_{j=-\alpha}^{\alpha-1} \frac{1}{\frac{1}{2} + m + j} < 2\alpha - m \cdot \frac{2\alpha}{m} = 0 .$$

This shows that the sign of $\bar{q}(\alpha)$ is equal to $\operatorname{sign}(-\cos(\alpha\pi)) = -(-1)^\alpha$ proving (ii). $\square$

Since the sign changes of $\bar{q}_m$ changes from α to $\alpha + \frac{1}{2}$, there must be a root in the interval $(\alpha, \alpha + \frac{1}{2})$ for each $\alpha = 1, \dots, m-1$. Considering the explicit formula of $\bar{q}_m$ we observe that $\bar{q}_m$ is an even function. To see this, note that the exact definitions of Γ and Ψ are not needed. Further, we know that the degree of $\bar{q}_m$ is equal to $2m$ as the characteristic polynomial q_m has degree m. This implies that there should be another positive root. The following lemma shows that this is, in fact, the case. Further, this root is the largest root of $\bar{q}_m$.

Lemma 5.27. *For each natural m we have:* $\operatorname{sign}(\bar{q}_m(\alpha m)) = (-1)^m$, *where α is the positive solution of $2\alpha = \ln(\frac{\alpha+1}{\alpha-1})$, i.e., $\alpha = 1.19968$.*

Proof. Using the reflection formula for the Gamma function, we obtain

$$\Gamma(\tfrac{1}{2} + m - \alpha m) = \frac{1}{\Gamma(\frac{1}{2} - m + \alpha m)} \cdot \frac{\pi}{\sin((\frac{1}{2} + m - \alpha m)\pi)} .$$

Since $\sin((\frac{1}{2} + m - \alpha m)\pi) = \cos((m - \alpha m)\pi) = (-1)^m \cos(\alpha m\pi)$ and $\alpha m > m$, the sign of $\bar{q}_m(\alpha m)$ is equal to the sign of $\frac{(-1)^m}{\cos(\alpha m\pi))} q_m^\Psi(\alpha m)$. From the reflection formula for the digamma function Ψ, we get

$$\begin{aligned} \Psi(\tfrac{1}{2} + m - \alpha m) &= -\pi\cot((\tfrac{1}{2} + m - \alpha m)\pi) + \Psi(\tfrac{1}{2} - m + \alpha m) \\ &= -\pi\frac{\cos((\frac{1}{2} + m - \alpha m)\pi)}{\sin((\frac{1}{2} + m - \alpha m)\pi)} + \Psi(\tfrac{1}{2} - m + \alpha m) \\ &= -\pi\frac{\sin(\alpha m\pi)}{\cos(\alpha m\pi)} + \Psi(\tfrac{1}{2} - m + \alpha m) . \end{aligned}$$

This leads to

$$q_m^{\Psi}(\alpha m) := \cos(\alpha m\pi)\left(2\alpha m + m\left(\Psi(\tfrac{1}{2} - m + \alpha m) - \Psi(\tfrac{1}{2} + m + \alpha m)\right)\right) .$$

Hence, canceling m, the sign of $\bar{q}_m(\alpha m)$ is equal to

$$(-1)^m\left(2\alpha + \left(\Psi(\tfrac{1}{2} + (\alpha-1)m) - \Psi(\tfrac{1}{2} + (\alpha+1)m)\right)\right) .$$

Defining

$$\Delta\Psi_m := -\Psi(\tfrac{1}{2} + (\alpha-1)m) + \Psi(\tfrac{1}{2} + (\alpha+1)m) = \sum_{j=0}^{2m-1} \frac{1}{\frac{1}{2} + (\alpha-1)m + j}$$

we know that $\Delta\Psi_m$ converges to $\ln(\frac{\alpha+1}{\alpha-1})$ if m goes to infinity. Because of the definition of α, this implies

$$\lim_{m\to\infty}(2\alpha - \Delta\Psi_m) = 2\alpha - \ln(\tfrac{\alpha+1}{\alpha-1}) = 0 .$$

Hence, the lemma follows if $2\alpha - \Delta\Psi_m$ converges to 0 from right. Thus, we have to show, $\Delta\Psi_m \le \ln(\frac{\alpha+1}{\alpha-1})$. In order to prove this, we show the following: For each $m \in \mathbb{N}$ the subsequence $\Delta\Psi_m, \Delta\Psi_{2m}, \Delta\Psi_{4m}, \Delta\Psi_{8m}, \ldots$ is nondecreasing. To see this, we use inequality (2.2) between the arithmetic and harmonic mean and obtain for each $m \in \mathbb{N}$

$$\begin{aligned}
\Delta\Psi_{2m} - \Delta\Psi_m &= \sum_{j=0}^{4m-1} \tfrac{1}{\frac{1}{2}+(\alpha-1)2m+j} - \sum_{j=0}^{2m-1} \tfrac{1}{\frac{1}{2}+(\alpha-1)m+j} \\
&= \sum_{j=0}^{2m-1} \tfrac{1}{\frac{1}{2}+(\alpha-1)2m+2j} + \tfrac{1}{\frac{1}{2}+(\alpha-1)2m+2j+1} - \tfrac{1}{\frac{1}{2}+(\alpha-1)m+j} \\
&\ge \sum_{j=0}^{2m-1} \tfrac{4}{\frac{1}{2}+(\alpha-1)2m+2j+\frac{1}{2}+(\alpha-1)2m+2j+1} - \tfrac{1}{\frac{1}{2}+(\alpha-1)m+j} \\
&= \sum_{j=0}^{2m-1} \tfrac{1}{\frac{1}{2}+(\alpha-1)m+j} - \tfrac{1}{\frac{1}{2}+(\alpha-1)m+j} \\
&= 0 .
\end{aligned}$$

□

Note that the proof shows that α is as small as possible, i.e., for each smaller constant the last lemma fails. Thus, we get the final result of this section:

Theorem 5.28. *Let $\alpha \in \mathbb{R}_+$ be the unique positive real solution of the equation $2\alpha = \ln(\frac{\alpha+1}{\alpha-1})$, i.e., $\alpha = 1.19968$. Then the direct working price of stability is bounded by $\alpha^2 = 1.43923$. Moreover, this value is tight.*

Proof. This follows directly from the discussion above. Because of the substitutions we know that $\bar{\alpha}$ is a root of $\bar{q}_m$ if and only if $\tilde{\beta} := \bar{\alpha}^2 - \frac{1}{4}$ is a root of $\tilde{q}_m$. Further, $\tilde{\beta}$ is a root of $\tilde{q}_m$ if and only if $\beta := \frac{\tilde{\beta}}{m^2} - 1$ is a root of q_m. Since αm is an upper bound on the largest root of $\bar{q}_m$, this shows that $\beta = \frac{\alpha^2 m^2}{m^2} - \frac{1}{4m^2} - 1 = \alpha^2 - \frac{1}{4m^2} - 1$ is an upper bound on the largest root of $\bar{q}_m$. Hence, as the largest eigenvalue of q_m plus 1 is an upper bound on the direct working price of anarchy of the instances $\mathcal{I}_m$, we know that $\rho^{\text{PoS}}_{\text{work}} \leq \alpha^2$. In addition, we know that α^2 cannot be replaced by a smaller constant in this estimation, which completes the proof. □

5.4.4 Completion Time Price of Stability

In this subsection we analyze the completion time price of stability ρ_{comp} of the direct flow model. It turns out that ρ_{comp} is unbounded in general, i.e., if we consider time-varying capacities. For that reason we only compute ρ_{comp} for constant scenarios where the capacity on each edge remains constant over time. We show that for such scenarios the completion time price of stability is equal to $\frac{3}{4}$. The following lemma shows that $\frac{3}{4}$ is an lower bound. Subsequently, in Example 5.30, we see that $\frac{3}{4}$ also works as an upper bound. In addition, Example 5.30 shows that ρ_{comp} is unbounded for time-varying capacities.

Lemma 5.29. *The completion time price of anarchy for networks consisting only of parallel s-t-edges is lower bounded by $\frac{3}{4}$ if all edge capacities are constant.*

Proof. Consider an instance $\mathcal{I}_m$ consisting of m parallel s-t-edges, and let $\mathcal{F}_{\text{O}}$ and $\mathcal{F}_{\text{N}}$ be the earliest arrival and the Nash flow over time, respectively. Further, let ℓ_{m+1} be a point in time such that the completion time price of anarchy of $\mathcal{I}_m$ is attained at the flow value $\text{val}(\mathcal{F}_{\text{N}})(\ell_{m+1}) = d\theta_{m+1}$ where we require $\theta_{m+1} \in \ell^{-1}(\ell_{m+1})$. Without loss of generality, we assume $\theta_{m+1} \geq \theta_m$. Otherwise we reduce transit times greater than $\ell_{m+1} - \theta_{m+1}$ to this value. On this new instance the Nash flow is not changed until time θ_{m+1}, but the earliest arrival flow is improved. Because of (5.26) and (5.27), the flow pattern of $\mathcal{F}_{\text{O}}$ and the flow pattern of $\mathcal{F}_{\text{N}}$ satisfy at time ℓ_{m+1}

$$\begin{aligned} \text{val}(\mathcal{F}_{\text{O}})(\ell_{m+1}) &= \sum_{i=1}^{m} u_i(\ell_{m+1} - \tau_i) \\ \text{val}(\mathcal{F}_{\text{N}})(\ell_{m+1}) &= \sum_{i=1}^{m} u_i(\ell_{m+1} - \ell_i) \\ \Rightarrow \qquad \text{val}(\mathcal{F}_{\text{O}})(\ell_{m+1}) - \text{val}(\mathcal{F}_{\text{N}})(\ell_{m+1}) &= \sum_{i=1}^{m} u_i\theta_i \ . \end{aligned} \tag{5.36}$$

Let τ_{m+1} be a point in time such that the earliest arrival flow pattern attains the value $\text{val}(\mathcal{F}_{\text{N}})(\ell_{m+1})$ at time τ_{m+1}, i.e., $\text{val}(\mathcal{F}_{\text{O}})(\tau_{m+1}) = \text{val}(\mathcal{F}_{\text{N}})(\ell_{m+1})$. Hence, in order to estimate the completion time price of anarchy on this instance, we have to bound $\frac{\tau_{m+1}}{\ell_{m+1}}$. Having in mind that $\mathcal{F}_{\text{O}}$ need not to send flow into t along *all* edges at time τ_{m+1}, equation (5.26) shows

$$\begin{aligned} \text{val}(\mathcal{F}_{\text{O}})(\tau_{m+1}) &\geq \sum_{i=1}^{m} u_i(\tau_{m+1} - \tau_i) \\ \Rightarrow \qquad \text{val}(\mathcal{F}_{\text{O}})(\ell_{m+1}) - \text{val}(\mathcal{F}_{\text{O}})(\tau_{m+1}) &\leq \sum_{i=1}^{m} u_i(\ell_{m+1} - \tau_{m+1}) \ . \end{aligned} \tag{5.37}$$

From the definitions above we know that the left hand sides of (5.36) and (5.37) coincide which shows

$$u_{1\dots m}(\ell_{m+1} - \tau_{m+1}) \;\leq\; \sum_{i=1}^{m} u_{1\dots i}\theta_i \;. \tag{5.38}$$

In the following we estimate the right hand side of (5.38). Since $\theta_1 = 0$ and $u_{m+1\dots m} = 0$, we have

$$\sum_{i=1}^{m} u_i\theta_i \;=\; \sum_{i=1}^{m}(u_{i\dots m} - u_{i+1\dots m})\theta_i \;=\; \sum_{i=1}^{m-1} u_{i+1\dots m}(\theta_{i+1} - \theta_i) \;. \tag{5.39}$$

Next, we apply the inequality (2.1) between the arithmetic and the geometric mean for two nonnegative real numbers. Since we assume $d \geq u_{1\dots m}$, this inequality shows $du_{1\dots m} \geq 4u_{1\dots i}u_{i+1\dots m}$ for all $i \in [m]$. Thus, using (5.29) we get for all $i \in [m]$

$$u_{i+1\dots m}(\theta_{i+1} - \theta_i) \;=\; \frac{u_{1\dots i}u_{i+1\dots m}}{d}(\ell_{i+1} - \ell_i) \;\leq\; \frac{1}{4}u_{1\dots m}(\ell_{i+1} - \ell_i) \;.$$

Inserting this inequality in (5.39) implies

$$\sum_{i=1}^{m} u_i\theta_i \;\leq\; \sum_{i=1}^{m-1}\frac{1}{4}u_{1\dots m}(\ell_{i+1} - \ell_i) \;=\; \frac{u_{1\dots m}}{4}(\ell_m - \ell_1) \;\leq\; \frac{u_{1\dots m}}{4}(\ell_{m+1} - \tau_1)$$

as $\ell_{m+1} \geq \ell_m$ and $\ell_1 = \tau_1$ hold by definition. Together with (5.38) this finally leads to

$$(\ell_{m+1} - \tau_1) - (\tau_{m+1} - \tau_1) \;\leq\; \frac{1}{4}(\ell_{m+1} - \tau_1) \qquad \Leftrightarrow \qquad \frac{\tau_{m+1} - \tau_1}{\ell_{m+1} - \tau_1} \;\geq\; \frac{3}{4} \;.$$

As $\tau_1 \geq 0$ holds, this completes the proof. □

Lemma 5.29 shows that the direct completion time price of stability for constant scenarios is bounded from below by $\frac{3}{4}$. The following example illustrated in Figure 5.10 shows that $\frac{3}{4}$ is the exact value of the completion time price of stability. In addition, the example observes that the completion time price of stability is unbounded for the general case of time-varying capacities.

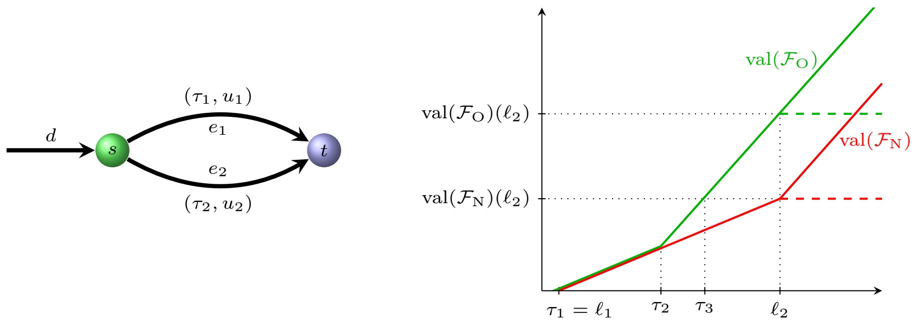
Example 5.30. As in Example 5.21, we consider a network consisting of two parallel s-t-edges (see Figure 5.10a). Using the notation of Example 5.21, we first observe that the completion time price of anarchy is attained for the flow value $\mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_2)$. Since ℓ_2 is the earliest point time at which arrives at t via e_2, we know that the flow patterns of the earliest arrival and the Nash flow over time behave the same way for larger flow values (see Figure 5.10b).

Using the notation of Lemma 5.29, we have to compute the ratio $\frac{\tau_3}{\ell_2}$. Following the corresponding prove, we observe that inequality (5.37) is satisfied with equality. Hence, (5.38) reads as

$$(u_1 + u_2)(\ell_2 - \tau_3) = u_2\theta_2 \;. \tag{5.40}$$

because $\theta_1 = 0$. Using equation (5.29) this shows

$$(u_1 + u_2)(\ell_2 - \tau_3) = \frac{u_1u_2}{d}(\ell_2 - \ell_1) \;.$$



(a) The network of $\mathcal{I}_2$.

(b) The earliest arrival and the Nash flow pattern. The dashed pattern refers to the instance where capacity is set to 0 after time ℓ_2.

Figure 5.10: Illustration of Example 5.30.

Setting $\ell_1 = \tau_1 := 0$, we obtain

$$(u_1 + u_2)(\ell_2 - \tau_3) \;=\; \frac{u_1 u_2}{d} \cdot \ell_2 \qquad \text{implying} \qquad \frac{\tau_3}{\ell_2} \;=\; 1 - \frac{u_1 u_2}{d(u_1 + u_2)} \;.$$

As implied by Theorem 5.20, this verifies that the completion time price of anarchy goes to 1 if d goes to ∞. Further, we see that the completion time price of anarchy increases if d decreases. Hence, in order to maximize the completion time price of anarchy, we set $d := u_1 + u_2$. This results in

$$\frac{\tau_3}{\ell_2} = \frac{3}{4} + \frac{1}{4} \cdot \frac{(u_1 - u_2)^2}{(u_1 + u_2)^2} \;. \tag{5.41}$$

Thus, if an instance satisfies $\tau_2 > \tau_1 = 0$, $d = u_1 + u_2$, and $u_1 = u_2$, the completion time price of anarchy of this instance is equal to $\frac{3}{4}$. This shows that the completion time price of anarchy is upper bounded by $\frac{3}{4}$.

In order to see that the completion time price of anarchy is unbounded for general time-varying capacities, we consider a point in time $\theta \in \mathbb{R}_+$ for which $\mathrm{val}(\mathcal{F}_\mathrm{O})(\theta) > \mathrm{val}(\mathcal{F}_\mathrm{N})(\theta)$ holds. Construct a new instance by reducing all capacities to zero after time θ. Hence, the Nash flow on this new instance never reach the value $\mathrm{val}(\mathcal{F}_\mathrm{O})(\theta)$ whereas the earliest arrival flow does. This shows that on the new instance the completion price of anarchy for the flow value $\mathrm{val}(\mathcal{F}_\mathrm{O})(\theta)$ is equal to ∞.

Summarizing the discussion in this subsection, we obtain the following theorem.

Theorem 5.31. *The completion time price of stability for the direct flow model is equal to $\frac{3}{4}$ if we restrict to networks with constant edge capacities.*

Proof. Using Theorem 5.19 this follows directly from Lemma 5.29 and Example 5.30. □

5.4.5 Average Arrival Time Price of Stability

This subsection is devoted to the analysis of the average arrival time price of stability. Unfortunately, we are not able to provide a complete analysis.

The following lemma shows that the lower bound of $\frac{3}{4}$ carries over from the completion time to the average arrival time price of stability whereas the rest of this subsection shall lead to the insight that this bound is not tight. In particular, we compute the exact price of anarchy for instances $\mathcal{I}_2$ consisting of two parallel edges for the flow value $\mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_2)$. Subsequently, we present some empirical results corroborating that $\frac{3}{4}$ underestimates the direct average arrival time price of stability. Because of the similarities it is quite natural to hope that the approach in Subsection 5.4.3 can be transferred to this case. Problems occurring in this connection are briefly discussed at the end of this subsection.

But before we start, note that the direct price of stability is in general unbounded. This is a particular consequence of the discussion at the end of Example 5.30. For this reason, we always consider constant scenarios in this subsection.

Lemma 5.32. *The average arrival time price of stability for the direct flow model is lower bounded by $\frac{3}{4}$ if edge capacities are constant.*

Proof. Recall the definition of the average arrival time price of stability in Section 4.3. Consider an instance with constant capacities and let $\mathcal{F}_\mathrm{O}$ and $\mathcal{F}_\mathrm{N}$ be an earliest arrival and a Nash flow over time. Because of Theorem 5.31 we obtain for each flow value $F \in \mathbb{R}_+$

$$\begin{aligned}\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{O})(F) &= \int_0^F \mathrm{val}_\mathrm{comp}(\mathcal{F}_\mathrm{O})(F')\,dF' \\ &\geq \int_0^F \frac{3}{4}\mathrm{val}_\mathrm{comp}(\mathcal{F}_\mathrm{N}))(F')\,dF' = \frac{3}{4}\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{N})(F)\,.\end{aligned}$$

This establishes the lemma. □

Lemma 5.32 shows that the lower bound for the completion time price of stability works also as a lower bound for the average completion time price of stability. Further, this is established in a very rudimentary manner. Recalling Example 5.30, we know that the bound is attained on an instance $\mathcal{I}_2$ for the flow value $\mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_2)$. Thus, a very natural question is whether or not this does also hold for the average arrival time price of anarchy on instances $\mathcal{I}_2$. The following lemma shows that this is not the case. In fact, the bound of $\frac{3}{4}$ for the average completion time price of anarchy on $\mathcal{I}_2$ can be significantly improved.

Lemma 5.33. *Consider an instance $\mathcal{I}_2$ consisting of two parallel s-t-edges. The average arrival time price of anarchy for the flow value $F_2 := \mathrm{val}(\mathcal{F}_N)(\ell_2)$ is equal to $\frac{23}{27} = 0.\overline{851}$ if we restrict to constant capacities.*

Proof. In order to compute the average arrival time price of anarchy for the flow value F_2, take a look at Figure 5.11. It shows that the value $\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{O})(F_2)$ is equal to the total area covered by the yellow, orange, green, and blue surfaces. For computing the value $\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{N})(F_2)$, we have to add the area of the gray surface to $\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{O})(F_2)$. Since the total area of the orange, green, blue, and gray surfaces is equal to the one of the red surface, we obtain

$$\begin{aligned}2\cdot\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{O})(F_2) &= u_1(\tau_2-\tau_1)^2 + (u_1+u_2)(\tau_3-\tau_2)^2 \\ &\quad \tau_1 u_1(\ell_2-\ell_1)^2 + 2(u_1+u_2)(\tau_3-\tau_2)(\tau_2-\tau_1) \\ 2\cdot\mathrm{val}_\mathrm{aver}(\mathcal{F}_\mathrm{N})(F_2) &= \tau_1 u_1(\ell_2-\ell_1)^2 + u_1(\ell_2-\ell_1)^2\,.\end{aligned}$$

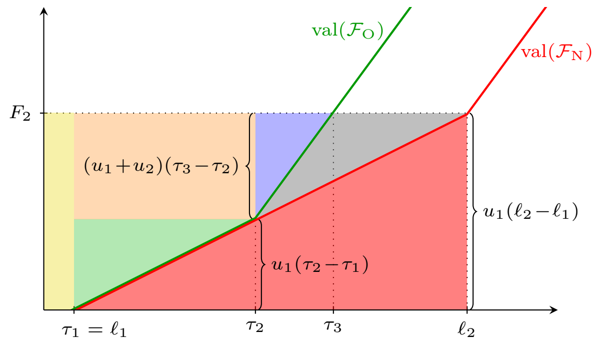


Figure 5.11: Computation of the average arrival objective function on an instance $\mathcal{I}_2$.

Since the yellow area occurs in both sums, the quotient $\frac{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{O}})(F_2)}{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{N}})(F_2)}$ becomes larger if the yellow area is reduced. This enables us to assume without loss of generality $\ell_1 = \tau_1 := 0$. This results in

$$\begin{aligned} 2\cdot\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{O}})(F_2) &= u_1\tau_2^2 + (u_1+u_2)(\tau_3-\tau_2)^2 + 2(u_1+u_2)(\tau_3-\tau_2)\tau_2 \;. \\ 2\cdot\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{N}})(F_2) &= u_1\ell_2^2 \;. \end{aligned} \tag{5.42}$$

From the equations (5.29) and (5.31) we get

$$\frac{\theta_2}{\ell_2} = \frac{u_1}{d} \qquad \text{and} \qquad \frac{\tau_2}{\ell_2} = \frac{d-u_1}{d} \;. \tag{5.43}$$

Since Example 5.30 analyzes the same scenario (but with a different objective), we know $u_2\theta_2 = (u_1+u_2)(\ell_2-\tau_3)$ because of (5.40) which shows

$$\tau_3-\tau_2 = \ell_2 - \frac{u_2}{u_1+u_2}\cdot\theta_2 - \tau_2 = \left(1-\frac{u_2}{u_1+u_2}\right)\cdot\theta_2 = \frac{u_1}{u_1+u_2}\cdot\theta_2$$

implying

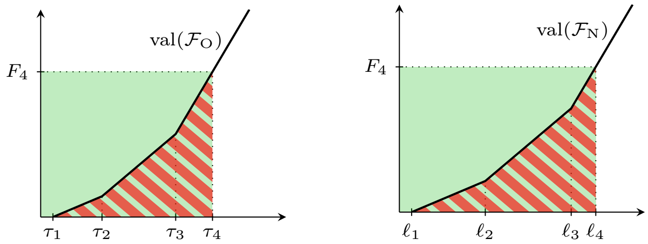
$$\frac{\tau_3-\tau_2}{\ell_2} = \frac{u_1}{u_1+u_2}\cdot\frac{\theta_2}{\ell_2} = \frac{u_1^2}{d(u_1+u_2)} \;. \tag{5.44}$$

A closer look at (5.42), (5.43), and (5.44) shows that $\frac{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{O}})(F_2)}{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{N}})(F_2)}$ becomes smaller if we increase d. Hence, without loss of generality, we assume $d = u_1+u_2$. By (5.42), (5.43), and (5.44), this shows

$$\begin{aligned} \frac{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{O}})(F_2)}{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{N}})(F_2)} &= \left(\frac{u_2}{u_1+u_2}\right)^2 + \frac{u_1+u_2}{u_1}\cdot\left(\frac{u_1^2}{(u_1+u_2)^2}\right)^2 \\ &\qquad + \frac{2(u_1+u_2)}{u_1}\cdot\frac{u_1^2}{(u_1+u_2)^2}\cdot\frac{u_2}{u_1+u_2} \\ &= \frac{u_2^2}{(u_1+u_2)^2} + \frac{u_1^3}{(u_1+u_2)^3} + \frac{2u_1u_2}{(u_1+u_2)^2} \;. \end{aligned}$$

Using the binomial identities in order to simplify the sum of the first and the third summand in the last term, we get

$$\frac{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{O}})(F_2)}{\mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{N}})(F_2)} = 1 + \frac{u_1^3}{(u_1+u_2)^3} - \frac{u_2^2}{(u_1+u_2)^2} = 1 - \frac{u_1u_2^2}{(u_1+u_2)^3}$$



(a) Earliest arrival flow over time pattern.

(b) Nash flow over time pattern.

Figure 5.12: Geometric computation of the average arrival time objective function on an instance $\mathcal{I}_m$.

Finally, by the inequality between the geometric and the arithmetic mean we know

$$u_1^2 u_2 \;=\; 4\frac{u_1}{2}\frac{u_1}{2}u_2 \;\leq\; 4\Big(\frac{\frac{u_1}{2}+\frac{u_1}{2}+u_2}{3}\Big)^3 \;=\; \frac{4}{27}(u_1+u_2)^3 \;.$$

where equality holds if and only if $u_1 = 2u_2$. This shows that we obtain a value of $1 - \frac{4}{27} = \frac{23}{27} = 0.\overline{851}$ for the average arrival time price of anarchy on this instance. Further, this bound is tight for all instances $\mathcal{I}_2$ satisfying $\tau_2 > \tau_1 = 0$, $u_1 = 2u_2$, and $d = u_1 + u_2$. □

Lemma 5.33 computes the average arrival time price of anarchy for $\mathcal{I}_2$ at the flow value $F_2 = \mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_2)$. Unfortunately, the average arrival time price of anarchy of $\mathcal{I}_2$ is in general not attained at F_2. Considering an instance leading to the upper bound $\frac{23}{27} = 0.\overline{851}$ we observe that $\frac{\tau_3}{\ell_2} = \frac{28}{36} = 0.\overline{7} < \frac{23}{27}$ holds by (5.41). Hence, increasing the flow value decreases the corresponding average arrival time price of anarchy. In fact, the exact price of anarchy of an instance $\mathcal{I}_2$ is equal to $\frac{1}{18}\cdot(9+\sqrt{33}) = 0.819142$. This value is attained if we set the parameters of $\mathcal{I}_2$, e.g., to

$$\begin{aligned} &u_1 = 8\;, \quad u_2 = 1+\sqrt{33} = 6.74446\;, \\ &\ell_1 = 0\;, \quad \ell_2 = 2\;, \qquad\qquad \text{and} \quad \ell_3 = -3+\sqrt{33} = 2.74456\;. \end{aligned}$$

These values are found by using a computer algebra system. With such a system we improve the current upper bound of $\frac{1}{18}\cdot(9+\sqrt{33})$ for the average arrival time price of anarchy on instances $\mathcal{I}_m$ systematically. For this consider an instance $\mathcal{I}_m$ consisting of m parallel s-t-edges and let ℓ_{m+1} be a point in time such that the average arrival time price of anarchy ρ_aver of $\mathcal{I}_m$ is attained at $F_{m+1} := \mathrm{val}(\mathcal{F}_\mathrm{N})(\ell_{m+1})$. As done before, we are able to assume that ℓ_{m+1} is not smaller than ℓ_m. Further, let τ_{m+1} be the time at which the earliest arrival flow pattern achieves a value of F_{m+1}, i.e., $\mathrm{val}(\mathcal{F}_\mathrm{O})(\tau_{m+1}) = F_{m+1}$. Without loss of generality, we assume $\tau_{m+1} \geq \tau_m$. Otherwise we delete the edges with a transit time greater than τ_{m+1}. This has no influence on the earliest arrival flow pattern until time τ_{m+1} but potentially weakens the performance of $\mathcal{F}_\mathrm{N}$. Taking a look at Figure 5.12, we see that the average arrival time objective value for F_4 of $\mathcal{F}_\mathrm{O}$ and $\mathcal{F}_\mathrm{N}$, respectively, equals the area of the green rectangle minus the the area of the red surface. Note that the area of the red surface equals the working objective value at τ_{m+1} and ℓ_{m+1}, respectively. Thus, from (5.26)/(5.27) and (5.32)/(5.33) we get

$$\begin{aligned}
2 \cdot \mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{O}})(F_{m+1}) &= 2\tau_{m+1} \cdot \mathrm{val}(\mathcal{F}_{\mathrm{O}})(\tau_{m+1}) - 2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{O}})(\tau_{m+1}) \\
&= 2\tau_{m+1} \cdot \sum_{i=1}^{m} u_i(\tau_{m+1} - \tau_i) - \sum_{i=1}^{m} u_i(\tau_{m+1} - \tau_i)^2 \\
&= \sum_{i=1}^{m} u_i(\tau_{m+1}^2 - \tau_i^2) \\
&= u_{1..m}\tau_{m+1}^2 - \sum_{i=1}^{m} u_i \tau_i^2 \ , \\
2 \cdot \mathrm{val}_{\mathrm{aver}}(\mathcal{F}_{\mathrm{N}})(F_{m+1}) &= 2\ell_{m+1}\mathrm{val}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1}) - 2 \cdot \mathrm{val}_{\mathrm{work}}(\mathcal{F}_{\mathrm{N}})(\ell_{m+1}) \\
&= 2\ell_{m+1} \sum_{i=1}^{m} u_i(\ell_{m+1} - \ell_i) - \sum_{i=1}^{m} u_i(\ell_{m+1} - \ell_i)^2 \\
&= \sum_{i=1}^{m} u_i(\ell_{m+1}^2 - \ell_i^2) \\
&= u_{1..m}\ell_{m+1}^2 - \sum_{i=2}^{m} u_i \ell_i^2 \ .
\end{aligned}$$

Now we try to proceed similarly to the approach in Subsection 5.4.3 where we compute the working price of stability for constant scenarios. Firstly, we want make the same simplifying assumption that all capacities are 1. Unfortunately, it turns out that optimizing the average arrival time price of anarchy on $\mathcal{I}_m$ causes several edges with zero transit times. In fact, these zero transit times where enforced by nonnegativity constraints which would directly imply that the approach of Subsection 5.4.3 does not work here.

However, several edges of zero transit times with a capacity of one can be interpreted as one edge of zero transit time with an arbitrary capacity. For this reason, we do not fix the capacity u_1 of e_1. Thus, setting the remaining capacities to 1, we have to minimize

$$\frac{(u_1 + m - 1)\tau_{m+1}^2 - \sum_{i=1}^{m} \tau_i^2}{(u_1 + m - 1)\ell_{m+1}^2 - \sum_{i=2}^{m} \ell_i^2} \ .$$

Using (5.38) which holds with equality because of the assumption $\tau_{m+1} \geq \tau_m$ and the equations (5.29)-(5.31), we are able to express the numerator in terms of $u_1, \ell_2, \ldots, \ell_{m+1}$. Since scaling of the variables $\ell_2, \ldots, \ell_{m+1}$ does not change the value of the fraction, we are able to assume that the denominator is 1. Hence, solving the optimization problem

$$\min\Big\{(u_1 + m - 1)\tau_{m+1}^2 - \sum_{i=1}^{m} \tau_i^2 \;\Big|\; (u_1 + m - 1)\ell_{m+1}^2 - \sum_{i=2}^{m} \ell_i^2 = 1\Big\} \qquad (5.45)$$

gives the right value for the average arrival time price of anarchy on instances $\mathcal{I}_m$. Table 5.1 shows the exact average arrival time price of anarchy for different $m \in$

m	1	2	3	4	5	10
ρ_{aver}	1	0.819142	0.807091	0.803768	0.802384	0.800707
$\frac{u_1}{m-1}$	∞	1.18614	0.94152	0.85533	0.81050	0.73221

m	15	20	25	50	75	100
ρ_{aver}	0.800423	0.800327	0.800283	0.800225	0.800215	0.800211
$\frac{u_1}{m-1}$	0.70884	0.69760	0.69099	0.67807	0.67384	0.67175

m	150	200	250	500	750	1000
ρ_{aver}	0.800209	0.800208	0.800207	0.800207	0.	0.
$\frac{u_1}{m-1}$	0.66966	0.66862	0.66800	0.66676	0.	0.

Table 5.1: The direct average arrival time price of anarchy of instances $\mathcal{I}_m$ for different $m \in \mathbb{N}$. The value u_1 of a tight instance is also shown.

$\mathbb{N}$ computed with a computer algebra system. It seems that the average arrival time price of anarchy converges to a value which lies *slightly above* 0.8. Further, we see that the capacity u_1 of e_1 converges to something around $\frac{2}{3}$ if we scale it down by $m-1$. I do *not* believe that this limit is exactly $\frac{2}{3}$. For the values of m shown in the Table 5.1 the corresponding values for $\ell_2, \dots, \ell_{m+1}$ are shown in Figure 5.13. In this figure each ℓ_i is scaled with $\sqrt{m}$ and this scaled value can be read off the point $\frac{i-1}{m}$. Hence, for a given m, the values ℓ_i correspond to the points $(\frac{i-1}{m}, \sqrt{m}\ell_i)$ which are linearly interpolated. We see that in this manner the ℓ_i's also converge nicely to some function if m goes to infinity.

Finally, we briefly discuss the problems in case we want to apply the approach of Subsection 5.4.3 to this case. As already mentioned, the first difference is that we cannot set all capacities to 1. Especially the capacity u_1 of e_1 cannot be fixed. Further, taking a look at (5.45) we see that we can easily normalize the set over which we optimize – substitute ℓ_{m+1} with $\frac{\tilde{\ell}_{m+1}}{\sqrt{u_1+m-1}}$. This shows that, in contrast to Subsection 5.4.3, we have to optimize over a standard hyperbel instead of a unit circle. Nevertheless, all this is manageable and makes not really a problem. The problem occurs after we simplify the matrix of the quadratic form. Instead of a 3-diagonal we obtain a 5-diagonal matrix. Also for the determinant of a 5-diagonal matrix a recursion exists. But the order of this recursion is 5. Finding the expicit formula of this recursion is the problem. The computer algebra system I used could not come up with a suitable result and so the approach stopped.

5.5 Your Comments

This chapter analyzes a path-based routing game over time which is based on the classical flow over time model where waiting at nodes is forbidden. A precise notion of this so-called direct flow over time model is given. Further, the Classical Successive Shortest Path algorithm is generalized to arbitrary time varying capacities such that it computes an earliest arrival flow. Besides, Nash equilibria are characterized and a nearly complete analysis of the performance of direct Nash flows over time is presented. This includes an approach to each of the four objective functions introduced in Section 4.3 with respect to the price

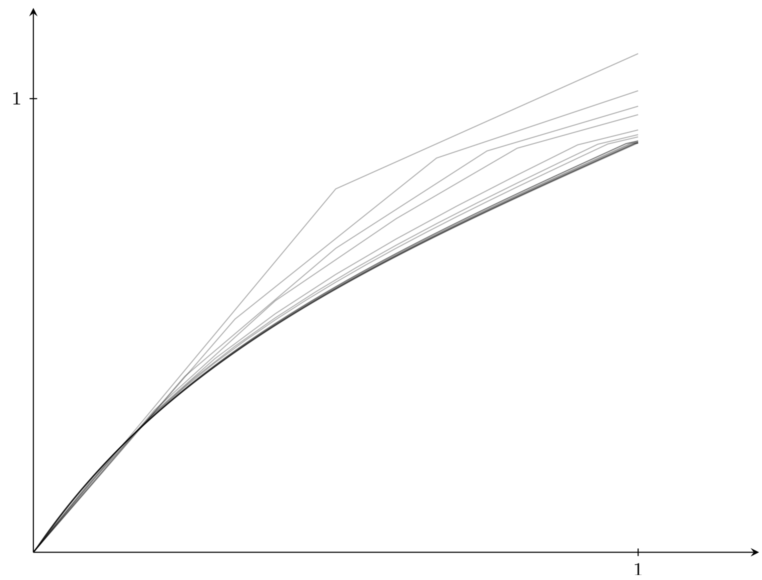


Figure 5.13: The arrival times $\ell_1, \ldots, \ell_{m+1}$ of the tight instances.

of anarchy just as the price of stability.

In literature, direct routing is rarely considered explicitly. In the area of packet routing, the question arises if given packets can be routed along predetermined paths without waiting at intermediate nodes. This problem is related to discrete flows over times and, firstly, considered by Symvonis [86] on trees. Other articles in this direction are, e.g., [13, 70]. Aspects of direct routing in the continuous model are considered by Fleischer and Skutella in [23, 24]. Here, the question is how much is lost if storage is forbidden. The answer, one loses some factor between $\frac{4}{3}$ and 2 for instances consisting of multiple commodities. By personal communication to the authors, you have heard that someone spends a bottle of champagne for closing this gap. Clearly, as you already mentioned, for s-t-scenarios nothing is lost as there exist an earliest arrival flow where storage does not occur (see, e.g., [56, 91] for the discrete setting and [25] for the continuous model). Maybe, this is one reason, why direct routing is not that often explicitly considered in classical flow over time theory.

For direct flows over time, this thesis provides the first study of Nash equilibria. Further, as far as you could observe, Section 4.3 provides the first complete nontrivial analysis of the performance of Nash equilibria over time for a given nonatomic routing game over time.

Concerning the approach presented in Subsection 5.4.3 where the direct working price of stability is evaluated, you find out that the recursion for computing the determinant of a 3-diagonal matrix is common knowledge. Nevertheless, you could not find out whether or not the presented approach for maximizing a quadratic form over the unit sphere were already used somewhere else. You decide that if you find an article using this approach, you will inform the author. Concerning the recursion formula for computing the determinant of a 5-diagonal matrix, you find the article of Sweet [85]. Maybe, this helps to evaluate the exact value of the direct average arrival time price of stability for constant capacities.

Turning your attention to the contribution of this chapter, you observe that equation (5.1) is quite similar to the *static* network loading constraint. Especially, if all transit times are 0, the corresponding notions coincide, i.e.,

$$\sum_{P,i|e_i^P=e} F_P = F_e$$

for all $e \in E$. Maybe, this simple relation is the very reason why Nash equilibria are characterizable and analyzable for the direct flow model. On the other hand, the price for this simple equation is a rather complex definition of the direct flow model. It seems to be consistent, but the $\mathcal{F}$-continuity is only partially proven. Nevertheless, you strongly believe that $\mathcal{F}$-continuity holds squarely.

In Section 5.2 it is shown that the concept of earliest arrival flows carries over from classical flows over time which is already observed by Philpott [71]. The interesting and novel thing is the approach for computing a direct earliest arrival flow with the SUCCESSIVE SHORTEST PATH algorithm. Clearly, it can be seen as the generalization of the CLASSICAL SUCCESSIVE SHORTEST PATH algorithm. But rather, it is a direct generalization of the STATIC SUCCESSIVE SHORTEST PATH algorithm to flows over time. Such an approach is also taken in [43] where the *static* MaxFlow-MinCut Theorem is directly extended to (measure-based) flows over time.

Direct Nash equilibria are considered in Section 5.3 and it is shown that they show an arbitrary bad performance compared with an earliest flow. Nevertheless, so called EA-Nash flows, which are constructed from an earliest arrival flow, show a quite good behavior. Especially, if there is a huge demand, a EA-Nash flow is nearly optimal. Possible consequences for real world applications, you discuss in Section 7.5 after you know something about the flow behavior with respect to the deterministic queuing model . However, you note already here that forbidding waiting seems to be a good tool for improving network performance.

Another even more promising feature of the EA-Nash flows is that they correspond to direct Nash flows on a network consisting only of parallel s-t-edges. Thus, every price of anarchy on these simple networks equals the corresponding direct price of stability which are analyzed in Section 5.4. For arbitrary Lebesgue integrable capacity functions, the evacuation just as the working pricing of stability is equal to 2. Further, the completion and average arrival time price of stability is unbounded. For constant capacity functions, the working price of stability decreases to $\alpha^2 = 1.439$ where $\alpha \in \mathbb{R}_+$ is a unique solution of $2\alpha = \log \frac{\alpha+1}{\alpha-1}$. Moreover, the completion and average arrival time price of anarchy are bounded by $\frac{3}{4}$ for constant scenarios. It is worth to mention that all bounds are tight – except for the average arrival time price of anarchy.

Clearly, a tight bound for average arrival time price of anarchy in constant scenarios would be great as this would close the last gap. Besides, simplifying the approach leading to the working price of anarchy would also be nice. Maybe, as you now know the exact value, a significantly shorter proof is possible. Also a simple sequence of instances converging to the working price of stability could be of interest. Having in mind traffic lights, you think that computing the prices of stability for periodic capacity functions seems to be promising. In fact, it would be nice to see how some price of stability tends from the value for constant scenarios to the value of arbitrary scenarios when increasing the period length.

Besides these questions concerning the performance of Nash flows, establishing $\mathcal{F}$-continuity not only partially would complete the basic treatment of the direct model. And, of course, one promising aspect for future research should not be forgotten: Generalizing the notion of direct flow over time models using flow measures instead of flow functions.

Chapter 6

Static Thin Flows

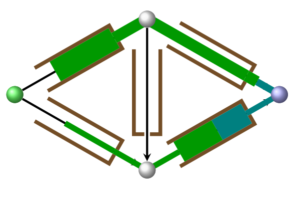
This chapter introduces a new class of *static* network flows called thin flows where flow is distributed evenly over the edges. They play a fundamental role in analyzing Nash flows over time for the deterministic queuing model as they occur as time derivatives of the underlying static flow (see Chapter 7). Beside this, due to its special structure, thin flows are worth to be studied by its own.

Before we start defining thin flows in Section 6.2, we provide all necessary basic knowledge of static flow theory in Section 6.1. Here, we present also network related notations used throughout this chapter. Section 6.2 also contains first technical properties of thin flows. At first glance, it is not clear whether or not thin flows exist in general. With this in mind, we present an interesting relationship of general thin flows to a subclass called thin flows without resetting. As we see in Section 6.3, this subclass is well-analyzable. In particular, thin flows without resetting are computable in polynomial time and are unique in some sense which we show in Subsection 6.3.1. Further, we provide a structural and sensitivity analysis for thin flows without resetting in Subsection 6.3.2 and 6.3.3, respectively. Finally, all of this work is needed to establish existence and uniqueness results presented in Section 6.4. We show that thin flows always exist and analyze under which circumstances they are unique.

6.1 Preliminaries

In this section we impart basic knowledge about static flow theory which we need throughout this chapter. Some very essential notations are already presented in Section 2.3. Recall that static flows act on directed graphs $G := (V, E)$ consisting of a finite set of nodes V and a finite set of edges E. Each edge connects some node $v \in V$ to another node $w \in V$. In this case we write $e = vw$ and call v the tail and w the head of e. A path is a sequence of edges where the head of one edge is the tail of the next. It is called v-w-path if v is the tail of the first edge and w is the head of the last edge. A v-v-path is called cycle and a graph G is called acyclic if it does not contain any cycle. For characterizing acyclic graphs by the subsequent well-known lemma, we need the following definition.

Definition 6.1 (Topological Order)**.** Let $G := (V, E)$ be a directed graph. An ordering $v_1, \dots, v_{|V|}$ of the nodes is called *topological order* if and only if $i < j$



holds for each edge $v_i v_j \in E$. This means there exists no edge leading from a node with a larger index to a node with lower index.

