

**Verspätungsfortpflanzung in Verkehrsnetzen -
Modellierung und Berechnung mit abgeschlossenen
Familien von Wahrscheinlichkeitsverteilungen**

D i s s e r t a t i o n

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt von
Dipl.-Math. Fabian Kirchoff
aus Bad Homburg

genehmigt von der Fakultät für
Mathematik / Informatik und Maschinenbau
der Technischen Universität Clausthal

Tag der mündlichen Prüfung
14. April 2015

Inhaltsverzeichnis

1. Einleitung	1
2. Einordnung in existierende Arbeiten	5
2.1. Verspätungsmodellierung in Verkehrsnetzen	5
2.1.1. Verspätungen als stetige Zufallsvariablen	5
2.1.2. Verspätungen als diskrete Zufallsvariablen	7
2.1.3. Das Max-Plus Modell	8
2.2. Anwendungsbereiche	11
2.2.1. Analytische Verspätungshochrechnung	11
2.2.2. Fahrplanoptimierung	11
2.2.3. Stochastische Simulation	11
3. Modellierung	13
3.1. Streckennetz	13
3.2. Ereignis-Aktivitätsnetzwerk	15
3.3. Verspätungsarten	19
3.4. Der Anschlussgraph	20
4. Empirische Analyse von Verspätungsdaten	23
4.1. Die Struktur der Daten	24
4.2. Extrahierung der benötigten Informationen	25
4.2.1. Mindestfahr- und Mindesthaltezeiten	25
4.2.2. Quellverspätungen	26
4.2.3. Ergebnis der Datenanalyse	27
5. Approximation von Phasentypverteilungen	31
5.1. Grundlagen	31
5.2. Literatur	38
5.3. Beispiele	41
5.3.1. Hyper-Exponentialverteilung	41
5.3.2. EM-Algorithmus für Hyper-Erlangverteilungen	44
5.4. Hyper-Erlangverteilung mit Gewichtung der Null	51
6. Bestimmung der Verspätungsverteilungen	55
6.1. Theta-Exponentialpolynome	56

6.2.	Einbettung der Hyper-Erlangverteilungen	59
6.3.	Grundlegende Operationen	59
6.3.1.	Excess Beyond	60
6.3.2.	Maximum	61
6.3.3.	Addition	61
6.4.	Anschlussicherung	62
6.4.1.	Einfache Anschlussicherung	63
6.4.2.	Vorausschauende Anschlussicherung	65
6.5.	Zugfolgerestriktionen	70
6.5.1.	Theoretische Umsetzung	70
6.5.2.	Berechnung - Subtraktion	74
6.6.	Komplexitätsreduzierung	76
6.6.1.	Metriken für Verteilungsfunktionen	77
6.6.2.	Approximation der Gestalt einer Verteilungsfunktion	80
6.6.3.	Verwendetes Verfahren	81
6.6.4.	Evaluation des Verfahrens	84
6.7.	Fehlerfortpflanzung	89
6.7.1.	Excess-Beyond	91
6.7.2.	Faltung	91
6.7.3.	Produkt	92
6.8.	Topologische Sortierung	92
6.9.	Stochastische Abhängigkeiten	93
7.	Verspätungfortpflanzung in kreisartigen Strukturen	99
7.1.	Modellierung	99
7.1.1.	Kreisartige Strukturen	99
7.1.2.	Verspätungfortpflanzung	104
7.2.	Verteilungskonvergenz	114
7.2.1.	Bedingungen für Verteilungskonvergenz	115
7.2.2.	Einfluss unterschiedlicher Sortierungen	120
7.3.	Endlichkeit der Verspätungen	123
7.3.1.	Übertragbare Ergebnisse aus der Warteschlangentheorie	123
7.3.2.	Stabilität in einem Kreis der Kreisvereinigung	126
7.3.3.	Stabilität in der Kreisvereinigung	135
7.4.	Pseudo-topologische Sortierung	138
7.5.	Empirische Resultate - Konvergenzgeschwindigkeit	138
8.	Ereignisgesteuerte stochastische Simulation	145
8.1.	Elemente der Simulation	147
8.2.	Der Simulationsalgorithmus	151
9.	Anwendung der Methoden im Rahmen der Optimierung	155

9.1. Generierung der Reisendenströme	155
9.2. Bewertung eines Fahrplans	156
9.2.1. Mittlere fahrplanmäßige Reisezeit	157
9.2.2. Mittlere Ankunftsverspätung	157
9.3. Fahrplanoptimierung in HiTT	158
10. Zusammenfassung	161
 Anhang	
A. Bezeichnungen	165
B. Beweise	167
B.1. Loynes	167
C. Algorithmen	171
C.1. Ereignisgesteuerte stochastische Simulation	171
C.2. Komplexitätsreduzierung	178
C.3. Optimierung	179
C.4. Verspätungsfortpflanzung	184
Literaturverzeichnis	185

1. Einleitung

Verspätungen werden sowohl von den Kunden als auch vom Dienstleister negativ bewertet. Bahnreisende nehmen Verspätungen als Abweichung vom Fahrplan wahr. Ist ein Fahrplan nicht verlässlich, so verringert dies die Akzeptanz der Bahn als Verkehrsmittel. Kunden werden nach Alternativen suchen. Aber auch für den Dienstleister verursachen Verspätungen Probleme. Neben unzufriedenen Kunden mindern Verspätungen die Produktivität: Züge und Personal können erst verspätet wieder eingesetzt werden, in Bahnhöfen werden Gleise blockiert. Eine Möglichkeit, den Folgen von Verspätungen entgegen zu wirken, wäre, in Personal, Fuhrpark oder Infrastruktur zu investieren. Dies wird in der Regel relativ kostspielig sein.

Zusätzlich bietet sich daher an, Fahrpläne zu entwickeln, die sich in der Realität als robust gegen mögliche Störungen im Betriebsablauf erweisen. Anschaulich erachten wir einen Fahrplan als robust, wenn sich das System trotz dieser Störungen weitestgehend fahrplangetreu verhält. Dies ist möglich, wenn der Fahrplan entsprechend großzügige Zeitpuffer enthält, die der technisch möglichen Mindestzeit der verschiedenen Vorgänge des Betriebsablaufs aufgeschlagen werden. Als Störungen werden in dieser Arbeit nur kleinere typische Verzögerungen betrachtet. Dies kann beispielsweise eine nicht öffnende Tür eines Zuges im Bahnhof sein. Es kann aber auch eine verlangsamte Fahrt wegen Tieren im Gleisbereich sein. Explizit nicht gemeint sind größere, katastrophale Verzögerungen bzw. Ereignisse wie ein Defekt des Triebwagens oder gar Kollisionen von Zügen. Solche Verzögerungen wird kein Fahrplan kompensieren können.

Es geht aber auch nicht nur darum, die erwähnten kleineren Verzögerungen zu kompensieren. Diese für sich allein stellen nicht das Hauptproblem dar. Es ist deren Fortpflanzung und Akkumulierung im Bahnnetz, welche zu nicht mehr kontrollierbaren Verspätungen führen können. Diese problematischen Verspätungen treten dann in der Regel an ganz anderer Stelle auf als ihre ursächlichen Verzögerungen im Betriebsablauf. Ein guter, d.h. robuster, Fahrplan enthält daher nicht nur eine angemessene Gesamtmenge zeitlichen Puffers, er weist auch den jeweiligen Stellen im Bahnnetz angemessene Anteile der Gesamtmenge zu, sodass eine kritische Fortpflanzung von Verspätungen möglichst vermieden werden kann.

Bisher haben wir von möglichst geringen Verspätungen als ein Kriterium für gute Fahrpläne gesprochen. Zum einen muss präzisiert werden, auf welcher Ebene die Verspätungen gemessen werden. Dies können direkt die Verspätungen der Züge sein, es können aber auch die Verspätungen aus Sicht der Kunden, also die Verspätungen bei Ankunft in der Zielstation der Kunden, betrachtet werden. Die Unterscheidung

ist wichtig. Ein insgesamt stark verspätetes System muss für die Reisenden keine größeren Folgen haben, solange sie zumindest ihre Anschlüsse erreichen. Der Dienstleister leidet jedoch, wie oben angesprochen, unter einer Produktivitätsabnahme.

Grundsätzlich kann geringe Verspätung aber nicht das einzige Kriterium sein. Dies führte zu Fahrplänen mit Zeitpuffern, die jede Störung im Betriebsablauf kompensieren können. Das bedeutete auf der einen Seite unverhältnismäßig große Reisezeiten für die Kunden. Auf der anderen Seite wäre auch hier die Produktivität des Dienstleisters beeinträchtigt. Es benötigt also ein Kriterium, welches dem entgegenwirkt. Ein mögliches Kriterium ist hier die, möglichst geringe, fahrplanmäßige Reisezeit der Kunden. Mittels der genannten beiden Kriterien können Fahrpläne optimiert werden.

In der Arbeitsgruppe „Stochastische Optimierung“ von Prof. Dr. Kolonko wird unter anderem am Software-Projekt HiTT gearbeitet. Im Rahmen dieses Projektes können Fahrpläne bezüglich verschiedener Kriterien optimiert werden. Der Einfluss von stochastischen Verspätungen sowie insbesondere deren Fortpflanzung konnte bereits vor Beginn dieser Arbeit per ereignisgesteuerter stochastischer Simulation untersucht werden. Die Simulation wird, abhängig von den gewählten Optimierungskriterien, für die Bewertung eines Fahrplans benötigt und stellt dabei einen wesentlichen, zeitkritischen Schritt dar. Alternativ sollte daher in dieser Arbeit nach analytischen Methoden gesucht werden, welche die (approximative) Bestimmung der entsprechenden Verspätungsverteilungsfunktionen sowie der mittleren, d.h. erwarteten, Verspätungen ermöglichen. Diese Herangehensweise erwies sich letztlich als deutlich schneller. So ist die Fahrplanoptimierung unter Berücksichtigung der Verspätungen nun auch für größere Probleminstanzen einsetzbar.

Die Idee, Verspätungen auf Ebene der Verteilungsfunktionen fortzupflanzen ist nicht neu. Diese Arbeit greift einen Ansatz aus dem Flugverkehr auf und führt ihn weiter fort. Auch im Anwendungsbereich des Bahnverkehrs gibt es vergleichbare Arbeiten. Die wesentliche Erweiterung und zugleich den Schwerpunkt dieser Arbeit bildet die Modellierung und Berechnung der Verspätungsfortpflanzung in Kreisen des betrachteten Ereignis-Aktivitätsnetzwerkes. Die Ereignisse repräsentieren in der vorliegenden Arbeit die Ankünfte und Abfahrten von Linien in ihren Stationen. Die Bestimmung der Verspätungsverteilungen all dieser Ereignisse setzt grundsätzlich eine topologische Sortierung voraus. So kann die Verspätungsverteilung eines Ereignisses erst dann bestimmt werden, wenn zuvor die Verspätungsverteilungen aller Vorgängerereignisse bestimmt wurden. Dieses Vorgehen ist genau dann möglich, wenn keine Kreise im Netzwerk existieren. In diesem Fall ist die jeweilige Verspätungsverteilungsfunktion bereits nach einmaliger Berechnung endgültig festgelegt.

Für Fälle, in denen das Ereignis-Aktivitätsnetzwerk nicht kreisfrei ist, wird in der vorliegenden Arbeit ein neues iteratives Verfahren zur Approximation der Verspätungsverteilungen in solchen Kreisen vorgestellt. Für Aussagen zur Konvergenz des Verfahrens werden Resultate aus der Warteschlangentheorie herangezogen. Es

wird untersucht, unter welchen Voraussetzungen die Verspätungen in Verteilung konvergieren, d.h. insbesondere wann die Verspätungen fast sicher endlich sind. Außerdem wird die Eindeutigkeit der Grenzverteilungen diskutiert. Dieser Themenkomplex wird in den Kapiteln 6 und 7 behandelt.

Zu Beginn dieser Arbeit war die Vorstellung, dass wir zu einem bestimmten Bahnnetz Daten über die im realen Betriebsablauf auftretenden Störungen erhalten. Mittels einer geeigneten Verteilungsfamilie sollten die zugehörigen empirischen Verteilungsfunktionen approximiert und dann so miteinander verknüpft werden, dass die reale Verspätungsfortpflanzung nachgebildet wird. Die Akquise aber auch die anschließende Analyse entsprechender Daten gestaltete sich jedoch schwierig. Das untersuchte Datenmaterial eignete sich letztlich nicht, um Rückschlüsse auf die Situation in anderen, größeren Netzen zu ziehen. Es unterstützt aber die spätere Wahl der Verteilungsfamilie. Daher beschäftigt sich Kapitel 4 mit der Analyse des Datenmaterials. In Kapitel 5 werden die Verteilungsfamilie und das Verfahren vorgestellt, die für die Approximation der empirischen Verteilungsfunktionen verwendet werden.

Zur Validierung der analytischen Methoden auf Ebene der Verteilungsfunktionen wurde im Rahmen dieser Arbeit eine neue ereignisgesteuerte stochastische Simulation entwickelt. Ziel war es, die Laufzeit der bisherigen Simulation zu reduzieren. Die Grundlage bildete eine studentische Arbeit, deren Kernidee eine besondere Repräsentation der Umsteigebeziehungen zwischen verschiedenen Linien darstellt. Hierauf aufbauend wurde die Simulation soweit modifiziert, dass sie, bei möglichst geringer Laufzeit, belastbare Ergebnisse für die Fahrplanbewertung liefert. Werden sowohl die Simulation als auch die Berechnung der Verteilungsfunktionen in HiTT für die Fahrplanbewertung verwendet, so stimmen die Ergebnisse (abhängig von den Vorgaben für die in der Simulation verwendeten Konfidenzintervalle) gut überein.

Die Optimierung der Fahrpläne findet in HiTT mittels eines genetischen Algorithmus statt. Die Weiterentwicklung dieses Verfahrens ist nicht Gegenstand dieser Arbeit. In Kapitel 9 wird gezeigt, wie die zuvor bestimmten Verspätungsverteilungen in HiTT verwendet werden, bevor in Kapitel 10 die Ergebnisse der Arbeit zusammengefasst werden.