Lemma 6.2. *A directed graph is acyclic if and only if it supports a topological order.*

In the following, we define cuts which play a fundamental role in static flow theory.

Definition 6.3 (Cut)**.** For a directed graph $G := (V, E)$, a node subset $X \subseteq V$ is called *cut.* The set $\delta^+(X) := \{vw \in E \mid v \in X, w \in V \setminus X\}$ contains all edges leaving the cut. In contrast, the set of edges entering the cut are denoted by $\delta^-(X) := \{wv \in E \mid v \in X, w \in V \setminus X\}$. Hence, $\delta(X) := \delta^+(X) \cup \delta^-(X)$ are all edges crossing the cut. A cut is called *directed* if $\delta^-(X) = \emptyset$ and *trivial* if $\delta(X) = \emptyset$. If $s \in X$ holds for some node $s \in V$, we call X an *s-cut.* If, in addition, $t \in V \setminus X$ holds for some other node $t \in V$, we call X an *s*-*t*-cut.

Usually, a cut is identified with its outgoing edges. So we also use $\delta^+(X)$ for denoting the cut X. If X consists only of node $v \in V$ we also write $\delta^+(v)$, $\delta^-(v)$, and $\delta(v)$ for $\delta^+(\{v\})$, $\delta^-(\{v\})$, and $\delta(\{v\})$, respectively.

Further, we use $E(X_1, X_2) := \{vw \in E \mid v \in X_1, w \in X_2\}$ to denote all edges leaving a node set X_1 and entering another node set X_2.

Throughout this chapter we often consider a static flow on a part of the underlying graph. Clearly, a subgraph can be constructed by deleting several edges and nodes. Beside this, contraction is also an important operation on graphs. The contraction of a part of a graph G shows how this part is connected to the rest of G. In the following, we define these operations formally leading to the notion of a minor of a graph.

Definition 6.4 ((Induced) Subgraph)**.** Let $G := (V, E)$ be some directed graph. Then H is a *subgraph* of G if and only if H is a directed graph with $V(H) \subseteq V$ and $E(H) \subseteq E$. In this case we write $H \subseteq G$. Further, H is an *induced subgraph* of G if and only if $V(H) \subseteq V$ and $E(H) = \{vw \in E \mid v, w \in V(H)\}$. In this case we also say that H is the subgraph of G induced by $X := V(H)$ and write $H = G[X]$.

Definition 6.5 (Removing Nodes and Edges)**.** Let $G := (V, E)$ be a directed graph and $E' \subseteq E$ be an edge subset. The subgraph $G \setminus E'$ resulting out of G by deleting all edges in E' is defined by $V(G \setminus E') := V$ and $E(G \setminus E') := E \setminus E'$. Given a node subset $X \subseteq V$ the graph $G \setminus X$ resulting out of G by deleting all nodes in X equals the subgraph of G which is induced by $V \setminus X$, i.e., $G \setminus X := G[V \setminus X]$. In case E' and X consists of a single edge $e \in E$ and of a single node $v \in V$ we also write $G \setminus e$ and $G \setminus v$ instead of $G \setminus \{e\}$ and $G \setminus \{v\}$, respectively

Definition 6.6 (Contraction)**.** Let $G := (V, E)$ be a directed graph and $X \subseteq V$ be a node subset. The *contraction* of G with respect to X is a directed graph H constructed as follows:

(i) Delete X from G and add a new node v_X to G.

(ii) For every edge vw and wv connecting V and $V \setminus X$ with $v \in X$ and $w \in V \setminus X$ insert an edge $v_X w$ and $w v_X$, respectively.

More formally, H is defined as follows:

$$\begin{aligned} V(H) &:= \quad (V \setminus X) \cup \{v_X\} \\ E(H) &:= E\big(G[V \setminus X]\big) \cup \big\{v_X w \;\big|\; vw \in \delta^+(X)\big\} \\ &\qquad\qquad\qquad\quad \cup \big\{w v_X \;\big|\; wv \in \delta^-(X)\big\}. \end{aligned}$$

Note that we do not merge parallel edges arising out of this construction.

Let $\mathcal{X}$ be a family of pairwise disjoint subsets of V, i.e., $X \subseteq V$ for all $X \in \mathcal{X}$ and $X_1 \cap X_2 = \emptyset$ for all $X_1, X_2 \in \mathcal{X}$ with $X_1 \neq X_2$. The contraction of G with respect to $\mathcal{X}$ is obtained by contracting each node subset $X \in \mathcal{X}$ consecutively.

Definition 6.7 (Minor)**.** Let $G := (V, E)$ be a directed graph. A directed graph H resulting out of G by deleting and contracting operations is called *minor*.

In the following, we define static flows as in Section 2.3 and discuss related concepts as the congestion of an edge and the capacity of a cut. In addition, we extend the notion of congestion to paths and cuts leading to the definition of the label of a node and of a sparsest cut. For formalizing these concepts, we use the following convention. Whenever a set of indices is assigned to a vector, this is interpreted as sum of the corresponding values over these indices. For example, given a static flow $x \in \mathbb{R}_+^E$ and subset $X \subseteq V$ of the nodes, $x(\delta^+(X)) := \sum_{e \in \delta^+(X)} x_e$ equals the amount of flow leaving X. Finally, we state the famous Max-Flow-Min-Cut Theorem which is one of the supporting pillars of static flow theory.

Definition 6.8 ((Static) Flow)**.** Let $G := (V, E)$ be a directed graph. Then a nonnegative real vector $x \in \mathbb{R}_+^E$ is called *static flow*. *Node balances* are defined by a real vector $b \in \mathbb{R}^V$ satisfying $b(V) := \sum_{v \in V} b_v = 0$. A node $v \in V$ is called *source* if $b_v > 0$ and *sink* if $b_v < 0$. We say that a static flow x *respects* node balances b if and only if the following *flow conservation constraint* is satisfied:

$$x\big(\delta^+(v)\big) - x\big(\delta^-(v)\big) = b_v \qquad\qquad \forall v \in V \ .$$

A flow respecting certain node balances is also called b-flow. Further, a b-flow x is called s-flow if $b_v \leq 0$ holds for all $v \in V \setminus \{s\}$ and s-t-flow if, in addition, $b_v = 0$ holds for all $v \in V \setminus \{s, t\}$. In this case s is the unique source and t is the unique sink. Further, the *flow value* $|x|$ of an s-flow x is defined as $|x| := b_s$.

Given a nonnegative real vector $u \in \mathbb{R}_+^E$ interpreted as *edge capacities* we say that x obeys *capacity constraints* if and only if $x \leq u$ holds. In this spirit a *maximum s-t-flow* is an s-t-flow x maximizing the flow value $|x|$ while satisfying the capacity constraints.

Definition 6.9 (Congestions and Labels)**.** Let x be a static flow on a directed graph G and let $u \in \mathbb{R}_+^E$ be edge capacities. The *congestion* of an edge $e \in E$ is defined as $\frac{x_e}{u_e}$. Further, the congestion of some path $P := (e_1, \ldots, e_{|P|})$ equals the maximum edge congestion over all of its edges, i.e., $\max\{\frac{x_{e_i}}{u_{e_i}} \,|\, i = 1, \ldots, |P|\}$.

Assuming that x is an s-flow for some given source $s \in V$ we define the *(flow) label* ℓ_v of a node $v \in V$ as the minimum congestion over all s-v-path. The *label* ℓ_e of an edge $e = vw \in E$ is given by $\max\{\ell_v, \frac{x_e}{u_e}\}$. Hence, the node labels are given by

$$\ell_s := 0 \qquad \text{and} \qquad \ell_w := \min\big\{\ell_e \;\big|\; e \in \delta^-(w)\big\} \qquad \forall w \in V \setminus \{s\} \ .$$

Further, we call a path P starting at s *thin* if the congestion of every s-v-subpath of P has a congestion of ℓ_v. If, in addition, a nonnegative real demand $d \geq 0$ is given, we initialize ℓ_s with $\ell_s := \frac{|x|}{d}$. Note that the labels are computable by a modified version of DJIKSTRA'S algorithm for computing shortest path.

Definition 6.10 (Minimum and Sparsest Cuts)**.** Let $G := (V, E)$ be a directed graph and $X \subseteq V$ be a cut. Then the capacity of X is given by $u\big(\delta^+(X)\big)$. For a network (G, u, s, t) a *minimum cut* is an s-t-cut minimizing the capacity. Given a b-flow instance (G, u, b) the *congestion* of a cut X is defined by $\frac{b(X)}{u(\delta^+(X))}$. Further, a cut maximizing the congestion is called *sparsest cut.*

Theorem 6.11 (Max-Flow-Min-Cut Theorem, [26])**.** *Let x be an s-t-flow and X be an s-t-cut on a network (G, u, s, t). Then the value of x is not greater than the capacity of X. Further, the maximum value of an s-t-flow equals the minimum capacity of an s-t-cut.*

Let x be a b-flow and X be a cut on a b-flow instance (G, u, b). Then the maximum edge congestion of an edge is not smaller than the congestion of X. Further, the minimum value of the maximum edge congestion of a b-flow equals the maximum congestion of a cut.

As a direct consequence of the Max-Flow-Min-Cut Theorem, the congestion q^* of a sparsest cut on a network (G, u, b) is computable with the following linear program:

$$\begin{aligned} \min \quad q^* \qquad \text{subject to} \qquad x\big(\delta^+(v)\big) - x\big(\delta^-(v)\big) &= b_v \qquad && \forall v \in V \\ x &\leq q^* \cdot u && \\ x &\in \mathbb{R}_+^E && \end{aligned}$$

In particular, this shows that the congestion of a sparsest is computable in polynomial time.

The Max-Flow-Min-Cut Theorem applied b-flows shows that there exists a b-flow x such that the maximum edge congestion equals the congestion of a sparsest cut. In this case we say that the congestion of a sparsest cut *is attained* by x or, similarly, that x *respects* the sparsest cut value. If x respects the sparsest cut value q^*, there is a nice relation between the flow labels of some nodes and q^*. This is the subject of the following proposition.

Proposition 6.12. *Let (G, u, s, b) be a b-flow instance where only one source s has a positive b-value. Further, let x be a b-flow respecting the congestion q^* of a sparsest cut $X \subsetneq V$. Then the label ℓ_v of every node $v \in V \setminus X$ equals q^*.*

Proof. As x respects the sparsest cut value, we know that the congestion of each edge is smaller than q^* implying $\ell_v \leq q^*$ for all nodes $v \in V$. Hence, it remains to show that $\ell_v \geq q^*$ holds for all $v \in V \setminus X$. From the definition of X and flow conservation we get:

$$\begin{aligned} q^* = \frac{b(X)}{u\big(\delta^+(X)\big)} &= \frac{x\big(\delta^+(X)\big) - x\big(\delta^-(X)\big)}{u\big(\delta^+(X)\big)} \\ &\leq \frac{x\big(\delta^+(X)\big)}{u\big(\delta^+(X)\big)} = \sum_{e \in \delta^+(X)} \frac{x_e}{u\big(\delta^+(X)\big)} = \sum_{e \in \delta^+(X)} \frac{u_e}{u\big(\delta^+(X)\big)} \cdot \frac{x_e}{u_e} \end{aligned}$$

Note that the right hand side is a convex combination of $\frac{x_e}{u_e}$, $e \in \delta^+(X)$. As we know $\frac{x_e}{u_e} \leq q^*$ for all $e \in E$, this shows $\frac{x_e}{u_e} = q^*$ for all $e \in \delta^+(V)$. This shows that every path connecting s with a node of $V \setminus X$ has a congestion of at least q^* implying $\ell_v \geq q^*$ for all $v \in V \setminus X$. □

In the following, consider the behavior of a static flow on a part of a given graph. Firstly, we define the restriction of a flow to a minor. This definition is motivated by intuition and straight forward. Based on this, we identify a suitable network such that the restriction of a flow becomes feasible.

Definition 6.13 ((Feasible) Restriction of a Flow)**.** Let x be some static flow on a directed graph G. The *restriction* x' of x to some subgraph $H \subseteq G$ is simply the restriction of x as a real vector to $E(H)$. In case H results out of G by contracting a node subset $X \subseteq V$ to a node v_X, the restriction x' of x to H is defined as follows:

$$\begin{array}{llll} x'_e := x_e & \forall e \in E\big(G[V \setminus X]\big), & x'_{v_X w} := x_{vw} & \forall e = vw \in \delta^+(X), \\ \text{and} & & x'_{w v_X} := x_{wv} & \forall e = vw \in \delta^-(X)\ . \end{array}$$

Since a flow x just as edge capacities u are elements of $\mathbb{R}^E_+$, the definition of a restricted flow is one-to-one transferable for defining the restriction of u.

Definition 6.14 (Contracted and Induced b-flow Instance)**.** Let (G, u, b) be a b-flow instance and $X \subseteq V$ be a node subset. Then the *contracted b-flow instance* (H, u^H, b^H) consists of the directed graph H arising out of G by contracting X to v_X and the restriction u^H of u to H. Further, the node balance b^H is defined by $b^H_{v_X} := b(X)$ and $b^H_v := b_v$ for all $v \in V \setminus X$.

Let x be a b-flow. Then the *induced b-flow instance* (H, u^H, b^H) with respect to x consists of the induced subgraph $H := G[X]$ and the restriction u^H of u to H. Further, the node balance b^H is defined as

$$b^H_v \ := \ b_v \ - \ x\big(\delta^+(v) \cap \delta^+(X)\big) \ + \ x\big(\delta^-(v) \cap \delta^-(X)\big) \qquad \forall v \in X\ .$$

Based on this, consider the case that X is a sparsest cut of (G, u, b) with congestion $q \in \mathbb{R}_+$ and x is a corresponding b-flow. As each outgoing edge of X must have congestion q and each incoming edge carries no flow, we know the flow crossing the cut X without knowing x. Hence, in this case we have:

$$b^H_v \ := \ b_v \ - \ q \cdot u\big(\delta^+(v) \cap \delta^+(X)\big) \qquad \forall v \in X\ .$$

In this manner, we define a *b-flow instance induced by* X if X is a sparsest cut.

It is not hard to see that the restriction of a flow is feasible on the corresponding contracted or induced b-flow instance. For later reference, this is put in the following proposition.

Proposition 6.15. *The restricted flow is a b-flow on corresponding induced or contracted b-flow instances.*

In the last part of this preliminary section, we consider an interesting result of Picard, Queyranne [72]. They identify a nice structure, which allows to identify all minimum s-t-cuts of a given maximum s-t-flow instance as directed cuts of another graph H. As each edge of a minimum cut is used up to its capacity, the graph H can be used to identify those edges. Thus, H represent the set of edges on which the flow value is constant over all maximum s-t-flow. The structure analysis of thin flows without resetting in Subsection 6.3.2 can be seen as a generalization of their approach.

Theorem 6.16 (Picard, Queyranne [72]). *Consider a maximum s-t-flow instance $\mathcal{I} := (G, u, s, t)$. Then there exists a graph H such that the directed cuts of H correspond one-to-one to the minimum cuts of $\mathcal{I}$. Further, H is acyclic and computable as follows:*

(i) Set $H := G$ and $\mathcal{I}' := \mathcal{I}$.

(ii) Compute an inclusionwise maximal minimum s-t-cut X of $\mathcal{I}'$ and delete all nodes from X which are not reachable from s in $G[X]$.

(iii) Delete $\delta^-(X)$ from H, contract $V \setminus X$ in H, and assign the result to H.

(iv) Set $\mathcal{I}'$ to the b-flow instance induced by X.

(v) Compute a sparsest cut X of $\mathcal{I}'$ and delete all nodes which are not reachable from s in $G[X]$ from X.

(vi) If the congestion of X equals 1 go to (iii). Otherwise contract the graph of $\mathcal{I}'$ in H and return H.

The following lemma is used for transferring Theorem 6.16 to b-flow instances. Besides, it serves as a deeper understanding into the structure of static b-flows. Internalizing the proof idea provides intuition for the most elementary nature of b-flows.

Lemma 6.17. *Let $\mathcal{I} := (G, u, b)$ be some b-flow instance, $q \in \mathbb{R}_+$ be the congestion of some sparsest cut, and x be a b-flow arising out of a sparsest cut computation. Further, let $\mathcal{I}'$ be a b-flow instance induced by some $X \subseteq V$ with respect to x and q' be the corresponding sparsest cut value. Then:*

(i) $q' \leq q$ holds in general,

(ii) $q' < q$ holds if X is an inclusionwise minimal sparsest cut or if $V \setminus X$ is an inclusionwise maximal sparsest cut,

(iii) $q' = q$ holds if there exists a sparsest cut X' of $\mathcal{N}$ with $X' \subseteq X$.

Proof. As x arises out of some sparsest cut computation, the maximum edge congestion caused by x equals q. Hence, as the restriction of x to $\mathcal{I}'$ satisfies the flow conservation constraints by Proposition 6.15, we obtain $q' \leq q$ by Theorem 6.11 which proves (i).

For proving (ii) in case of X is an inclusionwise minimal sparsest cut and (iii) let $X' \subsetneq X$ be some cut on $\mathcal{I}'$. Having in mind that X' is also a cut of $\mathcal{I}$, we partition $\delta^+(X')$ into the set $E_1 := \delta^+(X') \setminus \delta^+(X)$ of edges not crossing X and the set $E_2 := \delta^+(X') \cap \delta^+(X)$ of edges crossing X. As q is the value of a sparsest cut, we know

$$\begin{aligned} q \;\geq\; \frac{b(X')}{u\big(\delta^+(X')\big)} \qquad &\Leftrightarrow \qquad 0 \;\geq\; b(X') - q \cdot u\big(\delta^+(X')\big) \\ &\Leftrightarrow \qquad 0 \;\geq\; \Big(b(X') - q \cdot u\big(\delta^+(X')\big)\Big) \cdot u(E_2)\ . \end{aligned}$$

Note that equality holds if and only if X' is a sparsest cut of $\mathcal{I}$. Hence, in case of (ii) the last inequality is strict and in case of (iii) it is satisfied with equality. Now we add $b(X') \cdot u(E_1)$ to both sides of the last inequality. So after expanding the right hand side and using $u\big(\delta^+(X')\big) = u(E_1) + u(E_2)$ we get

$$\begin{aligned} b(X') \cdot u(E_1) &\geq b(X') \cdot u(E_1) + b(X') \cdot u(E_2) - q \cdot u\big(\delta^+(X')\big) \cdot u(E_2) \\ &= b(X') \cdot u\big(\delta^+(X')\big) - q \cdot u\big(\delta^+(X')\big) \cdot u(E_2) \\ &= u\big(\delta^+(X')\big) \cdot \big(b(X') - q \cdot u(E_2)\big) \end{aligned}$$

resulting in

$$\frac{b(X')}{u\big(\delta^+(X')\big)} \geq \frac{b(X') - q \cdot u(E_2)}{u(E_1)} .$$

As already mentioned, this inequality is satisfied strictly for (ii) and with equality for (iii). Let G' be the graph of $\mathcal{I}'$. Since E_1 are exactly the edges leaving nodes out of X' and entering nodes contained in $X \setminus X' = V(G') \setminus X'$, we have $E_1 = \delta^+_{G'}(X')$. Further, the definition of b' shows $b(X') - q \cdot u(E_2) = b'(X')$ as X is a sparsest cut of $\mathcal{I}$ in case of proving (ii) and X' is a sparsest cut of $\mathcal{N}$ in case of proving (iii). This shows

$$q \geq \frac{b(X')}{u\big(\delta^+(X')\big)} \geq \frac{b'(X')}{u'\big(\delta^+_{G'}(X')\big)}$$

where equality holds if and only if X' is a sparsest cut of $\mathcal{I}$. This proves (ii) in case of X is an inclusionwise minimal sparsest cut and because of (i) also (iii). For the case of $V \setminus X$ is an inclusionwise maximal sparsest cut, the proof of (ii) follows the same lines of arguments. Thus, we omit further details. □

As already mentioned, Lemma 6.17 enables us to reformulate Theorem 6.16 for b-flow instances. The following corollary represent all sparsest cuts of a given b-flow instance as directed cuts of a suitable graph H. Therefore, H identifies all edges on which the flow value is constant over all b-flows respecting the sparsest cut value. Using induction, the correctness is directly implied by Lemma 6.17.

Corollary 6.18. *Consider a b-flow instance $\mathcal{I} := (G, u, b)$. Then there exists a graph H such that the directed cuts of H corresponds one-to-one to the sparsest cuts of $\mathcal{I}$. Further, H is acyclic and computable as follows:*

(i) Set $H := G$ and $\mathcal{I}' := \mathcal{I}$.

(ii) Compute an inclusionwise maximal sparsest cut X of $\mathcal{I}'$ and delete all nodes from X which are not reachable from any source in $G[X]$. Set q to the congestion of X.

(iii) Delete $\delta^-(X)$ from H, contract $V \setminus X$ in H, and assign the result to H.

(iv) Set $\mathcal{I}'$ to the b-flow instance induced by X.

(v) Compute a sparsest cut X of $\mathcal{I}'$ and delete all nodes which are not reachable from any source in $G[X]$ from X.

(vi) *If the congestion of X equals q go to (iii). Otherwise contract the graph of $\mathcal{I}'$ in H and return H.*

Recording the contracted node sets during the algorithm in Corollary 6.18 leads to the following definition.

Definition 6.19 (Partition Encoding All Sparsest Cuts)**.** Consider a b-flow instance (G, u, b). A partition $\mathcal{X} := \{X_1, \dots, X_k\}$ of V *encodes all sparsest cuts* if and only if $\bigcup_{i:=1}^{j} X_i$ is a sparsest cut for all $1 \leq j < k$ and every sparsest cut X results out of a union of sets in $\mathcal{X}$, i.e., $X = \bigcup_{i \in I} X_i$ for some $I \subseteq \{1, \dots, k\}$. If $b = 0$ we set $\mathcal{X} := \{V\}$.

6.2 Thin Flows

In this section we introduce a special kind of static flows which we call thin flows. The definition of thin flows relies among other things on a given edge subset E_1 where some kind of resetting occurs. Depending on whether or not this set E_1 of resetting edges is empty, we speak of thin flows with and without resetting. It turns out that thin flows with resetting are much harder to analyze than such without resetting. Nevertheless, there exists a strong relation between these two kinds of thin flows which is discussed during this section.

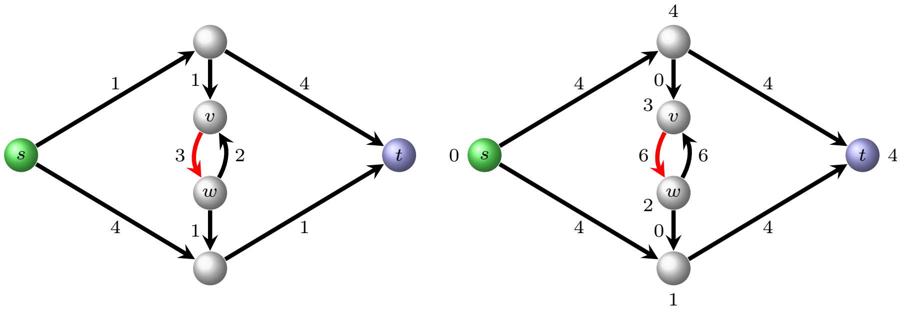
But first we start with some basic definitions. Although thin flows occur in Nash equilibria for the deterministic queuing model as static s-t-flows it is worth to generalize the definition to s-flows. Recalling Section 6.1, s-flows are b-flows where the entire flow is supplied by one source s. In fact, for analyzing reasons, this generalization turns out to be very useful. More precisely, we work with the following instances.

Definition 6.20 (Resetting Network)**.** Let (G, u, s) be a network consisting of a directed graph $G := (V, E)$, edge capacities $u \in \mathbb{R}_+^E$, and a source s. Further, let $E_1 \subseteq E$ be an edge subset. We call (G, u, s, E_1) a *resetting network* and an edge e in E_1 *resetting edge*.

The node s plays the role of a unique source and every other node can potentially be sink. This means that the definition of thin flows below is built upon the definition of s-flows. Note that we require only the flow conservation constraints for a feasible s-flow. In particular, an s-flow has not to obey capacity constraints. Nevertheless, capacities are a basic feature of thin flows and are used for computing flow labels. However, for defining thin flows we need generalized labels arising out of the special set E_1 of resetting edges.

Definition 6.21 (Resetting Congestion and Resetting Labels)**.** Consider a resetting network (G, u, s, E_1) and let $x \in \mathbb{R}^E$ be an s-flow. For some given path $P := (e_1, \dots, e_{|P|})$ starting at s, let $e_i \in E_1$ be the last resetting edge of P, i.e., $i := \max\{j \mid e_j \in E_1\}$. Then the *resetting congestion* of P is defined as $\ell_P := \max\{\frac{x_{e_j}}{u_{e_j}} \mid j \geq i\}$.

The *resetting label* ℓ_v of some node v equals the minimum resetting congestion over all s-v-path and the *resetting label* ℓ_e of an edge $e = vw \in E$ is given by $\max\{\ell_v, \frac{x_e}{u_e}\}$ or $\frac{x_e}{u_e}$ depending on whether e is a usual or a resetting edge, respectively.



(a) The resetting network. (b) A thin flow with resetting.

Figure 6.1: Example of a thin flow with resetting. The source is green and the sink is blue. Left the underlying network is shown, where the capacities are placed on the edges. Right a corresponding thin flow with resetting on the red edge is shown. Flow values and resetting labels are placed at the edges and at the nodes, respectively.

Note that in contrast the usual congestion of a path P equals the maximum congestion over all of its edges. Hence, if $E_1 = \emptyset$, the resetting congestion of P coincides with the natural congestion. The name "*resetting* congestion" refers to the special edges in E_1 which play the following role. Whenever a path starting at s traverses an edge $e \in E_1$, it "forgets" the congestion of all edges seen so far and "resets" its congestion to $\frac{x_e}{u_e}$. In this sense, the resetting congestion generalizes the standard definition of the congestion of a path. In this manner, resetting labels generalize the usual flow labels. However, there is another difference addressing computational aspects. For computing standard flow labels we can assume that the minimum congestion of an s-v-path is attained by a simple one. But this is no longer true for the resetting label as we see below in Example 6.23.

Further, the resetting congestion of a cycle C depends on the node $v \in V(C)$ which is used as the initial point for the computation. This is the reason that we speak of a "v-v-path" instead of a "cycle containing v" in the following definition.

Definition 6.22 (Thin Flow)**.** Let (G, u, s, E_1) be some resetting network and let $x \in \mathbb{R}_+^E$ be an s-flow with corresponding resetting node labels $\ell \in \mathbb{R}_+^V$. Then x is called *thin flow* if and only if every flow carrying s-v-path has a resetting congestion of ℓ_v and every flow carrying v-v-path has a resetting congestion of at most ℓ_v for all $v \in V$.

We call x thin flow *with resetting* if E_1 is nonempty. If E_1 is empty we call x thin flow *without resetting*.

Based on new insights presented in [18] by Cominetti, Correa, and Larré, this definition of thin flows differs slightly from that given in [44]. The following example showing a thin flow with resetting is consecutively extended during this chapter and illustrated in Figure 6.1. This example shows also that the resetting label of some node may be defined via nonsimple paths.

Example 6.23. Consider the resetting network depicted in Figure 6.1a consisting of the green source s, the blue sink t, and the red resetting edge vw where the capacities are shown at the edges. Note that capacities do not restrict feasible flows but are used for defining node labels. Figure 6.1b shows a thin flow x with resetting on $\{vw\}$ of value 8, i.e., $b_s = 8$, $b_t = -8$, and the node balance of every other node equals 0. The numbers at the edges represent the flow values of x whereas the numbers at nodes stand for corresponding labels ℓ. First observe that x and ℓ satisfy Definition 6.22. That is, the resetting congestion of every flow carrying s-v'-path equals $\ell_{v'}$ and the resetting congestion of the two flow carrying cycles $C_1 := (vw, wv)$ and $C_2 := (wv, vw)$ equals the label of v and w, respectively. Note that, although both cycles consist of the same edge set, we are enforced by Definition 6.22 to distinguish between these two as C_1 is a v-v-path and as C_2 is a w-w-path. Hence, the resetting congestion of C_1 must be compared with the label of v and the one of C_2 must be compared with w.

Thus, it remains to check whether ℓ coincides with the family of resetting labels. This is obviously the case for all nodes except v. Considering the unique simple s-v-path P_1 we see that P has a resetting congestion of 4 which is larger than the desired label of 3. However if we extend P by the two edges vw and wv we obtain a resetting congestion of 3 for this new path P_2. This shows that the resetting label of v is defined via the nonsimple path P_2.

A closer look at this example shows that we can vary the flow on the cycle given by vw and wv slightly without violating Definition 6.22. In fact, we can assign any flow value between 3 and 8 to both edges. This ensures that the resetting label of v does not exceed 4 and that the resetting label of w does not become smaller than 1. Hence, the resetting labels just as the thin flow with resetting itself are not unique in general. However, if we require that the resetting label of a node is always attained by some simple path, v must have a resetting congestion of 4. Hence, in this case there exists only one thin flow resetting which sends 8 flow units along the cycle (vw, wv).

In terms of Example 6.23, we potentially lose uniqueness of thin flows if the resetting label of a node v need not to be attained by a simple s-v-path. Unfortunately, this really happens as we see in Section 6.4. Nevertheless, there is an elementary reason for characterizing thin flows as they are defined. So far, a thin flow is defined path-based. However, like for generic static flows it is quite basic for many approaches to have an edge-based formulation.

Lemma 6.24. *Consider a resetting network (G, u, s, E_1) and let $x \in \mathbb{R}_+^{E(G)}$ be an s-flow. Then x is a thin flow if and only if the corresponding resetting labels satisfy:*

$$\ell_s = 0 \ , \tag{6.1}$$
$$\ell_w \geq \ell_v \qquad \forall e = vw \in E \setminus E_1, x_e = 0 \ , \tag{6.2}$$
$$\ell_w = \max\{\ell_v, \frac{x_e}{u_e}\} \qquad \forall e = vw \in E \setminus E_1, x_e > 0 \ , \tag{6.3}$$
$$\ell_w = \frac{x_e}{u_e} \qquad \forall e = vw \in E_1 \ . \tag{6.4}$$

Proof. First assume that x is an s-flow and let ℓ be corresponding labels satisfying (6.1)-(6.4). Further, let v be some node and P be some s-v-path. By traversing P from s to v, we obtain inductively that the resetting congestion of P is at least ℓ_v. Moreover, if P is flow carrying we need not to apply (6.2) during the induction steps implying that the resetting congestion of P equals ℓ_v.

Next assume that x is a thin flow, set $\ell_s := 0$, and let ℓ_v be the minimal resetting congestion over all s-v-path for each node v. So if $x_e = 0$ for some edge $e = vw$, we know $\ell_w \leq \ell_v$. Now consider an edge $e = vw$ with $x_e > 0$. Hence, e is contained in some flow carrying s-t-path or in a flow carrying cycle.

Assume that e lies on a flow carrying s-t-path and let P be a corresponding s-w-subpath containing e as the last edge. As each flow carrying s-w path has the same minimal resetting congestion ℓ_w, this shows $\ell_w = \max\{\ell_v, \frac{x_e}{u_e}\}$ or $\ell_w = \frac{x_e}{u_e}$ depending on whether $e \notin E_1$ or $e \in E_1$, respectively.

Finally, assume that e is the last edge of some flow carrying w-w-path C. If C contains no resetting edge we know $e \notin E_1$ and $\ell_w = \max\{\ell_v, \frac{x_e}{u_e}\}$. On the other hand if C contains a resetting edge we know that the resetting congestion of C is equal to ℓ_w because of Definition 6.22 as ℓ_w is the minimum resetting congestion over all s-w-path. Hence, we get $\ell_w = \max\{\ell_v, \frac{x_e}{u_e}\}$ or $\ell_w = \frac{x_e}{u_e}$ depending on whether $e \notin E_1$ or $e \in E_1$, respectively. □

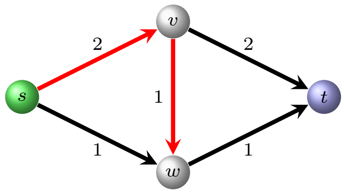
Definition 6.22 and Lemma 6.24 assume implicitly that the label of s is initialized with 0 for computing the resetting labels. However, for analyzing Nash equilibria for the deterministic queuing model in Section 7.3 we need the following modification when defining thin flows.

Remark 6.25. If, in addition to a resetting network, a supply $d \in \mathbb{R}_+$ is given, we initialize $\ell_s := \frac{|x|}{d}$ where $|x| := b_s$ is the flow value of x when computing the resetting labels with respect to some flow x. Then x is a thin flow if Definition 6.22 holds with respect to these new node labels.

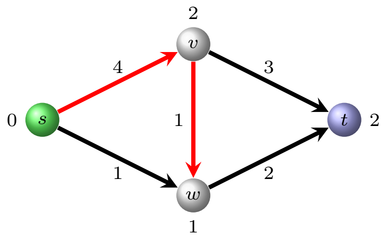
However, for studying thin flows we can restrict to instances with infinite supply d, i.e., $\ell_s = 0$. This is due to the fact that we can model a finite supply simply by adding a dummy source s_0 and an edge $s_0 s$ with capacity d to the network. Then of course, a thin flow on the new instance corresponds to a thin flow on the original instance and vice versa. As the entire outflow of s has to traverse the new edge $s_0 s$, this is directly implied by (6.1) and (6.3).

The following example illustrated in Figure 6.2 shows that introducing an additional supply in terms of Remark 6.25 leads to completely different thin flows. Note that, in contrast, scaling a thin flow leads again to a thin flow where the node balances just as the resetting labels are scaled by the same factor.

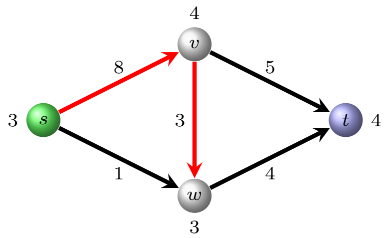
Example 6.26. Consider the resetting network depicted in Figure 6.2a consisting of the green source s, the blue sink t, and the red resetting edges sv and vw where the capacities are shown at the edges. In the following, we compute a thin flow with resetting x using Lemma 6.24. First note that because of the scaling argument we can assume that the label ℓ_w of w equals 1. Hence, condition (6.3) and (6.4) applied to edge sv and vw, respectively, imply $x_{sw} = 1 = x_{vw}$. Hence, from flow conservation we get $x_{wt} = 2$ leading to $\ell_t = 2$ because of (6.3). Also from flow conservation we know $x_{sv} \geq x_{vt}$. Since the capacities of sv and vt coincide, we obtain $\ell_v = \ell_t$ by using (6.4) and (6.3) for sv and vt, respectively. Thus, we get $x_{sv} = 4$ implying, finally, $x_{vt} = 3$. The entire flow of value 5 is shown in Figure 6.2b. Note that this thin flow is unique up to scaling as we had no choice at any point during the construction of x. Comparing this with the uniqueness discussion of Example 6.23, we see that the resetting labels of this example are always obtained by simple paths. On closer inspection we notice that the actual reason for this is the absence of a cycle containing a resetting edge. In fact, a thin flow on such instances is unique as shown in Section 6.4.



(a) The resetting network.



(b) Thin flow with resetting of value 5 in terms of Definition 6.22.



(c) Thin flow with resetting of value 9 in terms of Remark 6.25 with a supply of $d = 3$.

Figure 6.2: This example shows that thin flows vary quite arbitrarily if an additional support in terms of Remark 6.25 is given. The source is green and the sink is blue. The upper figure shows the underlying network with resetting on the red edges, where the capacities are placed on the edges. Right and left a thin flow with and without an additional supply is shown. Flow values and resetting labels are placed at the edges and at the nodes, respectively.

Next consider the case that an additional supply $d := 3$ is given. Then a thin flow x' of value 9 in terms of Remark 6.25 is depicted in Figure 6.2c. A closer look at this setting shows that x' is also unique. However, x and x' behave quite different. As we see in Subsection 6.3.2, where we analyze the structure of thin flows *without* resetting, the main problem is that the subsets of nodes marked by the same label differ. More precisely, the nodes s and w get a different label with respect to x but they have the same label in case x' is considered. Note that this not caused by the different flow values as this property is invariant under scaling.

As already mentioned, whenever a path traverses a resetting edge $e \in E_1$, it "forgets" the congestion of all edges seen so far. Hence, in order to study thin flows with resetting it could be promising to route flow on some resetting edge $e = vw \in E_1$ directly from s to w. That is, we anchor e not at v but at s, i.e., we switch the tail of e from v to s as illustrated in Figure 6.3. More formally:

Definition 6.27 (Anchored Network)**.** Let (G, u, s, E_1) be a resetting network and let $e = vw \in E_1$ be a resetting edge. We construct a new resetting network as follows. We delete edge $e = vw \in E_1$ from G and add a corresponding edge $e_s := sw$ with capacity u_e. We call e_s the *anchored edge* of e and the resulting network the e*-anchored network*. Note that we do not add the anchored edge to E_1.

Applying the procedure above subsequently to all resetting edges in E_1 results in a thin flow network *without* resetting. We call this network the *anchored network* of (G, u, s, E_1). Note that for a single node w there can be more than one incoming anchored edges.

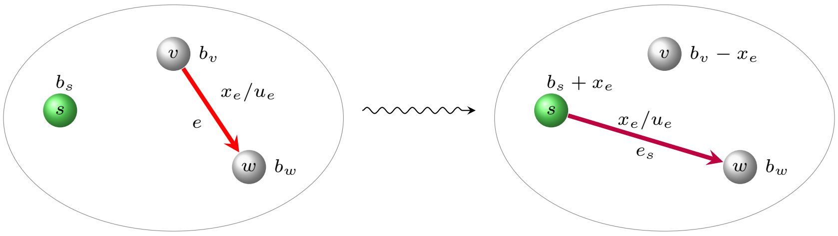
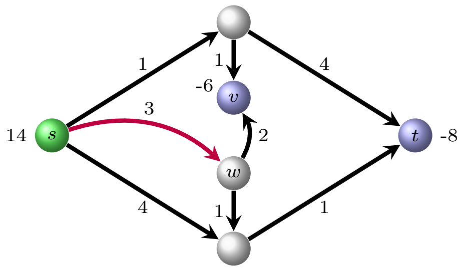
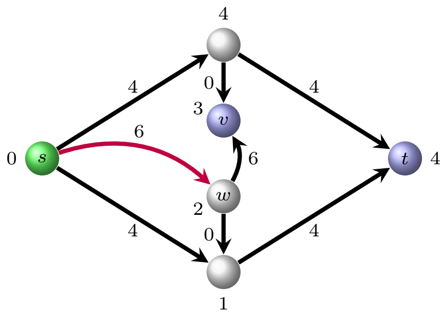


Figure 6.3: Construction of the e-anchored network $\mathcal{N}_e$ and transformation of a thin flow x to a thin flow on $\mathcal{N}_e$.



(a) The vw-anchored network with edge capacities and node balances.



(b) The transformed thin flow with corresponding node labels.

Figure 6.4: The vw-anchored network of Example 6.23 with the corresponding transformed flow arising out of the flow in Figure 6.1b.

Now thin flows on a resetting and on the corresponding anchored network are related as follows. If we anchor some resetting edge $e = vw \in E_1$ at s, we delete some flow leaving v. In fact, given some thin flow x we reduce the outgoing flow of v by a value of x_e and, hence, the flow conservation constraint is violated. But instead of decreasing the flow from s to v this is resolved by redefining the node balance of v. This has also the advantage that we do not lose any thin flow property (see Figure 6.3 for an illustration of this transformation).

Lemma 6.28. *Let $\mathcal{N} := (G, u, s, E_1)$ be a resetting network, $e = vw \in E_1$ be a resetting edge, and $\mathcal{N}_e := (G', u', s, E_1')$ be the corresponding e-anchored network. Let x be a thin flow with resetting on $\mathcal{N}$ respecting certain node balances b. Transform x to a flow x' on $\mathcal{N}_e$ by sending the flow x_e on e along e_s. That is, we set $x'_{e_s} := x_e$, $b'_s := b_s + x_e$ and $b'_v := b_v - x_e$ and let everything else unchanged. Then x' is a thin flow with resetting on the e-anchored network respecting node balances b'.*

Vice versa, let x' be a thin flow with resetting on $\mathcal{N}_e$ respecting node balances b'. Setting $x_e := x'_{e_s}$, $b_s := b'_s - x'_e$, and $b_v := b'_v + x'_e$ yields a thin flow with resetting on $\mathcal{N}$ if $b_v \leq 0$.

Proof. We only have to verify that flow conservation is satisfied at v and w. Further, we have to check condition (6.3) and (6.4) depending on which direction has to be proven. But the constructed flows obviously satisfy these constraints. □

Figure 6.4 shows how Lemma 6.28 works for the thin flow with resetting in Example 6.23. Applying the construction in Lemma 6.28 subsequently to all resetting edges results in a strong relation between thin flows with and without resetting.

Corollary 6.29. *Let $\mathcal{N} := (G, u, s, E_1)$ be a resetting and $\mathcal{N}_{\mathrm{a}} := (G^{\mathrm{a}}, u^{\mathrm{a}}, s)$ be the corresponding anchored network. Further, let x be a thin flow with resetting on the network $\mathcal{N}$ respecting certain node balances b. Set*

$$\begin{aligned} x'_{e_s} &:= x_e && \forall e \in E_1 \ , \\ x'_e &:= x_e && \forall e \in E(G) \setminus E_1 \ , \\ b'_s &:= b_s + x(E_1) \ , \\ b'_v &:= b_v - \sum_{e \in \delta^+(v) \cap E_1} x_e && \forall v \in V \setminus \{s\} \ . \end{aligned}$$

Then x' is a thin flow on the anchored network $\mathcal{N}_{\mathrm{a}}$ respecting node balances b'.

Vice versa, let x' be a thin flow without resetting on $\mathcal{N}_{\mathrm{a}}$ respecting node balances b'. Set

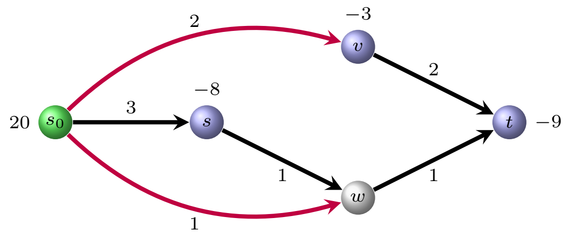
$$\begin{aligned} x_e &:= x'_{e_s} && \forall e \in E_1 \ , \\ x_e &:= x'_e && \forall e \in E(G) \setminus E_1 \ , \\ b_s &:= b_s - \sum_{e \in E_1} x'_{e_s} \ , \\ b_v &:= b'_v + \sum_{e \in \delta^+(v) \cap E_1} x'_{e_s} && \forall v \in V \setminus \{s\} \ . \end{aligned}$$

If $b_v \leq 0$ for all $v \in V$ then x is a thin flow with resetting on $\mathcal{N}$ respecting node labels b.

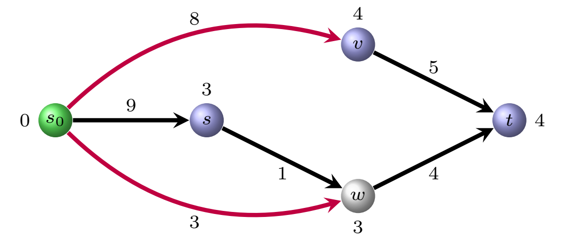
Proof. This follows directly from Lemma 6.28 □

Before we discuss an interesting consequence of this corollary we outline how it works for the case that an additional supply is given. For this we consider Example 6.26 and assume that we are given a supply of $d = 3$. Hence, Figure 6.2c shows the unique thin flow with resetting on this instance. Recalling Remark 6.25, we have to insert a dummy source s_0 and an edge $s_0 s$ with capacity 3 before we are able to apply Corollary 6.29. Thus, the corresponding anchored network is given in Figure 6.5a and the corresponding thin flow without resetting is given in Figure 6.5b. Note that s as the unique source of the original network becomes a sink of the anchored network as it is the tail of the resetting edge sv.

Although Corollary 6.29 shows a very easy relation between thin flows with and without resetting, it is essential for analyzing thin flows with resetting. To see this, assume for a moment that every node v being left by a resetting edge has exactly one outgoing resetting edge e^v and a node balance of 0. Then Corollary 6.29 establishes a one-to-one correspondence between thin flows x with resetting and thin flows x' without resetting where the flow on each anchored edge e^v_s equals the b-value of v, i.e., $x'_{e^v_s} = b'_v$. Hence, interpreting the $x'_{e^v_s}$'s as the image of some function depending on the b'_v's the thin flow x' without resetting is a fixed point with respect to this function. Thus, computing a thin flow with



(a) The anchored network after a dummy source s_0 is added.



(b) Thin flow without resetting on the anchored network and corresponding node labels.

Figure 6.5: The anchored network of Example 6.26 in case an additional supply of 3 is given with the corresponding transformed flow arising out of the flow in Figure6.2c.

resetting means finding a fixed point in the space of thin flows without resetting. This is even more interesting as thin flows without resetting are well-analyzable and computable in polynomial time, see the subsequent Section 6.3. In fact, we use this fixed point approach to prove existence and uniqueness results in Section 6.4. For this we also need the following observations. Recall that the resetting label ℓ_e of an edge $e = vw$ is defined as $\ell_e := \frac{x_e}{u_e}$ and $\ell_e = \max\{\ell_v, \frac{x_e}{u_e}\}$ in case of $e \in E_1$ and $e \notin E_1$, respectively. Hence, for a thin flow x we always have $x_e \leq \ell_e u_e$.

Lemma 6.30. *Let (G, u, s, E_1) be a resetting network, $v \in V \backslash \{s\}$ be some node, and $e \in \delta^+(s)$ be some outgoing edge of s occurring in some s-v-path P. Further, let x be some thin flow with resetting respecting node balances $b \in \mathbb{R}^V$. Then there exists a nonnegative real number $\lambda > 0$ such that $x_e \geq -\lambda b_v$. Moreover, this λ is independent on x and b.*

Proof. First we show that the label ℓ_v of v grows at least linearly with $-b_v$. If the label of v equals 0, it follows from the thin flow conditions that no flow is sent into v implying $b_v = 0$. Hence, we have $\ell_v > 0$ whenever $b_v < 0$. So let $b_v > 0$ and consider the subset $X := \{w \in V \mid \ell_w = \ell_v\}$ of nodes with label ℓ_v. As $\ell_v > 0$ and $\ell_s = 0$, we have $s \notin X$ in addition to $v \in X$. Thus, at least a flow of value $-b_v$ has to cross the cut $\delta^-(X)$, i.e., $x\big(\delta^-(X)\big) \geq -b_v$. Further, we know that the label of each flow carrying edge $e' \in \delta^-(X)$ is equal to ℓ_v as x is a thin flow. Defining E' as the set of flow carrying edges in $\delta^-(X)$ we also know $x_{e'} \leq \ell_e u_e$ for all $e' \in E'$. This shows:

$$\frac{-b_v}{u\big(\delta^-(X)\big)} \;\leq\; \frac{x\big(\delta^-(X)\big)}{u\big(\delta^-(X)\big)} \;\leq\; \frac{x(E')}{u(E')} \;\leq\; \frac{\sum\limits_{e' \in E'} \ell_{e'} u_{e'}}{u(E')} \;=\; \ell_v \;.$$

Thus, if we set λ_1 to the inverse of the maximum value of any s-v-cut we obtain

$$-\lambda_1 \cdot b_v \leq \ell_v \; . \tag{6.5}$$

Next, we show that the label of v is linearly bounded by the flow x_e on e. For this, consider an arbitrary edge $\tilde{e} = \tilde{v}w$ with $\tilde{v} \neq s$. Then we have $\ell_w \leq \ell_{\tilde{v}}$ or

$$\ell_w \;=\; \frac{x_{\tilde{e}}}{u_{\tilde{e}}} \;\leq\; \frac{x\big(\delta^-(\tilde{v})\big)}{u_{\tilde{e}}} \;\leq\; \frac{\ell_v \cdot u\big(\delta^-(\tilde{v})\big)}{u_{\tilde{e}}} \;\leq\; \frac{|E| \cdot u_{\max}}{u_{\min}} \cdot \ell_v$$

where $u_{\max} := \max\{u_{e'} \mid e' \in E\}$ and $u_{\min} := \min\{u_{e'} \mid e' \in E\}$ are the maximum and minimum edge capacity, respectively. Further, we know from the thin flow definition that $\ell_{\text{head}(e)} = \frac{x_e}{u_e}$. Hence, we obtain inductively by setting $\lambda_2 := \frac{|E| \cdot u_{\max}}{u_{\min}}$:

$$\ell_v \leq \lambda_2^{|V|} \cdot \frac{x_e}{u_e} \; . \tag{6.6}$$

Note that we assume without loss of generality that P is simple. Further, equation (6.6) is established without the assumption that $b_v < 0$ and, therefore, holds for every $v \in V$ which is reachable from s via e. For concluding this proof, we observe that (6.5) and (6.6) imply

$$x_e \geq -\lambda \cdot b_v$$

where $\lambda := \frac{\lambda_1 \cdot u_e}{\lambda_2^{|V|}}$ is independent on x and b. □

So far, we have a lower bound on the flow of the edges leaving s. The next lemma gives an upper bound on the flow of all edges.

Lemma 6.31. *Let (G, u, s, E_1) be a resetting network and x be some thin flow with resetting respecting node balances $b \in \mathbb{R}^V$. Then there exists a nonnegative real number $\lambda > 0$ such that $x_e \leq \lambda b_s$ holds for every $e \in E$. Moreover, this λ is independent on x and b.*

Proof. First consider an edge $e \in \delta^+(s)$. As the label of s equals 0 by definition, we know that no flow enters s. Hence, by flow conservation we know

$$b_s \;=\; x\big(\delta^+(s)\big) \;\geq\; x_e \; .$$

Now consider an arbitrary node $v \in V$ and let $e \in \delta^+(s)$ be some edge such that v is reachable from s via e. Since equation (6.6) holds, we know that there exists a $\lambda_v > 0$ which is independent on x and b such that

$$\ell_v \;\leq\; \lambda_v x_e \;\leq\; \lambda b_s \; .$$

Concluding this proof, observe that for each edge $e = vw \in E$ the thin flow with resetting conditions (6.2)-(6.4) imply $x_e \leq \ell_w u_e$. Therefore, defining $\lambda := u_{\max} \cdot \max\{\lambda_v \mid v \in V\}$, leads to

$$x_e \leq \lambda b_s \qquad\qquad \forall e \in E \; . \qquad □$$

Note that Lemma 6.31 implies that a sequence of thin flows $(x^k)_{k\in\mathbb{N}}$ respecting node balances $(b^k)_{k\in\mathbb{N}}$ is uniformly bounded if $(b_s^k)_{k\in\mathbb{N}}$ is uniformly bounded. Hence, $(x^k)_{k\in\mathbb{N}}$ contains a convergent subsequence as a bounded sequence of $\mathbb{R}^E$. The next lemma shows that this convergent subsequence is also a thin flow.

Lemma 6.32. *Let (G, u, s, E_1) be a resetting network and $(x^k)_{k\in\mathbb{N}}$ be a sequence of some thin flow with resetting respecting node balances $(b^k)_{k\in\mathbb{N}}$. Further, assume that $(x^k)_{k\in\mathbb{N}}$ converges to some x^*, i.e., $\lim_{k\to\infty} x^k = x^*$. Then x^* is also a thin flow with resetting respecting node balances $b^* := \lim_{k\to\infty} b^k$.*

Proof. First, we observe that x^* is a static flow respecting node balances b^*. Since each x^k satisfies flow conservation, we know

$$\begin{aligned} b_v^* \;=\; \lim_{k\to\infty} b_v^k \;&=\; \lim_{k\to\infty} x^k\big(\delta^+(v)\big) - x^k\big(\delta^-(v)\big) \\ &=\; x^*\big(\delta^+(v)\big) - x^*\big(\delta^-(v)\big) \qquad \forall v \in V\ . \end{aligned}$$

This shows that x^* is a static flow respecting node balances b^*. In particular, this also ensures the existence of $b^* = \lim_{k\to\infty} b^k$.

To see that x^* is a thin flow with resetting we first observe that ℓ^* exists. Note that ℓ_v^k equals the minimum resetting congestion over all s-v-paths P (see Lemma 6.24). Considering some s-v-path P we obtain inductively, while traversing P, that the resetting congestion of P converges. Hence, as the minimum operator works continuously, we know that ℓ_v^* exists for all $v \in V$.

In order to verify conditions (6.1)-(6.4) for x^*, let $e = vw \in E$ be some edge. If $e \in E_1$ is a resetting edge we know from (6.4) applied to each x^k

$$\ell_w^* \;=\; \lim_{k\to\infty} \ell_w^k \;=\; \lim_{k\to\infty} \frac{x_e^k}{u_e} \;=\; \frac{x_e^*}{u_e}$$

which shows that condition (6.4) holds also for x^*.

Next assume that $e \in E \setminus E_1$ is not a resetting edge. If $x_e^* > 0$ we know that there exist a $K \in \mathbb{N}$ such that $x_e^k > 0$ for all $k > K$. Hence, from (6.3) we get:

$$\ell_w^* \;=\; \lim_{k\to\infty} \ell_w^k \;=\; \lim_{k\to\infty} \max\{\ell_v^k, \frac{x_e^k}{u_e}\} \;=\; \max\{\ell_v^*, \frac{x_e^*}{u_e}\}$$

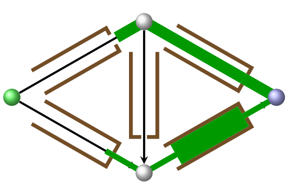
implying that (6.3) holds also for x^*. It remains to consider the case $x_e^* = 0$ for a nonresetting edge $e = vw \in E \setminus E_1$. If $\ell_w^* = 0$ condition (6.2) holds obviously. So assume $\ell_w^* > 0$ implying that there exists a $K \in \mathbb{N}$ such that $\frac{x_e^k}{u_e} \le \ell_w^k$ is valid for all $k > K$. Hence, independently on whether $x_e^k > 0$ or $x_e^k = 0$, we know $\ell_w^k \le \ell_v^k$ for all $k > K$. Thus, we get

$$\ell_w^* \;=\; \lim_{k\to\infty} \ell_w^k \;\le\; \lim_{k\to\infty} \ell_v^k \;=\; \ell_v^*$$

which proves (6.2) for x^*. Since (6.1) holds trivially, we know that x^* is a thin flow with resetting where corresponding node labels are given by ℓ^*. $\square$

6.3 Thin Flows without Resetting

Thin flows without resetting are the key for understanding general thin flows. As we see in Section 6.4, the understanding of thin flows without resetting enables



us to prove existence and uniqueness results for thin flows with resetting. In this sense the presented approach is strongly directed for achieving these results.

In Subsection 6.3.1 we establish basic properties and end up with a polynomial algorithm for thin flows. Beside this, we show that thin flows without resetting are unique in some sense. Based on the polynomial time algorithm, we exploit the structure of thin flows without resetting in Subsection 6.3.2 for providing a sensitivity analysis in Subsection 6.3.3. Thin flows with resetting can be seen as thin flows without resetting where the node balances of some nodes coincide with the flow values of certain edges. Hence, the sensitivity analysis is built upon small variations in the node balances.

6.3.1 Computation and Uniqueness

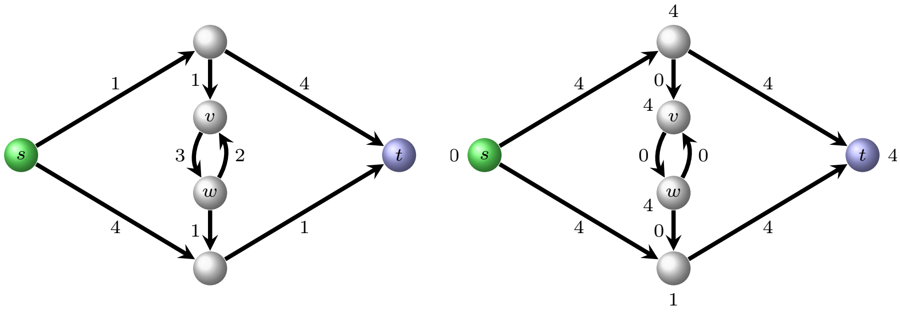
In this section we present some basic properties of thin flows without resetting. In particular, we show that these thin flows respect edge capacities if the underlying b-flow instance is feasible and that they are minimal in a sense motivating the "thin" in the name. Moreover, we give a polynomial time algorithm which is the key for obtaining a deep understanding of thin flows without resetting. The way this algorithm works shows that the node labels of a thin flow are unique.

Recalling Definition 6.22 we work on a network (G, u, s) consisting of a directed graph $G := (V, E)$, edge capacities $u \in \mathbb{R}_+^E$, and a source s. As there is no resetting edge, the resetting congestion of some path and the resetting label of some node coincide with the corresponding usual definitions. That is the congestion of a path P equals the maximum edge congestion over all of its edges and the label of a node v equals the minimum resetting congestion over all s-v-path. Hence, an s-flow x with corresponding node labels ℓ is a thin flow without resetting if and only if every flow carrying s-v path has a congestion of ℓ_v and every flow carrying cycle containing v has a congestion of at most ℓ_v for all $v \in V$. In this sense, every flow carrying path of a thin flow is thin. Note that unlike the resetting case the congestion of a cycle is independent on the starting node, i.e., if a v-v-path and a w-w-path define the same cycle, also the congestion of these two paths coincide. Another observation in this direction is that the label of a node v is always attained by the congestion of some simple s-v-path. Recall that this eliminates some reason why thin flows with resetting are in general not unique.

Next we reconsider Example 6.23 establishing that thin flows with resetting are not unique. Not surprisingly, a thin flow without resetting on the same network differs from one with resetting. However, for verifying a thin flow without resetting on this instance we take a look at Lemma 6.24 which shows that an s-flow is a thin flow without resetting if and only if the node labels satisfy conditions (6.1)-(6.3). The following example is depicted in Figure 6.6.

Example 6.33. Consider the network shown in Figure 6.6a which is similar to Example 6.23 except that each resetting edge is assumed to be a usual one. A thin flow x without resetting of value 8 with corresponding node labels ℓ is shown in Figure 6.6b. Note that, it is straightforward to verify that x satisfies conditions (6.1)-(6.3).

This thin flow is not unique as the flow on the cycle (vw, wv) can be set to an arbitrary value between 0 and 8 as this has no influence on the node labels. However, a closer look shows us that in this manner we obtain all feasible thin



(a) The underlying network. (b) A thin flow without resetting.

Figure 6.6: Example of a thin flow without resetting on the network shown in Figure 6.1a where each edge is assumed to be a nonresetting one. Left the resulting network is shown, where the capacities are placed on the edges. Right a corresponding thin flow without resetting. Flow values and resetting labels are placed at the edges and at the nodes, respectively.

flows of value 8. Hence, the node labels are unique for this example if the flow value is fixed. As we see subsequently, this is always the case for thin flows without resetting.

Example 6.33 deals with an instance where the node labels of any thin flow without resetting coincide. The next lemma is a first step in the direction of establishing this result for all thin flow without resetting instances.

Lemma 6.34. *Let x be a thin b-flow on a network (G, u, s) where only node s has a positive supply. Then the maximum label $\ell_{\max}$ of any edge is equal to the congestion q^* of a sparsest cut in $\mathcal{N}$, i.e., a node set X with $s \in X \subsetneq V$ maximizing $q^* := \frac{b(X)}{u(\delta^+(X))}$.*

Proof. The relation $\ell_{\max} \geq q^*$ is obvious because at least one edge in a sparsest cut must have a congestion of at least q^* in any b-flow. Thus, it remains to show $\ell_{\max} \leq q^*$. For this consider the cut $X := \{v \in V \mid \ell_v < \ell_{\max}\}$ consisting of all nodes with a label smaller than $\ell_{\max}$. In particular, this ensures $s \in X \subsetneq V$. Because the labels of the nodes in X are strictly smaller than the labels of the nodes not in X there is no flow on any edge in $\delta^-(X)$. Further, the congestion of any edge in $\delta^+(X)$ is at least $\ell_{\max}$. This leads to $\ell_{\max} \leq \frac{x(\delta^+(X))}{u(\delta^+(X))} = \frac{b(X)}{u(\delta^+(X))} \leq q^*$ because q^* is the congestion of a sparsest cut. □

A particular consequence of the Lemma 6.34 is that at least the maximum node label remains the same for any thin flow on some given instance. Furthermore, any thin flow respects the congestion of a sparsest cut. This implies that a thin s-t-flow where the maximum node label equals 1 is also a maximum s-t-flow. Recalling Example 6.33 this statement is true for the considered instance.

The following lemma considers the situation where restrictions of thin flows remain thin flows. Thus, this lemma can be used for proving statements about thin flows by induction over the number of nodes.

Lemma 6.35. *Let $\mathcal{I} := (G, u, s, b)$ be a s-flow instance and x be a thin flow without resetting. Further, let $X \subsetneq V$ be an s-cut such that $x\big(\delta^-(X)\big) = 0$ holds. Then the restriction of x to the b-flow instance $\mathcal{I}' = (G', u', s, b')$ induced by X with respect to x is a thin flow.*

Proof. First, note that the restriction x' of x to G' is a feasible b'-flow by Proposition 6.15. Consider a node $v \in X$ and let ℓ'_v and let ℓ_v be its label on G' and G with respect to x' and x, respectively. As every s-v-path in G' is also an s-v-path in G, we know $\ell'_v \geq \ell_v$. On the other hand we know that every flow carrying s-v-path in G is also a flow carrying s-v-path in G' as $x\big(\delta^-(X)\big) = 0$. This shows two things. First, $\ell'_v = \ell_v$ if there exists a flow carrying s-v-path and, second, every flow carrying s-v-path on G' has a congestion of ℓ'_v. Thus, it remains to prove that the congestion of every flow carrying cycle containing v is at most ℓ'_v. But as every flow carrying cycle in G' is also a flow carrying cycle in G this follows directly from $\ell'_v \geq \ell_v$. Hence, x' is a thin flow of $\mathcal{I}'$. □

Now we are ready to prove the main result in this subsection.

Theorem 6.36. *The node labels of a thin flow without resetting on any given instance (G, u, s, b) are unique. Moreover, a thin flow is computable in polynomial time.*

Proof. Consider two thin flows $x, \tilde{x} \in \mathbb{R}^E_+$. We show that the corresponding edge labels ℓ, $\tilde{\ell}$ are equal. Then this must also hold for the corresponding node labels. We prove this by induction over the number of nodes. If there is only one node s nothing has to be proven. Thus, assume that G consists of more than one node. Lemma 6.34 shows that the maximal edge label $\ell_{\max}$ is unique and, in fact, equals the congestion of a sparsest cut. Moreover, since x just as $\tilde{x}$ respect the sparsest cut value Proposition 6.12 shows that all edges having at least one incident node behind a sparsest cut admit this unique label, i.e., $\ell_e = \tilde{\ell}_e = \ell_{\max}$ for all edges $e \in \delta^+(X) \cup E\big(G[V \setminus X]\big)$ where X is an inclusionwise minimal sparsest cut with $s \in X \subsetneq V$.

Now we delete the node set $V \setminus X \neq \emptyset$ in order to construct the b-flow instance $\mathcal{I}'$ induced by X. As x and $\tilde{x}$ respect the sparsest cut value, we know $x\big(\delta^-(X)\big) = 0 = \tilde{x}\big(\delta^-(X)\big)$ and, hence, the corresponding restrictions to $\mathcal{I}'$ are thin flows by lemma 6.35. Because $G[X]$ has a smaller number of nodes than G, we apply the induction hypothesis and have finished this part of the proof.

To see that we can compute a thin flow on (G, u, s, b) in polynomial time note that the induction above is constructive and results in the Thin Flow algorithm stated directly after this proof. Because in any iteration we compute an inclusionwise minimal sparsest cut of a b-flow instance which can be done in polynomial time and because the number of iterations is bounded by the number of nodes this algorithm is polynomial. The correctness of the Thin Flow algorithm follows directly from Lemma 6.17 implying that the sparsest cut value in each iteration decreases strictly. □

The last theorem results in the following algorithm for computing a thin flow on a given network.

Thin Flow Algorithm	
Input:	A thin flow without resetting instance (G, u, s, b).
Output:	A thin flow $x \in \mathbb{R}_+^E$ without resetting and corresponding node labels $\ell \in \mathbb{R}_+^V$.

(1) Initialize $\mathcal{I}' := (G, u, s, b)$.

(2) Compute an inclusionwise minimal sparsest cut X on $\mathcal{I}' = (G', u', s, b')$ together with an s-flow x' respecting the sparsest cut value $q := \frac{b'(X)}{u'(\delta^+_{G'}(X))}$.

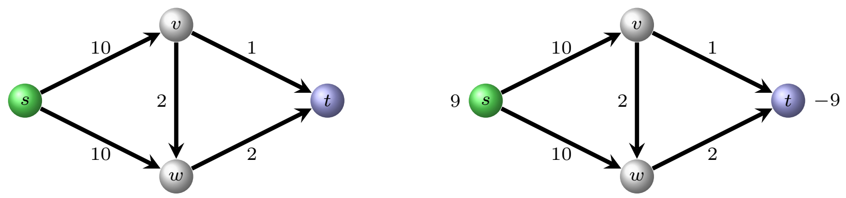
(3) Set $\ell_v := q$ for all $v \in V(G') \setminus X =: Y$ and $x_e := x'_e$ for all $e \in \delta^+_{G'}(X) \cup E\big(G'[Y]\big)$.

(4) Construct the b-flow instance induced by X and assign it to $\mathcal{I}'$.

(5) If $b'_s = 0$ set $\ell_v := 0$ for all $v \in V(G')$ and return x and ℓ. Otherwise go to (2).

The following example points out how the algorithm works.

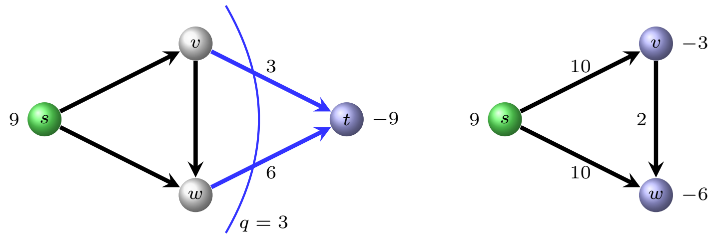
Example 6.37. In this example we compute a thin s-t-flow without resetting of value 9 on the network shown in Figure 6.7a. First, we transform this instance into an equivalent b-flow instance which is given to the Thin Flow algorithm. This b-flow instance illustrated in Figure 6.7b is assigned to $\mathcal{I}'$ in the initialization step (1).



(a) The original network. (b) Initialization of $\mathcal{N}'$.

Figure 6.7: The initialization step of the Thin Flow algorithm.

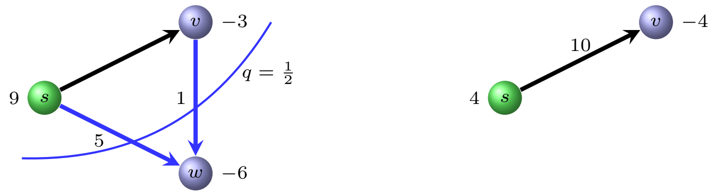
Now the first iteration of the Thin Flow algorithm starts. In Step (2) the sparsest cut shown in Figure 6.8a with a congestion of 3 is found. To see this enumerate all four s-cuts of $\mathcal{I}'$. Hence, any s-flow respecting the sparsest cut value sends 3 flow units along vt and 6 flow units along wt. So in Step (3) the label of t is set to 3 and the flow on vt and wt is set to 3 and 6, respectively (see Figure 6.10b where the complete thin flow is shown). Next, the b-flow instance induced by the sparsest cut is constructed in Step (4). For this the node t together with its two incoming edges is deleted. In order to catch the flow which would be lost by deleting vt and wt we introduce a node balance of -3 for v and -6 for w. Intuitively, we route the flow on these two edges back to their tails. In this manner, v and w become two new sinks. The final result of this construction is depicted in Figure 6.8b. This induced b-flow instance is assigned to $\mathcal{I}'$ in Step (4). As the node balances of this new $\mathcal{I}'$ are not zero, a new iteration is started in Step (5).



(a) Sparsest cut. (b) New induced b-flow instance $\mathcal{N}'$.

Figure 6.8: Iteration 1 of the Thin Flow algorithm.

Hence, the Thin Flow algorithm starts iteration 2. By enumerating the two s-cuts we know that the sparsest cut shown in Figure 6.9a with a congestion of $\frac{1}{2}$ is computed in Step (2). Thus, a corresponding s-flow sends flow of value 1 along vw and flow of value 5 along sw. Therefore, Step (3) sets the label of w to $\frac{1}{2}$ and the flow on vw and sw to 1 and 5, respectively. For computing the induced b-flow instance in Step (4) we route the flow on these two edges back to their tails. This implies new node balances of 4 and -4 for s and v (see Figure 6.10a). As the node balances of this new $\mathcal{I}'$ are not zero, a new iteration is started in Step (5).



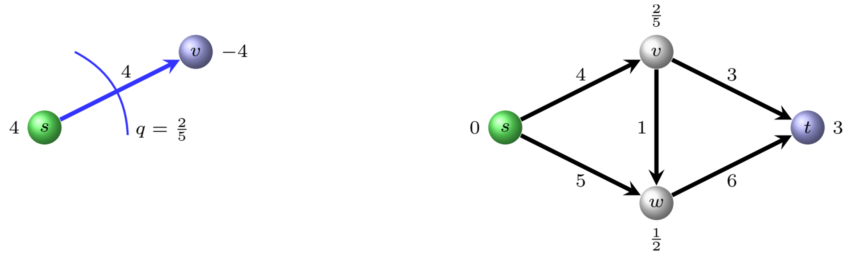
(a) The sparsest cut. (b) New induced b-flow instance $\mathcal{N}'$.

Figure 6.9: Iteration 2 of the Thin Flow algorithm.

The final iteration of the Thin Flow algorithm starts by computing the sparsest cut shown in Figure 6.10a with a congestion of $\frac{2}{5}$ in Step (2). As the unique flow sends 4 flow units along sv, Step (3) sets the label of w to $\frac{2}{5}$ and the flow on sv to 4. The induced b-flow instance constructed in Step (4) consists only of node s with a balance of 0. Thus, in Step (5) the label of s is set to 0 and the algorithm terminates by returning the computed thin flow without resetting with corresponding node labels. The final result is depicted in Figure 6.10b.

This example emphasizes also the following interesting aspect in terms of Lemma 6.17 which shows that sparsest cuts remain sparsest cuts on induced instances. Reconsidering the original network in Figure 6.7a we observe that the s-t-cut with the second smallest capacity is $\{s, w\}$ with a capacity of 12. However, the sparsest cut found in iteration 2 equals $\{s, v\}$. This shows that the second minimal cut of some instance is in general not the second sparsest cut found by the Thin Flow algorithm.

In the last part of this section we motivate the term "thin" in the name of thin flows. The following lemma shows that thin flows are exactly the flows where the edge labels are as small as possible.



(a) The sparsest cut.

(b) The final thin flow with resetting and corresponding node labels.

Figure 6.10: Iteration 3 and the output of the Thin Flow algorithm.

Lemma 6.38. *Let $x \in \mathbb{R}_+^E$ be a thin flow and $\tilde{x} \in \mathbb{R}_+^E$ be an arbitrary static flow on a given instance $\mathcal{I} := (G, u, s, b)$ with corresponding node labels ℓ and $\tilde{\ell}$, respectively. Further, let $e^* \in E$ be an edge where $\ell_{e^*} \neq \tilde{\ell}_{e^*}$ such that the label $\tilde{\ell}_{e^*}$ is maximal with this property. Then we have $\ell_{e^*} < \tilde{\ell}_{e^*}$ and, moreover, $\ell_e = \tilde{\ell}_e$ for all $e \in \delta^+(X) \cup E\big(G[V \setminus X]\big)$ where $X := \{v \mid \ell_v \leq \tilde{\ell}_{e^*}\}$.*

Proof. It follows from the thin flow condition (6.3) that $\ell_e = \frac{x_e}{u_e} > \tilde{\ell}_{e^*}$ for each edge $e = vw \in \delta^+(X)$ as the definition of X implies $\ell_v < \ell_w$. Vice versa from (6.2) we obtain $x_e = 0$ for each edge $e = vw \in \delta^-(X)$ as the definition of X implies $\ell_v > \ell_w$. Because of flow conservation this shows

$$\frac{x\big(\delta^+(X)\big)}{u\big(\delta^+(X)\big)} = \frac{b(X)}{u\big(\delta^+(X)\big)} = \frac{x\big(\delta^+(X)\big) - x\big(\delta^-(X)\big)}{u\big(\delta^+(X)\big)} \leq \frac{x\big(\delta^+(X)\big)}{u\big(\delta^+(X)\big)} .$$

The maximality of $\tilde{\ell}_{e^*}$ implies $\ell_e \geq \tilde{\ell}_e$ as $\ell_e > \tilde{\ell}_{e^*}$ for all $e \in \delta^+(X)$. Hence, we get $\frac{x_e}{u_e} \geq \frac{\tilde{x}_e}{u_e}$ for all $e \in \delta^+(X)$. Thus, for convexity reasons we have $x_e = \tilde{x}_e$ for all $e \in \delta^+(X)$. Hence, we know $\ell_e = \tilde{\ell}_e > \tilde{\ell}_{e^*}$ for all $e \in \delta^+(X)$. Thus, for all $e \in E\big(G[V \setminus X]\big)$ this shows $\tilde{\ell}_e > \tilde{\ell}_{e^*}$. Because of the maximality property of e^* this proves the "moreover"-part and, in addition, $e^* \in E\big(G[X]\big)$. However, as $\ell_e \leq \tilde{\ell}_{e^*}$ holds for all $e \in E\big(G[X]\big)$ by the definition of X, we are done. $\square$

Formalizing Lemma 6.38 we define a lexicographic order for static flows and observe that the minimal elements of this order are exactly the thin flows.

Definition 6.39. Let x be a b-flow on (G, u, s, b) where only one node s has a positive supply. The *edge label vector* $\ell_x(E) := (\ell_{e_1}, \dots, \ell_{e_m})$ is the vector of the edge label in nonincreasing order, i.e., $\{e_1, \dots, e_m\} = E$ and $\ell_{e_i} \geq \ell_{e_j}$ for $i < j$. On the set of static b-flows we define a total order $\prec$ by:

$$x \prec x' \quad :\Leftrightarrow \quad \ell_x(E) \prec_{\text{lex}} \ell_{x'}(E) .$$

The next theorem follows directly from Lemma 6.38. As a consequence, one could think of thin flows as flows distributing their flow evenly over all edges.

Theorem 6.40. *A static flow x is a thin flow on an instance (G, u, s, b) if and only if x is a minimal element regarding to $\prec$. Moreover, for minimal static b-flows, the corresponding edge label vectors are equal.*

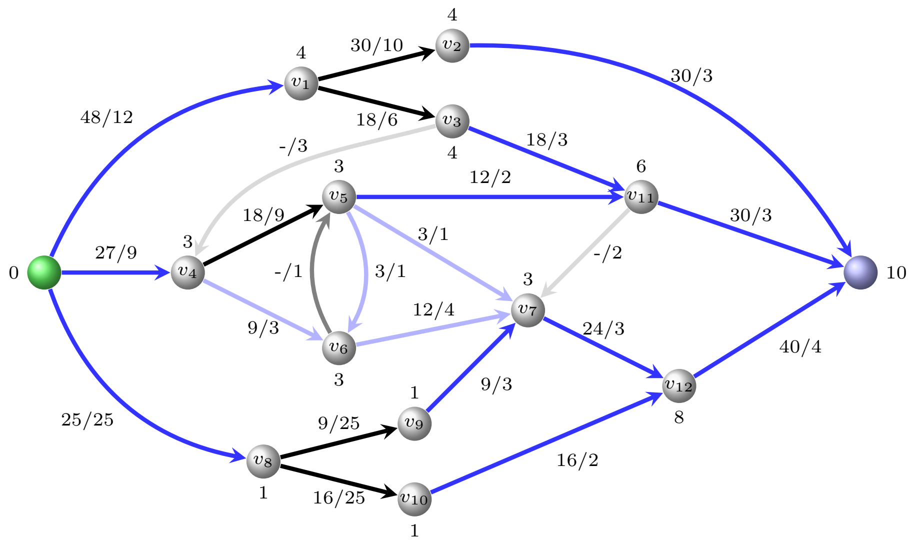


Figure 6.11: A thin flow with resetting. The quasi-critical edges are blue including the critical edges which are drawn dark blue. The quasi-empty edges carrying no flow in any thin flow are plotted gray including the light gray empty edges carrying still no flow if the instance is modified slightly. Black edges are special because they are not special. Each type of edges is explained in this subsection.

6.3.2 Structure Analysis

In this subsection we analyze the structure of a thin flow without resetting on an instance (G, u, s, b) where only node s has a supply, i.e., $b_s > 0$ and $b_v \leq 0$ for all $v \in V \setminus \{s\}$. As shown in Subsection 6.3.1, a thin flow is unique in some sense. Analyzing the underlying unique structure of a thin flow is part of this section. Note that Theorem 6.16 shows a one-to-one correspondence between the minimum s-t-cuts of some network (G, u, s, t) and the directed cuts of some graph arising out of (G, u, s, t). As each maximum s-t-flow has to use each edge of each minimum cut up to its capacity, Theorem 6.16 identifies the structure on which each maximum s-t-flow is unique. In this spirit this subsection is a generalization of Theorem 6.16 to thin flows without resetting. Further, it provides the basis for the sensitivity analysis in Subsection 6.3.3.

For explaining the important definitions and results in this and the subsequent subsection we establish in Figure 6.11 a running example. A thin flow without resetting of value 100 and corresponding node labels are also shown there. First consider the dark blue edges. These are all edges leading from a node with a lower label to a node with a higher label. Hence, recalling thin flow condition (6.3) the labels of these edges are uniquely determined by its congestion and, as the node labels are unique, they are the only edges having this property in any thin flow without resetting. Hence, it makes sense to collect these edges in some structure. Considering the two edges drawn in light gray, we observe that they connect a node with a higher label to a node with a lower label. Thus, condition (6.2) shows that these edges carry no flow in any thin flow as the node labels are unique. This motivates the following definition.

Definition 6.41 (Critical Network, Critical Edges)**.** Let $\mathcal{I} := (G, u, s, b)$ be a

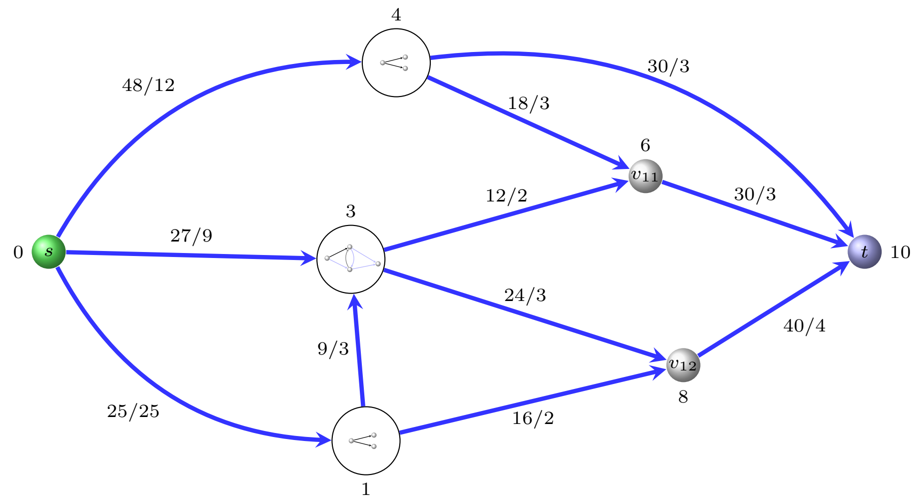


Figure 6.12: The critical network of the thin flow instance shown in Figure 6.11.

thin flow without resetting instance and $(\ell_v)_{v\in V} \in \mathbb{R}_+^V$ be the unique family of node labels. For each possible label, $\ell \in \mathbb{R}_+$ which is assigned to at least one node, let $X_\ell := \{v \in V \mid \ell_v = \ell\}$ be the set of all nodes marked with this label. Further, let $\mathcal{X}^{\mathrm{c}}$ be the node partition consisting of these X_ℓ's and E_0^{c} be the edge set containing all edges vw with $\ell_v > \ell_w$.

The *critical network* $\mathcal{N}_{\mathrm{c}} := (G^{\mathrm{c}}, u^{\mathrm{c}}, s^{\mathrm{c}})$ of $\mathcal{I}$ is defined as follows. The directed graph G^{c} arises out of G by first deleting all edges of E_0^{c} and subsequently contracting all node subsets in $\mathcal{X}^{\mathrm{c}}$. In this manner, the node subset $X_0 \in \mathcal{X}^{\mathrm{c}}$ consisting of all nodes with label 0 is contracted to s^{c}. Finally, restricting u to G^{c} results in u^{c}. Finally, we call each edge contained in $E(G^{\mathrm{c}})$ *critical* and each edge of E_0^{c} *empty*.

The critical network of the thin flow instance shown in Figure 6.11 is illustrated in Figure 6.12. In particular, the node sets $\{v_1, v_2, v_3\}$, $\{v_4, v_5, v_6, v_7\}$, and $\{v_8, v_9, v_{10}\}$ arising out of the node labels 4, 3, and 1 are contracted. As we see, the critical network contains exactly the dark blue edges as expected. Note that the set $E_0^{\mathrm{c}} = \{v_3v_4, v_{11}v_7\}$ of empty edges refers exactly to the light gray edges in Figure 6.11. That is, during the construction of the critical network we delete only edges which are not able to carry flow.

Recalling the THIN FLOW algorithm on page 214 we observe the following for some passed thin flow without resetting instance (G, u, s, b). In each iteration all nodes placed behind the sparsest cut X of the current $\mathcal{I}' = (G', u', s, b')$ are labeled with the congestion q of X in Step (3) and are deleted during the construction of the new induced b-flow instance in Step (4). Further, because of Lemma 6.17 the congestion q of a sparsest cut strictly decreases from one iteration to the next. Hence, in each iteration we delete the entire subset $V(G') \setminus X = \{v \in V \mid \ell_v = q\} = X_q$ of nodes labeled with q. Thus, we are able to construct the critical network of (G, u, s, b) if we delete $\delta^-(X)$ from G and contract $V(G') \setminus X$ in each iteration.

However, we can extract more information out of this construction. For this consider the fifth iteration of the THIN FLOW algorithm applied to the network

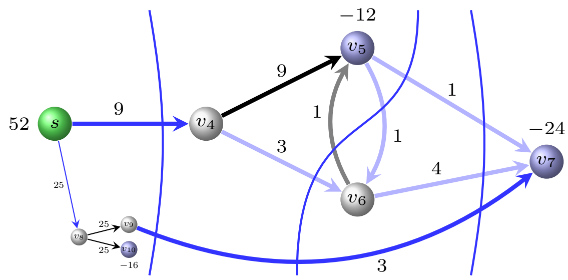


Figure 6.13: All sparsest cuts of the b-flow instance induced by the nodes with a label strictly smaller than 4 according to the thin flow shown in Figure 6.11. This instance occurs during an application of the THIN FLOW algorithm as the current $\mathcal{I}'$ in iteration 5. Note that only the capacities are shown on the edges. For the sake of clarity, unimportant parts are miniaturized.

shown in Figure 6.11 after the labels 10, 8, 6, and 4 are assigned to the nodes. That is, the current induced b-flow instance $\mathcal{I}'$ consists of all nodes with a label strictly smaller than 4. This instance together with all sparsest cuts is shown in Figure 6.13. Note that new node balances are obtained by routing the flow on outgoing critical edges back to their tails. In this manner three new sinks are created. Also note that this instance is unique as the node labels of a thin flow are unique and that the restriction of any thin flow respects the congestion of a sparsest cut on this instance. Hence, the flow on all edges crossing some sparsest cut in forward direction are unique as the congestion of these edges must be equal to the sparsest cut value. This shows with respect to the thin flow condition (6.3) that the label of these edges is defined by its congestion (but in general not uniquely). This is exactly the property of all light blue edges. Further, an edge crossing a sparsest cut in backward direction carries never flow which is the meaning of the dark gray edges. In total, this motivates the following definition collecting all light and dark blue edges similarly to the critical network. Note that the sparsest cuts behind the inclusionwise minimal sparsest cut are recognizable one-to-one in the b-flow instance induced by nodes with label 3. Recall that Definition 6.19 encodes all sparsest cuts of a b-flow instance.

Definition 6.42 (Quasi-critical Network and Edges)**.** Let $\mathcal{I} := (G, u, s, b)$ be thin flow without resetting instance and $(\ell_v)_{v \in V} \in \mathbb{R}_+^V$ be the unique family of node labels. For each possible label $\ell \in \mathbb{R}_+$, which is assigned to at least one node, let $X_\ell := \{v \in V \mid \ell_v = \ell\}$ be the set of all nodes marked with this label. For positive ℓ construct the b-flow instance induced by X_ℓ and let $\mathcal{X}_\ell := \{X_{\ell,1}, \dots, X_{\ell,k_\ell}\}$, $k_\ell \in \mathbb{N}$ be the partition encoding all sparsest cuts. For the label 0, sort the nodes in X_0 arbitrarily starting with s and let $\mathcal{X}_0$ be the node partition where each subset consists only of one node. Further, let $\mathcal{X} := \bigcup_\ell \mathcal{X}_\ell$ be the combined node partition of V and E_0^{qc} be the edge set containing all edges $vw \in E$ with $v \in X_{\ell_v,i}$ and $w \in X_{\ell_w,j}$ such that (ℓ_v, i) is lexicographically greater than (ℓ_w, j), i.e., $\ell_v > \ell_w$ or $\ell_v = \ell_w$ and $i > j$.

The *quasi-critical network* $\mathcal{N}_{\text{qc}} := (G^{\text{qc}}, u^{\text{qc}}, s)$ of $\mathcal{I}$ is defined as follows. The directed graph G^{qc} arises out of G by first deleting all edges of E_0^{qc} and subsequently contracting all node subsets in $\mathcal{X}^{\text{qc}}$. Moreover, restricting u to G^{qc} results in u^{qc}. Finally, we call each edge contained in G^{qc} *quasi-critical* and each

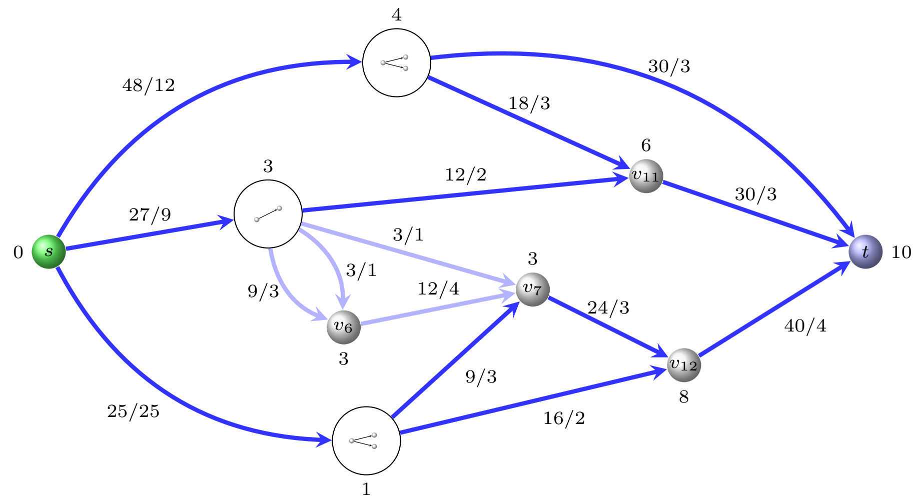


Figure 6.14: The quasi-critical network of the thin flow instance shown in Figure 6.11.

edge of E_0^{qc} *quasi-empty* if it is no element of $E\big(G[X_0]\big)$.