Die vorliegende Arbeit fand im Rahmen des Projekts *„Strukturuntersuchungen zur Entstehung und Fortpflanzung von Verspätungen in Verkehrsnetzen - Modellierung, Simulation und Optimierung eines stochastischen Netzwerks“* im Simulationswissenschaftlichen Zentrum Clausthal-Göttingen statt. Dies ist ein gemeinsames Projekt mit der Arbeitsgruppe „Optimierung“ von Prof. Dr. Schöbel an der Georg-August-Universität Göttingen. Im Rahmen dieser Zusammenarbeit wurden unter anderem die in den beiden Arbeitsgruppen existierenden Modelle verglichen. Die mit den jeweiligen Modellen hinsichtlich der Verspätungsfortpflanzung erhaltenen empirischen Ergebnisse stimmten dabei in hohem Maße überein.

2. Einordnung in existierende Arbeiten

Übersicht

2.1. Verspätungsmodellierung in Verkehrsnetzen	5
2.1.1. Verspätungen als stetige Zufallsvariablen	5
2.1.2. Verspätungen als diskrete Zufallsvariablen	7
2.1.3. Das Max-Plus Modell	8
2.2. Anwendungsbereiche	11
2.2.1. Analytische Verspätungshochrechnung	11
2.2.2. Fahrplanoptimierung	11
2.2.3. Stochastische Simulation	11

In diesem Kapitel soll ein wesentlicher Teil der Literatur diskutiert werden, die aus Bereichen stammt, welche auch Gegenstand der vorliegenden Arbeit sind. Der erste Abschnitt beschäftigt sich mit der Problemstellung, wie wir die im realen Betrieb in unterschiedlicher Ausprägung auftretenden Verspätungen in einem theoretischen Modell abbilden können. Der zweite Abschnitt zeigt dann, in welchen Bereichen bzw. auch bezüglich welcher Zielsetzung diese Modellierungen Anwendung finden.

2.1. Verspätungsmodellierung in Verkehrsnetzen

Die Modellierung von Verspätungen, nicht nur im Bahnverkehr, wurde in der Literatur bereits intensiv behandelt. Die verschiedenen Modellierungsansätze unterscheiden sich grundsätzlich darin, ob Verspätungen als stochastisch oder deterministisch angesehen werden. Im ersten Fall werden die Verspätungen als Zufallsvariablen modelliert und in der Regel einer bestimmten Verteilungsfamilie zugeordnet. So wird versucht, die in der Realität auftretenden Folgen der Verspätungen nachzubilden. Im zweiten Fall werden für die Verspätungen nur bestimmte Ausprägungen, sogenannte *Szenarien*, betrachtet. Für jedes Szenario lassen sich mögliche Folgen untersuchen.

2.1.1. Verspätungen als stetige Zufallsvariablen

Modelliert man Verspätungen als Zufallsvariablen, so kann man diese mittels Monte-Carlo Simulation reproduzieren und mit möglichen Pufferzeiten oder auch anderen (simulierten) Verspätungen verrechnen. Eine alternative Herangehensweise

stellen analytische Methoden zur Verknüpfung der zugehörigen Verteilungsfunktionen dar. In beiden Fällen ist die grundlegende Aufgabe, eine Verteilungsfamilie zu finden, welche die Verspätungen hinreichend gut, d.h. realistisch, charakterisieren kann. Im Falle der Verrechnung von Verspätungen auf Ebene der Verteilungsfunktionen, wird darüber hinaus die Abgeschlossenheit der Verteilungsfamilie unter den benötigten Operationen gefordert (s. Kapitel 6). Im Folgenden werden einige der in der Literatur zu findenden Verteilungsfamilien vorgestellt.

SCHWANHÄUSSER identifiziert in [Sch74] modifizierte Exponentialverteilungen als mögliche Verteilungsfamilie für Verspätungen. Für die Verteilungsfunktion einer Verspätung Y gilt dann für alle $t \in \mathbb{R}$

$$F_Y(t) = \left(1 - a \cdot e^{-\lambda t}\right) \cdot \mathbf{1}_{[0, \infty)}(t) .$$

Dabei lassen sich die Parameter a und λ dahingehend interpretieren, dass a den Anteil verspäteter Züge und $\frac{1}{\lambda}$ gerade die mittlere Verspätung dieser Züge repräsentiert. Allerdings kann deren Eignung nur für bestimmte Verspätungsarten nachgewiesen werden (s. [WN04, Yua06] sowie Abschnitt 4.2.3). Der Nachweis geschieht in der Regel mit empirischen Methoden. In [EF02] gelingt ein analytischer Nachweis unter Zuhilfenahme von Ergebnissen aus der Warteschlangentheorie.

Eine weitere Verteilungsfamilie, die in der Literatur (s. [HS92, Yua06, MM07]) für die Repräsentation von Verspätungsverteilungen verwendet wird, ist die Familie der Weibullverteilungen. Die Verteilungsfunktion einer Weibull-verteilten Zufallsvariablen Y ist für alle $t \in \mathbb{R}$ gegeben durch

$$F_Y(t) = \left(1 - e^{-(\lambda t)^k}\right) \cdot \mathbf{1}_{[0, \infty)}(t) ,$$

mit $\lambda, k \in \mathbb{R}_{>0}$. Sie erweist sich hinsichtlich der Eignung für die verschiedenen Verspätungsarten als flexibler. Wie auch im Falle der oben beschriebenen modifizierten Exponentialverteilungen, ist es bei Verwendung von Weibullverteilungen nicht möglich, die Ergebnisse angewandter Verspätungsoperationen wieder als Element der Verteilungsfamilie darzustellen. Die beiden Verteilungsfamilien sind also nicht *abgeschlossen* gegen die Verspätungsoperationen. Eine Möglichkeit ist hier, mittels numerischer Verfahren eine Approximation der Lösung zu erzeugen (s. [Mü90]).

Die Verwendung von Phasentypverteilungen wird unter anderem in [MM07] empfohlen. Ein Argument hierfür ist, dass die Familie der Phasentypverteilungen empirische Verspätungsverteilungen (auf den nicht-negativen reellen Zahlen) beliebig genau approximieren können. Des Weiteren wird auf die Abgeschlossenheit von Phasentypverteilungen gegen wesentliche Verspätungsoperationen hingewiesen. Letztlich erweist sich die Familie der allgemeinen Phasentypverteilungen jedoch als nicht flexibel genug, um sämtliche für die Verspätungsfortpflanzung in Frage kommenden Operationen anwenden zu können.

Die Repräsentation von Verspätungsverteilungen mittels sogenannter *Theta-Exponentialpolynome* wird unter anderem in [Fuh07] vorgeschlagen. Verteilungen, für die eine solche Repräsentation möglich ist, nennen wir im Folgenden auch *Theta-Exponentialverteilungen*. In [Fuh07] wird die Verspätungsfortpflanzung im Flugverkehr untersucht. Die Herangehensweise ist vergleichbar mit der in der vorliegenden Arbeit. Es wird eine bestimmte azyklische Phasentypverteilung für die Approximation von Verspätungsverteilungen verwendet. Deren Einbettung in die Familie der Theta-Exponentialpolynome ermöglicht, dank der dort gegebenen Abgeschlossenheit gegen benötigte Verspätungsoperationen, die Fortpflanzung der Verspätungen bzw. Verrechnung der Verteilungsfunktionen. Ein wesentlicher Unterschied zur vorliegenden Arbeit besteht in der speziellen Wahl der azyklischen Phasentypverteilung.

BÜKER und WENDLER stellen in [BW] eine Familie von Verteilungsfunktionen vor, mit denen es Ihnen möglich ist, Verspätungsfortpflanzungen in Bahnnetzen hinreichend genau und schnell zu berechnen. Die Verteilungsfunktion einer Zufallsvariablen Y wird gemäß

$$F_Y(t) = \sum_{j=1}^s \left(c_j - \sum_{i=1}^{n_j} a_{ij} \cdot e^{-\lambda_{ij} \cdot t} \right) \cdot \mathbf{1}_{I_j}(t)$$

repräsentiert, mit paarweise disjunkten Intervallen $I_j \subset \mathbb{R}$ und $\bigcup_{j=1}^s I_j = \mathbb{R}$. Diese Verteilungsfamilie ist abgeschlossen gegen alle benötigten Verspätungsoperationen. Darüber hinaus ist sie insgesamt flexibler als die Familie der Theta-Exponentialverteilungen. Dies betrifft im Wesentlichen die Bildung der Verteilungsfunktion von Differenzen zweier Zufallsvariablen, welche bei der Betrachtung sogenannter *Zugfolgerestriktionen* benötigt werden. Diese wurden in der vorliegenden Arbeit noch nicht in der Modellierung berücksichtigt. In Kapitel 6 schlagen wir trotzdem bereits eine mögliche alternative Bestimmung der resultierenden Verteilungsfunktion vor, welche auch für diesen Fall die Verwendung von Theta-Exponentialpolynomen ermöglicht.

2.1.2. Verspätungen als diskrete Zufallsvariablen

Ein anderer Weg ist es, die Verspätungen als diskrete Zufallsvariablen zu modellieren. Diese Herangehensweise verwenden BERGER ET AL. in [BGMHO11]. Die Verspätungsfortpflanzung, d.h. die Anwendung benötigter Operationen auf die entsprechenden Verteilungsfunktionen, gestaltet sich im diskreten Fall einfacher als im stetigen Fall. So spielt u.a. das Problem der Komplexitätsreduzierung (s. Kapitel 6) keine Rolle. BÜKER vergleicht in seiner Dissertation (s. [Bü10]) seinen Ansatz mit der Verspätungsmodellierung durch diskrete Zufallsvariablen. Insbesondere bezüglich der Genauigkeit schneidet seine auf stetigen Zufallsvariablen basierende Modellierung besser ab.

Auch SCHOEBEL und GOERIGK modellieren Quellverspätungen gewissermaßen als diskrete Verteilung (s. [Sch07], [GS11]). Sie gehen von einer diskreten Menge von Verspätungswerten aus, welche jene Quellverspätungen annehmen können, die den jeweiligen Betriebsabläufen zugeordnet wurden. Wurde jedem dieser Betriebsabläufe also ein Verspätungswert zugewiesen, so entspricht die Menge dieser Verspätungswerte einem sogenannten *Szenario*, dessen Auswirkungen untersucht werden soll. Die Auswahl eines Verspätungsszenarios unterliegt bestimmten Regeln. Werden sie beispielsweise zufällig gewählt, so entspricht dies gerade einer diskreten Gleichverteilung bezüglich der Menge aller Szenarien.

2.1.3. Das Max-Plus Modell

Einen hinsichtlich der Verspätungsförderung gänzlich verschiedenen Ansatz enthalten sogenannte Max-Plus (lineare) Modelle. Sie werden allgemein bei der Planung diskreter Ereignisse verwendet (*engl.* Discrete Event Scheduling, DES). Einen Überblick erhält man zum Beispiel in [BCOQ92]. Im Folgenden sollen die Grundlagen dieser Modellierung erläutert werden (s. [Gov07]).

Definition 2.1.1 (Max-Plus Algebra).

Es sei $\varepsilon := -\infty$ und $\mathbb{R}_\varepsilon := \mathbb{R} \cup \{\varepsilon\}$. Für $x, y \in \mathbb{R}_\varepsilon$ werden die additive bzw. multiplikative Verknüpfung

$$x \oplus y := \max\{x, y\} \quad \text{bzw.} \quad x \otimes y := x + y \quad (2.1)$$

definiert. Außerdem gelte für Matrizen $A, B \in \mathbb{R}_\varepsilon^{n \times m}$

$$[A \oplus B]_{ij} = a_{ij} \oplus b_{ij} = \max\{a_{ij}, b_{ij}\}$$

bzw. Matrizen $A \in \mathbb{R}_\varepsilon^{n \times k}$ und $B \in \mathbb{R}_\varepsilon^{k \times m}$

$$[A \otimes B]_{ij} = \bigoplus_{k=1}^n a_{ik} \otimes b_{kj} = \max\{a_{ik} + b_{kj} \mid 1 \leq k \leq n\}.$$

Dann wird $(\mathbb{R}_\varepsilon, \oplus, \otimes)$ eine Max-Plus Algebra genannt. Dabei ist ε das additive und 0 das multiplikative neutrale Element.

Hierauf aufbauend betrachtet man Folgen von Matrizen $(A_k)_{k \in \mathbb{N}}$ und Folgenvektoren $(x_k)_{k \in \mathbb{N}}$,

$$x_k = (x_k(1), x_k(2), \dots, x_k(n)) \in \mathbb{R}_\varepsilon^n$$

für alle $k \in \mathbb{N}$, welche der rekursiven Relation $x_{k+1} = A_k \otimes x_k$ genügen. Unter den Verknüpfungen (2.1) gehen die x_k also durch lineare Abbildungen rekursiv auseinander hervor. Eine wichtige Kenngröße der $(x_k)_{k \in \mathbb{N}}$ ist die *asymptotische Wachstumsrate*, definiert durch

$$\lim_{k \rightarrow \infty} \frac{x_k(i)}{k}, \quad (2.2)$$

$1 \leq i \leq n$, falls diese Grenzwerte existieren. Stimmen in dem Fall die Werte für alle $1 \leq i \leq n$ überein, so bezeichnet man diesen Wert in der Regel mit $\lambda = \lambda(A_k)$ und spricht vom *Lyapunov-Exponenten* der Folge $(A_k)_{k \in \mathbb{N}}$. Betrachtet man deterministische Systeme, d.h. $A_k = A \in \mathbb{R}_\varepsilon^{n \times n}$ für alle $k \in \mathbb{N}$, so entspricht λ gerade dem Eigenwert der Matrix A (s.u.). Im Falle von Systemen mit stochastischen Matrizen A_k ist die exakte Bestimmung des Lyapunov-Exponenten in der Regel nicht möglich. Es gibt in der Literatur allerdings effiziente Approximationsverfahren (s. [GHM08]).

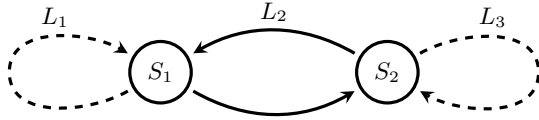
GOVERDE ET AL. haben, neben den bisher erwähnten, weitere Arbeiten veröffentlicht (s. [Gov98, GHM09]), in denen das Max-Plus Modell auf Verkehrsnetze angewendet wird. Der Zusammenhang soll anhand des Streckennetzes aus Abbildung 2.1 hergestellt werden. Wir nehmen an, dass für Linie L_1 und L_3 je ein Zug und für Linie L_2 zwei Züge zur Verfügung gestellt werden. Ein Zug der Linie L_2 bedient immer die Strecke von S_1 zu S_2 und der andere Zug die Rückrichtung. Die benötigten (deterministischen) Fahrzeiten seien f_1 für Linie L_1 auf der Streckenverbindung (S_1, S_1) , f_2 bzw. f_3 für Linie L_2 auf Streckenverbindung (S_1, S_2) bzw. (S_2, S_1) sowie f_4 für Linie L_3 auf der Streckenverbindung (S_2, S_2) .