Note that the quasi-critical network is unique up to the order of the nodes in X_0. The quasi-critical network of the thin flow instance shown in Figure 6.11 is illustrated in Figure 6.14 and consists of all blue edges. In particular, the node sets $\{v_1, v_2, v_3\}$, $\{v_4, v_5\}$, and $\{v_8, v_9, v_{10}\}$ are contracted. The set $E_0^{\mathrm{qc}} = \{v_3v_4, v_6v_5, v_{11}v_7,\}$ of quasi-empty edges refers exactly to the gray edges in Figure 6.11. Moreover, observe that the critical network is a minor of the quasi-critical network and that the corresponding sources coincide which is directly implied by the corresponding definitions.

Like the critical network also the quasi-critical network is computable during the Thin Flow algorithm. But instead of contracting the nodes behind an inclusionwise *minimal* sparsest cut Corollary 6.18 shows that we always have to contract the nodes behind an inclusionwise *maximal* sparsest cut. So if we change Step (2) for this purpose, we are able to iteratively construct the quasi-critical network of some thin flow without resetting instance. In fact, such an algorithm computing only the quasi-critical network is presented in Corollary 6.18 in case we change the termination condition to $q = 0$. That is, we iterate as long as the node balances of the induced b-flow instances are not equal to 0.

Based on the observation that quasi-critical edges occur in sparsest cuts and critical edges in inclusionwise minimal sparsest cuts, we are able to characterize these edges from another point of view.

Lemma 6.43. *Let (G, u, s, b) be some thin flow instance. Then an edge e is critical if and only if a small variation of u_e, in which direction however, causes that the node labels of a thin flow become different. Similarly, an edge is quasi-critical if and only if the node labels change in case u_e is decreased.*

Proof. This lemma is established inductively over the number of iterations of the Thin Flow algorithm using the following observations. The congestion of some cut behaves continuously in u. Therefore, changing the capacity of an edge which is not contained in any sparsest cut slightly does not cause a variation in the node labels. Further, decreasing the capacity of any edge in some sparsest cut leads to an increase in the sparsest cut value. In contrast, increasing the capacity of an edge e in some sparsest cut X means that X is no longer a sparsest cut or that the sparsest cut value decreases. Hence, if X is inclusionwise minimal this will decrease at least the label of the nodes in $X' \setminus X$ where X' is the inclusionwise minimal sparsest cut with $e \in E\big(G[X']\big)$ (or $X' = V$ if every sparsest cut contains e). $\square$

Note that something similar holds for the edges carrying no flow in any thin flow without resetting, i.e., the gray edges in Figure 6.11. In fact, light gray empty edges still carry no flow if the capacities are changed slightly. Further, dark gray quasi-empty edges still carry no flow if capacities on the corresponding sparsest cut are decreased slightly.

The following theorem provides the basis for the sensitivity analysis in the next subsection.

Theorem 6.44. *Let (G', u', s') be either the quasi-critical or the critical network of some thin flow instance (G, u, s, b). Then following statements hold:*

(i) The graph G' is acyclic.

(ii) Let $X \subseteq V$ be a node subset which is contracted to some node of G' and ℓ be the common label of all nodes in X. Then the congestion q of a sparsest on b-flow instance induced by X with respect to some thin flow is at most ℓ, i.e., $q \leq \ell$. Moreover, if (G', u', s') is the quasi-critical or the critical network then $q < \ell$.

(iii) Let b' be the restriction of b to G'. Then the restriction of a thin flow to (G', u', s') is unique. In particular, it is a thin flow respecting b' where the label of each edge is defined by its congestion.

Proof. The construction of (G', u', s') during the Thin Flow with Resetting algorithm computing either an inclusionwise minimal or maximal sparsest cut results in a reverse topological order for G'. Note that backward edges are deleted during this construction. Therefore, G' is acyclic. Recalling Lemma 6.17 statement (ii) is also directly implied by this construction.

Consider the restriction x' of a thin flow x to (G', u', s'). Hence, by Proposition 6.15, it respects node balances b'. Further, because of the definition of (G', u', s') we know that the congestion of each edge $e \in E(G')$ equals the common label of the nodes which are contracted to the head of e. In particular, this ensures that e is flow carrying and thus, satisfies thin flow condition (6.3) implying that x' is a thin flow without resetting. The uniqueness of the flow on the critical edges follows directly from the uniqueness of the node labels. Further, the uniqueness of the flow on the quasi-critical edges is established by incorporating Lemma 6.17, in addition. This shows (iii). $\square$

Before we discuss an interesting consequence of this theorem we remark the following.

Remark 6.45. As already mentioned, the critical and the quasi-critical network are computable during the Thin Flow algorithm by contracting the nodes behind the currently inclusionwise minimal and maximal sparsest cut, respectively.

However, assume that we contract the nodes behind an arbitrary sparsest cut in each iteration of the THIN FLOW algorithm. Then it follows directly from Corollary 6.17 that the critical network is a minor of the resulting network $\mathcal{N}^*$ which, in turn, is a minor of the quasi-critical network. On the other hand each such network $\mathcal{N}^*$ is obtainable as follows. Given a thin flow without resetting delete edges carrying no flow and contract nodes with the same label such that the resulting $\mathcal{N}^*$ is acyclic and the label of each remaining edge is defined by its congestion. To see this, note that every sparsest cut computed during the THIN FLOW algorithm corresponds to a directed cut in $\mathcal{N}^*$. Vice versa, every directed cut X of $\mathcal{N}^*$, where the maximum label of the nodes in X are at most as large as the minimum label of the nodes not in X, corresponds to a sparsest cut found by the THIN FLOW algorithm. To be more precise, let ℓ be the minimum label of the nodes not in X. Then X is a sparsest cut of the b-flow instance which is induced by all nodes with a label of at most ℓ.

Remark 6.46. Recall that given an s-flow a thin path P starts at s and the congestion of each s-v-subpath equals the label ℓ_v of v. In this sense, Theorem 6.44(iii) shows that every thin path of a thin flow without resetting corresponds to a flow carrying path in both the critical and the quasi-critical network. Note that the other direction is, of course, not true.

As the last part of this subsection, we mention the following interesting aspect of Theorem 6.44. Let (H, u, s, b) be the restriction of some thin flow instance to the (quasi-)critical network of a thin flow instance. In addition, let $V(H) := \{v_1, \dots, v_{|V(H)|}\}$ be a topological order and let $\ell_1, \dots, \ell_{V(H)}$ be the corresponding node labels. Further, let $b_i := b_{v_i}$ be the node balance of v_i and for $i < j$ let $u_{ij} := u\big(E(v_i, v_j)\big)$ be the total capacity of (quasi-)critical edges connecting v_i to v_j. Since the label of each node is defined by the common congestion of the incoming edges, we get the following equations from flow conservation:

$$b_j + \sum_{i|i<j} \ell_j u_{ij} = \sum_{i|i>j} \ell_i u_{ji} \qquad \forall j \in \big\{1, \dots, |V(H)|\big\} . \tag{6.7}$$

Observe that, up to scaling, $(\ell_1, \dots, \ell_{V(H)})$ is the unique solution to this equation system as the node labels of a thin flow are unique.

On the other hand, given a thin flow network (G, u, s, b), if we know the partition of the nodes with respect to the node labels and the corresponding topological order, we are able to compute the node labels with these equalities. Note that (6.7) is a recursion. Thus, solving the linear equality system (6.7) requires only computing the k recursion steps.

Hence, we are able to find a thin flow by guessing either the quasi-critical or the critical network. That is, we guess a partition of the nodes with the intention that each node of some subset has the same label in some thin flow. In addition, we guess an order of these unknown labels in order to build the equation system (6.7). Solving this equation system gives us candidates for the node labels which we have to verify. Here verifying means that we have to check if the computed node labels are nonnegative and obey the assumed order. Further, we have to check that each induced b-flow instance implies a sparsest cut value which is at most as large as the corresponding computed label. If this consistency check shows no contradiction, we have found a thin flow without resetting. Note that beside guessing either the quasi-critical or the critical network, it is enough to guess some minor of the quasi-critical network which in turn has the critical network as a minor.

As we have an polynomial time algorithm for finding a thin flow without resetting, this approach seems to be less promising. However, for finding a thin flows on similar instances this approach could be successful as the critical network of one instance could lead to a thin flow without resetting on the other instance. This approach is formalized in the next section in case we change only node balances.

6.3.3 Sensitivity Analysis

In this section we analyze the behavior of a thin flow without resetting in case node balances were changed slightly. Note that such an analysis is essential for studying fixed point iterations depending on node balances as discussed in Section 6.2 on page 208. In particular, interpreting thin flows with resetting as fixed points of thin flows without resetting this lays the foundation for existence and uniqueness results presented in Section 6.4. Further, this shows that under a weak assumption a simple fixed point iteration converges to a thin flow with resetting.

Consider some thin flow without resetting instance $\mathcal{I} := (G, u, s, b)$ and let $\Delta b \in \mathbb{R}^V$ with $\Delta b(V) = \sum_{v \in V} \Delta b_v = 0$ be the direction in which we want to change the node balances. Moreover, we assume $\Delta b_s > 0$ and $\Delta b_v < 0$ for all nodes $v \in V \setminus \{s\}$. That means, we only consider the case where additional flow should be send through the network. Besides, this ensures that we are able to interpret Δb as node balances of a thin flow. First, we consider the case where b is locally changed, i.e., we want to find a thin flow without resetting on (G, u, s) respecting $b + \epsilon \Delta b$ for some small $\epsilon > 0$.

Based on the edge classification presented in the previous section, the idea for solving this task is the following. Intuitively, edges which are neither critical nor quasi-critical with respect to $\mathcal{I}$ cannot become critical or quasi-critical if node balances where changed slightly. Hence, instead of considering the entire network (G, u, s) it is enough to construct the new flow on the quasi-critical network $\mathcal{N}_{\text{qc}} := (G^{\text{qc}}, u^{\text{qc}}, s)$.

So let b^{qc} and Δb^{qc} be the restrictions of b and Δb to $\mathcal{N}_{\text{qc}}$, respectively, and consider the unique restriction x^{qc} of a thin flow without resetting of $\mathcal{I}$. The task is to find some Δx^{qc} such that $x^{\text{qc}} + \Delta x^{\text{qc}}$ is a thin flow without resetting on $\mathcal{N}_{\text{qc}}$ respecting node balances $b^{\text{qc}} + \Delta b^{\text{qc}}$. For this consider some critical edge $e = vw \in E(G^{\text{qc}})$. Intuitively, e remains a critical edge if the node balance is changed only slightly. This means that the label of w is defined by the congestion e with respect to x^{qc} as well as with respect to $x^{\text{qc}} + \Delta x^{\text{qc}}$. Hence, we have $\frac{\Delta x^{\text{qc}}}{u_e} = \Delta \ell_w$ assuming that the label of w is changed by $\Delta \ell_w$. However, this can be interpreted as conditions (6.4) for a thin flow *with* resetting. In this sense, a critical edge can be considered as a resetting edge. In this manner, it would be nice to interpret Δx as a thin flow *with* resetting.

For this it remains to consider how the flow is changed on a quasi-critical edge $e = vw \in E(G^{\text{qc}})$ which is not critical. As the label of w is not uniquely defined by the congestion of e, the edge e could become a critical, a quasi-critical, or an usual edge if node balances are changed slightly. But as e remains flow carrying, we have to ensure that condition (6.3) remains valid for e. It turns out that requiring the same condition for e with respect to Δx^{qc} resolves

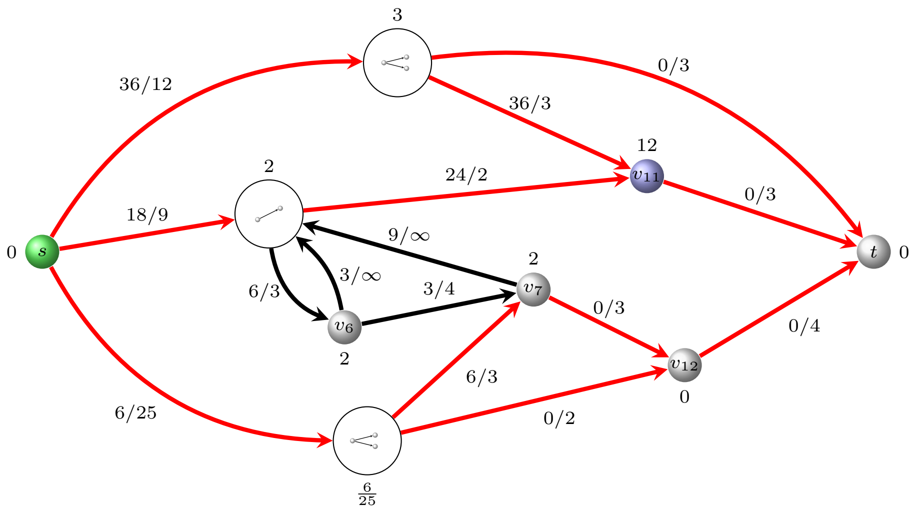


Figure 6.15: The sensitivity network of the thin flow instance shown in Figure 6.11 and a corresponding thin flow with resetting of value 60. Note that nonresetting edges are only drawn if they carry flow. In particular, the quasi-empty edge is missing. Thus, all edges with infinite capacity are backward edges of quasi-critical edges.

this problem in case $\Delta x_e^{\mathrm{qc}} \geq 0$. Unfortunately, in case of $\Delta x_e^{\mathrm{qc}} < 0$ this makes not really sense. On the other hand, interpreting $\Delta x_e^{\mathrm{qc}} < 0$ as sending flow on some backward edge $\overleftarrow{e}$, we have to ensure that the label of w becomes not smaller than the label of v as this would contradict (6.3). This can be achieved by requiring condition (6.3) for $\overleftarrow{e}$ if we set the capacity of $\overleftarrow{e}$ to ∞. In this manner, Δx^{qc} can be seen as thin flow *with* resetting on the following network. Note that we have to add also the quasi-empty edges to the quasi-critical network in order to ensure that they do not contradict condition (6.2).

Definition 6.47 (Sensitivity Network)**.** Let $\mathcal{I} := (G, u, s, b)$ be some thin flow without resetting instance. Then the *sensitivity network* $(G^{\mathrm{s}}, u^{\mathrm{s}}, s, E_1)$ results out of the quasi-critical network of $\mathcal{I}$ as follows: The resetting edges E_1 are the edges $E(G^{\mathrm{c}})$ of the critical network. Further, we insert for each nonresetting edge e a backward edge $\overleftarrow{e}$ of infinite capacity. That is, a backward edge is introduced for each edge which is quasi-critical but not critical. Finally, we add each quasi-empty edge again with a capacity of ∞ and all edges connecting nodes with a label of 0.

Before we observe in Lemma 6.48 that the sensitivity network works as desired we continue the example of the previous section shown in Figure 6.11 and ask the question how the thin flow without resetting is changed if we decrease the node balance of the node v_{11} with label 6 slightly. For this we consider a thin flow with resetting Δx of value 60 on the sensitivity network. Afterwards, we scale Δx such that the new flow is a feasible thin flow without resetting. Note that this approach is feasible as thin flows are scalable. The corresponding sensitivity network and a thin flow with resetting are depicted in Figure 6.15. So it remains to identify a suitable scaling factor $\epsilon > 0$. For this we first consider each type of edges with respect to the original instance and explain to what we

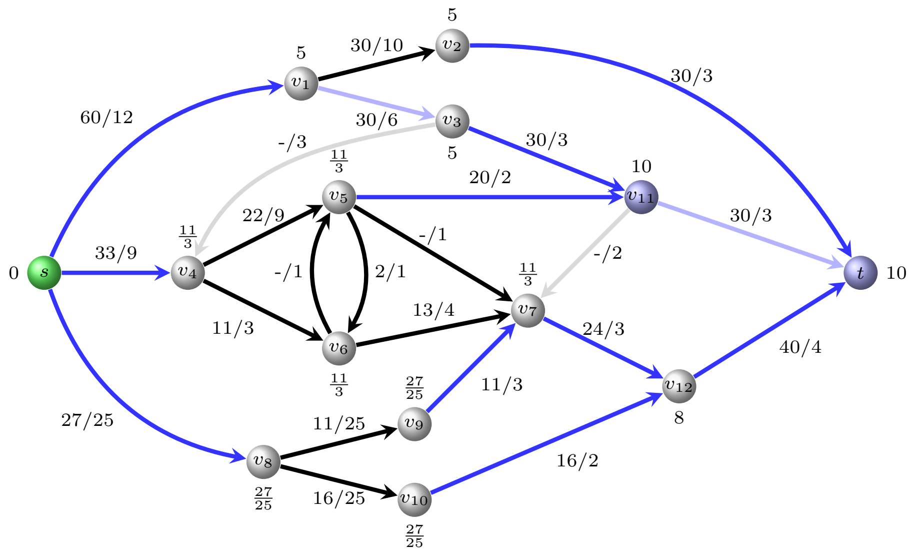


Figure 6.16: The thin flow without resetting on the with modified node balances.

have to pay attention.

First consider the critical edges. Our approach is based on the assumption that the label of critical edges is always defined by its congestion. That is, we have to ensure that the label of the tail node does not exceed the label of the head node. Consider the critical edge connecting the new sink v_{11} with t. The original label of v_{11} is 6 and the one of t is 10. The thin flow with resetting Δx shown in Figure 6.15 would increase the label of v_{11} by 12 whereas the label of t remains constant. Hence, we have to scale Δx with an ϵ of at most $\frac{1}{3}$.

With respect to the quasi-critical edges we first observe that everything is fine in case Δx uses only the forward edges. Unfortunately, if Δx uses also backward edges, we have to ensure that the new flow remains nonnegative. In this sense ϵ is again bounded by $\frac{1}{3}$ if we consider the quasi-critical edge from the contracted node $\{v_4, v_5\}$ to the node v_7.

On quasi-empty edges we have to ensure that the new congestion does not exceed the new flow label. However, Δx uses no quasi-empty edge implying that we get no bound on ϵ. The same holds for empty edges. Usually, we have to check that the new labels still satisfy condition (6.2).

It remains to consider the normal edges which are hidden in the contracted subsets. Here we have to ensure that the induced b-flow instances respect the label of the corresponding contracted node. That is, the new flow must routable over these subsets such that the maximum edge congestion is at most equal to this new label. Considering the upper contracted node $\{v_1, v_2, v_3\}$ in Figure 6.15 we observe that ϵ is again bounded by $\frac{1}{3}$.

Thus, we have to scale Δx with $\epsilon = \frac{1}{3}$ in order to obtain a feasible thin flow with resetting. The resulting flow, which sends 20 additional flow units to the new sink, is shown in Figure 6.16. Also the new types of the edges are recognizable there. In particular, we observe that critical just as normal edges could become quasi-critical. Further, we see that that quasi-critical and quasi-empty edges can change their type to normal edges.

Formalizing the discussion above leads to the following lemma. For this note that the new flow on the original network results out of a thin flow without resetting on an *extended* quasi-critical network which contains also the quasi-empty edges. Further, the thin flow $x^{\mathrm{qc}} + \Delta x^{\mathrm{s}}$ on the extended quasi-critical network is made up of the restriction x^{qc} of a thin flow without resetting on the original instance to the quasi critical network and the thin flow with resetting Δx^{s} on the sensitivity network. Hence, similarly to the definition of an augmented flow, the flow $x^{\mathrm{qc}} + \Delta x^{\mathrm{s}}$ is defined as

$$(x^{\mathrm{qc}} + \Delta x^{\mathrm{s}})_e := x^{\mathrm{qc}}_e + \Delta x^{\mathrm{s}}_e - \Delta x^{\mathrm{s}}_{\overleftarrow{e}}$$

for all edges e of the extended quasi-critical network, where we assume $\Delta x^{\mathrm{s}}_{\overleftarrow{e}} = 0$ if no backward edge is introduced for e.

Lemma 6.48. *Consider a thin flow without resetting instance $\mathcal{I} := (G, u, s, b)$ and let $\Delta b \in \mathbb{R}^V$ with $\Delta b(V) = 0$ be the direction in which the node balances should be changed. Moreover, assume that $\Delta b_s > 0$ and $\Delta b_v < 0$ holds for all nodes $v \in V \setminus \{s\}$. Further, let b^s and Δb^s be the restrictions of b and Δb to the sensitivity network $\mathcal{N}_s := (G^s, u^s, s, E_1)$ of $\mathcal{I}$, respectively.*

Consider the unique restriction x^{qc} of a thin flow without resetting of $\mathcal{I}$ to the quasi critical network and let Δx^s be a thin flow with resetting on $\mathcal{N}_s$ respecting node balances Δb^s. Then there exists an $\epsilon > 0$ such that the flow $x^{\mathrm{qc}} + \epsilon \cdot \Delta x^s$ on the extended quasi-critical network is a restriction of a thin flow without resetting on (G, u, s) respecting node balances $b + \epsilon \cdot \Delta b$.

Moreover, corresponding node labels are given by $\ell^{\mathrm{qc}} + \epsilon \cdot \Delta \ell^s$ where ℓ^{qc} and $\Delta \ell^s$ are the node labels corresponding to x^{qc} and Δx^s, respectively. In particular, this means that the new labels of all nodes which are contracted to a node v in the quasi-critical network coincide and are equal to $\ell^{\mathrm{qc}}_v + \epsilon \cdot \Delta \ell^s_v$.

Proof. First, we observe that $x^{\mathrm{qc}} + \epsilon \cdot \Delta x^{\mathrm{s}}$ is a thin flow without resetting on the extended quasi-critical network in case ϵ_0 is small enough. Moreover, we observe that corresponding node labels are given by $\ell^{\mathrm{qc}} + \epsilon \cdot \Delta \ell^{\mathrm{s}}$. Since $x^{\mathrm{qc}} + \epsilon \cdot \Delta x^{\mathrm{s}}$ respects node balances $b^{\mathrm{s}} + \epsilon \cdot \Delta b^{\mathrm{s}}$ by definition, we have to check the thin flow without resetting conditions (6.1)-(6.3). Consider an edge $e = vw$ of the extended quasi-critical network and distinguish the following cases:

Case 1: Let $e \in E$ be a critical edge of $\mathcal{I}$ and, hence, a resetting edge of the sensitivity network. Thus, it holds:

$$\frac{x^{\mathrm{qc}}_e}{u_e} = \max\left\{\ell^{\mathrm{qc}}_v, \frac{x^{\mathrm{qc}}_e}{u_e}\right\} = \ell^{\mathrm{qc}}_w \quad \text{and} \qquad \frac{\Delta x^{\mathrm{s}}_e}{u_e} = \Delta \ell^{\mathrm{s}}_w$$

$$\text{implying} \quad \frac{x^{\mathrm{qc}}_e + \epsilon \cdot \Delta x^{\mathrm{s}}_e}{u_e} = \ell^{\mathrm{qc}}_w + \epsilon \cdot \Delta \ell^{\mathrm{s}}_w.$$

In addition, we get $\dfrac{x^{\mathrm{qc}}_e + \epsilon \cdot \Delta x^{\mathrm{s}}_e}{u_e} = \max\left\{\ell^{\mathrm{qc}}_v + \epsilon \cdot \Delta \ell^{\mathrm{s}}_v, \dfrac{x^{\mathrm{qc}}_e + \epsilon \cdot \Delta x^{\mathrm{s}}_e}{u_e}\right\}$ if $\Delta \ell^{\mathrm{s}}_v \leq \Delta \ell^{\mathrm{s}}_w$ holds, or $\Delta \ell^{\mathrm{s}}_v > \Delta \ell^{\mathrm{s}}_w$ and $\epsilon \leq \dfrac{\ell^{\mathrm{qc}}_w - \ell^{\mathrm{qc}}_v}{\Delta \ell^{\mathrm{s}}_v - \Delta \ell^{\mathrm{s}}_w}$ are valid. Note that the upper bound on ϵ is strictly positive as e is critical implying $\ell^{\mathrm{qc}}_w > \ell^{\mathrm{qc}}_v$.

Case 2: Let e be a quasi-critical but not a critical edge of $\mathcal{I}$ with $\Delta x^{\mathrm{s}}_{\overleftarrow{e}} = 0$. Thus, it holds:

$$\frac{x^{\mathrm{qc}}_e}{u_e} = \ell^{\mathrm{qc}}_v = \ell^{\mathrm{qc}}_w \qquad \text{and} \qquad \max\Big\{\Delta\ell^{\mathrm{s}}_v, \frac{\Delta x^{\mathrm{s}}_e}{u_e}\Big\} = \Delta\ell^{\mathrm{s}}_w$$

$$\text{implying} \qquad \max\Big\{\ell^{\mathrm{qc}}_v + \epsilon\cdot\Delta\ell^{\mathrm{s}}_v, \frac{x^{\mathrm{qc}}_e + \epsilon\cdot\Delta x^{\mathrm{s}}_e}{u_e}\Big\} = \ell^{\mathrm{qc}}_w + \epsilon\cdot\Delta\ell^{\mathrm{s}}_w\ .$$

Case 3: Let e be a quasi-critical but not a critical edge of $\mathcal{I}$ with $\Delta x^{\mathrm{s}}_{\overleftarrow{e}} > 0$. Thus, since backward edges have infinite capacities, we obtain:

$$\frac{x^{\mathrm{qc}}_e}{u_e} = \ell^{\mathrm{qc}}_v = \ell^{\mathrm{qc}}_w \qquad \text{and} \qquad \Delta\ell^{\mathrm{s}}_v = \Delta\ell^{\mathrm{s}}_w$$

$$\text{implying} \qquad \ell^{\mathrm{qc}}_v + \epsilon\cdot\Delta\ell^{\mathrm{s}}_v = \ell^{\mathrm{qc}}_w + \epsilon\cdot\Delta\ell^{\mathrm{s}}_w\ .$$

Further, we know $\dfrac{\Delta x^{\mathrm{s}}_e}{u_e} \leq \ell^{\mathrm{qc}}_w = \ell^{\mathrm{qc}}_v$ and hence:

$$\frac{(x + \epsilon\cdot\Delta x^{\mathrm{s}})_e}{u_e} \leq \frac{x^{\mathrm{qc}}_e + \epsilon\cdot\Delta x^{\mathrm{s}}_e}{u_e} \leq \ell^{\mathrm{qc}}_v + \epsilon\cdot\Delta\ell^{\mathrm{s}}_v\ .$$

In order to obtain a feasible flow we note that $\dfrac{(x^{\mathrm{qc}} + \epsilon\cdot\Delta x^{\mathrm{s}})_e}{u_e} \geq 0$ holds for all $\epsilon \leq \dfrac{x^{\mathrm{qc}}_e}{\Delta x^{\mathrm{s}}_{\overleftarrow{e}}}$.

Case 4: Let $e \in E$ be a quasi-empty but not an empty edge of $\mathcal{I}$. Since e has infinite capacity on the sensitivity network, we know:

$$\ell^{\mathrm{qc}}_v = \ell^{\mathrm{qc}}_w \qquad \text{and} \qquad \Delta\ell^{\mathrm{s}}_v = \Delta\ell^{\mathrm{s}}_w$$

$$\text{implying} \qquad \ell^{\mathrm{qc}}_v + \epsilon\Delta\ell^{\mathrm{s}}_v = \ell^{\mathrm{qc}}_w + \epsilon\cdot\Delta\ell^{\mathrm{s}}_w\ .$$

Further, we have $\dfrac{(x + \epsilon\cdot\Delta x^{\mathrm{s}})_e}{u_e} = \dfrac{\epsilon\cdot\Delta x^{\mathrm{s}}_e}{u_e} \leq \ell^{\mathrm{qc}}_v + \epsilon\cdot\Delta\ell^{\mathrm{s}}_v$ if either $\dfrac{\Delta x^{\mathrm{s}}_e}{u_e} \leq \Delta\ell^{\mathrm{s}}_v$ holds, or $\dfrac{\Delta x^{\mathrm{s}}_e}{u_e} > \Delta\ell^{\mathrm{s}}_v$ and $\epsilon \leq \ell^{\mathrm{qc}}_v \cdot \Big(\dfrac{\Delta x^{\mathrm{s}}_e}{u_e} - \Delta\ell^{\mathrm{s}}_v\Big)^{-1}$ are valid.

This shows that $x^{\mathrm{qc}} + \epsilon\cdot\Delta x^{\mathrm{s}}$ is a thin flow without resetting on the extended quasi-critical network for small enough ϵ_0. It remains to verify the thin flow without resetting conditions for the empty and the normal edges. For an empty edge $e = vw$ we know that $\ell^{\mathrm{qc}}_v > \ell^{\mathrm{qc}}_w$ holds and that the flow on e remains 0. Thus, we have to ensure $\ell^{\mathrm{qc}}_v + \epsilon\cdot\Delta\ell^{\mathrm{s}}_v \geq \ell^{\mathrm{qc}}_w + \epsilon\cdot\Delta\ell^{\mathrm{s}}_w$ which is either implied by $\Delta\ell^{\mathrm{s}}_v \geq \Delta\ell^{\mathrm{s}}_w$ or $\Delta\ell^{\mathrm{s}}_v < \Delta\ell^{\mathrm{s}}_w$ and $\epsilon \leq \dfrac{\ell^{\mathrm{qc}}_v - \ell^{\mathrm{qc}}_w}{\Delta\ell^{\mathrm{s}}_w - \Delta\ell^{\mathrm{s}}_v}$.

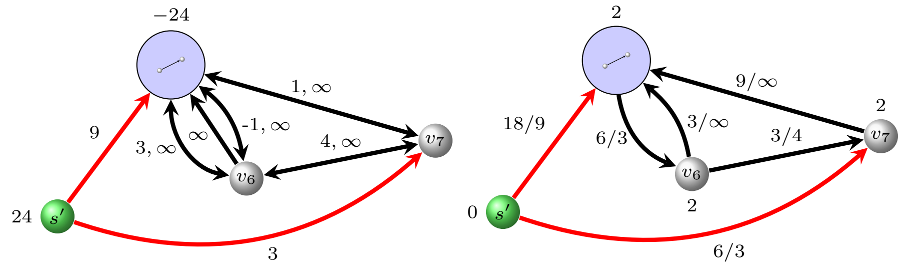
Finally, we show that b-flow instances induced by the contracted node subsets are feasible with respect to the node labels $\ell^{\mathrm{qc}} + \epsilon\cdot\Delta\ell^{\mathrm{s}}$. That is, we show that $x^{\mathrm{qc}} + \epsilon\cdot\Delta x^{\mathrm{s}}$ can be routed feasible over the contracted sets providing that ϵ_0 is small enough. Therefore, let $X \subsetneq V$ be some node subset which is contracted to some node v of the quasi-critical network. As ℓ^{qc}_v equals the common label which is assigned by the original thin flow without resetting to each node of X, we know $\ell^{\mathrm{qc}}_v > q$ by Lemma 6.44 where q is the congestion of a sparsest cut of the induced b-flow instance. Since we know that q changes continuously with the node balances, there exists an $\epsilon_X > 0$ such that $x^{\mathrm{qc}} + \epsilon_X \cdot \Delta x^{\mathrm{s}}$ can be routed over $G[X]$ with a congestion of at most $\ell_v^{\mathrm{qc}} + \epsilon_X \cdot \Delta\ell_v^{\mathrm{s}} \geq \ell_v^{\mathrm{qc}}$. Hence, ensuring $\epsilon \leq \min\{\epsilon_X \mid X$ is contracted to some quasi-critical node $\}$ gives the desired result. □

In the following we discuss how the type of an edge can be changed in case node balanced were increased. Considering the critical edges of some thin flow without resetting instance $\mathcal{I}$, which are handled in Case 1 of Lemma 6.48, we see that they either remain critical edges or become quasi-critical edges if they define ϵ. Similarly, empty edges either remain empty edges or become quasi-empty edges. Further, quasi-critical edges which are handled in Case 2 and 3 remain either quasi-critical edges or could become critical and normal edges. If ϵ is defined by Case 3 they can even become quasi-empty edges. However, on the way to quasi-empty edges, meanwhile, they must be normal. A similar behavior holds for quasi-empty edges. Finally, normal edges could remain normal or become quasi-critical and quasi-empty. To summarize this, assume that the types are ordered like empty, quasi-empty, normal, quasi-critical, and critical. Then decreasing node balances means that the type of an edge changes continuously with respect to this order.

Lemma 6.48 shows how a thin flow without resetting is changed if node balances were increased. Further, the variation of the thin flow without resetting depends on a thin flow *with* resetting. In Section 6.4 we use this result for establishing existence and uniqueness for arbitrary thin flow on arbitrary resetting instances. This seems to be circular logic. However, we break this circular argumentation by showing that a thin flow with resetting on a sensitivity network is computable. Here, we use the special structure of the quasi-critical and the critical network which are used for defining the sensitivity network. In particular, we need that the critical network is acyclic and that the quasi-critical network is a refinement of the critical network.

The next algorithm computes such a thin flow with resetting on the sensitivity network as follows. We traverse the critical network in reverse topological order and assign uniquely determined flow values to the edges. So consider a particular node v of the critical network and assume that we know the flow on the outgoing edges. Further, assume that the node balance of v is equal to 0 and that v occurs one-to-one in the sensitivity network, i.e., that v is not contracted. This means that the corresponding b-flow instance has a sparsest cut value which is strictly greater than the label defining v. Then all incoming edges are resetting edges. Therefore, we know that the flow on the incoming edges must amount to the flow on the outgoing edges on the one hand and must lead to the same congestion on the other hand. Clearly, from these two observations we are able to uniquely determine the flow on the incoming edges and, hence, we can proceed with the next node.

However, if v corresponds to more than one node of the quasi-critical network, it requires a little more work in order to obtain the flow on the incoming edges. In this case we have to solve the thin flow without resetting instance constructed in Step (4) of the following algorithm. Such an instance is exemplarily shown in Figure 6.17 for the nodes with a label of 3 given the quasi-critical network shown in Figure 6.14. Note that the quasi-empty edge leaving v_6 is inserted.



(a) The thin flow without resetting instance. The capacities are shown on the edges. All backward edges point to the left.

(b) A thin flow without resetting. Only flow carrying edges are shown.

Figure 6.17: The thin flow instance constructed in Step (4) of the Sensitivity Thin Flow with Resetting algorithm according to the nodes with label 3 in Figure 6.14.

Sensitivity Thin Flow with Resetting Algorithm

Input: A sensitivity network $\mathcal{N}_s := (G^{\mathrm{s}}, u^{\mathrm{s}}, s, E_1)$ and a node balance $b \in \mathbb{R}_+^{V(G^{\mathrm{s}})}$.

Output: A thin flow $x \in \mathbb{R}_+^{E(G^{\mathrm{s}})}$ with resetting respecting b.

(1) Set $x := 0$ to the zero flow.

(2) Let $s^{\mathrm{c}} = v_1, \ldots, v_k$ be a topological order of the critical network for some suitable $k \in \mathbb{N}$ and set $i := k$.

(3) Let $X \subsetneq V(G^{\mathrm{s}})$ be the set of nodes which are contracted to v_i.

(4) Construct a thin flow instance $\mathcal{I}' := (G', u', s', b')$ as follows:

- The graph G' arises out of the induced subgraph $G^{\mathrm{s}}[X]$ by adding a super source s' and an edge $e_{s'} := s'w$ for each edge $e = vw \in \delta^-_{G^{\mathrm{s}}}(X)$.
- For each edge in $E\big(G^{\mathrm{s}}[X]\big)$ set $u'_e := u_e$ and for each edge $e \in \delta^-_{G^{\mathrm{s}}}(X)$ set $u'_{e_{s'}} := u_e$.
- Set $b'_v := b_v - x\big(\delta^+_{G^{\mathrm{s}}}(v) \cap \delta^+_{G^{\mathrm{s}}}(X)\big)$ for all $v \in X$ and $b'_{s'} := -b(X)$.

(5) Set $x_e := x'_{e_{s'}}$ for all $e \in \delta^-_{G^{\mathrm{s}}}(X)$ where x' is a thin flow without resetting on $\mathcal{I}'$.

(6) If $i = 1$ return x. Otherwise set $i := i - 1$ and go to (2)

Before we show the correctness of the Sensitivity Thin Flow with Resetting algorithm we mention the following aspect. The basic feature, which allows us to compute a thin flow with resetting, is that due to the resetting edges the sensitivity network is decomposable into several thin flow *without* resetting instances. Of course, these nonresetting instances constructed in Step (4) are not independent but corresponds one-to-one to a node of the critical network. As the critical network is acyclic we calculate the thin flow without resetting iteratively. Once a thin flow x with resetting on the sensitivity network is found, the nonresetting instances are describable as follows: Let $s = v_1, \ldots, v_k$ be the topological order of the nodes of the critical network as in Step (2) and $X_1, \ldots, X_k$

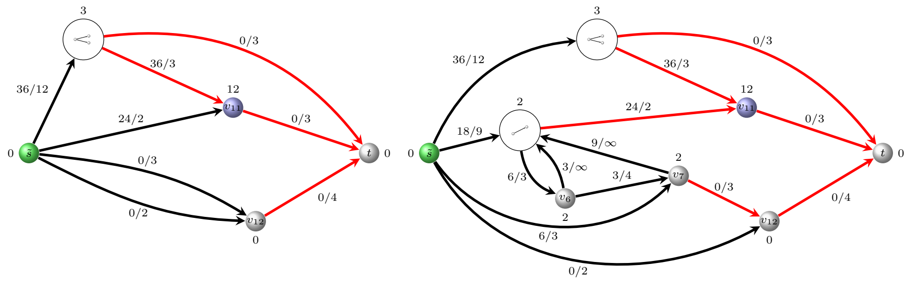


Figure 6.18: The invariant used in the proof of Lemma 6.49 before and after the quasi-critical nodes with label 3 are handled by the Sensitivity Thin Flow with Resetting algorithm. The topological order used in Step (2) corresponds to the reverse order of the labels.

be the corresponding contracted subsets of quasi-critical nodes. Then for every i the nonresetting instance of v_i coincides with the b-flow instance on the sensitivity network induced by $\bigcup_{j=1}^{i} X_i$ with respect to x where the set $\bigcup_{j=1}^{i-1} X_i$ is contracted.

Lemma 6.49. *The* Sensitivity Thin Flow with Resetting *algorithm works correctly.*

Proof. Considering x before Step (6) we show the following invariant. Let v_i be the node of the critical network which has been handled in the current iteration, i.e., we are in iteration $k-i+1$. Consider the b-flow instance $(\tilde{G}, \tilde{u}, \tilde{s}, \tilde{b})$ resulting out of $(G^{\mathrm{s}}, u^{\mathrm{s}}, s, b)$ where the set $\tilde{X} := \bigcup_{j \geq i} X_j$ is contracted. Then defining $\tilde{E}_1 := E_1 \cap E\big(G^{\mathrm{s}}[\tilde{X}]\big)$, the restriction of x is a thin flow with resetting on $\tilde{\mathcal{I}} := (\tilde{G}, \tilde{u}, \tilde{s}, \tilde{b}, \tilde{E}_1)$. This invariant is illustrated in Figure 6.18 for the sensitivity network in Figure 6.15.

The validity of this invariant follows directly from the induction principle. The thin flow with resetting conditions on $\tilde{\mathcal{I}}$ outside of the current $\mathcal{I}'$ follow from the induction assumption as all outgoing edges of X are resetting edges. Moreover, the thin flow with resetting conditions inside of $\mathcal{I}'$ are directly implied by Step (4) and (5) as x' is a thin flow without resetting on $\mathcal{I}'$. After the last iteration we get $\tilde{\mathcal{I}} = (G^{\mathrm{s}}, u^{\mathrm{s}}, s, E_1, b)$. Hence, this lemma is proven. ☐

In the following we estimate how the flow on the outgoing edges of s is changed if we increase node balances. As already mentioned, this is an important key in obtaining existence and uniqueness results for thin flow with resetting. For this we consider two thin flow without resetting instances $\mathcal{I} := (G, u, s, b)$ and $\tilde{\mathcal{I}} := (G, u, s, \tilde{b})$ with $\Delta b := \tilde{b} - b$. Further, we assume that Lemma 6.48 is applicable with $\epsilon = 1$. If not, decrease $\tilde{b}$ in the direction of $-\Delta b$ until this holds. In order to give an estimation for the corresponding unique node labels, we need a *common critical network*. This is a network which is a minor of the quasi-critical network of both instances and which has in turn the critical network of both instances as a minor. For example, consider the thin flow without resetting instance shown in Figure 6.11 and the one where the balance of the node v_{11} is decreased by 20 (see Figure 6.16 for a corresponding thin flow). Then a common critical network is the critical network of the original instance. This is illustrated in Figure 6.19.

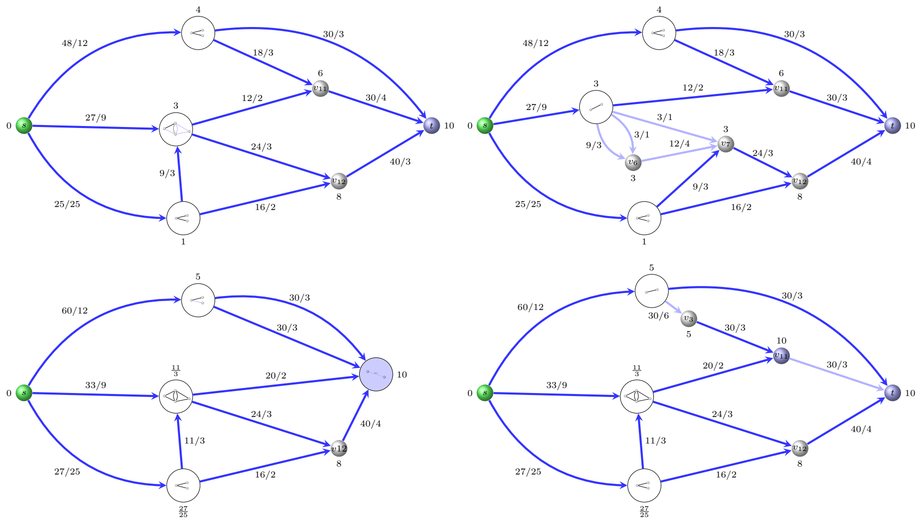


Figure 6.19: The critical and quasi-critical network of the thin flow instance shown in Figure 6.11 are depicted on top. Below, the critical and quasi-critical network for slightly changed node balances are drawn (see Figure 6.16 for the thin flow). A common critical network is the critical network of the original instance shown top left.

Lemma 6.50. *Let $\mathcal{I} := (G, u, s, b)$ and $\tilde{\mathcal{I}} := (G, u, s, \tilde{b})$ be two thin flow without resetting instances with $\Delta b := \tilde{b} - b$ such that Lemma 6.48 is applicable with $\epsilon = 1$. Then there exists a common critical network.*

Proof. Let $\mathcal{N}^{\mathrm{c}}$, $\mathcal{N}^{\mathrm{qc}}$, $\tilde{\mathcal{N}}^{\mathrm{c}}$, and $\tilde{\mathcal{N}}^{\mathrm{qc}}$ be the critical and quasi-critical network of $\mathcal{I}$ and $\tilde{\mathcal{I}}$, respectively. Next we construct a network $\mathcal{N}^*$ out of $\mathcal{N}^{\mathrm{c}}$ which is a common critical network for $\mathcal{I}$ and $\tilde{\mathcal{I}}$ as we see subsequently.

Apply the Sensitivity Thin Flow with Resetting algorithm on the sensitivity network of $\mathcal{I}$ with node balances Δb. In each iteration expand the critical node v_i with the critical network of the current $\mathcal{I}'$ constructed in Step (4). Note that we have to fan out s' in order to place the critical network of $\mathcal{N}'$ feasible within the critical network $\mathcal{N}^{\mathrm{c}}$ of $\mathcal{I}$. The resulting network is $\mathcal{N}^* = (H^*, u^*, s)$.

In order to show that $\mathcal{N}^*$ is a common critical network first observe that $\mathcal{N}^{\mathrm{c}}$ is obviously a minor of $\mathcal{N}^*$ as $\mathcal{N}^*$ results out of $\mathcal{N}^{\mathrm{c}}$ by expanding critical nodes. Moreover, the construction of $\mathcal{N}^*$ shows directly that $\mathcal{N}^*$ is a minor of $\mathcal{N}^{\mathrm{qc}}$. Note that $\mathcal{N}^*$ could also be obtained from $\mathcal{N}^{\mathrm{qc}}$ by contracting nodes with the same label based on the thin flow x' of the current $\mathcal{I}'$ in each iteration of the Sensitivity Thin Flow with Resetting algorithm.

It remains to show that $\mathcal{N}^*$ is a minor of $\tilde{\mathcal{N}}^{\mathrm{qc}}$ and has $\tilde{\mathcal{N}}^{\mathrm{c}}$ as a minor. For this, let $v \in V(H^*)$ be a node of H^* and let $X \subsetneq V$ be the node subset of G which is contracted to v. Note that X is first contracted to an induced subgraph of $\mathcal{N}^{\mathrm{qc}}$ and, subsequently, to v during the construction of $\mathcal{N}^*$. Considering a thin flow on $\tilde{\mathcal{I}}$, it follows from Lemma 6.48 and the uniqueness of the node labels that all nodes of X have the same label. This shows that $\tilde{\mathcal{N}}^{\mathrm{c}}$ is a minor of $\mathcal{N}^*$. Further, the label of every edge in $\delta^-_{G^*}(v) \cup \delta^+_{G^*}(v)$ must be defined by its congestion. Hence, as $\mathcal{N}^*$ is acyclic by construction, $\mathcal{N}^*$ is a minor of $\tilde{\mathcal{N}}^{\mathrm{qc}}$ (see Remark 6.45). $\square$

The existence of a common critical network of two thin flow instances allows us to compute new node labels in case the node balances increase slightly. For this, consider two thin flow without resetting instances with node balances b and $\tilde{b}$ such that Lemma 6.48 is applicable with $\epsilon = 1$ for $\Delta b := \tilde{b} - b$. Further, let $\mathcal{N}^* := (H^*, u^*, s)$ be the common critical network. For applying the equation system (6.7) let $V(H^*) := \{v_1, \dots, v_{|V(H^*)|}\}$ be a topological order of H^*. Moreover, let $\ell_1, \dots, \ell_{V(H^*)}$ and $\tilde{\ell}_1, \dots, \tilde{\ell}_{V(H^*)}$ be corresponding node labels and $b_i := b_{v_i}$ and $\tilde{b}_i := \tilde{b}_{v_i}$ be the corresponding restricted node balances with respect to b and $\tilde{b}$, respectively. Finally, for $i < j$ let $u_{ij} := u\big(E(v_i, v_j)\big)$ be the total capacity of quasi-critical edges connecting v_i to v_j. Applying the equation system (6.7) to b and $\tilde{b}$ we obtain:

$$\begin{aligned} b_j + \sum_{i|i<j} \ell_j u_{ij} &= \sum_{i|i>j} \ell_i u_{ji} && \forall j \in \big\{1, \dots, |V(H^*)|\big\} \\ \tilde{b}_j + \sum_{i|i<j} \tilde{\ell}_j u_{ij} &= \sum_{i|i>j} \tilde{\ell}_i u_{ji} && \forall j \in \big\{1, \dots, |V(H^*)|\big\} \end{aligned}$$

Subtracting the first equation system componentwise from the second equation system we get:

$$\Delta b_j + \sum_{i|i<j} \Delta\ell_j u_{ij} = \sum_{i|i>j} \Delta\ell_i u_{ji} \qquad \forall j \in \big\{1, \dots, |V(H^*)|\big\} \tag{6.8}$$

This equation is, in fact, a recursion as the common critical network is acyclic. This leads to the following lemma which bounds the flow in case node balances were increased slightly.

Lemma 6.51. *Let $\mathcal{I} := (G, u, s, b)$ and $\tilde{\mathcal{I}} := (G, u, s, \tilde{b})$ be two thin flow without resetting instances with $\Delta b := \tilde{b} - b$ such that Lemma 6.48 is applicable with $\epsilon = 1$. Further, let X be a directed cut of a common critical network (H^*, u^*, s) and $\bar{E} \subseteq \delta^+_{H^*}(X)$ be a subset of edges contained in this cut. Then we have*

$$\Delta x_e \geq 0 \qquad \forall e \in E(H^*) \tag{6.9}$$

$$\text{and} \qquad \|\Delta x(\bar{E})\|_1 = \Delta x(\bar{E}) \leq \Delta b_1 \,. \tag{6.10}$$

Moreover, (6.9) *is satisfied strictly by some $e \in E$ if and only if there exists a directed path in H^* from s to a node $v \in V(H^*)$ with $\Delta b_v < 0$ containing e. On the other hand, equality holds in* (6.10) *if and only if every directed path in H^* from s to a node $v \in V(G^*)$ with $\Delta b_v < 0$ contains an edge of $\bar{E}$.*

Proof. Transforming equation (6.8) we obtain:

$$\Delta\ell_j = \frac{1}{\sum_{i|i<j} u_{ij}} \big(-\Delta b_j + \sum_{i|i>j} \Delta\ell_i u_{ji}\big) \qquad \forall j \in \big\{2, \dots, |V(H^*)|\big\}$$

Visiting the nodes of H^* in reverse topological order this becomes a recurrence for the $\Delta\ell_j$'s. This shows inductively that $\Delta\ell_j > 0$ holds if and only if there exists a path in H^* from s to a node $v \in V(G^*)$ with $\Delta b_v < 0$ containing v_j and

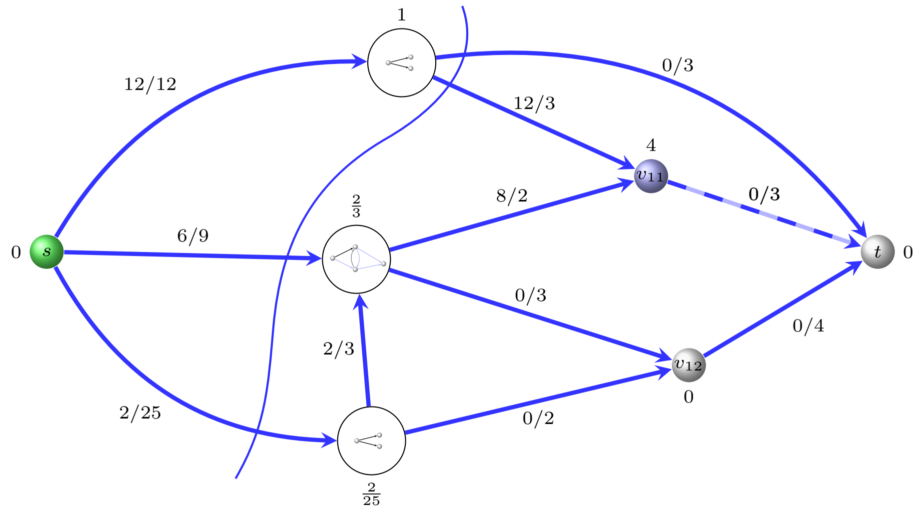


Figure 6.20: The flow $\Delta(x)$ with corresponding node labels $\Delta(\ell)$ on the common critical network for the instance depicted in Figure 6.11 where the node balance of v_{11} is decreased by 20. The sinuous line stands for a directed cut.

that $\Delta\ell_j > 0$ if and only if there exists *no* path in H^* from s to a node $v \in V(G^*)$ with $\Delta b_v < 0$ containing v_j. As the label of every edge $e \in E(H^*)$ is defined by its congestion with respect to $\delta(X)$, this proves (6.9) and the corresponding "moreover"-part. Further, (6.9) implies $\|\Delta x(\bar{E})\|_1 = \Delta x(\bar{E})$ and, as X is a directed cut, we know $\Delta x(\bar{E}) \leq \Delta b_1$ from flow conservation which proves (6.10). Finally, the corresponding "moreover"-part follows directly from the "moreover"-part of (6.9). □

For verifying Lemma 6.51 the common critical network together with the difference flow Δx for the scenario, which is exemplarily considered through this section, is shown in Figure 6.20. So consider the cut X consisting of s and the contracted node defined by $\{v_1, v_2, v_3\}$. If $\bar{E}$ contains only the edge sv_8 we see that (6.10) is satisfied strictly. On the other hand, if the edges sv_8, sv_3 and v_3v_{11} are contained in $\bar{E}$ then equality holds for (6.10) because every s-v_{11}-path in the common critical network traverses an edge of $\bar{E}$. However, in both cases there exists an s-v_{11}-path on the common critical network which contains an edge of $\bar{E}$. But in contrast, if we consider the original network where no node is contracted, we observe that only in the second case there exists a directed s-v_{11}-path which contains an edge of $\bar{E}$. This is due to the additional backward edges of infinite capacities. This observation is handled in the following corollary.

Corollary 6.52. *Using the notation of Lemma 6.51 assume that $\Delta x(\bar{E}) = \Delta b_1$ holds. Then for each node v with $\Delta b_v < 0$ there exists a directed s-v-path in G containing an edge of $\bar{E}$.*

Proof. Let $v \in V$ be some node with $\Delta b_v < 0$. Given a thin flow without resetting on $\tilde{\mathcal{I}}$ there exists a flow carrying s-v-path P. Further, P corresponds to a directed s-v^*-path P^* in H^*. So Lemma 6.51 shows that $\Delta x_e > 0$ holds for all edges e of P^*. Hence, $\Delta x(\bar{E}) = \Delta b_1$ implies that P^* crosses the cut X

along some edge $e^* \in \bar{E}$. Thus, since each edge of P^* is also contained in P, this proof is finished. □

Next we consider the flow behavior if we send an arbitrary amount of additional flow through the network. As we see, Lemma 6.51 carries directly over to this case. However, as the step size $\epsilon > 0$ in Lemma 6.48 depends on the current node balances, we cannot apply Lemma 6.51 inductively in order to change the node balances arbitrary.

Theorem 6.53. *Consider two thin flow without resetting instances with node balances b and $\tilde{b}$ such that $\Delta b := \tilde{b} - b \geq 0$. Then for each $\bar{E} \in \delta^+(s)$ we know that*

$$\Delta x_e \geq 0 \qquad \forall e \in E \tag{6.11}$$

$$\text{and} \qquad \|\Delta x(\bar{E})\|_1 = \Delta x(\bar{E}) \leq \Delta b_1 \;. \tag{6.12}$$

Proof. Assuming the opposite we increase b in the direction of Δb as long as this theorem holds. Then we show that this theorem must also hold at this point and, hence, we can increase b further because of Lemma 6.51. This shows that we can reach $\tilde{b}$ in such a manner proving this theorem.

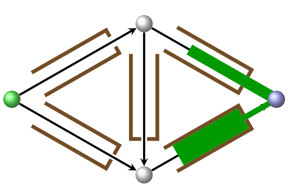
Let $\lambda \in [0,1]$ be maximal such that this theorem holds for node balances b and $b' := b + \lambda' \Delta b$ for all $\lambda' \in (0, \lambda)$. Lemma 6.51 ensures that λ is strictly positive. So consider a nondecreasing nonnegative sequence $(\lambda_k)_{k\in\mathbb{N}}$ converging to λ ($\lim_{k\to\infty} \lambda_k = \lambda$) and let $(x^k)_{k\in\mathbb{N}}$ be a corresponding sequence of thin flow respecting node balances $b^k := b + \lambda_k \Delta b$.

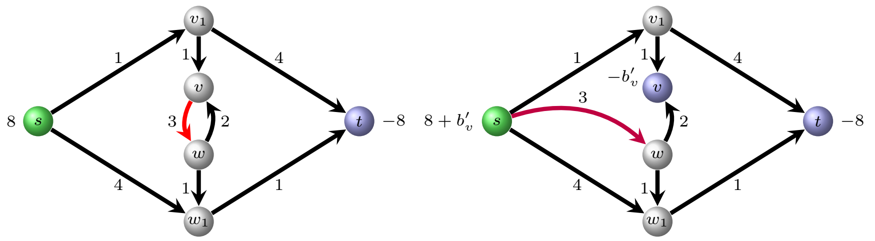
As the node labels do not decrease, we know that $(x^k)_{k\in\mathbb{N}}$ is bounded and, hence, there exists a converging subsequence. So assume without loss of generality that $(x^k)_{k\in\mathbb{N}}$ is converging to the limit point $x^* := \lim_{k\to\infty} x^k$. Since $(x^k)_{k\in\mathbb{N}}$ satisfies $0 \leq \Delta x^k(\bar{E}) \leq \Delta b_1^k$, we obtain $0 \leq \Delta x^*(\bar{E}) \leq \Delta b_1^*$. But this contradicts the maximality of λ because of Lemma 6.51. □

We conclude this subsection by presenting a complete sensitivity analysis for the anchored network of Example 6.23. As already mentioned, such an approach builds the basis for establishing uniqueness and existence of thin flows with resetting. It also suggests that increasing node balances leads only to a finite number of completely different thin flows without resetting. Here, completely different means that the corresponding critical and quasi-critical networks differ. In this sense the flow behavior can be described by a finite number of common critical networks.

Example 6.54. As already mentioned, this example provides a complete sensitivity analysis on the anchored network of the thin flow with resetting instance given in Example 6.23. The original instance and the corresponding anchored network are redrawn in Figure 6.21a and 6.21b. Initially we want to send eight flow units from s to t.

A thin flow without resetting which sends eight flow units from s to t is shown in Figure 6.22a. Note that corresponding labels are shown at all nodes except at s where the total flow is shown such that we can easily identify how many additional flow units are sent to v. Clearly decreasing the node balance of v leads to an increase of ℓ_v. Since the label of v_1 is strictly greater than ℓ_v, no additional flow is sent over $v_1 v$ in at first. Therefore, all additional flow traverses the edge wv and, hence, the label of w increases. Because ℓ_w is strictly smaller



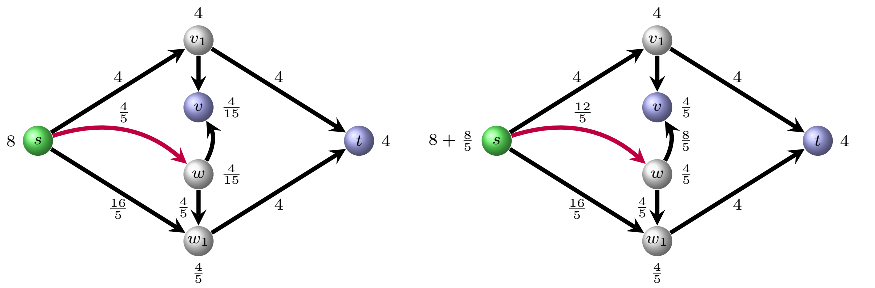


(a) The resetting network.

(b) The vw-anchored network with edge capacities and node balances.

Figure 6.21: The initial instances of the sensitivity analysis.

than ℓ_{w_1} flow which is additionally sent through w is supported by sw. Hence, in the first step additional flow traverses only the path swv as long as the label of v and w remains smaller than ℓ_{v_1} and ℓ_{w_1}, respectively. The resulting flow of is shown in Figure 6.22b where the label of w and w_1 coincide.



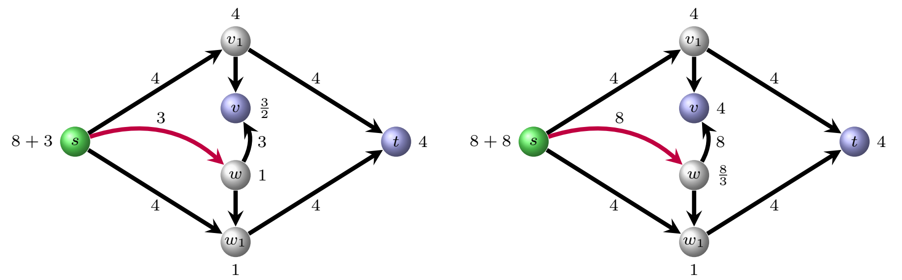
(a) Thin flow without resetting where no additional flow is sent to v.

(b) Thin flow without resetting where $\frac{8}{5}$ additional flow units are sent to v.

Figure 6.22: The initial thin flow without resetting and the first local increasing step.

In the first step we send $\frac{8}{5}$ additional flow units to v. As before, the further flow must traverse the edge wv. Since this increases the label of w and (initial) flow is sent along ww_1, the label of w_1 must also be increased. Hence, in the next step additional flow is sent along the edges sw and sw_1 to w. This is achieved by sending flow backwards over the edge w_1w implying that the flow on this edge decreases. Thus, in this manner we are able to sent flow to v as long as the flow on ww_1 remains nonnegative. Note that we also have to take into account that the label of w_1 remains smaller than the label of t. This leads to the flow drawn in Figure 6.23a where no flow is sent over the edge ww_1. So $\frac{7}{5}$ flow units are sent in this step.

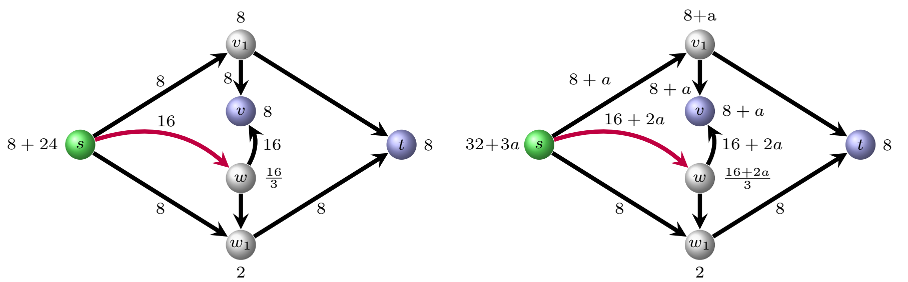
Considering the flow behavior of the next flow, we again observe that all flow has to use the edge wv which leads to an increase in the label of w. But unlike the previous case this causes no increase in the label of w_1 because the edge ww_1 carries no flow. In fact, it is a quasi-empty edge. Therefore, further flow is only send over the path swv as long as the label of v remains smaller than the label of v_1. This leads to the flow in Figure 6.23b where five additional flow units are sent. Also note that in this step b_v is decreased in such a manner that the flow on the anchored edge sw equals b_v. Thus, using the transformation in Lemma 6.28 results in a feasible thin flows with resetting on the original resetting instance.



(a) Thin flow without resetting where 3 additional flow units are sent to v.
(b) Thin flow without resetting where 8 additional flow units are sent to v.

Figure 6.23: The second and the third local increasing step.

For verifying how additional flow is sent in the fourth step we observe that the maximum label is also attained by ℓ_v. This implies that the next flow is sent over the entire network (except edge ww_1). In particular, flow on the edge v_1t is decreased as further flow has to use the corresponding backward edge. Hence, in this step we are able to send flow until the flow on v_1t vanishes. The resulting flow is depicted in Figure 6.24a showing that 16 additional units are sent.



(a) Thin flow without resetting where 24 additional flow units are sent to v.
(b) Thin flow without resetting where $24 + a$ additional flow units are sent to v.

Figure 6.24: The fourth local increasing step and behind.

For the next step we observe that additional flow traverses the paths swv and sv_1v. A difference flow Δx on the current sensitivity network, which consists of these two paths, is shown in Figure 6.24b by the multipliers of a. We also observe that there is no bound for scaling factor of this difference flow, i.e., for all $a \in \mathbb{R}_+$ we obtain a feasible thin flow with resetting for $b_v = -24-3a$. Hence, this concludes the sensitivity analysis for this scenario. Note that because of Lemma 6.51 thin flows without resetting for b_v-values, which are not shown in the figures, are obtainable by interpolating the thin flows of the corresponding surrounding b_v-values linearly.

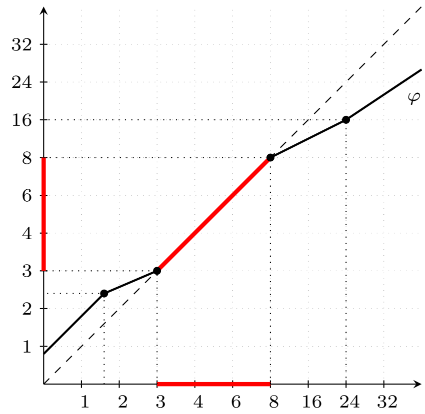


Figure 6.25: The function φ_e for the resetting edge in Example 6.23. The red color indicates the set of fixed points. Note that both axis are scaled such that slopes are preserved.

6.4 Existence and Uniqueness

In this section we establish the existence and uniqueness of thin flows with resetting. Whereas existence is obtained in general (if the tail of every resetting edge is reachable from s), we prove uniqueness for the case that every directed cycle contains no resetting edge. However, if we restrict the definition of the resetting labels to simple paths, thin flows with resetting are also unique in general.

For proving the existence result, we use the idea presented in Section 6.2 at page 208. In fact, we use induction over the number of resetting edges in order to prove that a thin flow with resetting exists for all node balances.

For this consider a resetting edge $e = vw \in E_1$ contained in some resetting network $\mathcal{N} := (G, u, s, E_1)$ and let $\mathcal{N}_e := (G', u', s, E_1')$ be the corresponding e-anchored network. Then Lemma 6.28 shows that, for given node balances b, we have to find Δb_v such that there exists a thin flow with resetting x' on $\mathcal{N}_e$ which sends Δb_v flow units over the anchored edge e_s of e, i.e., $x'_{e_s} = \Delta b_v$. Further, x' has to respect the node balances b' arising out of b where b_v is decreased and b_s is increased by Δb_v. Note that by the induction hypothesis there exists a thin flow on $\mathcal{N}_e$ respecting all possible node balance b'.

To put this approach in some functional analysis framework, we consider the set-valued function $\varphi_e : \mathbb{R}_+ \to 2^{\mathbb{R}_+}$ which maintains all possible flow values x'_{e_s} to a given Δb_v. That is, for every $a \in \varphi_e(\Delta b_v)$ there exists a thin flow with resetting x' on $(\mathcal{N}_e, b')$ with $x_{e_s} = a$ and vice versa. Hence, a fixed point $a \in \varphi_e(a)$ of φ_e exists if and only if there exists a thin flow x' on $(\mathcal{N}_e, b')$ with $x_{e_s} = a = \Delta b_v$. Thus, a fixed point of φ_e results in a thin flow with resetting on $(\mathcal{N}, b)$. Exemplarily, for the network consisting of one resetting edge e depicted in Figure 6.1 the function ϕ_e is drawn in Figure 6.25. Note that this function is obtained out of the sensitivity analysis for this scenario (see Example 6.54).

For proving the existence of a fixed point of φ_e we use the following one-dimensional version of the Kakutani Fixed Point Theorem.

Theorem 6.55 (Kakutani Fixed Point Theorem, [41]). *Let $I \subseteq \mathbb{R}_+$ be some closed interval of $\mathbb{R}_+$. Further, let $\phi : I \to 2^I$ be a set-valued function on I with a closed graph and the property that $\phi(a)$ is non-empty and convex for all $a \in I$. Then ϕ has a fixed point, i.e., there exist some $a \in I$ with $a \in \phi(a)$.*

For applying Kakutani's Fixed Point Theorem we show in the following:

- The graph of φ_e is closed.
- There exists an $M \in \mathbb{R}_+$ such that $\varphi_e(a) \cap [0, M] \neq \emptyset$.
- The image $\varphi_e(a)$ is convex for every $a \in \mathbb{R}_+$.

As we see below, this allows us to apply Kakutani's Fixed Point Theorem to φ_e restricted to $I := [0, M]$ which establishes the existence of thin flows with resetting. Note that for restricting φ_e to I we also have to restrict the image $\varphi_e(a)$ to $\varphi_e(a) \cap I$ for all $a \in I$. First, we prove the closeness property of φ_e.

Lemma 6.56. *Let $e = vw \in E_1$ be a resetting edge of some thin flow with resetting instance (G, u, s, E_1, b) and $S := \{(a^*, b^*) \mid a^* \in \varphi_e(b^*)\}$ be the graph of φ_e. Then the set S is closed.*

Proof. We show that each converging sequence of S attains its limit in S. So let $(a_i, b_i)_{i \in \mathbb{N}}$ be a sequence with $(a_i, b_i) \in S$ for all $i \in \mathbb{N}$ which converges in $\mathbb{R}^2_+$ to some (a^*, b^*). Further, let x^i be a thin flow with resetting on the e-anchored network with $x^i_{e_s} = a_i$ and $\Delta b_v = b_i$. Because of Lemma 6.31 we know that $(x^i)_{i \in \mathbb{N}}$ is bounded and, hence, there exists a converging subsequence. So assume without loss of generality that $(x^i)_{i \in \mathbb{N}}$ is converging and let x^* be the limit point, i.e., $x^* := \lim_{i \to \infty} x^i$. Now it follows from Lemma 6.32 that x^* is also a thin flow with resetting. Further, Lemma 6.32 shows that x^* respects the node balances arising out of b where b_v is decreased and b_s is increased by b^*. Since we also have $x^*_{e_s} = a^*$, this proof is finished. □

Next we show the second precondition which we need for applying Kakutani's Fixed Point Theorem. For this we need the assumption that the tail of every resetting edge is reachable from s in G.

Lemma 6.57. *Let $e = vw \in E_1$ be a resetting edge of some thin flow with resetting instance (G, u, s, E_1, b). Then there exists a positive real number $M \in \mathbb{R}_+$ such that $\varphi_e(a) \subseteq [0, M] \neq \emptyset$ holds for all $a \in [0, M]$.*

Proof. The basic observation which we use in this proof is Lemma 6.30. Since the tail of every resetting edge is reachable from s in G, there exists a original edge $e' \in \delta^+(s) \setminus \{e_s\}$ via which v is reachable from s. Now Lemma 6.30 ensures the existence of some real number $\lambda \in (0, 1)$ such that $x'_{e'} \geq -\lambda b'_v$ for every thin flow with resetting x' on $\mathcal{N}_e$ where b'_v is the node balance of v. Since the label of s is always equal to 0, no flow enters s implying

$$b_1 + \Delta b_v \;=\; b'_1 \;=\; x'\big(\delta^+(s)\big) \;\geq\; x'_{e_s} + x'_{e'} \;\geq\; x'_{e_s} - \lambda b'_v \;\geq\; x'_{e_s} + \lambda \Delta b_v$$

as $-b'_v \geq \Delta b_v$ holds. Hence, for $\Delta b_v \leq \frac{b_1}{\lambda} =: M$ we have

$$x'_{e_s} \;\leq\; b_1 + (1 - \lambda)\Delta b_v \;\leq\; b_1 + (1 - \lambda)M \;=\; M\ .$$

Note that the second inequality sign follows from $\lambda < 1$. Recalling the definition of φ_e this establishes the lemma. □

As the first two assumptions of Kakutani's Fixed Point Theorem are obtained quite directly, the third requirement needs more work. So far we consider thin flows with resetting where only one resetting edge is anchored. However, for obtaining the last missing detail we use Theorem 6.53 arising out of the sensitivity analysis for thin flows *without* resetting. Therefore, we consider thin flows *with* resetting as a fixed point in the space of thin flows *without* resetting as explained in Section 6.2 on page 208.

For formalizing this, let $\mathcal{N}_a := (G^{\mathrm{a}}, u^{\mathrm{a}}, s)$ be the anchored network of some resetting network $\mathcal{N} := (G, u, s, E_1)$. Then Corollary 6.29 shows that for given node balances b we have to find $\Delta b_{\mathrm{tail}(e)}$ for all $e \in E_1$ such that there exists a thin flow without resetting x^a on $\mathcal{N}_a$ which sends $\Delta b_{\mathrm{tail}(e)}$ flow units over the anchored edge e_s of e, i.e., $x^{\mathrm{a}}_{e_s} = \Delta b_{\mathrm{tail}(e)}$ for all $e \in E_1$. Further, x^{a} has to respect the node balances b^{a} arising out of b where $b_{\mathrm{tail}(e)}$ is iteratively decreased by $\Delta b_{\mathrm{tail}(e)}$ when considering all resetting edges $e \in E_1$. Note that for a particular node $v \in V$ the node balance b_v is decreased more than once in case v is the tail of multiple resetting edges. Further, b_s has to be increased by $\sum_{e \in E_1} \Delta b_{\mathrm{tail}(e)}$.

As before, we put this approach into a functional analysis framework. But note that thin flows without resetting are unique on critical edges. Since every outgoing edge of s must be critical (or carries zero flow) by definition, we do not need a set-valued function. Instead, we consider a ordinary function $\Phi : \mathbb{R}_+^{E_1} \to \mathbb{R}_+^{E_1}$ which assigns the flow values on the anchored edges to a given $\Delta b \in \mathbb{R}_+^{E_1}$. Thus, $a = \Phi(\Delta b)$ means that there exists a thin flow without resetting x^{a} on $(\mathcal{N}_a, b^{\mathrm{a}})$ with $x^{\mathrm{a}}_{e_s} = a_e$ for each resetting edge $e \in E_1$ and vice versa. Hence, a fixed point $a = \Phi(a)$ of Φ exists if and only if there exists a thin flow without resetting x^{a} on $(\mathcal{N}_a, b^{\mathrm{a}})$ with $x^{\mathrm{a}}_{e_s} = a_e = \Delta b_{\mathrm{tail}(e)}$ for all $e \in E_1$. In the following we use Φ in order to establish the last assumption which we need for applying Kakutani's Fixed Point Theorem. Beside this, subsequently, we use Φ also for proving the uniqueness result of thin flows without resetting. The next lemma gives a first impression how the fixed points of Φ are distributed.

Lemma 6.58. *Let $a \in \mathbb{R}_+^{E_1}$ be some real vector and $a^* \in \mathbb{R}_+^{E_1}$ be a fixed point of Φ. Then we know:*

(i) If $a \geq \Phi(a)$ then $a^*_{e^1} < a_{e^1}$ for all $e^1 \in E_1$ with $a_{e_1} < \Phi(a)$ and the componentwise minimum $a^{\min} := \min\{a, a^*\}$ is a fixed point of Φ.

(ii) If $a \leq \Phi(a)$ then $a^*_{e^1} > a_{e^1}$ for all $e^1 \in E_1$ with $a_{e_1} > \Phi(a)$ and the componentwise maximum $a^{\max} := \max\{a, a^*\}$ is a fixed point of Φ.

Proof. In order to prove (i), since $a^{\min} \leq a, a^*$, Theorem 6.53 shows

$$0 \leq \Phi(a) - \Phi(a^{\min}) \quad \text{and} \quad \|\Phi(a) - \Phi(a^{\min})\|_1 \leq \|a - a^{\min}\|_1 \tag{6.13}$$

$$\text{and} \quad 0 \leq \Phi(a^*) - \Phi(a^{\min}) \quad \text{and} \quad \|\Phi(a^*) - \Phi(a^{\min})\|_1 \leq \|a^* - a^{\min}\|_1 \ . \tag{6.14}$$

Now let E_1^1 be the set of resetting edges where the minimum in $a^{\min}$ is attained by a, i.e., $E_1^1 := \{e^1 \in E_1 \mid a^{\min}_{e^1} = a_{e^1}\}$. Then for all resetting edges not in E_1^1 the minimum in $a^{\min}$ is attained by a^*. Hence, we have because of $a \geq \Phi(a)$ and the first inequalities of (6.13) and (6.14)

$$\begin{aligned} &\Phi_{e^1}(a^{\min}) \;\leq\; \Phi_{e^1}(a) \;\leq\; a_{e^1} \;=\; a^{\min}_{e^1} && \forall e^1 \in E_1^1 \\ \text{and} \qquad &\Phi_{e^1}(a^{\min}) \;\leq\; \Phi_{e^1}(a^*) \;=\; a^*_{e^1} \;=\; a^{\min}_{e^1} && \forall e^1 \notin E_1^1 \ . \end{aligned}$$

On the other hand we obtain from (6.14) as a^* is a fixed point

$$\sum_{e^1 \in E_1} \Phi_{e^1}(a^{\min}) \;\geq\; \sum_{e^1 \in E_1} (\Phi_{e^1}(a^*) - a^*_{e^1} + a^{\min}_{e^1}) \;=\; \sum_{e^1 \in E_1^1} a^{\min}_{e^1} \;.$$

From this we are able to conclude $\Phi_{e^1}(a^{\min}) = a^{\min}_{e^1}$ for all $e^1 \in E_1$. Hence, $a^{\min}$ is a fixed point of Φ. Further, in case of $a_{e_1} < \Phi(a)$ we see that $\Phi_{e^1}(a^{\min}) = a^{\min}_{e^1}$ is only valid if $e^1 \notin E_1^1$ proving (i). The proof of (ii) follows the same lines of arguments. Therefore, the details are omitted. □

Lemma 6.58 has the following interesting consequence which shows that there are unique minimum and maximum fixed points.

Corollary 6.59. *Taking the componentwise minimum or maximum of two fixed points of Φ yields another fixed point.*

Proof. Let a^1 and a^2 be two fixed points of Φ. As $a^1 = \Phi(a^1)$ hold we get from Lemma 6.58(i) that $\min\{a^1, a^2\}$ is a fixed point of Φ and from Lemma 6.58(ii) that $\max\{a^1, a^2\}$ is also a fixed point of Φ. □

The next lemma concludes the preparatory work for applying Kakutani's Fixed Theorem. Note that we prove the existence of thin flows via induction over the number of resetting edges. Hence, we are able to assume that the e-anchored network $\mathcal{N}_e$ supports a thin flow with resetting for every node balance b'.

Lemma 6.60. *Let $e = vw \in E_1$ be a resetting edge of some thin flow with resetting instance (G, u, s, E_1, b). Then the set of fixed points of φ_e is convex, i.e., it is an interval.*

Proof. If φ_e supports only one fixed point we are done. Hence, let $a_e^1, a_e^2 \in \mathbb{R}_+$ be two fixed points of φ_e with $a_e^1 < a_e^2$ and $a_e \in (a_e^1, a_e^2)$ be a point in between. We have to show that a_e is again a fixed point of φ_e, i.e., $a_e \in \varphi_e(a_e)$.

Since a_e^1 and a_e^2 are fixed points of φ_e, there exist thin flows with resetting x^1 and x^2 on (G, u, s, E_1, b) sending a value of a_e^1 and a_e^2 over e, respectively. In addition, x^1 and x^2 imply two fixed points a^1 and a^2 of Φ where the value at the entry for e equals a_e^1 and a_e^2, respectively. Further, we are able to assume without loss of generality $a^1 \leq a^2$ because of Corollary 6.59.

Next consider the e-anchored network $\mathcal{N}_e$. Because of the induction hypothesis there exists a thin flow with resetting x' respecting node balances b' arising of b where b_v is decreased by $\Delta b_v = a_e$. Further, x' translates directly to a thin flow without resetting x^{a} on the anchored network $\mathcal{N}_{\mathrm{a}}$ of (G, u, s, E_1) which coincides with the anchored network of $\mathcal{N}_e$. On the one hand we know that x^{a} respects node balances b^{a} arising out of b where b_v is decreased by $\Delta b_v = a_e$ and for each resetting edge $e' \in E_1'$ different from e the node balance $b_{\mathrm{tail}(e')}$ is decreased iteratively by $\Delta b_{\mathrm{tail}(e')} = x'_{e'}$. On the other hand we know that the flow on e_s equals x'_{e_s} and for each resetting edge $e' \in E_1'$ the flow $x^{\mathrm{a}}_{e'_s}$ on each anchored edge is given by $x^{\mathrm{a}}_{e'_s} = x'_{e'}$. Hence, setting $a^{-e} := (x'_{e'})_{e' \in E_1 \setminus \{e\}}$ we know

$$\Phi\big((a^{-e}, a_e)\big) = (a^{-e}, x'_e)$$

by the definition of Φ. Therefore, Lemma 6.58(i) applied to (a^{-e}, x'_e) and a^2 shows $x'_e \geq a_e$. Similarly, Lemma 6.58(ii) applied to the vectors (a^{-e}, x'_e) and a^1

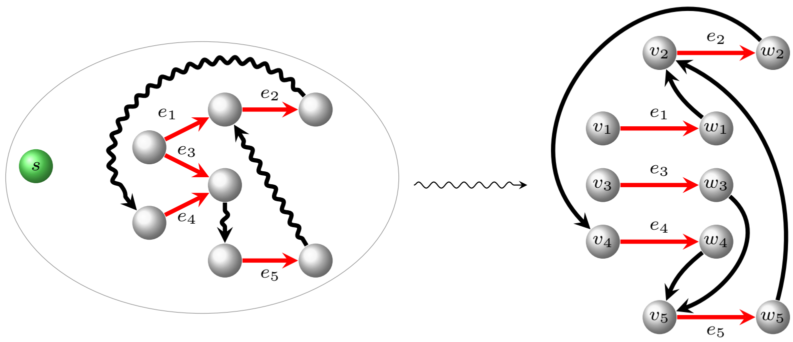


Figure 6.26: Construction of the graph H used in the proof of Theorem 6.62. Right we see the resetting instance together with all resetting edges of E_1^1. The sinuous lines correspond to paths in G. Left the graph H is shown where we draw only edges corresponding to paths in G traversing no resetting edge of E_1^1. Note that for this instance E_1^1 is actually a subset of $\{e_2, e_4, e_5\}$.

shows $x'_e \leq a_e$ implying $x'_e = a_e$. Hence, (a^{-e}, a_e) is a fixed point of Φ which in turn implies $a_e \in \varphi(a_e)$. □

Since with Lemma 6.60 all questions for applying the Kakutani's Fixed Point Theorem are answered, we are ready to establish the existence result for general thin flows with resetting.

Theorem 6.61. *Let (G, u, s, E_1) be a resetting network. Then for all node balances b a thin flow with resetting exists.*

Proof. Let $e \in E_1$ be some resetting edge and $I := [0, M]$ be an interval where M is taken out of Lemma 6.57. We show that Kakutani's Fixed Point Theorem is applicable to the function $\bar{\varphi}_e : I \to I$ defined by

$$\bar{\varphi}_e(a) := \varphi_e(a) \cap I \qquad\qquad \forall a \in I$$

Let S be the graph of φ_e. Since the graph of $\bar{\varphi}_e$ equals $S \cap I^2$, it is closed by Lemma 6.56. Further, the induction hypothesis and Lemma 6.57 show that $\bar{\varphi}_e(a)$ is nonempty for all $a \in I$. Finally, Lemma 6.60 shows that each set $\bar{\varphi}_e(a)$ is, in addition, convex. Hence, the assumptions of Kakutani's Fixed Point Theorem are satisfied and the existence of a fixed point of $\bar{\varphi}_e$ is established. As a fixed point of $\bar{\varphi}_e$ is trivially a fixed point of φ_e, this proof is done. □

After the existence is established for all thin flow with resetting instances, we address the question under which circumstances a thin flow with resetting is unique. Recalling Example 6.23 we see that thin flows with resetting are not unique, in general. This is mainly caused by the cycle which contains the resetting edge. Restricting to instances where no cycle contains a resetting edge, we are able to prove uniqueness of thin flows with resetting.

Theorem 6.62. *Let (G, u, s, E_1, b) be a thin flow with resetting instance. Then a thin flow with resetting respecting b is unique if every directed cycle contains no resetting edge.*

Proof. We prove this indirectly. So let $a^1 \in \mathbb{R}_+^{E_1}$ and $a^2 \in \mathbb{R}_+^{E_1}$ be two different fixed points of Φ. Because of Corollary 6.59 we are able to assume without loss of generality that $a^1 \lneq a^2$. Further, let $E_1^1 := \{e \in E_1 \mid a^1 < a^2\}$ be the set of indices where a^1 and a^2 differ.

Referring to Figure 6.26 we construct a directed graph H on the nodes which are connected via resetting edges in E_1^1, i.e., $V(H) := \bigcup_{e=vw \in E_1^1} \{v_e, w_e\}$. Note that a particular node $v \in V$ can have more than one copy in H in case it is the head or tail of multiple resetting edges. Further, for every resetting edge $e = vw \in E_1^1$ we add an edge $v_e w_e$ to H. Let $V_1 \subseteq V(H)$ be the set of nodes corresponding to the tail of a resetting edge and $V_2 \subseteq V(H)$ be the set of nodes corresponding to the head of a resetting edge, i.e., $V_1 := \bigcup_{e=vw \in E_1^1} \{v_e\}$ and $V_2 := \bigcup_{e=vw \in E_1^1} \{w_e\}$. Then every node in V_2 is reachable from a node in V_1. Finally, for $e, e' \in E_1^1$ we add an edge $w_e v_{e'}$ if and only if there exist a directed w-v-path in G. Now assume that every $v \in V_1$ gets an incoming edge in this manner. Hence, as in this case every node in H has at least one incoming edge, there must be a directed cycle C in H which could be obtained via backtracking. Further, it follows out of the construction of H that C contains at least one resetting edge. Thus, C corresponds to a directed cycle in G containing at least one resetting edge contradicting the assumption of this theorem.

Hence, it remains to prove that every node in V_1 admits an incoming edge in this manner. For this consider a resetting edge $e = vw \in E_1^1$ and increase the node balance of v by $a_e^2 - a_e^1$. That is, we consider the vector $a^e := (a^{1,-e}, a_e^2)$. As a^1 and a^2 are two fixed points of Φ, Theorem 6.53 shows

$$\Phi(a^1) \leq \Phi(a^e) \leq \Phi(a^e) \qquad \text{and} \qquad \|\Phi(a^e) - \Phi(a^1)\| = a_e^2 - a_e^1 \ .$$

Hence, the additional demand $a_e^2 - a_e^1$ of v is only resolvable on the anchored network of (G, u, s, E_1) by sending additional flow on the anchored edges of edges in E_1^1. Thus, by Corollary 6.52 at least one of these anchored edges must lie on an s-v-path. Let $e' = v'w' \in E_1^1$ be a corresponding resetting edge. Then this shows that there is a w'-v-path in G implying that v_e has at least one incoming edge in H. □

Theorem 6.62 shows on what kind of resetting networks thin flows are unique. However, as already mentioned, there is another aspect of thin flows which is responsible for nonuniqueness. This aspect goes back to the definition of thin flows and, in fact, to the definition of the resetting labels of a node. Here we do not assume that the resetting label is determined by the resetting congestion of a simple paths. If we change the definition in this manner, we obtain uniqueness in general. In fact, we consider the following definition of thin flows with resetting. Let x be an s-flow on a resetting network (G, u, s, E_1). Recalling Definition 6.9 of thin paths a path P starting at s is *a thin resetting path* if and only if the resetting congestion of every s-v-subpath equals the resetting label of v. In particular, this ensures that every thin resetting path is simple. Using this, x is called a *simple thin flow with resetting* if and only if every flow carrying s-v-path is a thin resetting path and the label of every node v is attained by some thin resetting path. The proof of the subsequent theorem uses the following variant of Brouwers fixed point theorem.

Theorem 6.63 (Brouwers Fixed Point Theorem, [12])**.** *Every continuous function from a compact subset of some Euclidean space to itself admits a fixed* *point.*

Theorem 6.64. *Let (G, u, s, E_1) be a resetting network. Then for all node balances a simple thin flow with resetting exists and is unique.*

Proof. Recalling Corollary 6.59 we show that a thin flow with resetting is simple if and only if it corresponds to the maximum fixed point of Φ. So consider a thin flow with resetting instance $\mathcal{I} := (G, u, s, E_1, b)$ and let a^* be the maximum fixed point of Φ with respect to $\mathcal{I}$.

First, we show that any thin flow x with resetting corresponding to a fixed point $a \lneq a^*$ is not simple. For this let $E_1^1 := \{e \in E_1 \mid a_e < a_e^*\}$ be the set of resetting edges e where the flow caused by x is strictly less a_e^*. Now let v be the tail of some resetting edge $e \in E_1^1$ and assume that there exists a thin resetting s-v-path P. In the following we derive a contradiction.

For this let $e' = v'w'$ be the first resetting edge of E_1^1 which is visited while traversing P from s to v and P' be the corresponding s-v'-path. Next consider the anchored flow x^{a} of x on the anchored network $\mathcal{N}_{\mathrm{a}}$ respecting node balances b^{a} resulting out of b and a in the usual manner. As P' is a thin resetting s-v'-path, there exists a thin s-v'-path P^{a} in $\mathcal{N}_{\mathrm{a}}$ using no anchored edge e_s with $e \in E_1^1$. In fact, P^{a} is obtainable as a subpath of P' starting at the last resetting edge. Let e^{a} be the first edge of P^{a}. Then e^{a} is either a usual edge contained in E or an anchored edge of a resetting edge not contained in E_1^1.

Now increasing the fixed point a slightly in the direction of $a^* - a$ we know that (6.10) of Lemma 6.51 is satisfied with equality where $\bar{E}$ consists of the anchored edges of E_1^1. On the other hand, Lemma 6.30 shows $\Delta x_{e^{\mathrm{a}}} > 0$ which is a contradiction. Hence, x is not a simple thin flow with resetting as there exists no thin resetting s-v-path.