Ziel ist es, für diese Linien einen periodischen Fahrplan zu finden. Wesentliche Voraussetzung ist immer, dass ein Ereignis erst dann beginnen kann, wenn alle dem Ereignis planmäßig vorangehenden Prozesse bereits abgeschlossen sind. Im Bahnverkehr hieße das beispielsweise, dass ein Zug erst dann abfahren kann, wenn alle Zubringer in der Abfahrtsstation angekommen sind. Daher lässt sich diese Modellierung auch mittels eines Petrinetzes veranschaulichen (s. Abbildung 2.1). Die beiden Transitionen repräsentieren die Abfahrten der Züge in den Stationen S_1 bzw. S_2 . Diese finden erst dann statt, wenn in den vorangehenden Stellen die Markierungen gesetzt sind. Die Fahrzeit der Züge lässt sich in den Stellen hinterlegen. Nach Aktivierung der zugehörigen Transition wird erst nach f_i Zeiteinheiten die Markierung in der Stelle gesetzt.

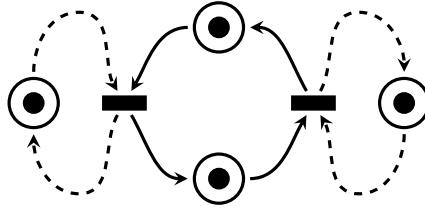
Im Folgenden steht $x_k(i)$ für die k -te Abfahrt, $k \in \mathbb{N}$, der Linie

- L_1 in Station S_1 , falls $i = 1$,
- L_2 in Station S_1 , falls $i = 2$,
- L_2 in Station S_2 , falls $i = 3$,
- L_3 in Station S_2 , falls $i = 4$.

Wir betrachten weder Mindestverweilzeiten noch Mindestzugfolgezeiten, obwohl dies grundsätzlich in einem Max-Plus Modell möglich wäre. In den Stationen existieren Umsteigebeziehungen zwischen den hier verkehrenden Linien. Die Abfahrt des Zuges der Linie L_1 hängt sowohl von der vorherigen Ankunft dieses Zuges als auch der des Zuges der Zubringerlinie L_2 ab. Es gilt $x_k(1) = \max\{x_{k-1}(1) + f_1, x_{k-1}(3) +$



(a) Beispielnetz mit Linien L_1, L_2, L_3 und Umsteigemöglichkeiten in den Stationen S_1, S_2



(b) Petrinetz

Abbildung 2.1.: Beispielnetz und zugehöriges Petrinetz

f_3 }. Die anderen beiden Abfahrtsereignisse spielen in diesem Fall keine Rolle. Bezüglich einer Max-Plus Algebra heißt das $x_k(1) = (x_{k-1}(1) \otimes f_1) \oplus (x_{k-1}(3) \otimes f_3)$. Insgesamt ergibt sich so

$$x_k = Ax_{k-1} = \begin{pmatrix} f_1 & \varepsilon & f_3 & \varepsilon \\ f_1 & \varepsilon & f_3 & \varepsilon \\ \varepsilon & f_2 & \varepsilon & f_4 \\ \varepsilon & f_2 & \varepsilon & f_4 \end{pmatrix} x_{k-1} \quad (2.3)$$

für alle $k \in \mathbb{N}$. Im Falle eines periodischen Fahrplans mit Takt τ gehen die x_k , mit x_0 als Startvektor der Abfahrtszeiten, aus der Beziehung $x_k = x_0 + k \cdot \tau e$ hervor. Dabei ist $e = (1, 1, \dots, 1)^T$. Im Max-Plus Modell suchen wir dementsprechend, bei gegebener Matrix A , gemäß (2.3) einen Vektor x_k und ein $\tau \in \mathbb{R}$ mit $x_{k+1} = \tau x_k = Ax_k$. Es ist dann x_k der Eigenvektor zum Eigenwert τ . In [Gov07] wird dies allgemeiner beschrieben. GOVERDE erläutert, dass der Eigenwert τ der minimalen Taktzeit entspricht. Es existiert mindestens ein Kreis im Netz, der sogenannte *kritische Kreis*, mit Zykluszeit τ . Im betrachteten Beispiel wurden deterministische Fahrzeiten verwendet. Möchte man stochastische Fahrzeiten verwenden, d.h. insbesondere Verspätungen zulassen, so führt dies zur Bestimmung des bereits oben erwähnten Lyapunov-Exponenten. Alternativ kann eine Monte-Carlo Simulation verwendet werden. Zu einem bestimmten Bahnverkehrsnetz werden dabei, durch Simulation der Fahrzeiten, verschiedene Max-Plus Modelle erzeugt (s. [BDS08]).

2.2. Anwendungsbereiche

Die verschiedenen vorgestellten Modellierungen besitzen Eigenschaften, die sich in Abhängigkeit vom jeweiligen Anwendungsbereich als vorteilhaft erweisen. Im Wesentlichen geht es immer darum, die Fortpflanzung von Verspätungen zu untersuchen und eine Aussage über die Größenordnung auftretender Verspätungen treffen zu können.

2.2.1. Analytische Verspätungshochrechnung

Der Ansatz von BÜKER in [Bü10] deckt sich in großen Teilen mit dem Ansatz dieser Arbeit. Ziel war es, Verspätungsfortpflanzung mittels geeigneter Operationen auf Ebene von stetigen Verspätungsverteilungen zu modellieren. Die Anwendung sollte jedoch, im Gegensatz zur Motivation dieser Arbeit, insbesondere in der Ermöglichung einer Echtzeitprognose von Ankunftsverspätungen sein. Hierfür wird auf die Genauigkeit der Ergebnisse besonderen Wert gelegt. Ebenfalls in diesen Anwendungsbereich fällt die Arbeit von BERGER ET AL. (s. [BGMHO11]).

FUHR modelliert in ihrer Dissertation die Verspätungsfortpflanzung im Flugverkehr. Sie verweist auf Ähnlichkeiten zur Situation im Bahnverkehr. Das Ziel war insbesondere, neben einer ereignisgesteuerten stochastischen Simulation, mittels analytischer Methoden eine Pünktlichkeitsbewertung von Flugplänen zu ermöglichen.

2.2.2. Fahrplanoptimierung

Die vorliegende Arbeit ist vergleichbar zum Modellierungsansatz der zuvor erwähnten Arbeiten. Bezüglich des Anwendungsbereichs ist er jedoch eher vergleichbar mit den Arbeiten von SCHOEBEL und GOERIGK in [Sch07] bzw. [GS11]. Die Verspätungsfortpflanzung dient hier als Werkzeug für die Bewertung von Fahrplänen hinsichtlich bestimmter Kriterien im Rahmen einer Optimierung von Fahrplänen. Der Unterschied liegt neben der Verspätungsmodellierung auch im Optimierungsverfahren. Während die vorliegende Arbeit ihre Anwendung im Rahmen einer Optimierung mittels eines genetischen Algorithmus findet, nutzen SCHOEBEL und GOERIGK lineare Programme für die Optimierung. Eine weitere, umfangreiche Arbeit steuern LIEBCHEN ET AL. in [LSS⁺07] bzw. [LSS⁺10] zu diesem Bereich bei.

2.2.3. Stochastische Simulation

Im Rahmen dieser Arbeit wurde, speziell zum Zwecke der Evaluation des auf Verteilungsebene operierenden Modells, eine ereignisgesteuerte stochastische Simulation entwickelt. Die Bewertung von Fahrplänen mittels Monte Carlo Simulation wurde zuvor bereits in der Literatur erprobt.

Ein erster Ansatz, auf den in der Literatur häufig verwiesen wird, ist die Dissertation [Noo96] von NOORDEEN. Analog zur vorliegenden Arbeit, verwendet er eine makroskopische Modellierung von Bahnnetzen. So wird nicht modelliert, welchen Weg Züge innerhalb von Stationen nutzen. Allerdings werden, im Gegensatz zur vorliegenden Arbeit, auch Kapazitätsbeschränkungen der Infrastruktur, beispielsweise die Anzahl der Gleise in Stationen, abgebildet. Statt Quellverspätungen als stochastische Größe zu betrachten, werden hier die Fahrzeiten von Zügen gemäß ihrer empirischen Verteilung erzeugt. Ziel war insbesondere die Möglichkeit zur Fortpflanzung von lokalen Verspätungen bzw. Störungen innerhalb eines realistischen, größeren Netzes. Die Ergebnisse der Arbeit von NOORDEEN finden ihre Anwendung im Softwareprojekt FASTA¹.

CAREY und CARVILLE stellen in [CC00] ein makroskopisches Simulationsmodell vor, mittels dessen sie Folgeverspätungen in größeren Stationen untersuchen. Simuliert werden dafür sogenannte exogene Verspätungen. Zu diesen Verspätungen zählen sowohl Quellverspätungen, wie sie in dieser Arbeit verwendet werden, als auch durch Folgeverspätungen (aus vorangegangenen Stationen) in der untersuchten (größeren) Station verursachte Verspätungen. Die Autoren betonen, dass sich ihr Modell für viele Zwecke verwenden lässt. So lässt sich beispielsweise auch die Güte von Heuristiken untersuchen, mit deren Hilfe versucht wird, die Stabilität von Fahrplänen in einem frühen Planungsstadium zu untersuchen (s. [Car99]). In [MB01] beschreiben MIDDELKOOP und BOUWMAN die Simulationsumgebung „Simone“. Hiermit ist es möglich, die Robustheit von Fahrplänen bezüglich aus Störungen resultierenden Folgeverspätungen zu simulieren und Verkehrsengpässe (engl. *bottleneck*) zu identifizieren. Die Stärke bzw. das Niveau der Störungen wird vor Start der Simulation festgelegt. Die Autoren betonen, dass die Simulation für realistische Bahnnetze verwendet werden kann.

TAKEUCHI ET AL. modellieren in [TTH07] die Interaktion von Zügen in einem Netzwerk mittels eines PERT Graphen. Sie führen ein neues Robustheitsmaß für Fahrpläne ein, welches die Kundenzufriedenheit berücksichtigt, und berechnen dieses simulativ. In [HMG⁺06] wird ein Simulationsmodell für den Nahverkehr in Kopenhagen vorgestellt. In [SBI12] wird ein Robustheitsmaß für Fahrpläne vorgestellt, welches sowohl analytisch als auch simulativ berechnet werden kann.

¹Swiss Federal Institute of Technology Lausanne EPFL
www.fasta.ch (18.04.2014)

3. Modellierung

Übersicht

3.1. Streckennetz	13
3.2. Ereignis-Aktivitätsnetzwerk	15
3.3. Verspätungsarten	19
3.4. Der Anschlussgraph	20

3.1. Streckennetz

Das Streckennetz stellt eine grobe Modellierung der Infrastruktur des *realen* Bahnnetzes dar.

Definition 3.1.1 (Streckennetz).

Das Streckennetz wird durch einen gerichteten Graphen $G = (\mathfrak{S}, \mathfrak{B})$ mit

- (i) der Menge \mathfrak{S} von **Stationen** $S \in \mathfrak{S}$ und
- (ii) der Menge der **Strecken** $\mathfrak{B} \subset \mathfrak{S} \times \mathfrak{S}$

repräsentiert. Strecken $(S, S') \in \mathfrak{B}$ verlaufen zwischen unmittelbar benachbarten Stationen $S, S' \in \mathfrak{S}$.

Die Information über die Nachfrage für Verbindungen mit Start in einer Station $S \in \mathfrak{S}$ und Ziel in $S' \in \mathfrak{S}$ ist durch eine Matrix gegeben.

Definition 3.1.2 (OD-Matrix).

Die Matrix

$$M = (M(S, S'))_{S, S' \in \mathfrak{S}} \tag{3.1}$$

heißt **OD-Matrix** (engl. Origin, Destination), wenn der Eintrag $M(S, S')$ gerade die Anzahl der Passagiere angibt, die während einer Zeiteinheit vom Start in Station $S \in \mathfrak{S}$ zur Zielstation $S' \in \mathfrak{S}$ fahren.

Definition 3.1.3 (Linien).

Es sei \mathfrak{L} die Menge der Linien. Eine **Linie** $L \in \mathfrak{L}$ ist ein Weg aus Strecken und besitzt einen **Takt** $\tau(L) \in \mathbb{N}$. Wir sprechen im Folgenden auch davon, dass Strecken $B \in \mathfrak{B}$ Element einer Linie $L \in \mathfrak{L}$ sind und schreiben in dem Falle dann $B \in L$.

Die Strecken einer Linie, sowie die darin implizit enthaltenen Stationen, bekommen einen eigenen Namen.

Definition 3.1.4 (Linienstrecke).

Eine **Linienstrecke** ist ein Tupel $(B, L) \in \mathfrak{B} \times \mathfrak{L}$. Die Menge aller Linienstrecken einer Linie $L \in \mathfrak{L}$ ist gegeben durch

$$\mathcal{B}_L = \{(B, L) \in \mathfrak{B} \times \mathfrak{L} \mid B \in L\} .$$

Definition 3.1.5 (Linienstation).

Eine **Linienstation** ist ein Tupel $(S, L) \in \mathfrak{S} \times \mathfrak{L}$. Die Menge aller Linienstationen einer Linie $L \in \mathfrak{L}$ ist gegeben durch

$$\mathcal{S}_L = \{(S, L) \in \mathfrak{S} \times \mathfrak{L} \mid (\exists B = (S, S') \in \mathfrak{B}) \vee (\exists B = (S', S) \in \mathfrak{B}) : B \in L\} .$$

Definition 3.1.6 (Restriktionen).

Für das Fahrgeschehen gibt es diverse Restriktionen. Dies sind

(i) die **Mindestfahrzeit**

$$d(S, S', L),$$

die Züge der Linie $L \in \mathfrak{L}$ auf der Strecke $B = (S, S') \in \mathfrak{B}$ mit $(B, L) \in \mathcal{B}_L$ benötigen,

(ii) die **Mindesthaltezeit**

$$h(S, L)$$

aller Züge der Linie $L \in \mathfrak{L}$ in Linienstation $(S, L) \in \mathcal{S}_L$,

(iii) die **Mindestumsteigezeit**

$$c(S, L, L'),$$

die ein Passagier in Station $S \in \mathfrak{S}$ beim Wechsel von Linie $L \in \mathfrak{L}$ zu Linie $L' \in \mathfrak{L}$ benötigt sowie

(iv) die **Mindestzugfolgezeiten**

$$head_{arr}(S, L, L') \quad \text{und} \quad head_{dep}(S, L, L')$$

für Züge der Linie $L \in \mathfrak{L}$ auf Züge der Linie $L' \in \mathfrak{L}$ bei Ankunft bzw. Abfahrt in Station $S \in \mathfrak{S}$.

3.2. Ereignis-Aktivitätsnetzwerk

Das im Folgenden definierte Ereignis-Aktivitätsnetzwerk (*engl.* activity-on-arc network) bildet die Grundlage der Modellierung in dieser Arbeit. In diesem Netzwerk werden Verspätungen fortgepflanzt. Die hier enthaltenen Informationen werden später in der Definition des Anschlussgraphen komprimiert.

Definition 3.2.1 (Ereignis-Aktivitätsnetzwerk).

Ein Ereignis-Aktivitätsnetzwerk $\mathfrak{N} = (\mathfrak{E}, \mathfrak{A})$ besteht aus den folgenden Elementen:

(i) Die Menge der Ereignisse $\mathfrak{E} = \mathfrak{E}^{arr} \cup \mathfrak{E}^{dep}$ mit

- der Menge der **Ankunftsereignisse** $\mathfrak{E}^{arr} = \{arr(S, L) \mid (S, L) \in \mathfrak{S} \times \mathfrak{L}, (S, L) \in S_L\}$ sowie
- der Menge der **Abfahrtsereignisse** $\mathfrak{E}^{dep} = \{dep(S, L) \mid (S, L) \in \mathfrak{S} \times \mathfrak{L}, (S, L) \in S_L\}$.

(ii) Menge der Aktivitäten $\mathfrak{A} = \mathfrak{A}^{stop} \cup \mathfrak{A}^{change} \cup \mathfrak{A}^{drive} \cup \mathfrak{A}^{head} \subset \mathfrak{E}^2$ mit

- der Menge der **Halteaktivitäten**

$$\mathfrak{A}^{stop} = \{stop(S, L) := (arr(S, L), dep(S, L)) \mid (S, L) \in \mathfrak{S} \times \mathfrak{L}, (S, L) \in S_L\},$$

die den Haltevorgang der Züge einer Linie in einer Station beschreiben,

- der Menge der **Umsteigeaktivitäten**

$$\mathfrak{A}^{change} = \{change(S, L, L') := (arr(S, L), dep(S, L')) \mid (S, L, L') \in \mathfrak{S} \times \mathfrak{L} \times \mathfrak{L}, L \neq L', (S, L) \in S_L, (S', L) \in S'_L\},$$

welche die Möglichkeit beschreiben, in einer Station von einer Linie in eine andere umzusteigen,

- der Menge der **Fahraktivitäten**

$$\mathfrak{A}^{drive} = \{drive(S, S', L) := (dep(S, L), arr(S', L)) \mid (S, S') \in \mathfrak{B}, L \in \mathfrak{L}, ((S, S'), L) \in \mathcal{B}_L\},$$

welche die Fahrt von Zügen einer Linie auf einer Strecke beschreiben.

- der Menge der **Zugfolgeaktivitäten**

$$\begin{aligned} \mathfrak{A}^{\text{head}} = & \{ \text{headarr}(S, L, L') := (\text{arr}(S, L), \text{arr}(S, L')) \mid \\ & (S, L, L') \in \mathfrak{S} \times \mathfrak{L} \times \mathfrak{L}, L \neq L', \\ & (S, L) \in \mathcal{S}_L, (S', L) \in \mathcal{S}'_L \} \\ \cup & \{ \text{headdep}(S, L, L') := (\text{dep}(S, L), \text{dep}(S, L')) \mid \\ & (S, L, L') \in \mathfrak{S} \times \mathfrak{L} \times \mathfrak{L}, L \neq L', \\ & (S, L) \in \mathcal{S}_L, (S', L) \in \mathcal{S}'_L \}, \end{aligned}$$

welche die einzuhaltenden Zugfolgezeiten für Abfahrten und Ankünfte verschiedener Linien in einer Station beschreiben.

Definition 3.2.2 (Fahrplan).

Es sei $\mathfrak{A} = (\mathfrak{E}, \mathfrak{A})$ ein Ereignis-Aktivitätsnetzwerk und Takt $\tau \in \mathbb{N}$ gegeben.

- (i) Ein (Takt-) **Fahrplan** ist ein Vektor $\pi = (\pi(e))_{e \in \mathfrak{E}}$ mit $0 \leq \pi(e) < \tau$ für alle $e \in \mathfrak{E}$. $\pi(e)$ ist die planmäßige Zeit für das Ereignis $e \in \mathfrak{E}$ in jedem Takt.
- (ii) Es seien $(l(a), u(a))$ mit $0 \leq l(a) \leq u(a) \leq \infty$ untere und obere Schranken für die Dauer der Aktivität $a \in \mathfrak{A}$. Dann ist der Fahrplan π **zulässig**, falls

$$[\pi(e_2) - \pi(e_1) - l(a)]_\tau \leq u(a) - l(a)$$

für alle Aktivitäten $a = (e_1, e_2) \in \mathfrak{A}$. Dabei ist $[r]_\tau := \min\{r + z\tau \mid z \in \mathbb{Z}, r + z\tau \geq 0\}$ die periodische Reduzierung modulo τ für $r \in \mathbb{R}$.

- (iii) Für eine Aktivität $a = (e_1, e_2) \in \mathfrak{A}$ bezeichnet

$$s(a) := [\pi(e_2) - \pi(e_1) - l(a)]_\tau$$

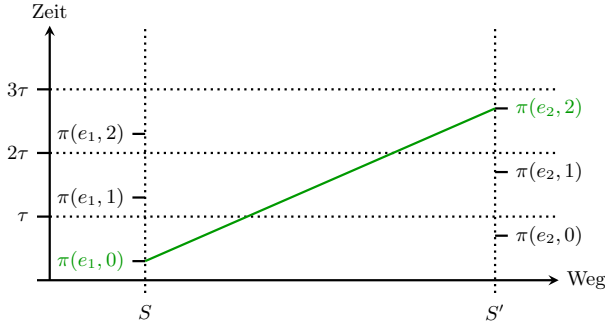
den im Fahrplan π enthaltenen **Zeitpuffer** der Aktivität a . Die Zulässigkeit des Fahrplans bedeutet daher $s(a) \leq u(a) - l(a)$.

Wir betrachten nun zu jedem periodischen Ereignis $e \in \mathfrak{E}$ die durch $z \in \mathbb{Z}$ nummerierten *Instanzen* eines jeden Taktes. Wir erweitern daher \mathfrak{E} zu $\widehat{\mathfrak{E}} := \mathfrak{E} \times \mathbb{Z}$. Die fahrplanmäßige absolute Zeit für die z -te Instanz des Ereignisses $e \in \mathfrak{E}$ ist

$$\pi(e, z) = z \cdot \tau + \pi(e). \tag{3.2}$$

Zu jeder periodischen Aktivität $a = (e_1, e_2) \in \mathfrak{A}$ und zu jedem $z \in \mathbb{Z}$ bzw. $(e_1, z_1) \in \widehat{\mathfrak{E}}$ gibt es ein $z_2 \in \mathbb{Z}$ mit $((e_1, z_1), (e_2, z_2)) \in \widehat{\mathfrak{A}}$ mit $\widehat{\mathfrak{A}} := \widehat{\mathfrak{E}}^2$ (vgl. Abbildung 3.1). Zulässig ist eine solche Aktivität, wenn gilt

$$l(a) \leq \pi(e_2, z_2) - \pi(e_1, z_1) \leq u(a)$$


 Abbildung 3.1.: Individualisierte Aktivität $\hat{a} = ((e_1, 0), (e_2, 2))$

bzw.

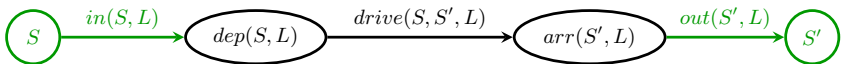
$$l(a) \leq (z_2 - z_1)\tau + \pi(e_2) - \pi(e_1) \leq u(a) .$$

Im Folgenden betrachten wir immer zulässige Aktivitäten $\hat{a} \in \widehat{\mathfrak{A}}$. Zur Abkürzung schreiben wir auch $\hat{a} = (i, j) \in \widehat{\mathfrak{A}}$ für $i = (e, z)$, $j = (e', z') \in \widehat{\mathfrak{E}}$ bzw.

$$\hat{a} = (i, j) = ((e, z), (e', z')) = ((e, e'), z, z') = (a, z, z') .$$

Die individualisierten Aktivitäten können daher auch als Elemente aus $\mathfrak{A} \times \mathbb{Z}^2$ aufgefasst werden.

Bevor wir in Abschnitt 3.4 den Anschlussgraphen einführen, definieren wir noch die erweiterte Version des Ereignis-Aktivitätsnetzwerkes, welches beispielsweise in Kapitel 9 für die Generierung der Reisendenströme benötigt wird. Das Ereignis-Aktivitätsnetzwerk \mathfrak{N} wird zunächst um sogenannte *Metaknoten* erweitert, welche die Modellierung der Aktivitäten „Einsteigen“ und „Aussteigen“ ermöglichen. Die Menge der Metaknoten, $\mathfrak{E}_{\mathfrak{S}}$, entspricht im Streckennetz gerade der Menge aller Stationen, \mathfrak{S} . Abbildung 3.2 zeigt einen Ausschnitt eines solchen Netzwerkes beispielhaft. Dort sind alle, im Vergleich zum Netzwerk \mathfrak{N} , hinzugekommenen Knoten und Kanten mit grün gekennzeichnet. Bei den Kanten handelt es sich um Einsteige- und Aussteigeaktivitäten aus neu hinzugekommenen Mengen \mathfrak{A}^{in} bzw. \mathfrak{A}^{out} .


 Abbildung 3.2.: Erweiterung des Ereignis-Aktivitätsnetzwerkes mit Metaknoten S, S' sowie Ein- und Aussteigeaktivitäten

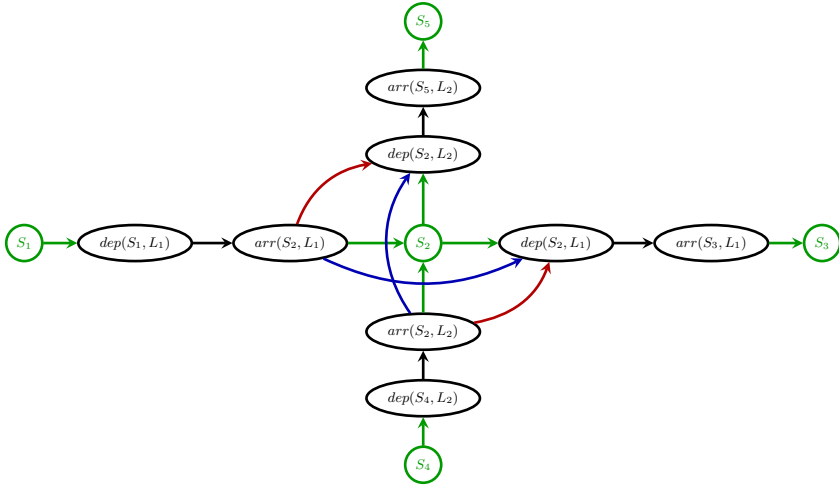


Abbildung 3.3.: Erweitertes Ereignis-Aktivitätsnetzwerk

Definition 3.2.3 (Metaknoten).

Ein **Metaknoten** ist Teil eines erweiterten Ereignis-Aktivitätsnetzwerkes. Er wird formal der Menge der Ereignisse zugeordnet. Für jede Station $S \in \mathfrak{S}$ des Streckennetzes existiert genau ein Metaknoten. Die Menge der Metaknoten ist definiert als

$$\mathfrak{E}^{meta} = \{meta(S) \mid S \in \mathfrak{S}\} .$$

Definition 3.2.4 (Ein- und Aussteigekanten).

Einsteige- bzw. Aussteigekanten stellen in einem erweiterten Ereignis-Aktivitätsnetzwerk die Verbindung zwischen den ursprünglichen Ereignissen des Ereignis-Aktivitätsnetzwerkes und den Metaknoten her. Die Mengen der **Ein-** und **Aussteigekanten** werden definiert als

$$\mathfrak{A}^{in} = \{in(S, L) = (meta(S), dep(S, L)) \mid (S, L) \in \mathcal{S}_L\}$$

bzw.

$$\mathfrak{A}^{out} = \{out(S, L) = (arr(S, L), meta(S)) \mid (S, L) \in \mathcal{S}_L\} .$$

Definition 3.2.5 (Erweitertes Ereignis-Aktivitätsnetzwerk).

Es sei $\mathfrak{N}(\mathfrak{E}, \mathfrak{A})$ ein Ereignis-Aktivitätsnetzwerk. Das zugehörige **erweiterte Ereignis-Aktivitätsnetzwerk** $\mathfrak{N}^*(\mathfrak{E}^*, \mathfrak{A}^*)$ besteht aus

- (i) der Knotenmenge $\mathfrak{E}^* = \mathfrak{E} \cup \mathfrak{E}^{meta}$ sowie
 (ii) der Kantenmenge $\mathfrak{A}^* = \mathfrak{A} \cup \mathfrak{A}^{in} \cup \mathfrak{A}^{out}$.

3.3. Verspätungsarten

Wir unterscheiden in dieser Arbeit im Wesentlichen zwei Arten von Verspätungen im Ereignis-Aktivitätsnetzwerk. Externe Störungen heißen *Quellverspätungen* und werden nur Aktivitäten zugeordnet. Demgegenüber werden *fortgepflanzte Verspätungen* direkt oder indirekt von Quellverspätungen verursacht. Diese sind mit den Ereignissen des Netzwerkes verknüpft. Mittels welcher Operationen die fortgepflanzten Verspätungen aus den Quellverspätungen hervorgehen, wird später auch nochmal in Kapitel 6 erläutert.