It remains to show that a thin flow x^* with resetting, which is obtained from a^*, is simple. As before, we assume the opposite. So let the set

$$X := \{v \in V \mid \text{there exists no thin resetting } s\text{-}v\text{-path}\}$$

be nonempty. Then we know that $\delta^-(X)$ contains no resetting edge. On the other hand, if $G[X]$ would contain no resetting edge we have $\ell_v = \ell_w$ for each edge $e = vw \in \delta^-(X)$ because only resetting edges can decrease node labels along a thin resetting path. Since $\ell_v = \ell_w$ for at least one edge $e = vw \in \delta^-(X)$ contradicts the definition of X, we know that there exists at least one resetting edge in $G[X]$ and that $\ell_v > \ell_w$ holds for all edges $e = vw \in \delta^-(X)$.

Let $E_1^1 := E_1 \cap E\big(G[X]\big)$ be the set of resetting edges contained in $G[X]$ and $V^1 := \{v \in V \mid \delta^+(V) \cap E_1^1\}$ be the set of corresponding tail nodes. Next, consider the anchored network $\mathcal{N}_{\mathrm{a}} := (G^{\mathrm{a}}, u^{\mathrm{a}}, s)$ and update the node balances to b^{a} such that x^* corresponds to a feasible thin flow without resetting on $\mathcal{N}_a$. Now decrease the node balances of the nodes in V^1 slightly in any direction. Because of $\ell_v > \ell_w$ for all edges $e = vw \in \delta^-(X)$, we know that all of this additional flow is send via the anchored edges of E_1^1. Since the set

$$B := \Big\{(-\Delta b_v)_{v \in V^1} \in \mathbb{R}_+^{V^1} \;\Big|\; \sum_{v \in V^1} -\Delta b_v = 1\Big\}$$

is compact and the node labels of a thin flow without resetting change continuously with the node balances, Lemma 6.48 shows that there exists a $\delta > 0$ such that Lemma 6.51 is applicable whenever $\sum_{v \in V^1} \Delta b_v \leq \delta$. Further, we know that (6.10) is satisfied with equality for $\bar{E} = E_1^1$.

Next, consider the set

$$A := \Big\{ a \in \mathbb{R}_+^{E_1} \;\Big|\; \begin{aligned} a - a^* &\geq 0, \\ a_e &= a_e^* \quad \forall e \notin E_1^1, \\ \|a - a^*\|_1 &= \delta \end{aligned} \Big\} .$$

consisting of vectors a arising out of a^* by increasing the entries corresponding to resetting edges in E_1^1 in total by δ. Then the above discussion shows $\Phi(a) \in A$ for all $a \in A$. Thus, Brouwers fixed point theorem shows that Φ has a fixed point in A contradicting the maximality of a^*. $\square$

6.5 Your Comments

This chapter introduces a new class of static flows called thin flows and is mainly devoted to prove their existence and uniqueness on a given network. The class of thin flows contains a simple subclass which are called thin flows without resetting. A nice relation between arbitrary thin flows and thin flows without resetting is presented motivating a deeper analysis of thin flows without resetting. It turns out that these special thin flows are easy characterizable via polynomial algorithm as optimal flows. Further, a structure just as a sensitivity analysis is presented for thin flows without resetting. Based on these insight, existence and uniqueness is proven for arbitrary thin flows.

Taking a look at the literature you observe that thin flows with resetting are a novel class of static flows. On the other hand, studying thin flows with resetting is motivated by their relation to Nash flows over time for the deterministinc queuing model (see Chapter 7). Searching for thin flows *without* resetting, you find the work of Cole, Dodis, and Roughgarden [17]. Among other things, they analyze so-called static bottleneck routing game where the cost of path is defined as the maximum cost of its edges. For this they consider subpath-optimal flows which are static Nash flows of this routing game. It turns out that taking the congestion of an edge as its cost function, subpath-optimal flows are exactly thin flows without resetting. Other articles where thin flows without resetting are also implicitly considered via generalized models are, e.g., [8, 33]. Interestingly, in [33], Georgiadis et al. consider a generalized version of the Thin Flow algorithm presented in Subsection 6.3.1. Nevertheless, you does not find any contribution which studies exactly thin flows without resetting.

Concerning the fixed point approach presented in Section 6.4, the function Φ is a Lipschitz-continuous mapping with constant 1. Such mappings are called nonexpanding in literature. A fixed point theorem which is applicable to Φ is introduced by, e.g., Shioji and Takahashi in [81]. Also a convergent fixed point iteration is presented in [81]. Clearly, this implies the existence of thin flow with resetting, in general. However, their approach works for a quite general class of norms. Especially, the existence of a unique maximum fixed point is not proven which is essential for proving the uniqueness of simple thin flows with resetting. Restricting to acyclic instances, the existence and uniqueness of thin flows with resetting is established by Cominetti, Correa, and Larré [18] in 2011.

The main open question of this chapter ask for an algorithm which computes a thin flow with resetting in polynomial time. Reading this chapter, you extract two promising candidates for such an algorithm. The first candidate arises out of the fixed point approach in Section 6.4. Turning this approach into a fixed point iteration could solve this task. Unfortunately, such an iteration only approximates the fixed point in most cases. To overcome this problem, one could use the following approach. Based on the current solution, simply guess subsets of nodes having the same resetting label and also guess an order of these label. As in the approach described at the end of Subsection 6.3.2, use this information to construct a linear equality system similarly to (6.7). If the solution of this linear equality system satisfies the thin flow with resetting constraints, you are done.

The other candidate for such a polynomial algorithm is illustrated in Example 6.54. The idea is to provide a more or less complete sensitivity analysis of thin flows without resetting on the anchored network of a given resetting instance. As before, you could use the information from an linear inequality system like (6.7) to get a promising direction in which the node balances should be changed. Thinking about these two methods, you would not be surprised if both algorithm find a thin flow with resetting in polynomial time. However, you need more intuition for the behavior of thin flows. Therefore, you decide to make a structure and sensitivity analysis similarly to the ones in Subsection 6.3.2 and 6.3.3 for thin flows with resetting at first. Maybe, you end up with another algorithm similarly to the THIN FLOW algorithm presented in Subsection 6.3.1.

Chapter 7

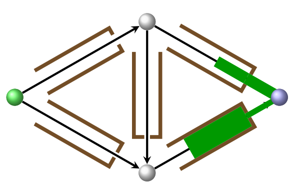
Nash Equilibria for the Deterministic Queuing Model

In this chapter we analyze routing games over time on the very popular deterministic queuing model. The deterministic queuing model works on a network where the flow behavior on each edge $e \in E$ is determined by a constant free flow transit time $\tau_e \in \mathbb{R}_+$ and a Lebesgue integrable capacity function $u_e : \mathbb{R}_+ \to \mathbb{R}_+$. As in the direct flow model, τ_e represents the time which a flow particle needs to traverse an edge e if e has capacity left. Further, $u_e(\theta)$ bounds the rate at which flow is able to leave e at time $\theta \in \mathbb{R}_+$. But in contrast to the direct flow model, waiting on edges is allowed which makes the model edge-based. In particular, a waiting queue builds up if more flow wants to traverse an edge e than e can handle. We assume that a waiting queue has no physical dimension and is placed directly in front of the head of the corresponding edge. Hence, if flow enters an edge e it takes τ_e time units to arrive at the end of the the waiting queue. In this manner the transit time experienced by a flow particle entering an edge e equals the sum of the free flow transit time τ_e plus the current waiting time.

In order to get familiar with the deterministic queuing model, we analyze a very specific scenario in Section 7.1. Here, we consider only networks where the transit time on each edge is zero. In addition, we assume that each capacity function is constantly equal to 1. We show how to compute atomic Nash equilibria on such instances. It turns out that each Nash equilibrium of a given instance is asymptotically representable by one maximal static s-t-flow. But however, as you will observe in Your Comments, all results of this section are generalizable to arbitrary constant capacity functions on so-called shortest path network, where the free flow transit times of all s-t-paths coincide.

In Section 7.2 we present the deterministic queuing model and observe that it is a consistent edge-based flow over time model. Subsequently, in Section 7.3 we characterize nonatomic Nash equilibria over time. It turns out that they are representable via a concatenation of *static* flows. In fact, these static flows are the thin flows analyzed in Chapter 6 which correspond to the underlying static flow defined in Section 4.2. This observation allows us to compute Nash equilibria especially for the case of constant transit times.

Finally, in Section 7.4 we compute the price of anarchy on a restricted class



of instances. We show that, on shortest path networks, a Nash flow over time is an earliest arrival flow. Unfortunately, we are not able to prove universally valid bounds on any kind of the prices of anarchy for arbitrary routing games over time. However, we show that the evacuation price of anarchy grows at least linearly in the number of edges in general. Surprisingly, for each other objective function, the price of anarchy on the corresponding class of instances is constant. This is even more interesting as we believe that the considered instances are, in fact, worst case instances for each price of anarchy.

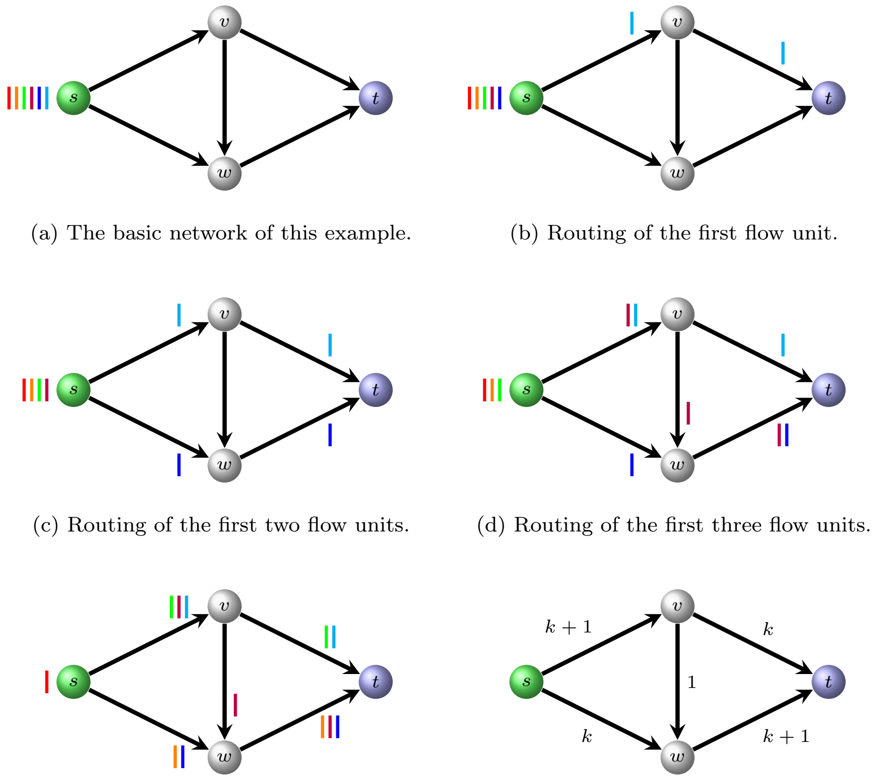
Throughout this section, the symbol ℓ always refers to a foresighted arrival time function defined in Section 3.4.

7.1 Atomic Case on Shortest Path Networks

In this section we analyze the atomic routing game over time on a restricted class of the discrete deterministic queuing model. Firstly, we describe the environment in which the players compete each other in order to send its flow unit through a given network. Subsequently, we state an algorithm which computes a possible Nash equilibrium arising out of the selfish decisions of the players. The algorithm does not terminate but provides a sequence of static flows encoding the selfish routing. This sequence is the object we study in more detail. In fact, we show that it converges to a maximum flow implying that a Nash equilibria in the presented scenario is optimal. Since players are identifiable via their controlled flow units, we do not distinguish between these two objects and only use the term flow units in the following.

As usual, the network where the routing occurs is encoded by a directed graph $G := (V, E)$ with node set V and edge set E. All flow units have a size of 1 and want to traverse the network from a given source $s \in V$ to a given sink t. Each edge $e \in E$ is determined by a capacity of 1 meaning that at over one time unit at most one flow unit is able to leave e. Hence, if too many flow units wants to traverse a certain edge e at the same point in time a waiting queue builds up at the end of e. In addition, a *free flow transit time* of 0 is assigned to each edge. That is, if no flow traverses an edge e at some point in time, an entering flow unit leaves e instantaneously. However, if there is waiting queue in front of the head of e, the flow unit has to line up first. In this manner the experienced transit time of a flow unit equals the current waiting time of the corresponding edge.

Since the flow units choose their paths through the network by themselves, we know that every flow unit chooses an s-t-path which is a currently shortest one regarding time (see Section 4.1). Of course, the choice of the flow units depends on their knowledge how the current flow situation evolves over time. We make the following assumptions. The flow units are located in a waiting queue in front of s and all together start traversing the network at time 0. In particular, all flow units are present at s at time 0 which intuitively leads to a demand of value "∞". Besides, every flow unit knows the routing of its predecessors in the initial waiting queue, i.e., the routing system has to obey past-orientation. Further, if two flow units arrive at one edge at the same point in time, the flow unit with the smaller position in the initial waiting queue traverses the edge first. And thirdly, every flow unit wants to arrive the sink t before their successors in the initial waiting queue which assures that the resulting routing



(a) The basic network of this example.

(b) Routing of the first flow unit.

(c) Routing of the first two flow units.

(d) Routing of the first three flow units.

(e) Routing of the first five flow units.

(f) Routing of the first $2k + 1$ flow units.

Figure 7.1: Routing behavior of Example 7.1. All transit times are 0 and all capacities are 1 and the flow behavior is given by the discrete deterministic queuing model.

is a Nash equilibrium. To ensure the latter assumption, one could think that every flow unit must only choose one path which leads as fast as possible to t taking into account the routing of the predecessors. But as already observed in Example 4.1, this is not enough in general. Since the routing in Example 4.1 is based on the discrete deterministic queuing model, we recall the basic insights for the subsequent approach.

Example 7.1. In order to follow this example, we redraw the routing behavior in Figure 7.1. Note that each edge has a free flow transit time of 0 and a capacity of 1. As in Example 4.1, we refer to the s-t-path svt as the upper path, to swt as the lower path, and to $svwt$ as the zigzag path.

The interesting aspect of this example results out of the routing of the third purple flow unit along the zigzag path which contains the edge vw (see Figure 7.1d). This causes that the current transit times of the all three paths coincide for the fourth green flow unit. Thus, we may conclude that the fourth green flow unit can take every path. But if the fourth green flow unit would take the zigzag path, it arrives at w at time 2, and therefore, the next orange flow unit would choose the lower path in order to be at w already at time 1. This means that the fifth orange flow unit would overtake the fourth green flow unit which cannot be tolerated by the fourth green flow unit. Therefore, flow unit 4 chooses either the upper or the lower path. If the fourth green flow unit takes the upper path, the fifth orange flow unit uses the lower path and vice versa (see Figure 7.1e).

As already observed in Example 4.1, the situation for the sixth red flow unit is the same as for the fourth green flow unit such that one of the next two flow units chooses the lower and the other the upper path in order to avoid being overtaken. This leads again to the same situation for the next two flow units, and from now on, the routing of the further flow units happens always in this manner. Thus, k of the first $2k + 1$ flow units take the upper path, one the zigzag, and the other k take the lower path (see Figure 7.1f).

We see that the flow units are routed such that they are at every node of the chosen path as early as possible. This ensures that in the resulting routing every flow unit arrives at the sink t as early as possible. Thus, we have a Nash equilibrium. On the other hand, simply choosing a currently quickest path (without being as fast as possible at every intermediate node) does not lead to an equilibrium as explained for the fourth green flow unit.

In addition to the observations in Example 4.1, we also see that the edge vw is only used by one flow unit, and it is not hard to see that the edge vw is always used at most once when preserving the nonovertaking condition. Considering the routing of the flow units until a certain position, we obtain a static flow where the flow on every edge is equal to the number of flow units traversing this edge (see Figure 7.1f for the static flow caused by the first $2k + 1$ flow units). Scaling down these static flows in order to obtain a feasible flow respecting the unit capacities, we get a sequence of static flows which converges to a maximal static flow on this network. In other words, having infinitesimal flow units, this implies that at every time the rate of flow arriving at t is equal to the maximal flow value of 2. Of course, a supervisor being able to choose the paths for the flow units would route half of the flow along the upper and half of the flow along the lower path which is optimal for this instance and leads again to an inflow rate of 2 into t. Thus, the Nash equilibrium and the system optimum in this example coincide if we consider infinitesimal flow units. As we see below in this section, this is always the case for routing scenarios on shortest paths networks.

In the last example the flow units are routed such that they are at every node of the chosen path as early as possible. The following generic algorithm computes the routing of flow units in an atomic Nash flow over time. For this the algorithm iteratively updates node labels $\ell_v \in \mathbb{R}_+$ and edge labels $\ell_e \in \mathbb{R}_+$ for all nodes $v \in V$ and all edges $e \in E$, respectively. A node label ℓ_v represents the earliest point in time at which the current flow unit is able to arrive at v. Similarly, an edge label ℓ_e denotes the earliest point in time at which the current flow unit is able to leave e. Further, in order to maintain always a static flow of value 1, we scale down the flow after each iteration.

Generic Iterative Algorithm

Input:	A network (G, s, t) with unit capacities and zero free flow transit times.
Output:	A sequence $(x^k)_{k\in\mathbb{N}}$of static s-t-flows $x^k \in \mathbb{R}^{E(G)}$ of value 1.

(1) Set $k := 0$, $x^k := 0$, and $\ell := 0$.

Figure 7.2: Example with zero transit times and unit capacities. The current labels are shown on the edges.

(2) Compute an s-t-path P such that, for all $v \in V(P)$, the s-v-subpath of P minimizes $\max_{e \in E(P')} \ell_e$ over all s-v-paths P' with a labeling algorithm, and let ℓ' be the computed labels.

(3) Set $\ell_e := \frac{k\ell'_e + 1}{k+1}$ for all $e \in E(P)$.

(4) Set $x^{k+1} = \frac{kx^k + \chi_P}{k+1}$, $k := k+1$ and go to (2).

An application of the Generic Iterative algorithm shows the behavior of the flow units of an atomic Nash flow. In fact, the static flow caused by the first k flow units in the initial waiting queue is given by kx^k. Therefore, we are able to identify the path chosen by the k-th flow unit as the unique flow carrying path of the static flow $kx^k - (k-1)x^{k-1}$. Note that in Step (3) we only update the labels on the edges and *not* the node labels. This is due to the fact that there could be several currently shortest s-t-path which traverse through a certain node v as early as possible. This means that the next flow unit could visit v at the same point in time as the current flow unit by choosing a different path. Increasing also the node labels in Step (3) would imply that such a flow behavior is no longer recognizable by the Generic Iterative algorithm. In order to get familiar with the Generic Iterative algorithm, we consider the following example.

Example 7.2. In Figure 7.2 we see a part of a network with zero transit times and unit capacities. On the edges the current labels are shown, i.e., the earliest points in time at which the current flow unit is able leave a certain edge. Thus, the current flow unit must wait one time unit before it can leave the first edge. Then it must wait again one time unit for arriving at v at time 2. At this point the flow unit must choose either the upper or the lower subpath and, of course, it would choose the upper subpath consisting of one edge in order to arrive at t at time 4. Choosing the lower v-t subpath would result in an arrival time of 5. Therefore, the current flow unit has an additional waiting time of 2 at the end of the edge vt.

Computing the labels for the next flow unit is done in two steps. Firstly, the labels on every edge of the chosen path must be increased by 1 because the next flow unit can only leave such an edge after the current flow unit has traversed this edge. Otherwise the capacity constraint would be violated. This is done in the current call of Step (3). For the second step, observe that this can cause an increase of the labels of further edges. In this example the increase of the label of the edge in $\delta^-(v)$ to 3 implies that the label of lower outgoing edge of v must also be increased to 3. Here, we assume that v has no further incoming edges. This is done in subsequent call of Step (2) where we could use a modified Dijkstra as the labeling algorithm.

In the last part of this section, we analyze the sequence $(x^k)_{k \in \mathbb{N}}$ of static flows computed during a run of the Generic Iterative algorithm. Having in mind the observations of Example 7.1, we show that this sequence becomes a maximal s-t-flow if i goes to ∞. This shows that an atomic Nash flow over time is an asymptotically optimal routing for the discussed scenarios. Further, we exploit the structure of such an optimal Nash flow. In fact, it is representable via a *static* thin flow without resetting. Note that such flows are scaled maximum s-t-flows by Lemma 6.34.

For this we introduce the following notation for the sequence $(x^k)_{k\in\mathbb{N}}$ of static flows computed by the GENERIC ITERATIVE algorithm. The corresponding labels are denoted by $(\ell^k)_{k\in\mathbb{N}}$ which coincide with the labels ℓ' computed during the call of Step (2) within the $(k+1)$-th iteration. That is, for all nodes $v \in V$ and edges $e \in E$, the values $k\ell_v^k$ and $k\ell_e^k$ represent the earliest point in time at which the $(k+1)$-th flow unit is able to arrive at v and to leave e, respectively. Note that the labels ℓ^k of a particular flow x^k do not coincide with the flow labels defined in Definition 6.9 in general. This is due to the fact that the label of an edge e can increase from one iteration to the next even if the current flow unit does not use e. The following Lemma shows some basic properties of the GENERIC ITERATIVE algorithm.

Lemma 7.3. *Let $(x^k)_{k\in\mathbb{N}}$ be the output of the* GENERIC ITERATIVE *algorithm and $(\ell^k)_{k\in\ell}$ be the sequence of the corresponding node labels. Then the following statements hold for all $k \in \mathbb{N}$, all nodes $v \in V$, and all edges $e \in E$:*

(i) The labels ℓ_v^k and ℓ_e^k are bounded by 1, i.e., we have $\ell_v^k \leq 1$ and $\ell_e^k \leq 1$.

(ii) The labels ℓ_v^k and ℓ_e^k are an upper bound on the flow labels of v and e with respect to x^k, respectively.

(iii) We have $\ell_v^k \leq \ell_t^k + \frac{1}{k}$ and $\ell_e^k \leq \ell_t^k + \frac{1}{k}$, respectively.

The following statements hold for all $k \in \mathbb{N}$ and all edges $e = vw \in E$:

(iv) We have $\max\{\ell_v^k, \ell_w^k\} \leq \ell_e^k \leq \max\{\ell_v^k, \ell_w^k\} + \frac{1}{k}$

(v) If $k\ell_e^k = (k-1)\ell_e^{k-1} + 1$ then $e \in E(P_k)$ or $\ell_e^{k-1} = \ell_v^{k-1}$.

Finally, it holds for all $k \in \mathbb{N}$:

(vi) We have $\ell^k(\delta^+(X)) := \sum_{e\in\delta^+(X)} \ell_e^k \geq 1$ for all $X \subset V \setminus \{t\}$ with $s \in X$.

Proof. Firstly, we show statement (i). Since the unscaled edge labels are increased by at most 1 during every call of Step (3), the time at which the next flow unit would arrive at a node v via some s-v-path is also increased by at most 1. Inductively, this shows that each unscaled label of $(k\ell^k)_{k\in\mathbb{N}}$ is bounded by k implying (i).

Because of Step (3) the unscaled label of an edge e is increased by 1 whenever a flow unit is sent through e. Hence, the unscaled flow on e is bounded by the unscaled label of e. Since we work with unit capacities, this shows (ii) because of Definition 6.9.

In order to prove statement (iii), observe that the current flow unit does not traverse an edge e and a node v if the current label of e and v is greater than the current label of t, respectively. Since each unscaled label is increased by at most 1, this shows that the maximum unscaled label is bounded by the unscaled label of t plus 1 implying (iii).

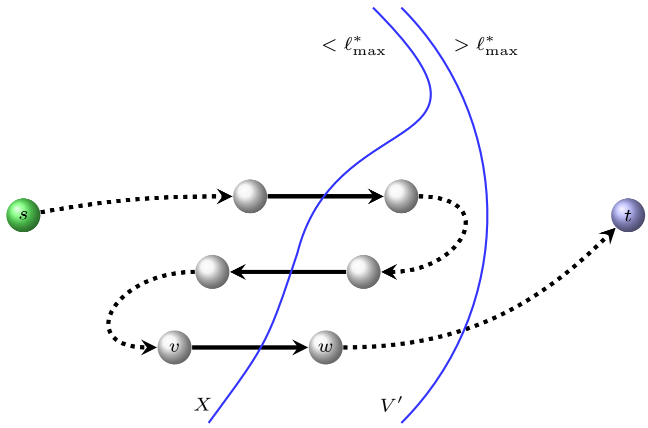


Figure 7.3: Scenario in the proof of Theorem 7.4.

For proving statements (iv) and (v), consider a particular iteration $k \in \mathbb{N}$ and an edge $e = vw \in E$. Because of the definition of the labels as earliest arrival times, we know $\max\{\ell_v^k, \ell_w^k\} \leq \ell_e^k$. Further, if a currently shortest s-v path uses edge e, we know $\ell_w^k = \ell_e^k$. So assume that for the next flow unit the edge e does not lie on a currently shortest s-v-path. If $\ell_v^k > \ell_w^k$ we must have $\ell_e^k = \ell_v^k$. On the other hand, if $\ell_v^k \leq \ell_w^k$ holds, we know $k\ell_e^k \leq k\ell_w^k + 1$. To see this, observe that under these circumstances the label of e is never increased if the latter inequality is satisfied with equality. This establishes (iv) inductively.

In order to prove (v), assume that $e \notin P_k$ and $\ell_e^{k-1} \neq \ell_v^{k-1}$. Hence, the label of e is not increased during Step (3). Further, we actually have $\ell_e^{k-1} > \ell_v^{k-1}$ in this case implying that the label of e is also not increased during the subsequent call of Step (2). This shows (v).

As already mentioned, the flow on an edge is bounded by the corresponding edge label. Hence, statement (vi) follows directly from the flow conservation constraint of x^k. $\square$

With this lemma we are able to prove that the selfish routing of the flow units is asymptotically optimal. Note that for zero free flow transit times an optimal routing is given by a maximal *static* s-t-flow. Routing flow in this manner implies that at each integral point in time the number of flow units arriving at t equals the maximum value of an static s-t-flow. This is of course optimal.

Scaling down a maximum s-t-flow to a value of 1 results in a maximum flow label of $\frac{1}{F}$ where F is the maximal flow value. This is a direct consequence of the Max-Flow-Min-Cut-Theorem which implies that on at least one cut each forward edge is used up to its capacity in every maximum s-t-flow. Since the Generic Iterative algorithm computes a sequence of static flows of value 1, we have to show that the maximum label of these flows converges to $\frac{1}{F}$. Even more, we show that the corresponding sequence of labels converges to the unique flow labels of a thin flow without resetting of value 1. Note that a thin flow without resetting represents a maximum s-t-flow by Lemma 6.34. For following the proof of the next theorem, it may be worth to take a look at Figure 7.3.

Theorem 7.4. *The sequence $(\ell^k)_{k\in\mathbb{N}}$ of the labels computed by the* Generic Iterative *algorithm satisfies*

$$\lim_{k\to\infty} \ell^k \;=\; \ell$$

where ℓ denotes the unique flow labels of a thin flow without resetting of value 1*.*

Proof. Consider the Generic Iterative algorithm. Because of Lemma 7.3(i) we know that each label is at most 1. Thus, ℓ^k is contained in the compact set $[0,1]^{V\cup E}$ for all $k \in \mathbb{N}$. This implies that $(\ell^k)_{k\in\mathbb{N}}$ has at least one accumulation point and this theorem states that all accumulation points are equal to ℓ.

In the following we establish this theorem indirectly. For this let $x \in \mathbb{R}_+^E$ be a thin flow without resetting of value 1 with corresponding labels ℓ and assume that there is an accumulation point of $(\ell^k)_{k\in\mathbb{N}}$ which is not a equal to ℓ. Among all such accumulation points consider one point ℓ^* with the following two properties:

(i) $\ell^*_{\max} := \max\{\ell^*_e \mid \ell^*_e > \ell_e\}$ is maximal over all accumulation points.

(ii) The node set $X := \{v \in V \mid \ell^*_v < \ell^*_{\max}\}$ is inclusionwise minimal over all accumulation points satisfying (i).

Before we proceed with the proof, note that Theorem 6.40 in combination with Lemma 7.3(ii) ensures the existence of $\ell^*_{\max}$ if ℓ^* does not coincide with the labels of a thin flow without resetting. Further, observe that the sink t is never contained in X. To see this, note that Lemma 7.3(iii) ensures that the maximum label of ℓ^* is not greater than the label ℓ^*_t of t.

In order to produce a contradiction, the main idea is to show that for every small $\epsilon > 0$ and every $K \in \mathbb{N}$ we have

$$\exists k \geq K, v^* \in V: \qquad \ell^k_v \;\geq\; \ell^*_{\max} - 4\epsilon \qquad \forall v \in (V \setminus X) \cup \{v^*\}\;. \tag{7.1}$$

Because the number of nodes in X is finite, we can conclude that there exists a node $v^* \in X$ such that for all $\epsilon > 0$ and all $K \in \mathbb{N}$ there exists a $k > K$ satisfying $\ell^k_v \geq \ell^*_{\max} - 4\epsilon$ for all $v \in (V \setminus X) \cup \{v^*\}$. Thus, an accumulation point $\bar{\ell}$ with $\bar{\ell}_v \geq \ell^*_{\max}$ for all $v \in (V \setminus X) \cup \{v^*\}$ exists. That means, if $\bar{\ell}$ does not contradict property (i), it must contradict (ii) implying that this theorem holds.

In order to prove (7.1), we need some inequalities which we discuss in following. The corresponding notation is adjusted to the subsequent use in this proof.

Consider the set $V' := \{v \in V(G) \mid \ell_v \leq \ell^*_{\max}\}$ of nodes having a thin flow label being at most $\ell^*_{\max}$ (see Figure 7.3). Because of (i) and Lemma 6.38, we know that, for every accumulation point, the flow outside $G[V']$ corresponds to a thin flow. Formalizing this, let $\bar{\ell}$ be an arbitrary accumulation point of $(\ell^k)_{k\in\mathbb{N}}$. Because of the maximality of $\ell^*_{\max}$ stated in (i), we know $\bar{\ell}_e \leq \ell^*_{\max}$ for all $e \in E(G[V'])$. Thus, for all $\epsilon_1 > 0$ there exists a $K \in \mathbb{N}$ such that

$$\ell^k_e \;\leq\; \ell^*_{\max} + \epsilon_1 \qquad\qquad \forall e \in E\big(G[V']\big), k \geq K\;. \tag{7.2}$$

Since $\bar{\ell}_v > \ell^*_{\max}$ holds for all $v \in V \setminus V'$ by the definition of V, we get $\bar{\ell}_e \geq \ell^*_{\max}$ for all $e \in \delta^+(V')$ from Lemma 7.3(iv) using a limit approach. This shows that for every $\epsilon > 0$ there exists another $K \in \mathbb{N}$ such that

$$\ell_e^k \;\geq\; \ell^*_{\max} - 3\epsilon \qquad\qquad \forall e \in \delta^+(X) \setminus \delta^+_{G[V']}(X), k \geq K \; . \tag{7.3}$$

Further, we know $\bar{\ell}_e = \ell_e$ for all $e \in E \setminus E(G[V'])$. In particular, this ensures that the sequence $(\ell_e^k)_{k\in\mathbb{N}}$ converges for all edges $e \in \delta^+(V')$. Since the set $\delta^+(X) \setminus \delta^+_{G[V']}(X)$ contains exactly those edges in $\delta^+(X)$ which also belong to $\delta^+(V')$, this shows that, for every $\epsilon_2 > 0$, there exists a $K \in \mathbb{N}$ such that

$$\ell^{k_2}\big(\delta^+(X) \setminus \delta^+_{G[V']}(X)\big) - \ell^{k_1}\big(\delta^+(X) \setminus \delta^+_{G[V']}(X)\big) \;\geq\; -\big|\delta^+_{G[V']}(X)\big| \cdot \epsilon_2 \tag{7.4}$$

holds for all $k_1, k_2 \geq K$.

Now let $q \in \mathbb{R}_+$ be the value of a sparsest cut Y on the b-flow instance induced by V' with respect to the thin flow without resetting x (see Definition 6.14). Hence, q as the largest node label occurring in a thin flow is smaller than $\ell^*_{\max}$. In fact, the maximality property of $\ell^*_{\max}$ stated in (i) shows $\ell^*_{\max} > q$ implying

$$\ell^*_{\max} = \frac{\ell^*(\delta^+_{G[V']}(X))}{|\delta^+(X)|} > \frac{b(Y)}{|\delta^+(Y)|} = q \; .$$

Since Y is a sparsest cut, we know $\frac{b(Y)}{|\delta^+(Y)|} \geq \frac{b(Y)}{|\delta^+(Y)|}$. Thus, the last inequality implies $\ell^*(\delta^+_{G[V']}(X)) > b(X)$. Further, we know $\ell^*_e = \ell_e = x_e$ for all $e \in \delta^+(V')$ because we deal with unit capacities and V' is a critical cut of the thin flow x. This shows

$$\begin{aligned}\ell^*\big(\delta^+(X)\big) \;&=\; \ell^*\big(\delta^+_{G[V']}(X)\big) + \ell^*\big(\delta^+(X) \setminus \delta^+_{G[V']}(X)\big)\\ &>\; b(X) + x\big(\delta^+(X) \setminus \delta^+_{G[V']}(X)\big) \; .\end{aligned}$$

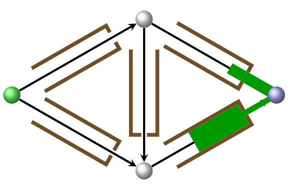
Because no flow of x travels backwards over the critical cut V', Definition 6.14 directly shows

$$\ell^*\big(\delta^+(X)\big) > 1 \; . \tag{7.5}$$

Using (7.5) we show that for every $K \in \mathbb{N}$ there exists an iteration $k_1 \geq K$ at which the edge label of at least two edges in the cut $\delta^+(X)$ is increased by 1. Assume the opposite. Since every s-t-path computed by the Generic Iterative algorithm has to cross $\delta^+(X)$ at least once, we know that there exists an iteration $K_1 \in \mathbb{N}$ after which the label of exactly one edge $\delta^+(X)$ is increased during each subsequent iterations. This shows for all $k \geq K_1$

$$\begin{aligned} k \cdot \ell^k\big(\delta^+(X)\big) \;&=\; K_1 \cdot \ell^{K_1}\big(\delta^+(X)\big) + k - K_1\\ \text{implying} \qquad \ell^*\big(\delta^+(X)\big) \;&=\; \lim_{k\to\infty} \ell^k\big(\delta^+(X)\big) \;=\; 1 \; .\end{aligned}$$

But this is a contradiction to (7.5). Hence, for every $K \in \mathbb{N}$ there exists an iteration $k_1 \geq K$ at which the edge label of at least two edges in the cut $\delta^+(X)$ is increased by 1.



The last observations for proving (7.1) directly follows from the fact that ℓ^* is an accumulation point of the sequence $(\ell^k)_{k\in\mathbb{N}}$. By the definition of X and V' we know $\ell^*_v < \ell^*_{\max}$ for all $v \in X$ and $\ell^*_v = \ell^*_{\max}$ for all $v \in V' \setminus X$. Thus, using a limit approach we obtain $\ell^*_e = \ell^*_{\max}$ for all $e \in \delta^+_{G[V']}(X)$ because of Lemma 7.3(iv). This shows

$$\ell^*_{\max} = \frac{\ell^*(\delta^+_{G[V']}(X))}{|\delta^+_{G[V']}(X)|} \; . \tag{7.6}$$

Finally, since x^* is an accumulation point, we know that, for every $\epsilon_2 > 0$ and every $k_1 \in \mathbb{N}$, there exists a $k_2 > k_1$ such that:

$$\ell^{k_2}\big(\delta^+_{G[V']}(X)\big) \;\geq\; \ell^*\big(\delta^+_{G[V']}(X)\big) - \big|\delta^+_{G[V']}(X)\big| \cdot \epsilon_2 \; . \tag{7.7}$$

In the following we show that (7.1) holds for all $\epsilon > 0$ and all sufficiently large $K \in \mathbb{N}$. More precisely, we choose a $K \geq \frac{1}{\epsilon}$ large enough in order to ensure the validness of (7.2) and (7.4) for

$$\epsilon_1 \;:=\; \frac{\epsilon}{|\delta^+_{G[V']}(X)| - 1} \qquad \text{and} \qquad \epsilon_2 \;:=\; \frac{\epsilon}{|\delta^+_{G[V']}(X)|} \; , \tag{7.8}$$

respectively. These assumptions are needed for applying Lemma 2.1 below. Moreover, we assume that K is large enough for the validity of (7.3).

So fix an $\epsilon > 0$ and such a K, and let $k_1 > K$ be an iteration where the computed path causes the increase of the label of at least two edges in $\delta^+(X)$. Further, let $k_2 > k_1$ be an iteration such that (7.7) holds. Finally, let $k-1$ be the last iteration smaller than k_2 where the unscaled labels of at least two edges in $\delta^+(X)$ were increased because of the chosen s-t-path P. This shows that the unscaled labels $k\ell^k$ and $k_2\ell^{k_2}$ differ on $\delta^+(X)$ in total by a value of $k_2 - k$ which leads to

$$\ell^{k_2}\big(\delta^+(X)\big) \;=\; \frac{1}{k_2}\Big(k\ell^k\big(\delta^+(X)\big) + k_2 - k\Big) \;=\; \frac{k}{k_2} \cdot \ell^k\big(\delta^+(X)\big) + \Big(1 - \frac{k}{k_2}\Big) \cdot 1 \; .$$

Since $\ell^{k_2}(\delta^+(X)) \geq 1$ holds by Lemma 7.3(vi) and the left hand side is a convex combination of $\ell^k(\delta^+(X))$ and 1, we obtain

$$\ell^k\big(\delta^+(X)\big) \;\geq\; \ell^{k_2}\big(\delta^+(X)\big)$$

which shows

$$\begin{aligned}\ell^k\big(\delta^+_{G[V']}(X)\big) \;\geq\; & \ell^{k_2}\big(\delta^+_{G[V']}(X)\big) + \ell^{k_2}\big(\delta^+(X) \setminus \delta^+_{G[V']}(X)\big) \\ & - \ell^k\;\big(\delta^+(X) \setminus \delta^+_{G[V']}(X)\big) \; .\end{aligned}$$

Applying (7.4) and (7.7) to the corresponding terms on the right hand side of this inequality leads to

$$\ell^k\big(\delta^+_{G[V']}(X)\big) \;\geq\; \ell^*\big(\delta^+_{G[V']}(X)\big) - 2\big|\delta^+_{G[V']}(X)\big| \cdot \epsilon_2 \; .$$

Further, dividing both sides by $|\delta^+_{G[V']}(X)|$ and using (7.6) results in

$$\frac{\ell^k(\delta^+_{G[V']}(X))}{|\delta^+_{G[V']}(X)|} \;\geq\; \ell^*_{\max} - 2\epsilon_2 \; .$$

Hence, because of (7.2) and (7.8), Lemma 2.1 is applicable and, in combination with (7.3), this shows

$$\ell_e^k \;\geq\; \ell^*_{\max} - 3\epsilon \qquad\qquad \forall e \in \delta^+(X)\ . \tag{7.9}$$

Because of the definition of k, we know $k\ell_e^k = (k-1)\ell_e^{k-1}+1$ holds for at least two edges $e \in \delta^+(V)$. Of course, this holds for the edge e_1 along which P traverses the cut $\delta^+(X)$ for the first time. However, let $e = vw \in \delta^+(X)$ be another edge for which this holds, i.e., $e \neq e_1$ and $k\ell_e^k = (k-1)\ell_e^{k-1} + 1$. Then the label of e is increased either during the call of Step (3) in iteration $k-1$ or during the call of Step (2) in iteration k. In the first case, which is depicted in Figure 7.3, we know $\ell_v^{k-1} \geq \ell_{e_1}^{k-1}$ because the edge labels do not decrease when traversing P. In the second case, we know $\ell_v^{k-1} = \ell_e^{k-1}$ because of Lemma 7.3(v). Hence, using (7.9) we obtain in both cases

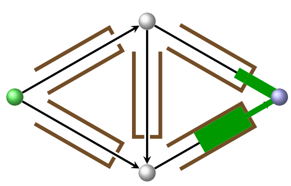
$$\ell^{k-1}(v) \;\geq\; \frac{k(\ell^*_{\max} - 3\epsilon) - 1}{k-1} \;\geq\; \ell^*_{\max} - 4\epsilon\ .$$

because k and K are chosen such that $k - 1 \geq K \geq \frac{1}{\epsilon}$ holds. □

"Oh!", you think, "this section ends already here." So you find it a good place for your observation that the entire approach is generalizable to arbitrary constant capacities. Arbitrary integral capacities work quite well as they bound the number of flow units which are able to leave an edge over one time unit. In this case you assume that several flow units traverse an edge side by side. Hence, replacing a single edge by several parallel edges of unit capacities directly yields the desired result. Further, you interpret arbitrary real capacities as follows. If a flow unit leaves an edge e at time θ, the next flow unit is able to enter the head of e not before time $\theta + \frac{1}{u_e}$ where $u_e \in \mathbb{R}_+$ denotes the capacity of e. In this sense, the capacity of an edge is rather used to determine current transit times instead as a bound on the number of leaving flow units. Generalizing the approach to this setting also requires a more general version of Lemma 2.1. Surprisingly, this generalization is already part of Remark 2.2. You wonder, why the author knew that you will need this generalization. However, with respect to zero transit times, you further observe that this approach is also generalizable to so-called shortest path networks, where each s-t-path has the same free flow transit time. As you note, the only property implied by zero free flow transit times which is used in this approach is that a flow unit always arrives at the same point in time at some node v if no waiting queue occurs. But this condition holds also for shortest path networks. Now you proceed with reading the next section.

7.2 Deterministic Queuing Model

As already mentioned in the introduction to this chapter, the deterministic queuing model is an edge-based flow over time model where current edge transit times consists of a constant free flow transit time and a waiting time. The waiting time is part of the model as the rate at which flow is able to leave a certain edge is bounded. Hence, if too much flow wants to traverse an edge, a waiting queue is built up. Further, we assume that the waiting queue has



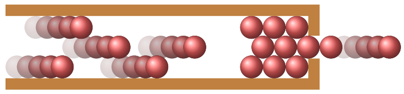


Figure 7.4: If more flow particles want to leave an edge than its capacity allows, they form a waiting queue.

no physical dimension. This implies that the time needed for arriving at the end of the waiting queue is independent on the size of the waiting queue. This constant time is given by the free flow transit time. Note the similarities to the direct flow model presented in Chapter 5. For observing this, assume that in some direct flow over time there is an s-t-path P consisting of one s-t-edge which is not used by flow particles traversing any other path. Then the direct flow behavior on P is described in the same manner as the flow behavior on an edge in the deterministic queuing model.

In this section we give a formal description of the deterministic queuing model and show its consistency. It turns out that the deterministic queuing model is equivalent to the classical flow over time model with time-varying capacity functions. In fact, the deterministic queuing model arises out of the classical flow over time model where waiting at nodes is modeled via waiting on the edges.

The deterministic queuing model works on a network $(G,\mathcal{U},\mathcal{T},s,t)$ consisting of a directed graph $G := (V,E)$, a family $\mathcal{U} := (u_e)_{e\in E}$ of capacity functions $u_e : \mathbb{R}_+ \to \mathbb{R}_+$, a family $\mathcal{T} := (\tau_e)_{e\in E}$ of *free flow transit times* $\tau_e \in \mathbb{R}_+$, a source $s \in V$, and a sink $t \in V$. The capacity $u_e : \mathbb{R}_+ \to \mathbb{R}_+$ of an edge e is a Lebesgue integrable function, and the value $u_e(\theta)$ bounds the rate at which flow is able to leave e at a certain point in time θ. Hence, a waiting queue builds up in front of the head of an edge if, at some point in time, more flow particles want to leave an edge than the capacity allows. For computing the waiting time in front of the head of an edge, we assume that no flow particle waits longer than is has to. That is, a waiting flow particle immediately uses free capacity to leave the corresponding edge instantaneously. The free flow transit time of an edge determines the time for traversing an edge if the waiting queue is empty. Since we assume that a waiting queue has no physical dimension, the free flow transit time also determines the time a flow particle has to take when traveling from the tail of an edge to the end of the waiting queue of that edge. Thus, the current (flow-dependent) transit time on an edge is the sum of the free flow transit time and the current waiting time. We think of the edges as corridors with large entries and small exits, which are wide enough for storing all waiting flow particles (see Figure 7.4).

Since the deterministic queuing model is an edge-based flow over time model, we consider an edge-based flow over time $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-)$. Recalling Definition 3.9, the family $\mathcal{F}^+ := (F_e^+)_{e\in E}$ represents the inflow functions. Further, we assume for the moment that $\mathcal{F}^- := (F_e^-|_{\leq\theta})_{e\in E,\theta\in\mathbb{R}_+}$ denotes the entire family of restricted outflow functions. Since it turns out that the deterministic queuing flow model satisfies FiFo, the flow behavior is actually completely described by the family of nonrestricted outflow functions. Along the same lines as for direct flows over time, we discuss the aspects which makes $\mathcal{F}$ feasible with respect to the deterministic queuing model in the following.

Since the capacity bounds the rate at which flow is able to leave an edge, the outflow $\mathcal{F}^-$ has to obey the following capacity constraints:

$$f_e^-|_{\leq\theta} \;\leq\; u_e \qquad\qquad \forall e \in E,\ \theta \in \mathbb{R}_+ \ . \tag{7.10}$$

The remaining capacity $u_e|_{\leq\theta} : \mathbb{R}_+ \to \mathbb{R}_+$ shows how much capacity is left for flow particles entering e after time θ, i.e.,

$$u_e|_{\leq\theta} \;:=\; u_e - f_e|_{\leq\theta} \qquad\qquad \forall e \in E,\ \theta \in \mathbb{R}_+ \ . \tag{7.11}$$

In order to formalize current waiting times, consider an edge e and a point in time $\theta \in \mathbb{R}_+$. Clearly, if the restricted capacity $u_e|_{\leq\theta}$ is positive at some point in time θ', flow entering e after time θ is able to leave e at time θ'. Further, flow shall leave an edge e as soon as there is some capacity left. Therefore, we define the current waiting time by

$$q_e(\theta) = \max\left\{\Delta \geq 0 \;\middle|\; \int_{\theta+\tau_e}^{\theta+\tau_e+\Delta} u_e|_{\leq\theta}(\vartheta)\; d\vartheta \leq 0\right\} \ . \tag{7.12}$$

Since restricted outflow functions obey capacity constraints, the remaining capacity is always nonnegative. Therefore, we are allowed to replace the "$\leq$"-sign by an "="-sign. Hence, over the interval $[\theta, \theta + q(\theta)]$, no flow entering e after time θ is able to leave e because $u_P|_{\leq\theta}$ is essentially 0 on this interval. Further, from time $\theta + q(\theta)$, flow entering e directly after time θ is able to leave P as

$$\int_{\theta+\tau_e+q_e(\theta)}^{\theta+\tau_e+q_e(\theta)+\epsilon} u_e|_{\leq\theta}(\vartheta)\; d\vartheta \;>\; 0$$

holds for all $\epsilon > 0$. Similarly to the direct flow model, we have to pay attention if the capacity function of an edge e is strictly positive. In this case setting current waiting times always to 0 results in a feasible q_e implying that the outflow function equals the inflow function. Therefore, the capacity constraint is violated by the outflow function if the inflow is larger than the capacity. This shows that the capacity condition (7.10) must be an explicit constraint of the deterministic queuing model.

Summarizing the above discussion, the routing on an edge e occurs as follows. After a flow particle enters e at some point in time θ, it directly traverses e to arrive at the head of e or at the end of waiting queuing of e after τ_e time units. If there is a waiting queue, it has to wait for $q_e(\theta)$ time units before it arrives at the head of e. Thus, it leaves e at time

$$\ell_e(\theta) := \theta + q_e(\theta) + \tau_e \ . \tag{7.13}$$

This motivates the following definition of a feasible flow over time for the deterministic queuing model.

Definition 7.5 (Feasible Flows over Time)**.** Let $\mathcal{F} := (\mathcal{F}^+, \mathcal{F}^-)$ be an edge based flow over time where $\mathcal{F}^+ := (F_e^+)_{e\in E}$ denotes the family of inflow functions and $\mathcal{F}^- := (F_e^-|_{\leq\theta})_{e\in E,\theta\in\mathbb{R}_+}$ the entire family of restricted outflow functions. Then $\mathcal{F}$ is a feasible flow over time respecting the deterministic queuing model if and only if ℓ_e satisfies (3.9), i.e,

$$F_P^-|_{\leq\theta}(\theta') = \int_{\ell^{-1}([0,\theta'])} f_P^+|_{[0,\theta]}(\vartheta)\; d\vartheta \tag{7.14}$$

for all times $\theta \in \mathbb{R}_+$ and all paths $P \in \mathcal{P}$. Moreover, each $F_P|_{\leq\theta}$ hat to satisfy the capacity constraint (7.10).

In the following, we show that the deterministic queuing model is consistent. Firstly, we prove that it satisfies FiFo and is past-oriented. This implies that the arrival times defined by (7.13) are foresighted. Based on this, we characterize the deterministic queuing model using only outflows. This allows us to establish $\mathcal{F}$-continuity. Clearly, most of this follows directly from the direct flow model as, from an abstract point of view, the flow behavior on an edge is a special case of the direct flow behavior on a path. Nevertheless, the corresponding proofs for the deterministic queuing model are simpler.

Lemma 7.6. *The deterministic queuing model satisfies the FiFo-principle and is past-oriented. Furhter, foresighted arrival times are given by* (7.13)*.*

Proof. To see that the deterministic queuing model is past-oriented, observe that the definition of $\ell_e(\theta)$ only depends on the restricted flow $F_e^+|_{\leq\theta}$ for all edges e and all points in time $\theta \in \mathbb{R}_+$. This shows that the restricted outflow $F_e^-|_{\leq\theta}$ is completely defined by $F_e^+|_{\leq\theta}$ implying that past-orientation holds.

For proving the FiFo principle, let e be an edge and θ_1, $\theta_2 \in \mathbb{R}_+$ be two points in time with $F_e^+(\theta_1) < F_e^+(\theta_2)$. Recalling Definition 3.16, we have to show $\ell_e(\theta_1) < \ell_e(\theta_2)$. In case of $\theta_1 + q_e(\theta_1) < \theta_2$, we obtain $\ell_e(\theta_1) < \ell_e(\theta_2)$ as q_e is nonnegative by definition. Therefore, we assume $\theta_2 \leq \theta_1 + q_e(\theta_1)$. As F_e^+ is nondecreasing, we know $\theta_1 < \theta_2$ implying $F_e^-|_{\leq\theta_1} \leq F_e^-|_{\leq\theta'} \leq F_e^-|_{\leq\theta_2}$ for all $\theta' \in (\theta_1, \theta_2)$. Because of (7.11), this shows $u_e|_{\leq\theta_1} \geq u_e|_{\leq\theta'} \geq u_e|_{\leq\theta_2}$ for all $\theta' \in (\theta_1, \theta_2)$. Since each remaining capacity function is nonnegative by (7.10), this shows for all $\theta' \in (\theta_1, \theta_2)$

$$0 \;\leq\; \int_{\theta'+\tau_e}^{\theta'+\tau_e+(\theta_1-\theta'+q_e(\theta_1))} u_e|_{\leq\theta'}(\vartheta)\, d\vartheta \;\leq\; \int_{\theta_1+\tau_e}^{\theta_1+\tau_e+q_e(\theta_1)} u_e|_{\leq\theta_1}(\vartheta)\, d\vartheta \;=\; 0\;.$$

Because of the maximum in (5.6), this leads to $\theta_1 - \theta' + q_e(\theta_1) \leq q_e(\theta')$ implying $\theta_1 + q_e(\theta_1) \leq \theta' + q_e(\theta')$ for all $\theta' \in (\theta_1, \theta_2)$. Along the same lines, we obtain $\theta' + q_e(\theta') \leq \theta_2 + q_e(\theta_2)$ for all $\theta' \in (\theta_1, \theta_2)$. In particular, this shows $\ell_e(\theta_1) \leq \ell_e(\theta') \leq \ell_e(\theta_2)$ for all $\theta' \in (\theta_1, \theta_2)$. From this we are able to deduce $(\theta_1, \theta_2) \subseteq \ell_e^{-1}([0, \ell(\theta_1))) \setminus \ell_e^{-1}([0, \ell(\theta_2)])$. Thus, by the the definition of restricted outflow functions in (3.9), we obtain

$$F_e^-|_{\leq\theta_2}(\theta_2) - F_e^-|_{\leq\theta_2}(\theta_1) \;\geq\; \int_{\theta_1}^{\theta_2} f^+(\vartheta)\, d\vartheta \;>\; 0\;.$$

As $F_e^-|_{\leq\theta_2}$ is nondecreasing, this shows $\ell(\theta_1) < \ell(\theta_2)$. Hence, the deterministic queuing model satisfies FiFo.

To see that the foresighted arrival time function of the deterministic queuing model is defined by (7.13), simply note that, because of the maximum in (7.12), additional flow entering an edge e at time $\theta \in \mathbb{R}_+$ would leave e from time $\theta + \tau_e + q_e(\theta)$ on. $\square$

In the following we give an alternative characterization of the flow behavior on edge. This characterization defines the outflow of an edge directly without using arrival time functions. Note that this completely defines the deterministic queuing model because of the FiFo principle.

Lemma 7.7. *Let F^+ be an inflow function of a certain edge e. Then F^- is the outflow of e with respect to the deterministic queuing model if and only if F^- is maximal among all functions satisfying*

$$F^- + \tau_e \;\leq\; F^+ \qquad\qquad \text{and} \qquad\qquad f^- < u_e \;. \tag{7.15}$$

Proof. First note the zero function obviously satisfies (7.15). Further, if for two functions (7.15) holds, also their componentwise maximum obeys (7.15). This shows that F^- is uniquely defined as the maximum element satisfying (7.15).

In order to prove that F^- coincide with the outflow function on e with respect to the deterministic queuing model, we have to show $F^+(\theta) = F^-(\ell(\theta))$ for all $\theta \in \mathbb{R}_+$. Fix some $\theta \in \mathbb{R}_+$. Since F^- satisfies the capacity constraint (7.10), we know for each $\theta' \in \mathbb{R}_+$ with $F_P^-(\theta' + \tau_e) > F_P^+(\theta)$

$$\begin{aligned}\int_{\theta+\tau_e}^{\theta'+\tau_e} u_e|_{\leq\theta}(\vartheta)\, d\vartheta \;&=\; \int_{\theta+\tau_e}^{\theta'+\tau_e} \big(u_e - f_e^-|_{\leq\theta}\big)(\vartheta)\, d\vartheta \\ &\geq\; F_e^-(\theta'+\tau_e) - F_e^+(\theta) \;>\; 0\end{aligned}$$

implying $\theta + q_e(\theta) < \theta'$ by the definition of q_e in (7.12). Because F^- is nondecreasing, this shows $F^+(\theta) \geq F^-(\ell_e(\theta))$.

Thus, it remains to prove $F^+(\theta) \leq F^-(\ell_e(\theta))$. For this let $\check{\ell}_e$ be the lower arrival time function of e with respect to F^+ and F^-. We show $\check{\ell}_e(\theta) \leq \ell_e(\theta)$. If $\check{\ell}_e(\theta) \leq \theta + \tau_e$, nothing has to be proven as $\ell_e(\theta) \geq \theta + \tau_e$ holds by definition. So let $\check{\ell}_e(\theta) > \theta + \tau_e$. Because of the minimum in the definition of the lower arrival time function, this implies

$$F^-(\theta') \;<\; F^-\big(\check{\ell}_e(\theta)\big) \;=\; F^+(\theta) \;\leq\; F^+(\theta' - \tau_e) \qquad \forall \theta' \in \big[\theta + \tau_e, \check{\ell}(\theta)\big) \;.$$

Thus, because of the maximality of F^-, the edge e is used up to its capacity over $[\theta + \tau_e, \check{\ell}(\theta))$ implying

$$\int_{\theta+\tau_e}^{\check{\ell}_e(\theta)} u_e|_{\leq\theta}(\vartheta)\, d\vartheta \;=\; 0 \;.$$

By the definition of q_e in (7.12), this shows $\check{\ell}_e(\theta) \leq \ell_e(\theta)$ as desired. Thus, we have $F^+(\theta) = F^-(\ell_e(\theta))$ and F^- is the outflow function of F^+ with respect to the deterministic queuing model.

For proving the converse, let F^- be the outflow function with respect to the deterministic queuing model. Consider a certain point in time $\theta \in \mathbb{R}_+$. If $F^+(\theta) = F^-(\theta + \tau_e)$ is valid, the maximality property of F^- is satisfied for θ. So assume $F^+(\theta) > F^-(\theta + \tau_e)$ implying $\ell(\theta) > \theta + \tau_e$ and

$$\int_{\theta+\tau_e}^{\ell_e(\theta)} u_e|_{\leq\theta}(\vartheta)\, d\vartheta \;=\; 0$$

because of (7.12). This shows that e is essentially used up to its capacity over $(\theta + \tau_e, \ell_e(\theta)) \neq \emptyset$. By Corollary 2.8, this shows that, for almost all points in time $\theta \in \mathbb{R}_+$, either $F^-(\theta + \tau_e) = F^+(\theta)$ or $f_e^-(\theta + \tau_e) = u_e(\theta + \tau_e)$ holds. Thus, F^- is the maximal function satisfying (7.15). □

Lemma 7.7 defines the deterministic queuing model as an explicit outflow model. Note that Definition 7.5 is an implicit formulation. Further, Lemma 7.7 shows directly that the deterministic queuing model is $\mathcal{F}$-continuous resulting in the following theorem.

Theorem 7.8. *The deterministic queuing model is consistent.*

Proof. Because of Lemma 7.6, it remains to verify the $\mathcal{F}$-continuity. That is we have to show that the outflow function of an edge behaves continuous with respect to the inflow function. But this follows directly from the observation in Lemma 7.7. □

7.3 Characterization of Nash Equilibria

In this section we characterize Nash equilibria for the deterministic queuing model equivalently to Definition 4.8. More precisely, we observe that Nash flows over time are determined by a sequence of thin flows introduced in Chapter 6. For this we analyze the underlying static flow (see Definition 4.15). Conversely, we show that a flow over time where the flow behavior is given by static thin flows is a Nash flow. We only observe this for the case of constant edge capacities. Since a Nash flow only sends flow along currently shortest paths, it makes sense to establish the following definition.

Definition 7.9 (Current Shortest Paths Network)**.** Consider a flow over time on a network $(G,\mathcal{U},\mathcal{T},s,t,D)$. For all $\theta \in \mathbb{R}_+$, the *current shortest path network* is the static network $\mathcal{N}_\theta := (G_\theta,\mathcal{U}_\theta,s,t,d_\theta)$ induced by the edges occurring in currently shortest paths, i.e., edges $e = vw$ with $\ell_e(\ell_v(\theta)) = \ell_w(\theta)$. Further, we set the real edge capacity $u_e^\theta \in \mathbb{R}_+$ of such an edge $e = vw$ to $u_e(\ell_w(\theta))$ and define the static supply by $d_\theta := d(\theta)$.

Note that all nodes v are contained in any current shortest path network if v is reachable from s in G. But in general, there are edges reachable from s in G which are not contained in a current shortest path network. In the following we consider thin flows on current shortest path networks. Note that a component of a current shortest path network $\mathcal{N}_\theta$ is the static supply d_θ. Hence, because of Remark 6.25, we initialize the resetting label $\ell_s \in \mathbb{R}_+$ of s by $\ell_s := \frac{|x|}{d_\theta}$ where $|x|$ is the value of a thin flow on $\mathcal{N}_\theta$.

Next we show that, for a Nash flow over time, the derivatives of the label functions and of the underlying static flow define a thin flow with resetting. The following theorem is only applicable if the derivatives of the label and the underlying static flow functions exist (from right). However, both the label functions and the underlying static flow functions are nondecreasing implying that both families of functions are differentiable almost everywhere.

Theorem 7.10. *Consider a Nash flow over time on a network $(G,\mathcal{U},s,t,\mathcal{T},D)$. Let $(\ell_v)_{v\in V}$ be the family of node label functions, $(q_e)_{e\in E}$ be the family of edge waiting time functions, and, for each $\theta \geq 0$, let $x(\theta)$ be the underlying static flow. Let $\theta \in \mathbb{R}_+$ be such that $\frac{dx_e}{d\theta}(\theta)$ and $\frac{d\ell_v}{d\theta}(\theta)$ exist (from right) for all $e \in E$ and $v \in V$, respectively. Then, on the current shortest path network $\mathcal{N}_\theta$, the derivatives (from right) $(\frac{dx_e}{d\theta}(\theta))_{e\in E(G_\theta)}$ form a thin flow of value $d(\theta)$ with* *resetting on the* waiting edges $E_1 := \{e \in E \mid q_e(\theta) > 0\}$. *Corresponding node labels are given by* $(\frac{d\ell_v}{d\theta}(\theta))_{v \in V(G_\theta)}$.

Proof. Firstly, we observe that $(\frac{dx_e}{d\theta}(\theta))_{e \in E(G_\theta)}$ is a static flow of value $d(\theta)$. Theorem 4.13 implies that $x(\theta+\epsilon)-x(\theta)$ is a static flow of value $D(\theta+\epsilon)-D(\theta)$ for all $\epsilon > 0$. Dividing both sides of each flow conservation constraint for static flows by ϵ shows that $(\frac{dx_e}{d\theta}(\theta))_{e \in E(G_\theta)}$ is a static flow of value $d(\theta)$.

Further, the label of s is set appropriately with respect to Remark 6.25 because equation (3.42) sets $\ell_s := \mathrm{id}$ implying

$$\frac{d\ell_s}{d\theta}(\theta) \;=\; 1 \;=\; \frac{d(\theta)}{d(\theta)} \;.$$

Next, we show that $(\frac{dx_e}{d\theta}(\theta))_{e \in E(G_\theta)}$ and the labels $(\frac{d\ell_v}{d\theta}(\theta))_{v \in V(G_\theta)}$ satisfy the thin flow with resetting conditions (6.2)-(6.4) with respect to the set of resetting edges $E_1 := \{e \in E \mid q_e(\theta) > 0\}$. We distinguish three cases and show that each of these conditions is satisfied in every case. For the rest of this proof let $e = vw \in E(G_\theta)$ be an edge which is contained in a currently shortest s-t-path at time $\theta \in \mathbb{R}_+$.

Case 1: Edge e fits this case if there is an $\epsilon > 0$ where $\ell_v(\theta')) + \tau_e < \ell_w(\theta')$ holds for all $\theta' \in (\theta, \theta+\epsilon]$. This means that a waiting queue is built, occurs already, or the capacity vanishes. In particular, if $e \in E_1$ is valid, e belongs to this case because foresighted node label functions are right-continuous. Since e is used up to its capacity in this case, we get

$$x_e(\theta+\epsilon) - x_e(\theta) \;=\; \int_{\ell_w(\theta)}^{\ell_w(\theta+\epsilon)} f_e^-(\vartheta)\, d\vartheta \;=\; \int_{\ell_w(\theta)}^{\ell_w(\theta+\epsilon)} u_e^-(\vartheta)\, d\vartheta \;.$$

Dividing both sides of the last equation by ϵ and letting ϵ tend to zero leads to

$$\frac{d\ell_w}{d\theta}(\theta) \;=\; \frac{dx_e}{d\theta}(\theta) \cdot \frac{1}{u_e(\ell_w(\theta))} \;.$$

Therefore, condition (6.4) is satisfied in this case. Further, condition (6.2) is satisfied because the label functions are nondecreasing. In order to show that condition (6.3) is valid in this case, we have to verify $\frac{d\ell_v}{d\theta}(\theta) \leq \frac{d\ell_w}{d\theta}(\theta)$ assuming $\ell_v(\theta) + \tau_e = \ell_w(\theta)$. Because we know that e is contained in a shortest path for all times in $(\theta, \theta+\epsilon]$, we can conclude that

$$\ell_v(\theta+\epsilon) - \ell_v(\theta) \;=\; \ell_w(\theta+\epsilon) - \ell_w(\theta) - q_e\big(\ell_v\big(\theta+\epsilon\big)\big) \;\leq\; \ell_w(\theta+\epsilon) - \ell_w(\theta) \;.$$

This yields the desired result if we divide both sides by ϵ and let ϵ tend to zero.

Case 2: Here, we consider the case that there exists a small $\epsilon > 0$ such that $\ell_v(\theta') + \tau_e > \ell_w(\theta')$ holds for all $\theta' \in (\theta, \theta+\epsilon]$. Further, we only have to consider scenarios which do not satisfy Case 1 implying $\ell_v(\theta) + \tau_e = \ell_w(\theta)$ because foresighted arrival time functions are right continuous. Therefore, the inequality $\ell_v(\theta+\epsilon) - \ell_v(\theta) > \ell_w(\theta+\epsilon) - \ell_w(\theta)$ is valid. Moreover, e is not contained in currently shortest path for all $\theta' \in (\theta, \theta+\epsilon]$. Hence, no flow is assigned to e during the time interval $(\ell_v(\theta), \ell_v(\theta+\epsilon)]$, i.e., $x_e(\theta+\epsilon) - x_e(\theta) = 0$. Thus, dividing both sides of the last inequality and of the last equation by ϵ and letting ϵ tend to zero yield

$$\frac{d\ell_w}{d\theta}(\theta) \;\leq\; \frac{d\ell_v}{d\theta}(\theta) \qquad \text{and} \qquad \frac{dx_e}{d\theta}(\theta) \;=\; 0 \;.$$

Thus, condition (6.2) is satisfied and the two other conditions are not relevant in this case.

Case 3: In this case we consider the complement of Case 1 and Case 2. Let $\epsilon > 0$. As we are in the complement of Case 1, we know $q_e(\theta) = 0$ implying $\ell_v(\theta) + \tau_e = \ell_w(\theta)$. Further, this ensures the existence of a $\theta_1 \in (\theta, \theta + \epsilon]$ such that $q_e(\ell_v(\theta_1)) = 0$ implying $\ell_v(\theta_1) + \tau_e \geq \ell_w(\theta_1)$. Since we need not consider situations which fall in Case 2, there exists an $\theta_2 \in (\theta, \theta + \epsilon]$ such that e is contained in a currently shortest path at time θ_2 implying $\ell_v(\theta_2) + \tau_e \leq \ell_w(\theta_2)$.

Hence, on the one hand, we have $\ell_v(\theta_1) - \ell_v(\theta) \geq \ell_w(\theta_1) - \ell_w(\theta)$ and, on the other hand, $\ell_v(\theta_2) - \ell_v(\theta) \leq \ell_w(\theta_2) - \ell_w(\theta)$. Dividing both inequalities by $\theta_1 - \theta$ and $\theta_2 - \theta$, respectively, and decreasing ϵ to 0, we obtain

$$\frac{d\ell_w}{d\theta}(\theta) \;=\; \frac{d\ell_v}{d\theta}(\theta) \ .$$

assuming that $\frac{d\ell_w}{d\theta}(\theta)$ exists. Therefore, condition (6.2) is satisfied. Because condition (6.4) does not apply in this case, it remains to show that condition (6.3) is valid. For this we observe $\frac{dx_e}{d\theta}(\theta) \cdot \frac{1}{u_e} \leq \frac{d\ell_w}{d\theta}(\theta)$. From the capacity constraint (7.10) for the deterministic queuing model, we get

$$x_e(\theta + \epsilon) - x_e(\theta) \;=\; \int_{\ell_w(\theta)}^{\ell_w(\theta+\epsilon)} f_e^-(\vartheta)\, d\vartheta \;\leq\; \int_{\ell_w(\theta)}^{\ell_w(\theta+\epsilon)} u_e^-(\vartheta)\, d\vartheta \ .$$

If we divide both sides by ϵ and let ϵ tend to 0, we get the desired result. This completes the proof. $\square$

The reverse direction of Theorem 7.10 also holds. If, for all times θ, the derivatives of the underlying static flow functions and the label functions of a flow over time are thin flows with resetting in the current shortest paths network, the flow over time is a Nash flow over time.

In the following, we turn our attention to constant scenarios, i.e., networks where each capacity function is constant and, therefore, representable by some nonnegative real number. The next Theorem 7.11 is not the direct conversion of Theorem 7.10 but rather a corollary of the reverse direction. It shows that a Nash flow over time can be computed as the concatenation of thin flows with resetting, which are *static flows*.

The situation is the following. We assume that we already have a restricted flow over time $\mathcal{F}|_{\leq\theta} := (f^+|_{\leq\theta}, f^-|_{\leq\theta})$ showing the selfish routing behavior of flow particles originating at s until a certain time θ. As we deal with constant scenarios, $\mathcal{F}|_{\leq\theta}$ is a Nash flow for the supply function

$$d(\vartheta) = \begin{cases} d & \text{for } \vartheta < \theta \\ 0 & \text{for } \vartheta \geq \theta \end{cases} .$$

Note that this does not really fit the model with a constant supply rate. But since the deterministic queuing model is consistent and, especially, past-oriented, this does not lead to any conflict.

In order to extend $\mathcal{F}|_{\leq\theta}$, we compute a thin flow x' on the current shortest path network $\mathbb{N}_\theta := (G_\theta, s, t, \mathcal{U}, d)$ of value $d \in \mathbb{R}_+$ with resetting on the waiting edges given by $E_1 := \{e \in E \mid q_e(\ell(v)) > 0\}$. Let ℓ' be the corresponding node labels of the thin flow x'. For a given $\alpha > 0$, we extend the node label functions ℓ_v of $\mathcal{F}|_{\leq\theta}$ by

$$\ell_v(\vartheta) \ := \ \ell_v(\theta) + (\vartheta - \theta) \cdot \ell'_v \qquad \forall v \in V \text{ and } \vartheta \in [\theta, \theta + \alpha) \ . \tag{7.16}$$

Then we also extend the flow rate functions by

$$f_e^+(\vartheta) \ := \ \frac{x'_e}{\ell'_v} \qquad \forall e = vw \in E \text{ and } \vartheta \in \big[\ell_v(\theta), \ell_v(\theta + \alpha)\big) \ , \tag{7.17}$$

$$f_e^-(\vartheta) \ := \ \frac{x'_e}{\ell'_w} \qquad \forall e = vw \in E \text{ and } \vartheta \in \big[\ell_w(\theta), \ell_w(\theta + \alpha)\big) \ . \tag{7.18}$$

The result is called *α-extension of $\mathcal{F}|_{\leq\theta}$*.

We want to show that an α-extension is a feasible Nash flow over time $\mathcal{F}|_{\leq\theta+\alpha}$ for flow particles entering the network until time $\theta + \alpha$. For this we have to choose α appropriately such that

$$\ell_w(\theta) - \ell_v(\theta) + \alpha(\ell'_w - \ell'_v) \ \geq \ \tau_e \qquad \forall e \in E_1 \ , \tag{7.19}$$

$$\ell_w(\theta) - \ell_v(\theta) + \alpha(\ell'_w - \ell'_v) \ \leq \ \tau_e \qquad \forall e \in E \setminus E(G_\theta) \ . \tag{7.20}$$

It is essential to show that α can be chosen strictly positive. In order to prove this, first observe that $\ell_w(\theta) - \ell_v(\theta) - \tau_e = q_e(\ell_v(\theta)) > 0$ for all edges $e \in E_1$. Hence, there exists an $\alpha_1 > 0$ such that (7.19) is satisfied for all $\alpha \leq \alpha_1$. The second condition (7.20) refers to edges $e = vw$ which are not contained in the current shortest path network at time θ, i.e., $\ell_w(\theta) < \ell_v(\theta) + \tau_e$. Thus, there exists an $\alpha_2 > 0$ such that the second condition is satisfied for all $\alpha \leq \alpha_2$. This shows the existence of an $\alpha > 0$ satisfying both conditions simultaneously.

In order to extend a restricted Nash flow as far as possible, α should be chosen as large as possible. That is, $\alpha > 0$ can be interpreted as the largest number such that no waiting queue decreases to 0 and no new edge is added to the current shortest path network. In particular, all edges satisfying (7.19) with equality are removed from the set of resetting edges E_1, and all edges satisfying (7.20) with equality are added to the current shortest path network at time $\theta + \alpha$.

If we insert the extension of the node labels (7.16) in the two conditions (7.19) and (7.20) on α, we obtain the following for all $\vartheta \in [\theta, \theta + \alpha)$ (using (6.3) in addition):

$$\ell_w(\vartheta) \ \geq \ \ell_v(\vartheta) + \tau_e \qquad \forall e = vw \in E(G_\theta) \text{ with } x'_e > 0 \ , \tag{7.21}$$

$$\ell_w(\vartheta) \ \leq \ \ell_v(\vartheta) + \tau_e \qquad \forall e = vw \in E \setminus E(G_\theta) \ . \tag{7.22}$$

This is used in the proof of the following theorem.

Theorem 7.11. *Let $\mathcal{F}|_{\leq\theta}$ be a restricted Nash flow over time on $[0, \theta)$ and let $\alpha > 0$ be a positive real number satisfying* (7.19) *and* (7.20)*. Then the α-extension of $\mathcal{F}|_{\leq\theta}$ is a restricted Nash flow over time $\mathcal{F}|_{\leq\theta+\alpha}$ on $[0, \theta + \alpha)$.*

Proof. Firstly, we show that the α-extension of $\mathcal{F}|_{\leq\theta}$ is a *feasible* flow over time. Afterwards, we show that the α-extension is also a *Nash* flow over time.

In order to show that the α-extension is a feasible flow over time, we first observe that Lemma 7.7 is applicable to each edge $e \in E$. Subsequently, we verify that the α-extension satisfies the strict flow conservation constraint (3.7) of Definition 3.9. Since $\mathcal{F}|_{\leq\theta}$ is a restricted Nash flow over time on $[0,\theta)$, it is enough to check these conditions for flow particles origanting at s from time θ on.

For using Lemma 7.7, we verify condition (7.15) in the following. If $x'_e = 0$, nothing has to be proven. So let $e = vw$ be an edge with $x'_e > 0$ and $\vartheta \in [\theta, \theta+\alpha)$ be a point in time. The thin flow conditions (6.3) and (6.4) show $\ell'_w \geq \frac{x'_e}{u_e}$ implying

$$f_e^-\big(\ell_v(\vartheta)\big) \;=\; \frac{x'_e}{\ell'_w} \;\leq\; u_e \;.$$

This proves (7.10). In order to prove $F^+ > F_e^- + \tau_e$, we observe from (7.17) and (7.18) that

$$\begin{aligned} F_e^+\big(\ell_v(\vartheta)\big) \;&=\; F_e^+\big(\ell_v(\theta)\big) + \big(\ell_v(\vartheta) - \ell_v(\theta)\big)\cdot\frac{x'_e}{\ell'_v} \\ \text{and} \qquad F_e^-\big(\ell_w(\vartheta)\big) \;&=\; F_e^-\big(\ell_w(\theta)\big) + \big(\ell_w(\vartheta) - \ell_w(\theta)\big)\cdot\frac{x'_e}{\ell'_w} \end{aligned}$$

are valid. Since $\mathcal{F}|_{\leq\theta}$ is a feasible Nash flow over time, $F_e^+(\ell_v(\theta)) = F_e^-(\ell_w(\theta))$ holds. In addition, we obtain from (7.16) that

$$\begin{aligned} \big(\ell_v(\vartheta) - \ell_v(\theta)\big)\cdot\frac{x'_e}{\ell'_v} \;&=\; (\vartheta-\theta)\cdot\ell'_v\cdot\frac{x'_e}{\ell'_v} \\ &=\; (\vartheta-\theta)\cdot\ell'_w\cdot\frac{x'_e}{\ell'_w} \;=\; \big(\ell_w(\vartheta) - \ell_w(\theta)\big)\cdot\frac{x'_e}{\ell'_w} \;. \end{aligned}$$

This yields $F_e^+(\ell_v(\vartheta)) = F_e^-(\ell_w(\vartheta))$ and proves $F^+ > F_e^- + \tau_e$ because of (7.21) and the monotonicity of F_e^-. Further, (7.16) shows $\ell_w(\vartheta) = \ell_v(\vartheta) + \tau_e$ in case of $\ell'_v = \ell'_w$ and $\ell_v(\vartheta) = \vartheta + \tau_e$. In each other case we have $\frac{x'_e}{u_e} = \ell'_w$ implying $f^-(\ell_w(\vartheta)) = u_e$ by (7.18). Hence, at least one inequality of (7.15) is satisfied with equality. Thus, Lemma 7.7 shows that the flow behavior on e is feasible with respect to the deterministic queuing model. Finally, the α-extension satisfies flow conservation because x' satisfies the static flow conservation constraints implying

$$\sum_{e\in\delta^-(v)} f_e^-\big(\ell(\vartheta)\big) \;=\; \sum_{e\in\delta^-(v)} \frac{x'}{\ell'_v} \;=\; \sum_{e\in\delta^+(v)} \frac{x'}{\ell'_v} \;=\; \sum_{e\in\delta^+(v)} f_e^+\big(\ell(\vartheta)\big) \;.$$

because of (7.17) and (7.18). Hence, the α-extension of f is a feasible flow over time.

It remains to show that the α-extension is also a Nash flow over time. For this, we show that the extended node label functions coincide with node label functions defined by (3.42) and that condition (ii) of Theorem 4.13 is satisfied. Note that condition (ii) of Theorem 4.13 is already proven if $x'_e > 0$, because we know that

$$x_e^+(\vartheta) \;:=\; F_e^+\big(\ell_v(\vartheta)\big) \;=\; F_e^-\big(\ell_w(\vartheta)\big) \;=:\; x_e^-(\vartheta) \;.$$

If $x'_e = 0$, this condition is still valid because $\mathcal{F}|_{\leq\theta}$ is a restricted Nash flow.

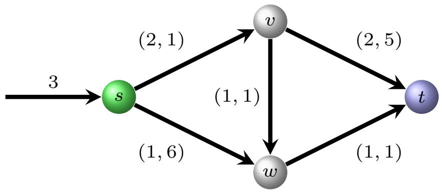


Figure 7.5: The network used in Example 7.12 for which we construct a Nash flow over time. On the incoming edge of the source s the supply rate is shown. For all other edges e, the pair (u_e, τ_e) denotes the capacity u_e and the free flow transit time τ_e.

It remains to consider the node label functions. Since $\ell'_s = 1$, the label of s is set correctly. If $x'_e > 0$ holds for an edge e, the equation $F_e^+(\ell_v(\vartheta)) = F_e^-(\ell_w(\vartheta))$ implies $\ell_v(\vartheta) + \tau_e + q_e(\ell_v(\vartheta)) = \ell_w(\vartheta)$ for all $\vartheta \in [\theta, \theta + \alpha)$. Next, we consider edges $e = vw$ contained in the shortest path network at time θ with $x'_e = 0$. It is not hard to see that we have $q_e(\ell_v(\theta)) = 0$ for these edges. But then condition (6.2) implies that $\ell_v(\vartheta) + \tau_e \geq \ell_w(\vartheta)$ for $\vartheta \geq \theta$. Hence, it remains to show $\ell_v(\vartheta) + \tau_e \geq \ell_w(\vartheta)$ for all $\vartheta \geq \theta$ for edges which are not in the shortest path network at time θ. But this follows directly from the definition of α. Hence, the α-extension of f is a restricted Nash flow over time on $[0, \theta + \alpha)$. □

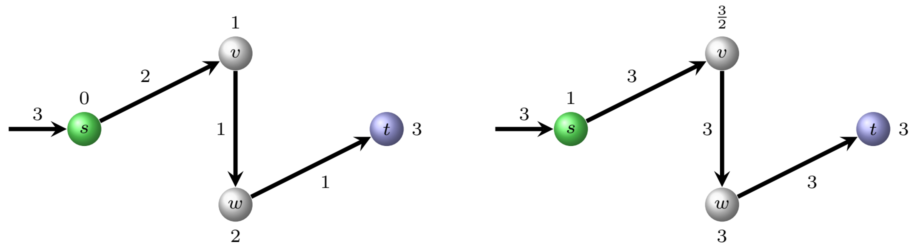
Theorem 7.11 shows that a restricted Nash flow over time is strictly extendable by a particular thin flow in case of constant capacities. Hence, a Nash flow over time can be seen as the concatenation of such static flows. In what follows, we exemplarily show how the α-extension can be used in order to construct a Nash flow over time.

Example 7.12. Consider the network shown in Figure 7.5. In order to construct a Nash flow over time, we first start at time 0 with the zero flow as the restricted flow until time 0 and compute an α-extension for some suitable $\alpha > 0$. Recalling Theorem 7.12, the α-extension results in a restricted Nash flow over time $\mathcal{F}|_{\leq\alpha}$. Next, we compute again an α-extension for $\mathcal{F}|_{\leq\alpha}$ resulting again in a restricted Nash flow, but now on a longer time interval. In the following we do not evaluate the inflow and outflow rate function explicitly because we want to focus on the interesting parts of this Nash flow computation. Note that they can be easily computed using (7.17) and (7.18).

Before computing an α-extension for the zero flow, we have to identify the current shortest path network $\mathcal{N}_0$ and need to evaluate the node label functions at time 0. It is quite obvious that $\mathcal{N}_0$ is given by the zigzag path *svwt* and that $(\ell_s(0), \ell_v(0), \ell_w(0), \ell_t(0)) = (0, 1, 2, 3)$. Since no waiting queue exists at this initial state, we have to find a thin flow x' of value $d = 3$ on $\mathcal{N}_0$ without resetting, i.e., $E_1 = \emptyset$. Figure 7.6 shows x' together with the corresponding node labels ℓ'.

Next, we have to choose an α satisfying (7.19) and (7.20). Since there is no waiting edge, we have to verify condition (7.20) for the edges *sw* and *vt* which are not contained in G_0. Hence, α has to satisfy:

$$\begin{aligned} 2 + 2\alpha &= \ell_w(0) - \ell_s(0) + \alpha(\ell'_w - \ell'_s) \leq \tau_{sw} = 6 \, , \\ 2 + \frac{3}{2}\alpha &= \ell_t(0) - \ell_v(0) + \alpha(\ell'_t - \ell'_v) \leq \tau_{vt} = 5 \, . \end{aligned}$$



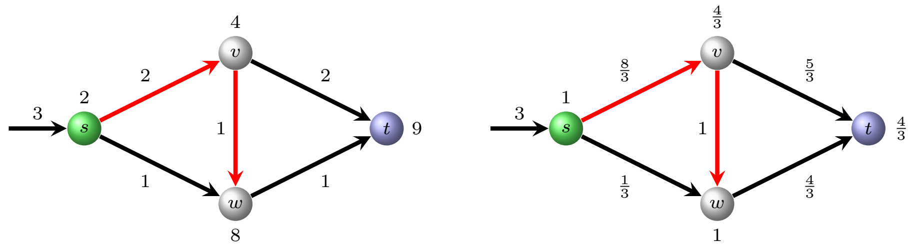
(a) The current shortest path network $\mathcal{N}_0$ with current node labels.

(b) A thin flow on $\mathcal{N}_0$.

Figure 7.6: Computation of the first α extension for the restricted flow $\mathcal{F}|_{\leq 0}$ until time 0.

Of course, we choose α as large as possible. Therefore, we set $\alpha := 2$. Thus the α-extension results in a restricted Nash flow on the time interval $[0, 2)$.

In order to extend this flow further, first note that both inequalities are satisfied with equality and hence, at time 2 both edges, sw and vt, enter the current shortest path network, i. e., $G_2 = G$. Using (7.16), the label functions at time 2 are $(\ell_s(2), \ell_v(2), \ell_w(2), \ell_t(2)) = (2, 4, 8, 9)$. This shows that on sv and vw the experienced travel time is greater than the free flow transit time implying that the current waiting time on these edges is strictly positive. Hence, sv and vw become resetting edges. The current shortest path network just as the corresponding thin flow are shown in Figure 7.7.



(a) The current shortest path network $\mathcal{N}_2$ with current node labels.

(b) A thin flow on $\mathcal{N}_2$.

Figure 7.7: Computation of the second α-extension for the restricted flow $\mathcal{F}|_{\leq 2}$ until time 2.

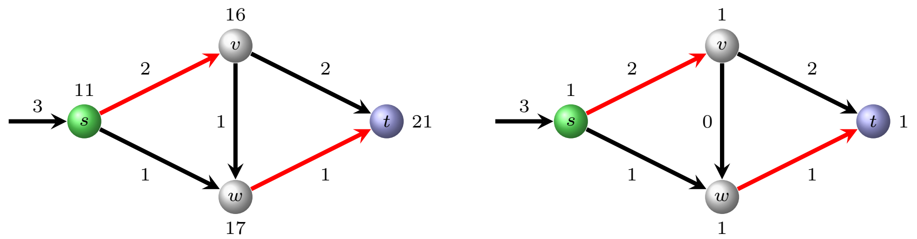
As in the first stage, we have to choose an α satisfying (7.19) and (7.20). Since $G_2 = G$, we have to verify condition (7.19) for the edges sv and vw. Hence, α must satisfy

$$\begin{aligned} 2 + \frac{1}{3}\alpha &= \ell_v(2) - \ell_s(2) + \alpha(\ell'_v - \ell'_s) \geq \tau_{sv} = 1 \,, \\ 4 - \frac{1}{3}\alpha &= \ell_w(2) - \ell_v(2) + \alpha(\ell'_w - \ell'_v) \geq \tau_{vw} = 1 \,. \end{aligned}$$

Note that the first inequality is valid for all $\alpha \in \mathbb{R}_+$. Hence, we set α such that the second inequality is satisfied with equality, i.e., $\alpha := 9$. Therefore, the α-extension is a restricted Nash flow over time up to time 11. Using (7.16), the label functions at time 11 are $(\ell_s(11), \ell_v(11), \ell_w(11), \ell_t(11)) = (11, 16, 17, 21)$.

This shows that the flow particles which start traversing the network at time 11 has to wait on sv and wt. On the other hand, the waiting queue on vw disappears completely within the time interval $[2, 11)$.

In the last iteration, we have to consider the current shortest path network $\mathcal{N}_{11}$ at time 11 where the underlying graph equals G. Further, the resetting edges are sv and wt. The network $\mathcal{N}_{11}$ and the corresponding thin flow x' are shown in Figure 7.8.



(a) The current shortest path network $\mathcal{N}_{11}$ with current node labels.

(b) A thin flow on $\mathcal{N}_{11}$.

Figure 7.8: Computation of the third α-extension for the restricted flow $\mathcal{F}|_{\leq 11}$ until time 11.

We see that x' is equal to the maximum s-t-flow where corresponding node labels are all equal to 1. Without going into details, we note that we can set α to ∞ without violating (7.19) and (7.20). Hence, this α-extension results in a Nash flow over time on the network shown in Figure 7.5 (defined for all nonnegative points in time). Since all node labels of x' are equal to 1, the experienced travel time of each edge remains constant from time 11 on. In particular, the lengths of the waiting queues on sv and wt do not vary over time.

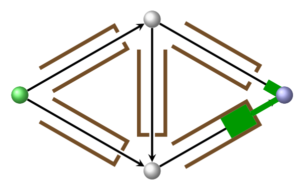
Beside this example, an animation of a Nash equilibrium for the deterministic queuing model is placed as a flip-book at the lower right boarder of this thesis.

7.4 Price of Anarchy

In this section we discuss the price of anarchy for the deterministic queuing model in constant scenarios. The characterization of Nash flows over time via thin flows with resetting enables us to completely analyze shortest paths networks where every s-t-path has the same total free flow transit time. Specifically, we show that such Nash equilibria are optimal and computable in polynomial time. That is, the price of anarchy is 1 for each of the objective functions discussed in Section 4.3.

For arbitrary free flow transit times, we analyze the performance of Nash equilibria on particular instances. As a result, we see that the evacuation price of anarchy increases linearly in the number of edges. Surprisingly, for all other objective functions, the prices of anarchy on these instances are constant. This is even more interesting as the considered networks seem to be worst case instances.