Die Zufallsvariable $X(e, z)$, $(e, z) \in \widehat{\mathfrak{E}}$, steht allgemein für die tatsächliche Eintrittszeit der z -ten Instanz des Ereignisses e . Die folgende Definition präzisiert dies weiter.

Definition 3.3.1 (Tatsächliche Eintrittszeit).

Für $z \in \mathbb{Z}$ sei

- (i) $X(arr, S, L, z)$ die z -te **Ankunftszeit** eines Zuges der Linie $L \in \mathfrak{L}$ in Linienstation $(S, L) \in S_L$,
 (ii) $X(dep, S, L, z)$ die z -te **Abfahrtszeit** eines Zuges der Linie $L \in \mathfrak{L}$ in Linienstation $(S, l) \in S_L S$.

Bevor wir zur ersten Definition des Begriffs *Verspätung* kommen, sei gesagt, dass in dieser Arbeit keine *Verfrühungen* betrachtet werden. Für $i \in \widehat{\mathfrak{E}}$ gilt also immer $X(i) \geq \pi(i)$. Unabhängig vom Fahrplan genügt für die Ausführung einer Aktivität $a = (e_1, e_2) \in \mathfrak{A}^{drive} \cup \mathfrak{A}^{stop}$ die Mindestdauer $l(a)$. Es kann bei einer individuellen Aktivität $\hat{a} = (a, z_1, z_2) \in \widehat{\mathfrak{A}}$ allerdings zu Verzögerungen im Betriebsablauf kommen, in Folge derer die *tatsächliche Dauer* $R(\hat{a})$ über der Mindestdauer liegt.

Definition 3.3.2 (Quellverspätung).

Sei $a = (e_1, e_2) \in \mathfrak{A}^{drive} \cup \mathfrak{A}^{stop}$ und $\hat{a} = (a, z_1, z_2) \in \widehat{\mathfrak{A}}$. Die **Quellverspätung**

$$D(\hat{a}) := R(\hat{a}) - l(a) \quad (3.3)$$

bezeichnet die Überschreitung der Mindestdauer einer individualisierten Aktivität \hat{a} . Dabei sind $D(\hat{a})$ und $R(\hat{a})$ Zufallsvariablen auf einem gemeinsamen Wahrscheinlichkeitsraum.

Es sei nun $j = (e_2, z_2) \in \widehat{\mathfrak{E}}$ das Ereignis am Ende der Aktivität $\hat{a} = (i, j) = (a, z_1, z_2) \in \widehat{\mathfrak{A}}^{drive} \cup \widehat{\mathfrak{A}}^{stop}$. Die frühest mögliche Eintrittszeit von j ist definiert als

$$X_0(j, \hat{a}) := \max \{ X(i) + l(a) + D(\hat{a}), \pi(j) \}. \quad (3.4)$$

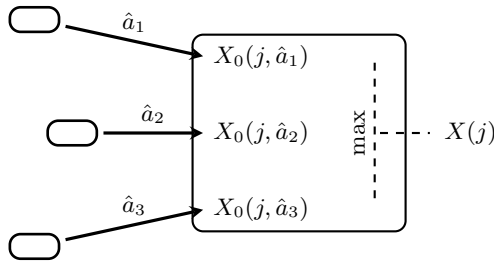


Abbildung 3.4.: Die tatsächliche Eintrittszeit $X(j)$ des Ereignisses $j \in \widehat{\mathfrak{E}}$ als Maximum der frühestmöglichen $X_0(j, \hat{a})$ bzgl. verschiedener Aktivitäten $\hat{a} \in \widehat{\mathfrak{A}}$

Ist j das Ende mehrerer Aktivitäten (vgl. Abbildung 3.4), so gilt

$$X(j) = \max \left\{ X_0(j, \hat{a}) \mid \exists i \in \widehat{\mathfrak{E}} : \hat{a} = (i, j) \in \widehat{\mathfrak{A}} \right\} . \quad (3.5)$$

Überschaubar ist die Situation bei Ankunftsereignissen. Dort ist neben der Fahraktivität eventuell die Zugfolgezeit zu beachten. Bei Halteaktivitäten muss zusätzlich auch noch auf möglicherweise verspätete Anschlüsse gewartet werden. Wie im Falle der Anwendung sogenannter *Anschlussicherungsstrategien* verfahren wird, zeigt Abschnitt 6.4.

Definition 3.3.3 (Fortgepflanzte Verspätung).

Die (**fortgepflanzte**) **Verspätung** eines Ereignisses $i \in \widehat{\mathfrak{E}}$ wird definiert als Zufallsvariable

$$Y(i) := X(i) - \pi(i) \geq 0 . \quad (3.6)$$

Dies entspricht der Abweichung des tatsächlichen Ereigniszeitpunktes vom fahrplanmäßigen.

In dieser Arbeit untersuchen wir für alle $e \in \mathfrak{E}$ die langfristige Verteilung ($n \rightarrow +\infty$) der fortgepflanzten Verspätung $Y(e, n)$, sofern diese überhaupt existiert. Aussagen über die Existenz dieser Verteilungen werden in Kapitel 7 behandelt. Die Quellverspätungen stellen dabei die Eingangsgrößen dar, die fortgepflanzten Verspätungen sind die Ergebnisse.

3.4. Der Anschlussgraph

Der Anschlussgraph fasst Teile eines Ereignis-Aktivitätsnetzwerkes so zusammen, dass er die für die Verspätungsfortpflanzung auf Ebene der Verteilungsfunktionen

wesentlichen Aktivitäten widerspiegelt. Abbildung 3.5 zeigt den zum (erweiterten) Ereignis-Aktivitätsnetzwerk aus Abbildung 3.3 gehörenden Anschlussgraphen. Für eine präzisere Beschreibung benötigen wir eine Klassifikation der Stationen. Es sei

- $\mathcal{S}_{Start} \subset \mathfrak{S}$ die Menge aller Stationen, welche Startstation einer Linie $L \in \mathcal{L}$ sind,
- $\mathcal{S}_{End} \subset \mathfrak{S}$ die Menge aller Stationen, welche Endstation einer Linie $L \in \mathcal{L}$ sind,
- $\mathcal{S}_{change} \subset \mathfrak{S}$ die Menge aller Umsteigestationen.

Definition 3.4.1 (Anschlussgraph).

Der **Anschlussgraph** $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ zu gegebenem Ereignis-Aktivitätsnetzwerk $\mathfrak{N} = (\mathfrak{E}_{\mathfrak{N}}, \mathfrak{A}_{\mathfrak{N}})$ enthält als

(i) **Knotenmenge** \mathfrak{V} für jede Linie $L \in \mathcal{L}$ die Tupel

$$V = (dep(S_i, L), arr(S_{i+1}, L), dep(S_{i+1}, L), \dots, dep(S_{j-1}, L), arr(S_j, L)) \in \mathfrak{V}$$

mit $(S_i, L) \in \mathcal{S}_L$ für $i \leq l \leq j$. Dabei gilt stets $S_i \in \mathcal{S}_{Start} \cup \mathcal{S}_{change}$, $S_j \in \mathcal{S}_{End} \cup \mathcal{S}_{change}$ sowie $S_l \in \mathfrak{S} \setminus (\mathcal{S}_{Start} \cup \mathcal{S}_{End} \cup \mathcal{S}_{change})$ für $i < l < j$.

(ii) **Kantenmenge** \mathfrak{E} Elemente $e = (V, V') \in \mathfrak{E} \subset \mathfrak{V}^2$ mit

$$\begin{aligned} V &= (dep(S_i, L), arr(S_{i+1}, L), dep(S_{i+1}, L), \dots, dep(S_{j-1}, L), arr(S_j, L)) \\ V' &= (dep(S_j, L'), arr(S_{j+1}, L'), dep(S_{j+1}, L'), \dots, dep(S_{k-1}, L'), arr(S_k, L')) \end{aligned}$$

für die eine (eindeutige) Aktivität $a = (arr(S_j, L), dep(S_j, L')) \in \mathfrak{A}_{\mathfrak{N}}^{change} \cup \mathfrak{A}_{\mathfrak{N}}^{stop}$ existiert.

Ein Knoten des Anschlussgraphen fasst also im Wesentlichen alle Ereignisse zusammen, die zwischen einem Abfahrtsereignis und einem Ankunftsereignis in (jeweils) Umsteigestationen liegen. Die Kanten des Anschlussgraphen sind anschaulich gerade die Umsteige- bzw. Halteaktivitäten des Ereignis-Aktivitätsnetzwerk, welche das letzte und erste Ereignis zweier Knoten des Anschlussgraphen miteinander verbinden.

Definition 3.4.2 (Vorgängerknoten, Nachfolgeknoten).

Es seien $V_1, V_2 \in \mathfrak{V}$ Knoten des Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$. V_2 ist ein **Nachfolgeknoten** von V_1 , falls eine Kante $e \in \mathfrak{E}$ existiert mit $e = (V_1, V_2)$. In diesem Fall ist zugleich V_1 **Vorgängerknoten** von V_2 . Mit $\text{Pre}_{\mathfrak{G}}(V)$ bezeichnen wir die Menge aller Vorgängerknoten des Knotens $V \in \mathfrak{V}$. Entsprechend bezeichnen wir mit $\text{Suc}_{\mathfrak{G}}(V)$ die Menge aller Nachfolgeknoten des Knotens V .

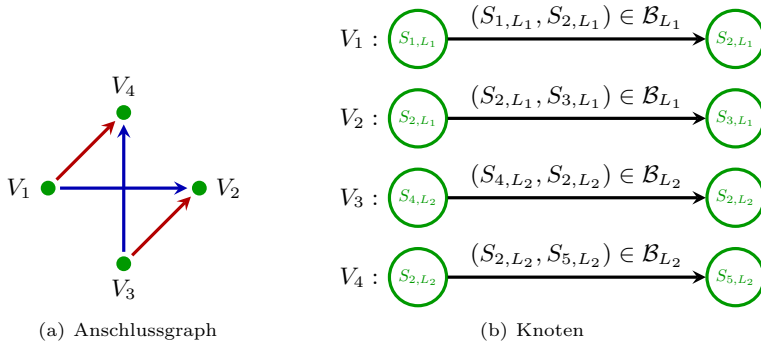


Abbildung 3.5.: Anschlussgraph zum (erweiterten) Ereignis-Aktivitätsnetzwerk aus Abbildung 3.3

Definition 3.4.3 (Vereinigungsknoten).

Vereinigungsknoten des Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ sind Knoten $V \in \mathfrak{V}$ für die gilt

$$|\mathfrak{Pre}(V)| > 1 .$$

Die Menge der Vereinigungsknoten des Anschlussgraphen \mathfrak{G} wird mit $\mathcal{J}_{\mathfrak{G}}$ bezeichnet.

Definition 3.4.4 (Verzweigungsknoten).

Verzweigungsknoten des Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ sind Knoten $V \in \mathfrak{V}$ für die gilt

$$|\mathfrak{Suc}(V)| > 1 .$$

Die Menge der Verzweigungsknoten des Anschlussgraphen \mathfrak{G} wird mit $\mathcal{B}_{\mathfrak{G}}$ bezeichnet.

4. Empirische Analyse von Verspätungsdaten

Übersicht

4.1. Die Struktur der Daten	24
4.2. Extrahierung der benötigten Informationen	25
4.2.1. Mindestfahr- und Mindesthaltezeiten	25
4.2.2. Quellverspätungen	26
4.2.3. Ergebnis der Datenanalyse	27

Die Datenakquisition wurde durch die DB NETZ AG am Standort Leipzig unterstützt. Ziel war es ursprünglich, zu einem relevanten Streckennetz (z.B. aktuelles Fernverkehrsnetz) realistische Daten über Quellverspätungen zu bekommen. Genauer heißt das, dass wir für jede Linienstation und jede Linienstrecke aller Linien eine Stichprobe der auftretenden Quellverspätungen benötigen. Dieses Ziel konnte mit den zur Verfügung gestellten Daten nicht erreicht werden. Es war uns jedoch möglich, einen grundsätzlichen Eindruck vom Aufbau solcher Verspätungsdatensätze sowie einen Einblick in Eigenarten möglicher Verspätungsverteilungen zu erhalten. Dies bestätigte letztlich die spezielle Wahl der Verteilungsfamilie, die wir für die Approximation bzw. Repräsentation der empirischen Verspätungsverteilungen verwenden.

In einem ersten Schritt musste die Frage beantwortet werden, welches reale Netz sich für die statistische Untersuchung von Quellverspätungen anbietet. Dafür muss klar sein, in welcher Form die Verspätungen festgehalten werden. Die Quellverspätungen selbst werden nicht gemessen. Stattdessen werden die (absoluten) Verspätungen nur indirekt, als Abweichung der sogenannten *Ist-Zeit* von der fahrplanmäßigen *Soll-Zeit*, festgehalten. Ein Kriterium war daher, dass wir Linien bzw. Linienabschnitte betrachten, die von möglichst wenig anderen Linien beeinflusst werden können. Ein weiteres Kriterium war, mit welchen Werkzeugen und mit welcher Genauigkeit die Verspätungen gemessen werden. Die Entscheidung fiel auf einen Streckenabschnitt des sächsischen Nahverkehrsnetzes (s. Abbildung 4.1). Da mit der DB NETZ AG die Vertraulichkeit der Daten vereinbart wurde, kann der Name des Streckenabschnittes hier nicht genannt werden. Eine Besonderheit war, dass die Verspätungen nicht per Hand sondern automatisiert und mit relativ hoher Genauigkeit (6 Sekunden) gemessen werden. Weitere Besonderheiten waren

- die Zweigleisigkeit des Streckenabschnittes,

4. Empirische Analyse von Verspätungsdaten

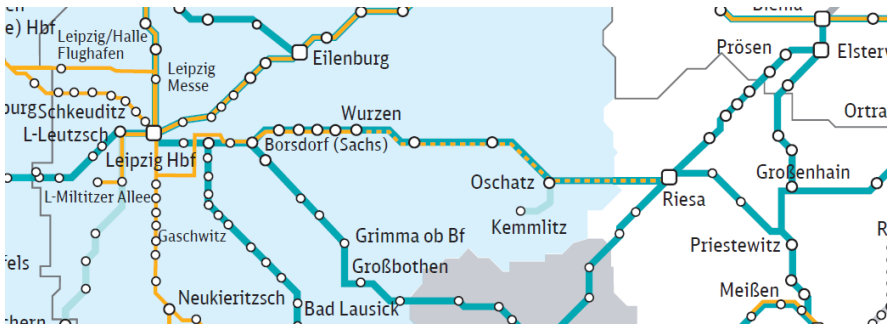


Abbildung 4.1.: Ausschnitt einer Streckenkarte für Thüringen, Sachsen-Anhalt und Sachsen¹

- keine Wartezeiten auf Zubringer, d.h. keine „echten“ Umsteigestationen, sowie
- die relativ geringe Verkehrsdichte.

Dies stützt die Zulässigkeit der in Abschnitt 4.2.2 beschriebenen Klassifikation der Datensätze. Durch die Besonderheiten des Streckenabschnittes wird die Wahrscheinlichkeit reduziert, dass die weitere Verspätungsentwicklung eines Zuges von der bereits erfahrenen Verspätung abhängt. Ist er moderat verspätet, wird er auf dem Streckenabschnitt nicht in Konflikt mit einem anderen Zug (Zugfolgerestriktionen) geraten. Er erfährt auch keine von Zubringern übertragene Verspätung. Wächst seine Verspätung an, so wird dies in aller Regel tatsächlich auf eine externe Störung des Betriebsablaufes (Quellverspätung) zurückzuführen sein. Weiteres Datenmaterial, d.h. insbesondere zu anderen Streckennetzen, konnte nicht mehr untersucht werden.

4.1. Die Struktur der Daten

Bevor die Ergebnisse vorgestellt werden, wird zunächst die Struktur der von der DB NETZ AG an uns übergebenen Daten diskutiert. Die Daten wurden seitens der DB NETZ AG in Form einer Tabelle in Textdateien (.txt-Dateien) gespeichert. Bis auf einige für uns nicht relevante Informationen gibt Tabelle 4.1 die Struktur dieser Dateien wieder. Im Folgenden wird die Bedeutung der einzelnen Spalten erklärt.

BST steht für Betriebsstelle.

¹Streckenkarte für Thüringen, Sachsen und Sachsen-Anhalt
<http://www.bahn.de/p/view/buchung/karten/streckennetz.shtml> (05.02.2014)

BST	BTG	ZG	ZN	SOLLAN	ISTAN	SOLLAB	ISTAB	NBSTNAME
-	28.02.10	14	-	hh:mm:ss	hh:mm:ss	hh:mm:ss	hh:mm:ss	-
-	01.03.10	40	-	hh:mm:ss	hh:mm:ss	hh:mm:ss	hh:mm:ss	-
⋮			⋮					⋮
-	05.03.10	11	-	hh:mm:ss	hh:mm:ss	hh:mm:ss	hh:mm:ss	-

Tabelle 4.1.: Die Struktur der Daten

BTG gibt den Betriebstag des Zuges an. Dies ist der Tag, an dem der Zug fahrplanmäßig die in der entsprechenden Zeile angegebene Betriebsstelle passieren soll.

ZG steht für Zuggattung. Die verschiedenen Zuggattungen wie ICE (InterCity-Express, 14), IC (InterCity, 11) oder RE (Regional-Express, 40) stehen hier als natürliche Zahl kodiert.

ZN steht für Zugnummer.

SOLLAN / SOLLAB repräsentieren die fahrplanmäßigen Ankunftszeiten und Abfahrtszeiten.

ISTAN / ISTAB zeigen die tatsächlich gemessenen Zeitpunkte für Ankünfte und Abfahrten.

NBSTNAME steht für den Namen der nächsten Betriebsstelle.

Für die Analyse des vorhandenen Datenmaterials haben wir das Statistiksytstem **R** verwendet². Das letztlich untersuchte Datenpaket umfasste einen Zeitraum von zwei Monaten und enthielt gut 38000 Datensätze.

4.2. Extrahierung der benötigten Informationen

Nachdem wir uns einen groben Überblick über die Daten verschafft haben, versuchen wir Stichproben der Quellverspätungsverteilungen zu extrahieren.

4.2.1. Mindestfahr- und Mindesthaltezeiten

Kenntnis der Mindestfahr- sowie Mindesthaltezeiten sind Voraussetzung für die Extrahierung der Quellverspätungen. Der Grund hierfür wird im folgenden Abschnitt

²The R Project for Statistical Computing
<http://www.r-project.org/>

näher erläutert. Entscheidend ist an dieser Stelle, wie wir diese Information erhalten. Die Mindestfahrzeiten müssen nicht grundsätzlich mit der jeweils geringsten, technisch möglichen Zeit übereinstimmen. Selbst wenn dem so wäre, hängen diese Zeiten sowohl von der speziellen Strecke als auch vom eingesetzten Zug ab. Dagegen lassen sich für Mindesthaltezeiten durchaus feste Regeln vorgeben. Uns lagen jedoch keine verlässlichen Informationen hierzu vor. Daher wurden sowohl Mindestfahr- als auch Mindesthaltezeiten aus dem Datenmaterial gewonnen. Dabei sind wir jeweils so vorgegangen, dass ein bestimmter, niedriger Prozentsatz der kürzesten tatsächlichen Halte- bzw. Fahrzeiten (Ist-Zeiten) einer festen Zuggattung (z.B. RE) extrahiert wurde. Der Mittelwert wurde dann schließlich als Mindestfahr- bzw. Mindesthaltezeit verwendet. Durch die Durchschnittsbildung einer kleineren Auswahl von Datensätzen werden etwaige Ungenauigkeiten in den Daten berücksichtigt. Die Stichproben der Quellverspätungen werden dann ebenfalls hinsichtlich dieser Zuggattung ausgewählt.

4.2.2. Quellverspätungen

Die grundsätzliche Schwierigkeit der Extrahierung von Quellverspätungen aus Daten zu fortgepflanzten Verspätungen wird unter anderem in [WN04] beschrieben. Es ist im Allgemeinen schwierig, zwischen Quell- und Folgeverspätungen zu unterscheiden, d.h. die genauen Ursachen der Verspätung aufzuschlüsseln. Wir nehmen an, dass die Besonderheiten des gewählten Streckenabschnittes (s.o.) unser Vorgehen rechtfertigt. Wir erinnern uns an die Definition der Quellverspätung in (3.3). Die Quellverspätung $D(\hat{a})$ bezeichnet die Überschreitung der Mindestdauer $l(a)$ einer individualisierten Aktivität $\hat{a} = (a, z, z') \in \hat{\mathfrak{A}}$. $D(\hat{a})$ ist eine Zufallsvariable. In den zugrunde liegenden Datensätzen sehen wir Realisierungen $d(\hat{a})$ dieser Zufallsvariablen. Wir setzen voraus, dass wir die Mindestfahr- und Mindesthaltezeiten sowie die zugehörigen Puffer kennen bzw. bereits aus den Daten extrahiert haben. Des Weiteren betrachten wir für die weitere Veranschaulichung die individualisierte Fahraktivität

$$\hat{a} = (\text{dep}(S, L, z_1), \text{arr}(S', L, z_2)) =: (i, j) \in \hat{\mathfrak{A}}^{\text{drive}} .$$

Die naive Herangehensweise wäre, für diese Aktivität die tatsächliche Fahrzeit zu berechnen, und anschließend die Differenz mit der Mindestfahrzeit zu bilden. Das Problem ist, dass Züge sich in der Realität in Abhängigkeit von ihrer aktuellen Verspätungen verhalten. So wird ein Zug bei pünktlicher Abfahrt eher die fahrplanmäßige Fahrzeit statt der Mindestfahrzeit nutzen. Mit der naiven Herangehensweise beobachteten wir eine positive Quellverspätung, obwohl es eventuell gar keine Störung im Betriebsablauf gab. Die Datensätze beinhalten also zensierte Information. Wir versuchen das Problem dadurch zu lösen, dass wir nur bestimmte Datensätze für die Auswertung der Quellverspätungen verwenden. Dies sind im Falle der auf Strecken auftretenden Quellverspätungen jene Datensätze, bei denen die

Abfahrtsverspätung mindestens so groß ist wie der zur Verfügung stehende Puffer. Wir unterscheiden dazu im Folgenden zwei Fälle. Dabei ist $y(i)$ die in den Daten enthaltene Realisierung der Zufallsvariablen $Y(i)$ (vgl.(3.6)), die (fortgepflanzte) Verspätung der Abfahrt in Station S .

- (i) Ist $0 \leq y(i) < s(a)$, so gibt eine Überschreitung der Mindestfahrzeit keinen Aufschluss darüber, ob eine Störung zur positiven Quellverspätung geführt hat. Im Extremfall einer pünktlichen Abfahrt, wird ein Zug eher die fahrplanmäßige Fahrzeit statt der Mindestfahrzeit nutzen. Auf diese Weise verursachte Quellverspätungen sollen nicht betrachtet werden, da sie nicht Folge einer Störung im Betriebsablauf wären.
- (ii) Ist stattdessen $s(a) \leq y(i)$, so wird der Zug, wenn möglich, die Mindestfahrzeit nutzen. Nur so würde der vorhandene Puffer in vollem Umfang zum Abbau der Verspätung genutzt. Ist die Quellverspätung hier positiv, hat der Zug also mehr Zeit als die Mindestfahrzeit benötigt, so muss es während der Fahrt eine Störung im Betriebsablauf gegeben haben.

Die Verteilung der Quellverspätungen sollte davon kaum beeinflusst werden. Es werden für die Quellverspätungen keine Wertebereiche ausgeschlossen. Es wird lediglich zwischen zulässigen und unzulässigen Datensätzen unterschieden.

4.2.3. Ergebnis der Datenanalyse

Das Ergebnis unserer Datenanalyse ist in den Abbildungen 4.2 und 4.3 graphisch veranschaulicht. Die oben genannten Besonderheiten des Streckenabschnittes resultieren, mit Ausnahme einer Strecke, in erwartungsgemäß geringen Quellverspätungen. Die Stichproben sind sicherlich nicht repräsentativ für andere Streckenabschnitte, welche die Besonderheiten nicht teilen. Wir halten aber fest, dass die empirischen Verteilungen alle eine recht einfache Struktur besitzen. In den Histogrammen fällt die Intensität entweder monoton, oder sie steigt zunächst monoton an, um anschließend monoton zu fallen. Es gibt für größere Verspätungswerte keine weiteren nennenswerten Anstiege der Intensität. Derartige empirische Verteilungen können mit dem in Kapitel 5 erläuterten Verfahren besonders gut approximiert werden. Auch das im Rahmen der Komplexitätsreduzierung in Kapitel 6 erklärte Verfahren eignet sich gut für die Approximation solcher Verteilungen.

4. Empirische Analyse von Verspätungsdaten

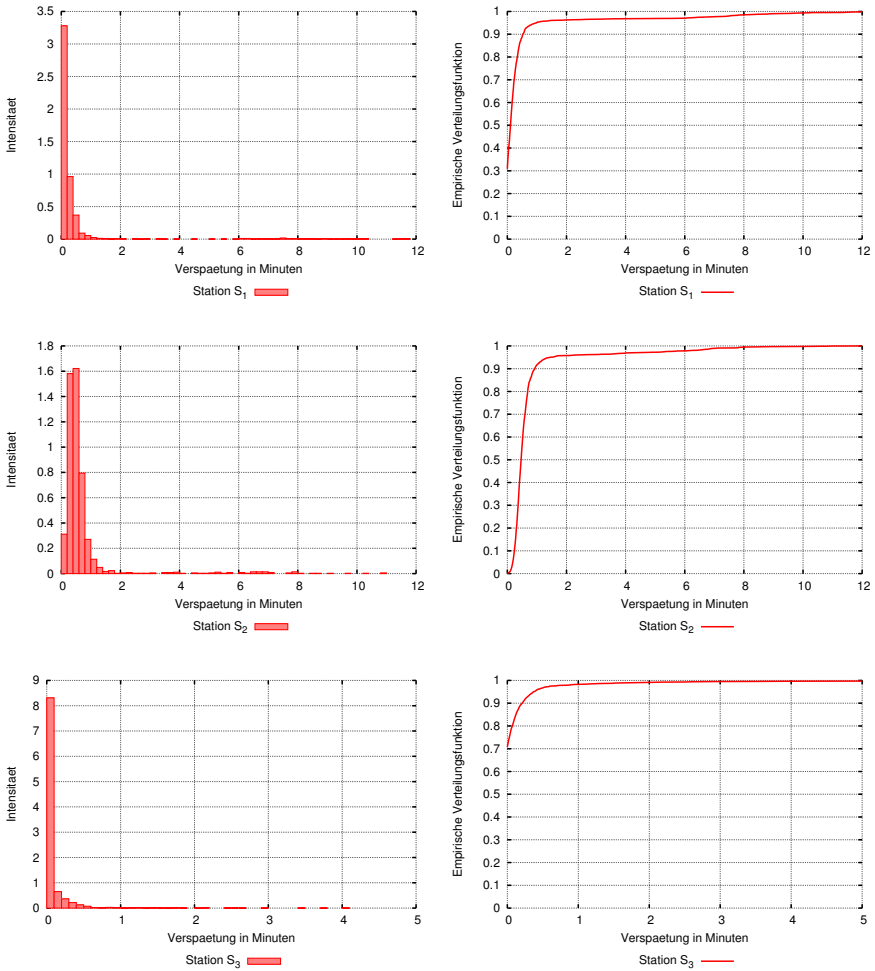


Abbildung 4.2.: Quellverspätungen in Stationen - Histogramm und empirische Verteilungsfunktion