In a *shortest path network* each edge is contained in a shortest s-t-path with respect to free flow transit times. The next Lemma shows that the price of anarchy on shortest path networks is equal to 1 for each of the objective



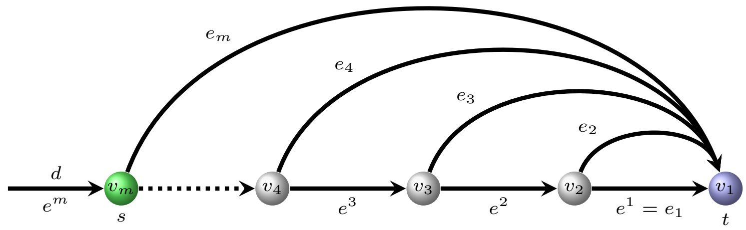


Figure 7.9: A family of instances with a nonconstant evacuation price of anarchy.

functions presented in Section 4.3. Moreover, it shows that a Nash equilibria is representable by one thin flow without resetting which computable in polynomial time as explained in Subsection 6.3.1. Note that we limit to constant edge capacity functions as already mentioned.

Theorem 7.13. *For shortest paths networks, each Nash flow over time is an earliest arrival flow. Moreover, a Nash flow over time can be computed in polynomial time.*

Proof. First note that the current shortest path network at time 0 is the entire network. Since at time 0 no waiting occurs, Theorem 7.11 is valid for $\alpha = \infty$ and shows that a thin flow x' without resetting on the shortest path network results in a Nash flow over time $\mathcal{F}$ (defined for all times). Lemma 6.34 implies that f sends at each point in time the maximum possible flow rate into t. This can be seen as follows. Let ℓ' be the corresponding node labels of x' and let F^* be the maximum static flow value on (G, u, s, t). Then Lemma 6.34 shows that the label of t equals $\ell'_t = \frac{d}{F^*}$. Hence, the definition of the flow rate functions in an α-extension shows that the inflow rate in t is $\frac{|x'|}{\ell'_t} = F^*$. This implies that f is an earliest arrival flow because this inflow rate is achieved as early as possible. Theorem 6.36 shows that this Nash flow over time can be computed in polynomial time. □

Theorem 7.13 shows that, on shortest path networks, a Nash flow over time is an earliest arrival flow assuming constant capacity functions. This does not hold for arbitrary transit times. In this case there even exist instances of the routing game over time where the evacuation price of anarchy is not constant. This is shown in the next lemma.

For the rest of this section, we consider the following class of instances which are illustrated in Figure 7.9. Regarding the capacities $u_{e_k} := u_k$ of the bow edges, let $u_k > 0$ be an arbitrary positive real number for all $k = 1, \ldots, m$. The capacities $u_{e^k} := u^k$ of the lower straight edges are given by $u^k := \sum_{i=1}^k u_i$ for all $k = 1, \ldots, m$. Note that the supply is represented by the edge e^m. Hence, the supply rate is equal to u^m.

The transit times of all lower straight edges are equal to 0, i. e., $\tau_{e^k} := 0$ for all $k = 1, \ldots, m$. Further, for $k = 2, \ldots, m$, the transit time $\tau_{e_k} := \tau_k$ of the bow edge e_k is set to

$$\tau_k := \alpha \cdot u^m \Big(\frac{1}{u^1} - \frac{1}{u^k}\Big) \tag{7.23}$$

for some given $\alpha > 0$. Note that, for $k = 2, \dots, m$, the definition of the capacities implies $u^1 < u^{k-1}$ and, hence, $\tau_k > 0$. Therefore, the current shortest path network at time 0 is the lower straight path P_1.

In order to compute a Nash flow over time, let x' be the thin flow on the lower straight path, i.e, $x'_{e^k} = u^m$ for all $k = 1, \dots, m$. Then corresponding node labels are given by $\ell'_{v_k} := \frac{u^m}{u^k}$. We observe that

$$\alpha'(\ell'_{v_1} - \ell'_{v_k}) \leq \tau_k \qquad\qquad \forall \alpha' \in [0, \alpha],\ k = 2, \dots, m$$

is valid where equality holds for $\alpha' = \alpha$. Since $\ell_{v_k}(0) = 0$ for all $k = 2, \dots, m$, condition (7.20) is satisfied with equality on all bow edges for α. Hence, in terms of Theorem 7.11, the α-extension of the zero flow over time is a restricted Nash flow over time on $[0, \alpha)$. In addition, at time α, the entire network is the current shortest path network and a thin flow is given by the edge capacities. Hence, from time α on, the Nash flow over time uses the entire network. Note that an earliest arrival flow uses each edge of the network up to its capacity right from the beginning. Further, considering the label functions of the node v_k, we obtain

$$\alpha \frac{u^m}{u^k} \;=\; \ell_{v_k}(\alpha) \;=\; \ell_t(\alpha) - \tau_k \ .$$

The left hand side follows from the definition of an α-extension and the right hand side from the definition of the label functions. We rewrite this equality as

$$\alpha \cdot u^m \;=\; \big(\ell_t(\alpha) - \tau_k\big) \cdot u^k \ . \tag{7.24}$$

Using these instances we are able to prove the following lemma.

Lemma 7.14. *There exists a family of instances for which the evacuation price of anarchy is $\Omega(m)$, where m is the number of edges. If all edge capacities are equal to 1, there exists a family of instances for which the evacuation price of anarchy is still $\Omega(\log(m))$.*

Proof. In order to get a lower bound on the evacuation price of anarchy, we consider the point in time $\ell_t(\alpha)$. In the Nash flow described above, the total amount of flow reaching the sink until this point in time is $u^m \cdot \alpha$. Since an earliest arrival flow uses the entire network right from the beginning, the total amount of flow at the sink in an earliest arrival flow is $\sum_{k=1}^{m}(\ell_t(\alpha) - \tau_k) \cdot u_k$. Thus, with (7.24) we obtain the following lower bound on the evacuation price of anarchy:

$$\frac{\sum_{k=1}^{m}(\ell_t(\alpha) - \tau_k)u_k}{u^m \alpha} \;=\; \sum_{k=1}^{m} \frac{(\ell_t(\alpha) - \tau_k)u_k}{u^m \alpha} \;=\; \sum_{k=1}^{m} \frac{u_k}{u^k} \ .$$

In fact, this is the exact price of anarchy for this example.

This shows that the price of anarchy increases linearly in the number of edges (set, e.g., $u^k := 2^k$). If we restrict to instances with unit edge capacities, the price of anarchy still increases logarithmically in the number of edges – set $u_k := 1$ and replace e^k by k parallel edges. Then the sum on the right hand side is equal to the harmonic series and the number of edges is quadratic in m. □

Lemma 7.14 shows that the evacuation price of anarchy grows linearly with the number of edges in general. Surprisingly, for each other objective function, the price of anarchy of the instances shown in Figure 7.9 are constant. To see this, we need the following estimation, for all $k = 1, \ldots, m$, which follows from the inequality (2.1) between the geometric and arithmetic mean

$$\frac{u_1}{u^k} = \prod_{i=2}^{k} \frac{u^{i-1}}{u^i} \leq \left(\frac{1}{k-1} \sum_{i=2}^{k} \frac{u^{i-1}}{u^i} \right)^{k-1} = \left(1 - \frac{1}{k-1} \sum_{i=2}^{k} \frac{u_i}{u^i} \right)^{k-1}$$

Since it is basic knowledge that, for each $\gamma \in \mathbb{R}_+$, the sequence $((1 - \frac{\gamma}{k})^k)_{k \in \mathbb{N}}$ converges from left to $e^{-\gamma}$, this shows

$$\frac{u_1}{u^k} \leq e^{-\sum_{i=2}^{k} \frac{u_i}{u^i}} \tag{7.25}$$

for all $k = 1, \ldots, m$. This equation is used in the proofs of the subsequent lemmas.

Lemma 7.15. *The working price of anarchy on instances depicted in Figure 7.9 is upper bounded by* $2 + \frac{2}{\sqrt{e}} = 3,2131$.

Proof. For some $m \in \mathbb{N}$, let $\mathcal{F}_m$ be the Nash flow on the instance consisting of $2m$ edges. First, we show that the working price of anarchy $\rho_{\text{work}}(\mathcal{F}_\text{N})(\ell_t(\alpha))$ at time $\ell_t(\alpha)$ is bounded by 2. For this we have to estimate

$$\rho_{\text{work}}(\mathcal{F}_\text{N})\big(\ell_t(\alpha)\big) := \frac{\text{val}_{\text{work}}(\mathcal{F}_\text{O})(\theta)}{\text{val}_{\text{work}}(\mathcal{F}_\text{N})(\theta)} = \frac{\sum_{k=1}^{m} \frac{1}{2}(\ell_t(\alpha) - \tau_k)^2 u_k}{\frac{1}{2} u^m \alpha(\ell_t(\alpha) - \tau_1)} .$$

Using equation (7.24) twice, we simplify this to

$$\rho_{\text{work}}(\mathcal{F}_\text{N})\big(\ell_t(\alpha)\big) = \sum_{k=1}^{m} \frac{(\ell_t(\alpha) - \tau_k)^2 u_k}{\frac{1}{u_1}(u^m \alpha)^2} = \sum_{k=1}^{m} \frac{u_1 u_k}{(u^k)^2} . \tag{7.26}$$

For proving $\rho_{\text{work}}(\mathcal{F}_\text{N})(\ell_t(\alpha)) \leq 2$ we set

$$\beta_k := \frac{u_k}{u^k} \qquad \text{implying} \qquad (1 - \beta_k) u^k = u^{k-1} \tag{7.27}$$

for all $k = 1, \ldots m$. Further, because of $(1 + \beta_k)(1 - \beta_k) = 1 - \beta_k^2 \leq 1$, this shows

$$\frac{\beta_k}{u^k} + \frac{1}{u^k} = (1 + \beta_k)(1 - \beta_k)\frac{1}{u^{k-1}} \leq \frac{1}{u^{k-1}} \tag{7.28}$$

for all $k = 1, \ldots m$. Using (7.27) and, inductively, (7.28), we get

$$\begin{aligned} \sum_{k=1}^{m} \frac{u_k}{(u^k)^2} = \sum_{k=1}^{m} \frac{\beta_k}{u^k} &\leq \frac{1}{u^m} + \sum_{k=1}^{m} \frac{\beta_k}{u^k} \\ &\leq \frac{1}{u^{m-1}} + \sum_{k=1}^{m-1} \frac{\beta_k}{u^k} \leq \cdots \leq \frac{1}{u^1} + \frac{\beta_1}{u^1} . \end{aligned}$$

Since $\beta_1 = 1$, this together with equation (7.26) leads to

$$\rho_{\text{work}}(\mathcal{F}_\text{N})\big(\ell_t(\alpha)\big) \leq 2 . \tag{7.29}$$

It is not hard to observe that this bound also holds also for all smaller values α' of α. To see this, compute new transit times using (7.23). On this new instance the transit times of the bow edges are reduced. Hence, the earliest arrival flow is improved, but the restricted Nash flow until time α' shows the same behavior as before. Thus, the working price of anarchy on this new instance at time $\ell_t(\alpha')$ is an upper bound for $\rho_{\text{work}}(\mathcal{F}_{\text{N}})\big(\ell_t(\alpha')\big)$ but is itself bounded by 2.

So it remains to consider the working price of anarchy for larger values than $\ell_t(\alpha)$. Assume that we increase $\ell_t(\alpha)$ by θ. Since after time $\ell_t(\alpha)$ both the Nash flow just as the earliest arrival flow sends flow into t at a rate of u^m, we have to bound

$$\rho_{\text{work}}(\mathcal{F}_{\text{N}})\big(\ell_t(\alpha)+\theta\big) = \frac{\text{val}_{\text{work}}(\mathcal{F}_{\text{O}})(\ell_t(\alpha)) + \theta\cdot\text{val}_{\text{evac}}(\mathcal{F}_{\text{O}})(\ell_t(\alpha)) + \frac{1}{2}\cdot u^m\theta\cdot\theta}{\text{val}_{\text{work}}(\mathcal{F}_{\text{N}})(\ell_t(\alpha)) + \theta\cdot\text{val}_{\text{evac}}(\mathcal{F}_{\text{N}})(\ell_t(\alpha)) + \frac{1}{2}\cdot u^m\theta\cdot\theta}\ .$$

Because of 7.29 we know

$$\frac{\text{val}_{\text{work}}(\mathcal{F}_{\text{O}})(\ell_t(\alpha)) + \frac{1}{2}\cdot u^m\theta\cdot\theta}{\text{val}_{\text{work}}(\mathcal{F}_{\text{N}})(\ell_t(\alpha)) + \frac{1}{2}\cdot u^m\theta\cdot\theta} \ \le\ 2\ .$$

Further, the proof of Lemma 7.14 shows $\text{val}_{\text{evac}}(\mathcal{F}_{\text{O}})(\ell_t(\alpha)) = u^m\alpha\sum_{k=1}^m \frac{u_k}{u^k}$. Hence, using $\text{val}_{\text{work}}(\mathcal{F}_{\text{O}})(\ell_t(\alpha)) = \frac{(u^m\alpha)^2}{2u_1}$ as shown for (7.26), we obtain

$$\frac{\theta\cdot\text{val}_{\text{evac}}(\mathcal{F}_{\text{O}})(\ell_t(\alpha))}{\text{val}_{\text{work}}(\mathcal{F}_{\text{N}})(\ell_t(\alpha)) + \frac{1}{2}\cdot u^m\theta\cdot\theta} \ =\ \frac{\theta\cdot u^m\alpha\cdot\sum_{k=1}^m\frac{u_k}{u^k}}{\frac{(u^m\alpha)^2}{2u_1}+\frac{1}{2}u^m\theta^2} \ =\ \frac{\sum_{k=1}^m\frac{u_k}{u^k}}{\frac{1}{2}(\frac{u^m}{u_1}\cdot\frac{\alpha}{\theta}+\frac{\theta}{\alpha})}\ .$$

Defining $\gamma := \sum_{k=1}^m \frac{u_k}{u^k}$, equation (7.25) together with the arithmetic-geometric-mean-inequality (2.1) leads to

$$\begin{aligned}\rho_{\text{work}}(\mathcal{F}_{\text{N}})\big(\ell_t(\alpha)+\theta\big) \ &\le\ 2+\frac{\sum_{k=1}^m\frac{u_k}{u^k}}{\frac{1}{2}(\frac{u^m}{u_1}\cdot\frac{\alpha}{\theta}+\frac{\theta}{\alpha})}\\ &\le\ 2+\frac{\gamma}{\frac{\alpha}{\theta}\cdot e^{\gamma-1}+\frac{\theta}{\alpha}} \ \le\ 2+\frac{\gamma}{e^{\frac{1}{2}(\gamma-1)}} \ \le\ 2+\frac{2}{\sqrt{e}}\end{aligned}$$

where the last "$\le$"-sign follows from standard curve sketching. □

We conclude this section by showing that the completion and, hence, the average arrival time price of anarchy on this special class of instances is constant.

Lemma 7.16. *The completion time price of anarchy on instances depicted in Figure 7.9 is upper bounded by $\frac{e-1}{e} = 0,6321$ implying that this also holds for the average arrival time price of anarchy.*

Proof. Firstly, observe without loss of generality that the completion time price of anarchy is attained for the flow value $\text{val}_{\text{evac}}(\mathcal{F}_{\text{N}})(\ell_t(\alpha))$. For smaller values we argue along the same lines as in Lemma 7.15. For larger values, the completion time price of anarchy decreases because the earliest arrival just as the Nash flow send the maximum possible rate into t.

In order to compute the completion time price of anarchy for the flow value $\text{val}_{\text{evac}}(\mathcal{F}_{\text{N}})(\ell_t(\alpha)) = u^m\alpha$, we have to find a $\theta\in\mathbb{R}_+$ and a $k\in\{1,\dots,m\}$ such that:

$$u^m\alpha \ =\ \sum_{i=1}^k(\theta-\tau_i)u_i \ =\ \text{val}_{\text{evac}}(\mathcal{F}_{\text{O}})(\theta)\ .$$

Rewriting the sum and using (7.23) leads for all $j \in \{1, \dots, k\}$ to

$$u^m \alpha \;=\; (\theta - \tau_j)u^k + \sum_{i=1}^{k} (\tau_j - \tau_i)u_i \;=\; (\theta - \tau_j)u^k + \sum_{i=1}^{k} u^m \alpha \Big(\frac{1}{u^i} - \frac{1}{u^j}\Big) u_i \;.$$

Separating $(\theta - \tau_j)u^k$ in this formula results in

$$(\theta - \tau_j)u^k = u^m \alpha \bigg(1 - \sum_{i=1}^{k} \Big(\frac{u_i}{u^i} - \frac{u_i}{u^j}\Big)\bigg) = u^m \alpha \bigg(1 + \frac{u^k}{u^j} - \sum_{i=1}^{k} \frac{u_i}{u^i}\bigg) \;.$$

Hence, for $j = k$ we obtain

$$0 \;\le\; (\theta - \tau_k)u^k \;=\; u^m \alpha \bigg(2 - \sum_{i=1}^{k} \frac{u_i}{u^i}\bigg) \qquad \Leftrightarrow \qquad 2 \;\ge\; \sum_{i=1}^{k} \frac{u_i}{u^i} \;. \tag{7.30}$$

Further, for $j = 1$ we get

$$(\theta - \tau_1)u^k \;=\; u^m \alpha \bigg(1 + \frac{u^k}{u_1} - \sum_{i=1}^{k} \frac{u_i}{u^i}\bigg) \;=\; (\ell_t(\alpha) - \tau_1)u_1 \bigg(1 + \frac{u^k}{u_1} - \sum_{i=1}^{k} \frac{u_i}{u^i}\bigg)$$

which we equivalently restate as

$$\frac{\theta - \tau_1}{\ell_t(\alpha) - \tau_1} \;=\; 1 + \frac{u_1}{u^k} - \frac{u_1}{u^k} \sum_{i=1}^{k} \frac{u_i}{u^i} \;=\; 1 - \frac{u_1}{u^k} \sum_{i=2}^{k} \frac{u_i}{u^i} \;.$$

Next, we set $\gamma := \sum_{i=2}^{k} \frac{u_i}{u^i}$. Because of (7.30), we know $0 < \gamma \le 1$. Hence, from (7.25) and standard curve sketching, we finally obtain

$$\frac{\theta - \tau_1}{\ell_t(\alpha) - \tau_1} \;\ge\; 1 - \frac{\lambda}{e^\lambda} \;\ge\; 1 - \frac{1}{e} = \frac{e-1}{e} \;=\; 0,6321 \;. \qquad \square$$

7.5 Your Comments

In this chapter the very popular deterministic queuing model is considered. This model is introduced by Vickrey [89]. Twelve years later it is rediscovered by Hendrickson and Kocur [38] and Henderson [37] note two years later that this model is already known. In 1984, Smith [83] considers cases where there exists a single bottleneck which every flow particle has to traverse. For this special case, he shows the existence of a Nash equilibrium. Akamatsu [2] presents an edge-based formulation of the deterministic queuing model on restricted single-source-instances. Mounce [58, 59] considers the case where the edge capacities can vary over time and states some existence results on limited instances. In 2009 he extends the existence result to arbitrary instances (see [60]).

Recently, Cominetti, Correa, and Larré [18] prove the uniqueness of Nash equilibria in constant scenarios assuming that inflow rates are piecewise constant. Bhaskar, Fleischer, and Anshelevich [11] show that the completion time price of anarchy is lower bounded by $\frac{e-1}{2e}$ if the earliest arrival flow saturates the network.

You observe that this chapter is based on [44] by Koch and Skutella. In addition, to the contribution of [44] a corresponding atomic routing game over time for the deterministic queuing model is considered and the continuous approach is extended from constant to time varying capacity function. Further, bounds on the working, completion time, and average arrival time price of anarchy are presented.

In Section 7.1 an atomic routing game for the deterministic queuing model is analyzed where each edge has a free flow transit time of 0 and a constant capacity function being equal to 1. The essential feature for the correctness of the presented approach is that every s-t-path has the same free flow transit time of 0. That is, in a uncongested network, flow particles starting at the same point in time are able to reach every node at the same point in time. But this aspect is also satisfied by shortest path network where each s-t-path admit the same minimal free flow transit time. In this manner, as you note already at the end of Section 7.1, the presented approach is generalizable to such networks. There you also observe that the contribution of this section remains valid, if you consider arbitrary constant capacity functions.

Recalling Section 4.2, where a Generic Iterative algorithm is used to prove the existence of Nash equilibria for consistent flow over time models, you see that the Generic Iterative algorithm of this chapter is a specialized version of that one suitable for computing nonatomic Nash equilibria of the deterministic queuing model on shortest path networks. However there is one aspect of this Generic Iterative algorithm leading to two very interesting open question. Theorem 7.4 shows that the output sequence of such an algorithm converges to a thin flow without resetting which is a maximum s-t-flow in particular. Hence, the Generic Iterative algorithm can be used to compute a maximum static s-t-flow without using a residual network. Because of the proof of Theorem 7.4, you recognize that you can interrupt the Generic Iterative algorithm after an exponential number of iteration, to obtain a maximum static s-t-flow after some standard rounding technique. However, you have the feeling that this also works after a polynomial number of iteration which would result in a polynomial maximum s-t-flow algorithm which does not act on the residual graph. The second question refers to a Nash flow computation. As in Section 4.2, the Generic Iterative algorithm can be used to compute Nash flow. Hence, in view of the first question, the second question is whether the Generic Iterative algorithm can be used to compute a nonatomic Nash equilibria efficiently. By personal communication to the author, you find out that this scenario is the origin of this thesis.

Clearly, a polynomial time algorithm for computing thin flows also leads to an efficient algorithm for finding Nash equilibria. This is based on the characterization of Nash flows as the concatenation of thin flows presented in Section 7.3. Unfortunately, so far it is even not clear whether the resulting algorithm terminates. However, you strongly believe that and if this is the case, another question ask for the final thin flow. Does this thin flow imply that the Nash flow sends flow at the same maximum rate into t as the earliest arrival flow? Thinking about this problem you come to the conclusion: "Yes, this must be the case!"

In Section 7.4 the performance of Nash equilibia for the deterministic queuing model is considered. On so-called shortest path networks, a Nash equilibrium is an earliest arrival flow implying that each price of anarchy is equal to 1.

Unfortunately, for arbitrary networks you does not find universally valid bounds are found. Instead, the performance of Nash equilibria is analyzed on a special class of instances. For these instances the evacuation price of anarchy grows linearly in the number of edges. Surprisingly, each other price of anarchy is constant. This is interesting as you share the opinion of the author that the analyzed class contains the worst case instances. You observe that the presented bounds for the working and average arrival time price of anarchy are not tight. So if it turns out that these instances really show the worst case, computing the tight bounds is important. May be, this works along the same approach presented in Subsection 5.4.3 which computes the exact working price of anarchy for the direct flow model.

As for the direct flow model, promising future research could be the analysis of periodic capacity functions. Besides, as always, also a measure-based approach to the deterministic queuing model seems to be interesting.

Chapter 8

My Conclusion

In his famous work, Nash [62] shows that every game admits an equilibrium which is stable in the sense that no player has an incentive to change his strategy for improving his personal situation. Roughgarden and Tardos [77] analyze the performance of these Nash equilibria for routing games based on static flows. One main drawback of this class of routing games is the nonconsideration of time. Ford and Fulkerson [27] introduce flows over time representing time-varying flow behavior. Combining game theory and flows over time leads to the notion of routing games over time. Since the seminal work of Friesz et al. [29], the number of publications in this area increases enormously. Nevertheless, compared with the static counterpart, quite less is still known about the flow behavior in routing games over time.

Using common knowledge, the theory of routing games over time is built from scratch in this thesis. Even the theory of flows over time is completely revised resulting in a quite general model. Imposing natural assumptions on the flow behavior, we give a precise notion of routing games over time. In this setting, we characterize and analyze Nash equilibria for the direct flow over time and the popular deterministic queuing model. Because of the high complexity of this topic, we focus on single commodity scenarios where each flow particle wants to travel from a unique source to a unique sink. Throughout this thesis, we make use of many different mathematical fields; including Lebesgue measurable functions, functional analysis, linear algebra, static flow theory, and game theory.

The general flow over time model presented in Chapter 3 defines a flow over time as a family of inflow functions and uses time-varying transit time functions for representing the flow behavior. A transit time function is given by an arbitrary Lebesgue measurable function and may depend on the entire inflow. In particular, past, present, and future flow situations may influence current transit times. Further, we show how the complete flow behavior is representable via outflow functions leading to a novel, more general notion of outflow models. For the sake of simplicity and compatibility, we impose natural assumptions on the flow behavior such as the FiFo principle and the past-orientation which are motivated by real world applications. Because of the generality of the flow model, we are able to give a precise definition for each of these assumptions. Such a well-behaved flow over time model is called consistent. For such consistent flow models, we identify a class of so-called foresighted transit time functions which seem to be as made for representing flow behavior.

It turns out that foresighted transit time functions show the intuitively correct selfish routing behavior in a routing game over time. In Chapter 4, we give a precise definition of Nash equilibria over time which is a direct generalization of the corresponding static definition. Further, we characterize Nash flows via underlying static flows. This characterization is the fundamental feature for establishing the existence of Nash equilibria for all consistent flow over time models. The existence proof is based on an algorithm which, in addition, shows that a nonatomic Nash equilibrium is obtainable as a limit of Nash equilibria on corresponding atomic routing games. In this sense, the definition of Nash flows over time is well-motivated.

Besides, we discuss four possibilities to measure the performance of Nash equilibria. The evacuation price of anarchy is based on the amount of flow which arrives at the sink until some given point in time. Further, the working price of anarchy measures the cumulative amount of time until some given point in time after a flow particle arrives at the sink. The time needed to send a given flow value to the sink is considered by the completion time price of anarchy. Finally, the average arrival time of a given amount of flow is analyzed by the average arrival time price of anarchy.

In Chapter 5 we analyze a routing game over time where the underlying flow over time model forbids waiting. In fact, we consider the continuous version of the flow over time model introduced by Ford and Fulkerson [27] where we require strict flow conservation at each intermediate node of the network. We allow time-varying capacities and show how to compute an earliest arrival flow in this case generalizing the well-known SUCCESSIVE SHORTEST PATH algorithm for computing either static minimum cost flows or earliest arrival flows for constant capacities. It turns out that the price of anarchy is unbounded for each of the four objective functions.

However, using earliest arrival flows, we identify Nash equilibria which show a quite good performance. Moreover, these so-called EA-Nash flows allow us to compute the price of stability by considering only networks consisting of parallel s-t-arcs. More precisely, the evacuation and, hence, the working price of stability is equal to 2. Further, the completion and average arrival time price of stability is unbounded for arbitrary capacity functions. Restricting to constant capacity functions, we reduce the working price of stability to α^2 where α is the unique solution of $2\alpha = \log \frac{\alpha+1}{\alpha-1}$. The completion time price of anarchy evaluates to $\frac{3}{4}$ which also bounds the average arrival time price of anarchy. Moreover, each bound is tight except the bound for the average arrival time price of anarchy in constant scenarios.

The very popular deterministic queuing model introduced by Vickrey [89] is discussed in Chapter 7. We define this model using arbitrary Lebesgue measurable capacity functions. However, the performance of Nash equilibria is only discussed for the case of constant capacities. With respect to the corresponding atomic routing game, we show that a Nash equilibrium is asymptotically optimal on networks where each edge has a free flow transit time of 0 and a capacity of 1. Similarly, nonatomic Nash equilibria over time are optimal on so-called shortest path networks where each s-t-path admits the same minimal free flow transit time. Unfortunately, we are not able to bound the prices of anarchy for arbitrary free flow transit times. On the other hand, we show that the evacuation price of anarchy grows at least linearly with the number of edges. Surprisingly, the working price of anarchy for the considered class of instances is constant. Finally, these scenario admits a completion time price of anarchy of $\frac{e}{e-1}$ which serves also an upper bound for the average arrival time price of anarchy. These results are based on a characterization of Nash flows for the deterministic queuing model which are representable via a sequence of special static flows.

These special static flows are called thin flows and are analyzed in Chapter 6. It is shown that arbitrary thin flows are computable via so-called thin flows without resetting building a subclass of thin flows. Thin flows without resetting are computable in polynomial time and, in some sense, unique and optimal. Analyzing thin flows without resetting, we prove the existence of arbitrary thin flows. Uniqueness results are also given.

In the following, we discuss directions of future research and open questions. So far, the general flow over time model is explained for continuous and discrete representations of time. A measure-based approach would combine both models in one. Besides, a measure-based approach could help to understand where basic limitations of flows over time are located in comparison with static flow theory. Concerning the direct flow over time model, the main open question is the exact average arrival time price of stability assuming constant capacity functions. Clearly, bounds on each price of anarchy are also an interesting open question for the deterministic queuing model. Another direction for future research is the efficient computation of Nash equilibria for the deterministic queuing model. Here, there are two possibilities. Firstly, a polynomial time algorithm for the computation of thin flow would solve the problem. Secondly, a more detailed analysis concerning the running time of the Generic Iterative algorithm could help. In this sense, as a first step, it is worth to show that the Generic Iterative algorithm applied to shortest paths networks leads to a polynomial time algorithm for the maximum s-t-flow problem which does not use residual networks.

Bibliography

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993. 4

[2] T. Akamatsu. An efficient algorithm for dynamic traffic equilibrium assignment with queues. *Transportation Science*, 35(4):389–404, 2001. 137, 274

[3] E. J. Anderson and A. B. Philpott. Optimization of flows in networks over time. In F. P. Kelly, editor, *Probability, Statistics and Optimization: A Tribute to Peter Whittle*, chapter 27, pages 369–382. Wiley, 1994. 46

[4] E. Anshelevich and S. Ukkusuri. Equilibria in dynamic selfish routing. In M. Mavronicolas and V. Papadopoulou, editors, *Algorithmic Game Theory*, volume 5814 of *Lecture Notes in Computer Science*, pages 171–182. Springer, 2009. 137

[5] C. Arzelá. Sulle funzioni di linee. *Memorie della Accademia delle Scienze dell'Istitutio di Bologna Classe di Scienze Fisiche, Matematiche e Naturali*, 5(5):55–74, 1895. 33

[6] G. Ascoli. Le curve limiti di una varietà data di curve. *Atti della Reale Accademia dei Lincei Memorie della Classe di Scienze Fisiche, Matematiche e Naturali*, 18(3):521–586, 1883-1884. 33

[7] M. Balmer, M. Rieser, K. Meister, D. Charypar, N. Lefebvre, and K. Nagel. Matsim-t: Architecture and simulation times. In A. Bazzan and F. Klügl, editors, *Multi-Agent Systems for Traffic and Transportation Engineering*, pages 57–78. IGI Global, 2009. 137

[8] R. Banner and A. Orda. Bottleneck routing games in communication networks. *IEEE Journal on Selected Areas in Communications*, 25(6):1173–1179, 2007. 245

[9] M. Beckmann, C. B. Mc Guire, and C. B. Winsten. *Studies in the Economics of Transportation*. Yale University Press, New Haven, 1956. 7, 45

[10] M. J. Ben-Akiva, M. Bierlaire, H. N. Koutsopoulos, and R. Mishalani. Dynamit: A simulation-based system for traffic prediction and guidance generation. In *Proceedings of the 3rd Triennial Symposium on Transportation Systems*, 1998. 137

[11] U. Bhaskar, L. Fleischer, and E. Anshelevich. A stackelberg strategy for routing flow over time. In D. Randall, editor, *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 192–201, 2011. 274

[12] L. E. J. Brouwer. Über Abbildungen von Mannigfaltigkeiten. *Mathemathische Annalen*, 71(1):97–115, 1911. 243

[13] C. Busch, M. Magdon-Ismail, M. Mavronicolas, and P. Spirakis. Direct routing: Algorithms and complexity. *Algorithmica*, 45(1):45–68, 2006. 192

[14] M. Carey. Link travel times i: Properties derived from traffic-flow models. *Networks and Spatial Economics*, 4(3):257–268, 2004. 5, 8, 111, 112

[15] M. Carey. Link travel times ii: Properties derived from traffic-flow models. *Networks and Spatial Economics*, 4(4):379–402, 2004. 111

[16] H. K. Chen and C. F. Hsueh. A model and an algorithm for the dynamic user-optimal route choice problem. *Transportation Research Part B: Methodological*, 32(3):219–234, 1998. 137

[17] R. Cole, Y. Dodis, and T. Roughgarden. Bottleneck links, variable demand, and the tragedy of the commons. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 668–677. ACM, 2006. 245

[18] R. Cominetti, J. Correa, and O. Larré. Existence and uniqueness of equilibria for flows over time. In Luca Aceto, Monika Henzinger, and Jirí Sgall, editors, *Automata, Languages and Programming*, volume 6756 of *Lecture Notes in Computer Science*, pages 552–563. Springer, 2011. 7, 9, 203, 245, 274

[19] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Selfish routing in capacitated networks. *Mathematics of Operations Research*, 29(4):961–976, 2004. 46

[20] C. F. Daganzo. Queue spillovers in transportation networks with a route choice. *Transportation Science*, 32(1):3–11, 1998. 6

[21] P. Dubey. Inefficiency of nash equilibria. *Math. Oper. Res.*, 11:1–8, 1986. 7

[22] L. Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii academiae scientiarum Petropolitanae*, 8:128–140, 1741. 3

[23] L. Fleischer and M. Skutella. Minimum cost flows over time without intermediate storage. In *Proceedings of the 14th Annual ACM–SIAM Symposium on Discrete Algorithms*, pages 66–75, Baltimore, MD, 2003. 139, 192

[24] L. Fleischer and M. Skutella. Quickest flows over time. *SIAM Journal on Computing*, 36(6):1600–1630, 2007. 6, 192

[25] L. Fleischer and É. Tardos. Efficient continuous-time dynamic network flow algorithms. *Operations Research Letters*, 23(3–5):71–80, 1998. 46, 192

[26] L. R. Ford, jr. and D. R. Fulkerson. Maximal flow through a network. *Canadian Journal of Mathematics*, 8(3):399–404, 1956. 3, 198

[27] L. R. Ford, jr. and D. R. Fulkerson. Constructing maximal dynamic flows from static flows. *Operations Research*, 6(3):419–433, 1958. 4, 41, 42, 46, 277, 278

[28] L. R. Ford, jr. and D. R. Fulkerson. *Flows in Networks*. Princeton University Press, 1962. 4

[29] T. L. Friesz, D. Bernstein, T. E. Smith, R. L. Tobin, and B. W. Wie. A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research*, 41(1):179–191, 1993. 4, 7, 111, 112, 137, 277

[30] T. L. Friesz, J. Luque, R. L. Tobin, and B.-W. Wie. Dynamic network traffic assignment considered as a continuous time optimal control problem. *Operations Research*, 37(6):893–901, 1989. 112, 137

[31] D. Fudenberg and J. Tirole. *Game Theory*. The MIT Press, 1991. 6

[32] D. Gale. Transient flows in networks. *Michigan Mathematical Journal*, 6(1):59–63, 1959. 5, 46

[33] L. Georgiadis, P. Georgatsos, K. Floros, and S. Sartzetakis. Lexicographically optimal balanced networks. *IEEE/ACM Transactions on Networking*, 10(6):818–829, 2002. 245

[34] B. D. Greenshields. A study of traffic capacity. *Highway Research Board Proceedings*, 14:448–477, 1935. 5

[35] S. Han and B G. Heydecker. Consistent objectives and solution of dynamic user equilibrium models. *Transportation Research Part B: Methodological*, 40(1):16–34, 2006. 137

[36] E. Helly. über lineare funktionaloperationen. *Sitzungsberichte der Akademie der Wissenschaften in Wien Mathematische-naturwissenschaftliche Klasse, Abteilung 2a*, 121:165–297, 1912. 33

[37] J. V. Henderson. On a paper by c. hendrickson and g. kocur entitled "schedule delay and departure time decisions in a deterministic model". *Transportation Science*, 17(3):358–359, 1983. 274

[38] C. Hendrickson and G. Kocur. Schedule delay and departure time decisions in a deterministic model. *Transportation Science*, 15(1):62–77, 1981. 137, 274

[39] M. Hoefer, V. S. Mirrokni, H. Röglin, and S.-H. Teng. Competitive routing over time. *Theoretical Computer Science*, 412(39):5420–5432, 2011. 138

[40] J. J. Jarvis and H. D. Ratliff. Some equivalent objectives for dynamic network flow problems. *Jarvis and H. Donald Ratliff*, 28(1):106–109, 1982. 46

[41] S. Kakutani. A generalization of brouwer's fixed point theorem. *Duke Mathematical Journal*, 8(3):457–459, 1941. 239

[42] R. Koch. *Routing Games over Time*. PhD thesis, TU Berlin, 2012. This is what you are reading. 7

[43] R. Koch, E. Nasrabadi, and M. Skutella. Continuous and discrete flows over time. *Mathematical Methods of Operations Research*, 73(3):301–337, 2011. 111, 114, 193

[44] R. Koch and M. Skutella. Nash equilibria and the price of anarchy for flows over time. *Theory of Computing Systems*, 49:71–97, 2011. 7, 203, 275

[45] E. Köhler and M. Skutella. Flows over time with load-dependent transit times. *SIAM Journal on Optimization*, 15(4):1185–1202, 2005. 111

[46] B. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*. Springer Publishing Company, Incorporated, 4th edition, 2010. 4

[47] E. Koutsoupias and C. H. Papadimitriou. Worst-case equilibria. In *Proceedings af the 16th annual symposium on theoretical aspects of computer science*, pages 404–413. Springer, 1999. 7

[48] W. H. K. Lam and H.-J. Huang. Dynamic user optimal traffic assignment model for many to one travel demand. *Transportation Research Part B: Methodological*, 29(4):243–259, 1995. 111

[49] H. Lebesgue. *Intégrale, Longueur, Aire*. PhD thesis, Université de Paris, 1902. 20

[50] H. Lebesgue. Intégrale, longueur, aire. *Annali di Matematica Pura ed Applicata (1898 - 1922)*, 7(1):231–359, 1902. Nancy. 45

[51] M. J. Lighthill and G. B. Whitham. On kinematic waves. ii. a theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 229(1178):317–345, 1955. 5

[52] W. H. Lin and H. K. Lo. Are the objective and solutions of dynamic user-equilibrium models always consistent? *Transportation Research Part A: Policy and Practice*, 34(2):137–144, 2000. 137

[53] H. Mahmassani and R. Herman. Dynamic user equilibrium departure time and route choice on idealized traffic arterials. *Transportation Science*, 18(4):362–384, 1984. 137

[54] H. S. Mahmassani, H. A. Sbayti, and X. Zhou. *DYNASMART-P Version 1.0 User's Guide*. Maryland Transportation Initiative, College Park, Maryland, 2004. 137

[55] D. K. Merchant and G. L. Nemhauser. A model and an algorithm for the dynamic traffic assignment problems. *Transportation Science*, 12:183–199, 1978. 5, 111

[56] E. Minieka. Maximal, lexicographic, and dynamic network flows. *Operations Research*, 21(2):517–527, 1973. 46, 192

[57] P. R. de Montmort. *Essay d'analyse sur les jeux de hazard*. Paris, 2nd edition, 1713. 6

[58] R. Mounce. Convergence in a continuous dynamic queueing model for traffic networks. *Transportation Research Part B: Methodological*, 40(9):779–791, 2006. 274

[59] R. Mounce. Convergence to equilibrium in dynamic traffic networks when route cost is decay monotone. *Transportation Science*, 41(3):409–414, 2007. 274

[60] R. Mounce. Existence of equilibrium in a continuous dynamic queueing model for traffic networks with responsive signal control. In W. H. K. Lam, S. C. Wong, and H. K. Lo, editors, *Transportation and Traffic Theory 2009: Golden Jubilee*, pages 327–344. Springer, 2009. 274

[61] J.-S. Mun. Traffic performance models for dynamic traffic assignment: An assessment of existing models. *Transport Reviews*, 27(2):231–249, 2007. 5, 111

[62] J. Nash. Non-cooperative games. *The Annals of Mathematics*, 54(2):286–295, 1951. 4, 6, 45, 277

[63] J. von Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928. 6

[64] N. Noam Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007. 6

[65] A. Orda and R. Rom. Shortest-path and minimum-delay algorithms in networks with time-dependent edge-length. *Journal of the ACM*, 37:607–625, 1990. 97, 113

[66] M. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press, 1994. 6

[67] A. de Palma, M. Ben-Akiva, C. Lefèvre, and N. Litinas. Stochastic equilibrium model of peak period traffic congestion. *Transportation Science*, 17(4):430–453, 1983. 137

[68] C. H. Papadimitriou. Algorithms, games, and the internet. In *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, STOC '01, pages 749–753, New York, NY, USA, 2001. ACM Press. 7

[69] S. Peeta and A. K. Ziliaskopoulos. Foundations of dynamic traffic assignment: The past, the present and the future. *Networks and Spatial Economics*, 1(3–4):233–265, 2001. 7, 137

[70] Britta Peis, Martin Skutella, and Andreas Wiese. Packet routing: Complexity and algorithms. In E. Bampis and K. Jansen, editors, *Approximation and Online Algorithms*, volume 5893 of *Lecture Notes in Computer Science*, pages 217–228. Springer, 2010. 192

[71] A. B. Philpott. Continuous-time flows in networks. *Mathematics of Operations Research*, 15(4):640–661, 1990. 5, 46, 114, 193

[72] J.-C. Picard and M. Queyranne. On the structure of all minimum cuts in a network and applications. *Mathematical Programming Studies*, 13:8–16, 1980. 199, 200

[73] A. C. Pigou. *The Economics of Welfare*. 1920. 7, 45, 137

[74] B. Ran, D. E. Boyce, and L J. Leblanc. A new class of instantaneous dynamic user-optimal traffic assignment models. *Operations Research*, 41(1):192–202, 1993. 137

[75] B. Ran, R. W. Hall, and D. E. Boyce. A link-based variational inequality model for dynamic departure time/route choice. *Transportation Research Part B: Methodological*, 30(1):31–46, 1996. 137

[76] P. I. Richards. Shock waves on the highway. *Operations Research*, 4(1):42–51, 1956. 5

[77] T. Roughgarden and É. Tardos. How bad is selfish routing? *Journal of the ACM*, 49(2):236–259, 2002. 7, 45, 277

[78] H. L. Royden and P. M. Fitzpatrick. *Real Analysis*. Prentice Hall, 4th edition, 2009. 20, 45

[79] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, New York, 3rd edition, 1976. 20, 45

[80] A. Schrijver. *Combinatorial Optimization*. Springer, 2003. 4

[81] N. Shioji and W. Takahashi. Strong convergence of approximated sequences for nonexpansive mappings in banach spaces. *Proceedings of the American Mathematical Society*, 125:3641–3645, 1997. 245

[82] M. Skutella. An introduction to network flows over time. In William Cook, László Lovász, and Jens Vygen, editors, *Research Trends in Combinatorial Optimization*, pages 451–482. Springer, 2009. 5

[83] M. J. Smith. The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck. *Transportation Science*, 18(4):385–394, 1984. 7, 274

[84] M. J. Smith. A new dynamic traffic model and the existence and calculation of dynamic user equilibria on congested capacity-constrained road networks. *Transportation Research Part B: Methodological*, 27(1):49–63, 1993. 137

[85] R. A. Sweet. A recursive relation for the determinant of a pentadiagonal matrix. *Communications of the ACM*, 12(6):330–332, 1969. 192

[86] A. Symvonis. Routing on trees. *Information Processing Letters*, 57(4):215–223, 1996. 6, 192

[87] W. Y. Szeto and H. K. Lo. A cell-based simultaneous route and departure time choice model with elastic demand. *Transportation Research Part B: Methodological*, 38(7):593–612, 2004. 137

[88] N. B. Taylor. The contram dynamic traffic assignment model. *Networks and Spatial Economics*, 3(3):297–322, 2003. 137

[89] W. S. Vickrey. Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260, 1969. 4, 6, 137, 274, 278

[90] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1:325–378, 1952. 4, 7, 45

[91] W. L. Wilkinson. An algorithm for universal maximal dynamic flows in a network. *Operations Research*, 19(7):1602–1612, 1971. 46, 192

[92] Y. W. Xu, J. H. Wu, M. Florian, P. Marcotte, and D. L. Zhu. Advances in the continuous dynamic network loading problem. *Transportation Science*, 33(4):341–353, 1999. 111, 112, 113

[93] S. Yagar. Dynamic traffic assignment by individual path minimization and queuing. *Transportation Research*, 5(3):179–196, 1971. 7, 137

[94] D. Zhu and P. Marcotte. On the existence of solutions to the dynamic user equilibrium problem. *Transportation Science*, 34(4):402–414, 2000. 138