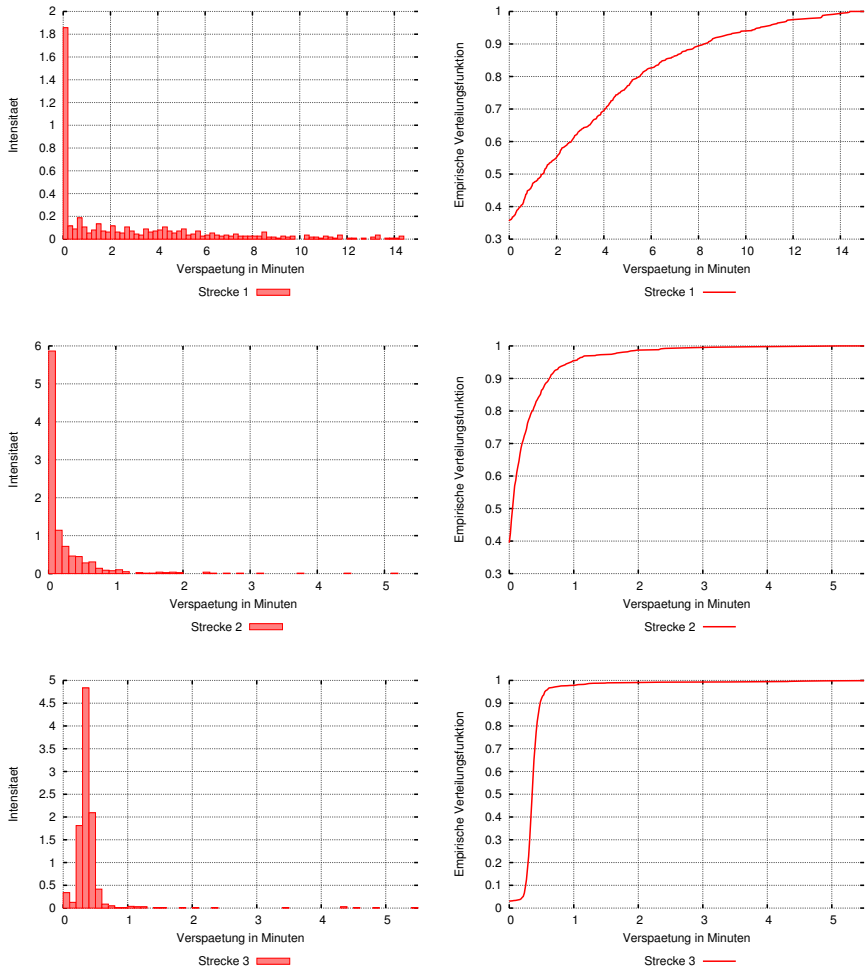


Abbildung 4.3.: Quellverspätungen in Stationen - Histogramm und empirische Verteilungsfunktion

5. Approximation von Phasentypverteilungen

Übersicht

5.1. Grundlagen	31
5.2. Literatur	38
5.3. Beispiele	41
5.3.1. Hyper-Exponentialverteilung	41
5.3.2. EM-Algorithmus für Hyper-Erlangverteilungen	44
5.4. Hyper-Erlangverteilung mit Gewichtung der Null	51

In diesem Kapitel beschäftigen wir uns mit der Approximation von Verteilungen bzw. Verteilungsfunktionen durch Phasentypverteilungen. Wir benötigen die Möglichkeit, Verteilungen zu approximieren, an zwei Stellen in dieser Arbeit. Auf der einen Seite wollen wir die gegebenen empirischen Verteilungen von Quellverspätungen approximieren. Auf der anderen Seite sind wir mit dem Problem konfrontiert, dass im Zuge der Anwendung der erwähnten Operationen, die Komplexität der Darstellung der jeweiligen Ergebnisse stark ansteigt. Daher müssen wir durch eine Approximation der Ergebnisse, die Komplexität der Darstellung wieder reduzieren, ohne dabei wesentliche Information zu verlieren.

Im ersten Abschnitt dieses Kapitels stellen wir die benötigten Grundlagen über Momente von Verteilungen und über die Familie der Phasentypverteilungen zusammen. Im zweiten Abschnitt wird Literatur zum Bereich der Approximation von Verteilungen mittels Phasentypverteilungen diskutiert. In Abschnitt drei werden zwei Verfahren aus der Literatur detaillierter beschrieben. Das letztere dieser beiden Verfahren wird von uns im vierten Abschnitt schließlich so modifiziert, dass wir es für unsere Zwecke einsetzen können.

5.1. Grundlagen

Die Momente von Wahrscheinlichkeitsverteilungen sind in dieser Arbeit von wesentlicher Bedeutung. Es wird ausgenutzt, dass die Momente, falls sie existieren, der Charakterisierung einer Verteilung dienen. Warum dies so ist, wird im Folgenden deutlich.

Definition 5.1.1 (Laplace-Stieltjes-Transformierte).

[Bau13] X sei eine (1-dimensionale) stetige Zufallsvariable mit Verteilungsfunktion F_X . Dann heißt

$$\Phi(s) := \mathbf{E} \left[e^{-sX} \right] = \int_{\mathbb{R}} e^{-st} dF_X(t), \quad t \in \mathbb{R}, \quad (5.1)$$

die **Laplace-Stieltjes-Transformierte (LST)** von X bzw. F_X .

Satz 5.1.2 (Eindeutigkeit).

[Bau13] Existieren zu den stetigen Zufallsvariablen X und Y , mit Verteilungsfunktionen F_X bzw. F_Y , die Laplace-Stieltjes-Transformierten Φ_X und Φ_Y in den Intervallen I_X bzw. I_Y mit $I_X \cap I_Y \neq \emptyset$, und gilt $\Phi_X(s) = \Phi_Y(s)$ für alle $s \in I_X \cap I_Y$, so gilt $F_X(t) = F_Y(t)$ in allen Punkten, in denen sowohl F_X als auch F_Y stetig sind.

Satz 5.1.3 (Momente).

[Bau13] Es sei X eine integrierbare stetige Zufallsvariable mit Verteilungsfunktion F_X . Des Weiteren existiere ihre Laplace-Stieltjes-Transformierte Φ_X im Intervall I_X . Dann ist Φ_X in I_X analytisch mit Ableitungen beliebiger Ordnung der Form

$$\frac{d^k}{ds^k} \Phi_X(s) = (-1)^k \cdot \int_{\mathbb{R}} t^k \cdot e^{-st} dF_X(t) .$$

In [Bau13] wird daraus gefolgert, dass sich die LST einer integrierbaren stetigen (reellwertigen) Zufallsvariablen X durch die folgende Potenzreihe darstellen lässt.

$$\Phi_X(s) = \sum_{k=0}^{\infty} \frac{(-s)^k}{k!} \mathbf{E} \left[X^k \right] \quad (5.2)$$

Existieren also die k -ten Momente einer Verteilung für alle $k \in \mathbb{N}$, so ist sie durch ihre Momente eindeutig charakterisiert. Abhängig vom Anwendungsfall, genügt es häufig, Verteilungen nur mittels der ersten zwei oder drei Momente zu charakterisieren. Dem ist auch in der vorliegenden Arbeit so.

Es folgt nun eine Charakterisierung von Phasentypverteilungen. In dieser Arbeit konzentrieren wir uns insbesondere auf eine Teilfamilie der Phasentypverteilungen, die Hyper-Erlangverteilungen.

Definition 5.1.4 (Phasentypverteilung).

[LR99] Es sei $Y = \{Y_t \mid t \in \mathbb{R}_{\geq 0}\}$ ein endlicher Markovprozess mit Zustandsmenge $\mathcal{S} = \{0, \dots, n\}$, einem Vektor $(\alpha_0, \alpha) \in [0, 1]^{n+1}$ bestehend aus den Startwahrscheinlichkeiten der einzelnen Zustände und Übergangsmatrix

$$Q = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{t} & T \end{pmatrix} .$$

Dabei ist T eine $n \times n$ -Matrix und \mathbf{t} entsprechend ein n -dimensionaler Spaltenvektor und es gilt

$$(i) \quad T_{ii} < 0, t_i \geq 0, T_{ij} \geq 0 \quad \text{für } 1 \leq i \neq j \leq n,$$

$$(ii) \quad T\mathbf{1} + t = \mathbf{0} \text{ und}$$

$$(iii) \quad \alpha_0 + \alpha\mathbf{1} = \mathbf{1}.$$

Die durch das Parametertupel (α, T) repräsentierte Wahrscheinlichkeitsverteilung der Zeit X bis zum Übergang in den absorbierenden Zustand 0 wird **Phasentypverteilung** (PH-Verteilung) n -ter Ordnung genannt (Schreibweise: X ist $PH(\alpha, T)$ -verteilt.).

Satz 5.1.5.

[LR99] Eine $PH(\alpha, T)$ -verteilte Zufallsvariable X besitzt die Verteilungsfunktion

$$F(x) = 1 - \alpha e^{Tx} \mathbf{1} \quad \text{für } x \geq 0, \quad (5.3)$$

und die Dichtefunktion

$$f(x) = \alpha e^{Tx} \mathbf{t}. \quad (5.4)$$

Die Matrix-Exponentialfunktion ist definiert als $e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n$.

Korollar 5.1.6 (Momente einer Phasentypverteilung).

[BB05] Es sei X eine $PH(\alpha, T)$ -verteilte Zufallsvariable. Dann sind deren Momente für alle $k \in \mathbb{N}$ durch

$$\mathbf{E}[X^k] = (-1)^k \cdot k! \cdot \alpha T^{-k} \mathbf{1} \quad (5.5)$$

gegeben.

Die Familie der allgemeinen Phasentypverteilungen besitzt eine wichtige, auch für diese Arbeit bedeutende Eigenschaft. Sie ist sowohl gegen die Faltungsoperation als auch gegen die Multiplikation abgeschlossen. Dies ist dann von Bedeutung, wenn wir ein Verfahren implementieren wollen, welches diese Operationen verwendet. Die Ergebnisse sind dann stets wieder Element der Verteilungsfamilie und daher darstellbar.

Satz 5.1.7 (Abgeschlossenheit gegen Faltung).

[Neu81] Sind F und G Verteilungsfunktionen stetiger PH-Verteilungen mit Parametern (α, T) bzw. (β, V) und Ordnung n bzw. m , so ist ihre Faltung $F * G$ eine PH-Verteilung mit Parameterdarstellung (γ, L) , gegeben durch $\gamma = [\alpha, \alpha_0\beta]$ und

$$L = \begin{pmatrix} T & \mathbf{t}\beta \\ \mathbf{0} & V \end{pmatrix}.$$

Bemerkung 5.1.8 (Das Kronecker Produkt).

Sind $A := (a_{ij})$ und $B := (b_{kl})$ $n \times m$ - bzw. $r \times s$ -Matrizen, so ist deren Kronecker Produkt $A \otimes B$ definiert als die $nr \times ms$ -Matrix mit Einträgen

$$(A \otimes B)_{(i_1, i_2)(j_1, j_2)} := a_{i_1, j_1} \cdot b_{i_2, j_2}.$$

Die so definierte Matrix kann als Blockmatrix dargestellt werden:

$$(A \otimes B) = \begin{pmatrix} a_{11}B & \dots & a_{1m}B \\ \vdots & & \vdots \\ a_{n1}B & \dots & a_{nm}B \end{pmatrix}$$

Satz 5.1.9 (Abgeschlossenheit gegen Maximumsbildung).

[Neu81] Sind F und G Verteilungsfunktionen stetiger PH-Verteilungen mit Parametern (α, T) bzw. (β, V) und Ordnung n bzw. m , so ist F_{max} die Verteilungsfunktion einer PH- (γ, L) verteilten Zufallsvariablen mit Ordnung $nm + n + m$. Für die Parameter gilt $\gamma = (\alpha \otimes \beta, \beta_0 \alpha, \alpha_0 \beta)$ bzw.

$$L = \begin{pmatrix} T \otimes I_m + I_n \otimes V & I_n \otimes \mathbf{v} & \mathbf{t} \otimes I_m \\ \mathbf{0} & T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & V \end{pmatrix},$$

wobei I_n die n -dimensionale Einheitsmatrix ist.

Im folgenden Beispiel werden einige wichtige Teilfamilien von azyklischen Phasentypverteilungen aufgeführt und durch das zugehörige Phasendiagramm veranschaulicht.

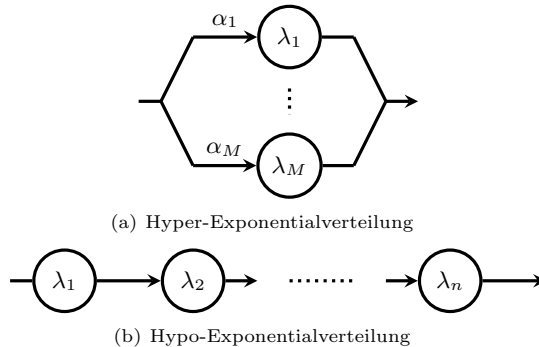
Beispiel 5.1.10.(a) **Hypo- / Hyper-Exponentialverteilungen:**

Abbildung 5.1.: Erweiterungen der Exponentialverteilung

Das einfachste Beispiel für eine Phasentypverteilung ist die Exponentialverteilung. Aus ihr lassen sich verschiedenste andere Phasentypverteilungen zusammensetzen. Ein Beispiel hierfür ist die Hyper-Exponentialverteilung. Dabei handelt es sich um eine Konvexkombination von Exponentialverteilungen mit unterschiedlichen λ -Werten (s. Abbildung 5.1 (a)). Für die Dichte- bzw. Verteilungsfunktion einer Hyper-Exponentialverteilten Zufallsvariablen X gilt

$$f_X(t) = \sum_{m=1}^M \alpha_m \lambda_m e^{-\lambda_m t} \quad \text{bzw.} \quad F_X(t) = 1 - \sum_{m=1}^M \alpha_m e^{-\lambda_m t} \quad (5.6)$$

mit $\sum_{m=1}^M \alpha_m = 1$ und $\lambda_m \in \mathbb{R}_{>0}$ für alle $1 \leq m \leq M$. Dagegen sind Hypo-Exponentialverteilungen Verteilungen der Summe stochastisch unabhängiger Zufallsvariablen mit möglicherweise unterschiedlichen λ -Werten. Im Phasendiagramm (s. Abbildung 5.1 (b)) wird dies durch eine Serie von Exponentialverteilungen erkennbar. Sind alle λ -Werte identisch, so handelt es sich um eine Erlangverteilung.

(b) **Verallgemeinerte Erlangverteilung:**

Die verallgemeinerte Erlangverteilung unterscheidet sich von einer Erlangverteilung dadurch, dass der zugehörige stochastische Prozess nach dem ersten Zustand mit Wahrscheinlichkeit α direkt in den absorbierenden Zustand übergeht und nur mit Gegenwahrscheinlichkeit $1 - \alpha$ in den nächsten nichtabsorbierenden Zustand wechselt.

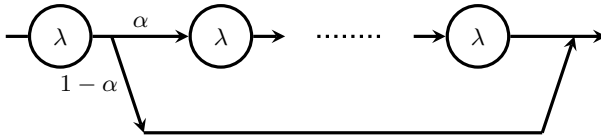


Abbildung 5.2.: Verallgemeinerte Erlangverteilung

(c) **Coxverteilungen:**

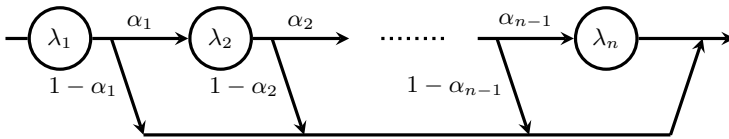


Abbildung 5.3.: Coxverteilung

Die Struktur von Coxverteilungen ähnelt im Phasendiagramm (s. Abbildung 5.3) der Struktur von Hypo-Exponentialverteilungen. Der Unterschied liegt darin, dass Übergänge in den nächsten Zustand nicht deterministisch erfolgen. Stattdessen kann der Prozess mit der jeweiligen Gegenwahrscheinlichkeit direkt in den absorbierenden Zustand übergehen. Sie spielen in der Literatur bei der Approximation von Verteilungen eine wesentliche Rolle. COX zeigt in [Cox55], dass jede Verteilung mit rationaler LST durch eine Coxverteilung dargestellt werden kann.

(d) **Erlang-Coxverteilungen:**

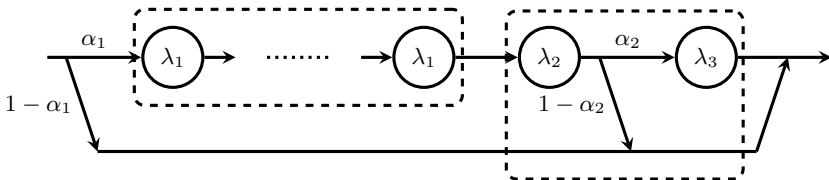


Abbildung 5.4.: Erlang-Coxverteilung

Erlang-Coxverteilungen sind eine Kombination aus einer Erlang- und einer Coxverteilung, wobei der Prozess auch direkt in den absorbierenden Zustand übergehen kann.

(e) **Erlang-Exp- / Exp-Erlangverteilungen:**

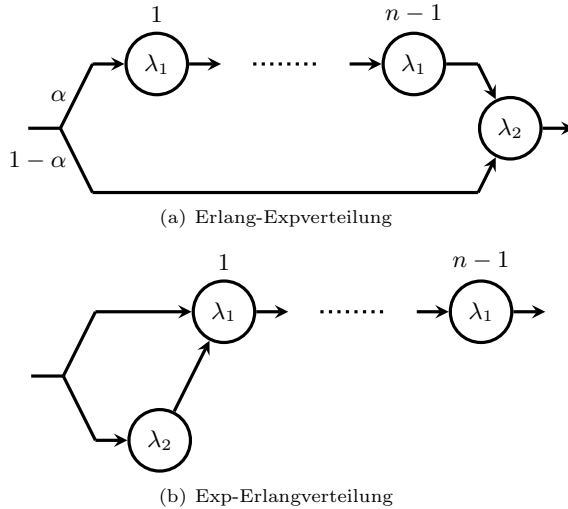


Abbildung 5.5.: Kombinationen von Exponential- und Erlangverteilungen

Erlang-Exp bzw. Exp-Erlangverteilungen sind Kombinationen aus einer Exponential- und einer Erlangverteilung (s. Abbildung 5.5). Der Prozess entscheidet hier probabilistisch, ob er nur die Zustände der Exponentialverteilung bzw. Erlangverteilung oder hintereinander die Zustände beider Verteilungen durchläuft.

(f) **Hyper-Erlangverteilungen:**

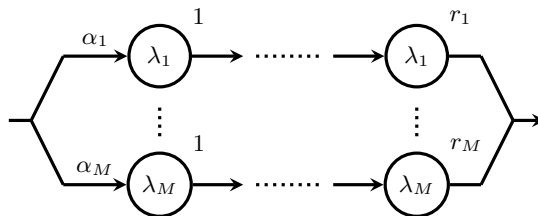


Abbildung 5.6.: Hyper-Erlangverteilung

Die Hyper-Erlangverteilung gehört ebenfalls zu den azyklischen Phasentypverteilungen. Eine Beschreibung der Beziehung zwischen Hyper-Erlangverteilung-

en, azyklischen und allgemeinen Phasentypverteilungen ist in [TBT06] in Abschnitt 2 zu finden. Gemäß Abbildung 5.6 ist eine Hyper-Erlangverteilung eine Konvexkombination von M unabhängigen, mit Wahrscheinlichkeiten $\alpha_1, \dots, \alpha_M$ gewichteten Erlangverteilungen. Dabei gilt $\sum_{m=1}^M \alpha_m = 1$. Die Anzahl der Phasen der m -ten Erlangverteilung ist gleich r_m , und es gilt ohne Beschränkung der Allgemeinheit $1 \leq r_1 \leq \dots \leq r_M$. Des Weiteren ist λ_m der Parameter der Exponentialverteilungen aus denen sich die m -te Erlangverteilung zusammensetzt. Die entsprechende Wahrscheinlichkeitsdichte (einer Zufallsvariablen X) ist gegeben durch

$$f_X(t) = \sum_{m=1}^M \alpha_m \frac{(\lambda_m t)^{r_m-1}}{(r_m-1)!} \lambda_m e^{-\lambda_m t} \quad (5.7)$$

und die zugehörige Verteilungsfunktion durch

$$F_X(t) = 1 - \sum_{m=1}^M \alpha_m \sum_{i=0}^{r_m-1} \frac{(\lambda_m t)^i}{i!} e^{-\lambda_m t}.$$

Außerdem gilt für die Momente

$$E[X^k] = \sum_{m=1}^M \alpha_m \frac{(r_m + i - 1)!}{\lambda_m^i (r_m - 1)!}, \quad k \in \mathbb{N}, \quad (5.8)$$

und die Anzahl aller Zustände $N = \sum_{m=1}^M r_m$.

5.2. Literatur

Die Approximation von (empirischen) Verteilungsfunktionen durch Phasentypverteilungen wurde in der Literatur bereits intensiv behandelt. Die vielen unterschiedlichen Verfahren lassen sich in drei Gruppen einteilen. Die meisten versuchen entweder, mit einer bestimmten Teilfamilie der Phasentypverteilungen (mehrere) Momente der zu approximierenden Verteilung zu treffen, oder Sie versuchen die Form bzw. den Verlauf der zugehörigen Verteilungsfunktion möglichst genau widerzuspiegeln. Letztlich betrachtet man in diesen beiden Gruppen bei der Approximation zwei verschiedene Arten des Abstandes zwischen Verteilungen bzw. Verteilungsfunktionen und versucht bezüglich dieses Abstandbegriffes eine möglichst genaue Approximation zu erreichen. Verfahren, welche eine Approximation bezüglich beider Abstandsbegriffe versuchen, bilden die dritte Gruppe.

Bevor man sich für ein bestimmtes Verfahren oder zumindest eine der drei Gruppen von Verfahren entscheidet, muss man sich damit beschäftigen, in welchem Rahmen die approximierten Verteilungsfunktionen angewendet und welche Fragen mittels ihrer beantwortet werden sollen. Versucht man allein die Momente einer

Verteilung anzunähern, so gibt einem dies keine Auskunft darüber, wie genau man den Verlauf der zugehörigen Verteilungsfunktion trifft. Andererseits führt eine gute Anpassung des Verlaufs einer Verteilungsfunktion im Allgemeinen nicht gleichzeitig zur Übereinstimmung der Momente der Verteilung. Dies ist dann nicht erforderlich, wenn man beispielsweise weniger an den Momenten sondern an den Quantilen einer Verteilung interessiert ist. Ideal scheinen also Verfahren, welche bezüglich beider Abstandsbegriffe approximieren. Welche Nachteile diese Verfahren jedoch teilweise haben, wird im Folgenden noch dargestellt.

Wir beschäftigen uns zunächst mit der großen Gruppe von Verfahren zur Anpassung der Momente von Phasentypverteilungen. Weit verbreitet sind in der Literatur Verfahren, welche sich auf die ersten drei Momente konzentrieren. Es ist naheliegend, dass man bei der Approximation nicht mehr Aufwand betreiben möchte als notwendig. Wieviele Momente man zur Übereinstimmung bringen muss, hängt vom Umfeld ab, in dem die angepassten Phasentypverteilungen eingesetzt werden sollen. Dies sind in der Literatur in der Regel Warteschlangensysteme. Dabei stellte sich heraus (s. [Whi84, JT91a]), dass es bei Warteschlangensystemen oft nicht ausreicht, nur die ersten beiden Momente zu berücksichtigen, da auch das dritte Moment ihr Verhalten stark beeinflussen kann. Das Verhalten misst man mit bestimmten Leistungsmerkmalen eines Warteschlangensystems wie zum Beispiel der mittleren Warteschlangenlänge. [JT91b] ist eine Fortsetzung der Arbeit in [JT91a]. Es wird allerdings auch festgestellt, dass bei Warteschlangensystemen die Anpassung von drei Momenten der von zwei Momenten nicht in jedem Fall überlegen sein muss. Statt sich nur auf die Anzahl anzupassender Momente zu konzentrieren, ist es manchmal wichtiger, das Augenmerk auf das gewählte Approximationsverfahren bzw. die benutzte Verteilungsfamilie zu legen. In [OHB03, OHB04] werden vier Kriterien für die Güte von Verfahren zur Momentenanpassung aufgelistet.

1. **Die Anzahl übereinstimmender Momente:** Obwohl die Anpassung für nur drei Momente weit verbreitet ist, ist es im Allgemeinen durchaus wünschenswert, mehr als drei Momente zu treffen.
2. **Die Effizienz des Verfahrens:** Die Effizienz betrifft vor allem die Laufzeit. Daneben ist es wünschenswert, wenn es für die Bestimmung der Parameter der Phasentypverteilung eine Lösung in geschlossener Form gibt.
3. **Die Anwendbarkeit des Verfahrens:** Positiv zu bewerten sind auch Verfahren, mit denen man eine möglichst große Klasse von Verteilungsfunktionen approximieren kann.
4. **Die Anzahl der Phasen,** welche für die Darstellung der Phasentypverteilung benötigt werden, sollte möglichst gering sein.

Zu Beginn dieses Abschnitts wurde bereits gezeigt, warum es plausibel ist, sich bei der Approximation von Verteilungen auf die Anpassung hinreichend vieler Momente

der Verteilungen zu konzentrieren. Notwendige Voraussetzung ist natürlich, dass die Momente überhaupt existieren. Ein klassisches Beispiel für eine Verteilung, die diese Voraussetzung nicht erfüllt, ist die Cauchyverteilung. In ihrem Fall existiert gar kein Moment. Viele für die Praxis relevante Verteilungen werden jedoch durch ihre Momente vollständig bestimmt. Es stellt sich auch die Frage, wie man die Approximationsfähigkeit einer Verteilungsfamilie beurteilt. In der Literatur zieht man hierfür in der Regel den *quadrirten Variationskoeffizienten* heran. Dieser ist für eine gegebene Zufallsvariable X definiert als

$$C_X^2 = \frac{\mathbf{Var}(X)}{\mathbf{E}[X]^2} = \frac{\mathbf{E}[X^2]}{\mathbf{E}[X]^2} - 1.$$

Für Exponentialverteilungen gilt immer $C^2 = 1$. Ein Verfahren zur Momentenannäherung, in denen für die Approximation diese Verteilungen verwendet werden, macht nur für zu approximierende Verteilungen Sinn, für die ebenfalls $C^2 = 1$ gilt. Zuletzt kann man sich fragen, wann die Anzahl der für die Darstellung der approximierten Verteilungen benötigten Phasen als gering erachtet werden kann. Bei der Beantwortung dieser Frage fällt der Begriff der *minimalen Phasenzahl*. Hiermit beschäftigen sich unter anderem BOBBIO ET AL. in [BHT05]. In der vorliegenden Arbeit wird die Fragestellung nicht detaillierter betrachtet.

Eine Vielzahl in der Literatur zu findender Verfahren hat gemein, dass sie bei mindestens einem der Kriterien schlecht abschneiden. Die folgenden beiden Verfahren haben den Makel, dass sie bei der Approximation nur die ersten beiden Momente zur Übereinstimmung bringen. In [SC75] nutzen SAUER und CHANDY für die Anpassung der Momente in Abhängigkeit vom quadrierten Variationskoeffizienten entweder eine Hyper-Exponentialverteilung (s. Abbildung 5.1 (a)) mit zwei Zweigen im Falle $C^2 > 1$ oder eine verallgemeinerte Erlangverteilung (s. Abbildung 5.2) im Falle $C^2 < 1$. Ebenfalls die verallgemeinerte Erlangverteilung nutzt für diesen Fall MARIE in [Mar80]. Für den Fall $C^2 > 1$ nutzt MARIE eine Coxverteilung (s. Abbildung 5.3). Des Weiteren findet man in der Literatur Arbeiten, die sich bei der Menge approximierbarer Verteilungen auf solche mit speziellen Eigenschaften zurückziehen, zum Beispiel von ALTIOK in [Alt85] oder von WHITT in [Whi82]. Letztere Arbeit, bzw. eine Fortsetzung dieser Arbeit, wird weiter unten noch detaillierter betrachtet.

JOHNSON und TAAFFE kommen dem Ziel ziemlich nahe, in allen vier Kriterien gut abzuschneiden. Ihre Beiträge sind auch für diese Arbeit von größerer Bedeutung. In [JT89] stellen Sie eine Lösung in geschlossener Form für die Anpassung dreier Momente vor. Dafür nutzen sie Hyper-Erlangverteilungen, bei denen die Ordnungen der Komponenten identisch sind. OSOGAMI und HARCHOL-BALTER verweisen in [OHB04] darauf, dass das Verfahren lediglich beim vierten Kriterium nicht optimal ist. Darauf aufbauend entwickeln JOHNSON und TAAFFE in [JT90b] Verfahren für drei verschiedene Familien von Phasentypverteilungen. Dies sind, neben

allgemeinen Phasentypverteilungen auf $\mathbb{R}_{>0}$, Coxverteilungen und aus zwei Zweigen bestehende Hyper-Erlangverteilungen. Die Verfahren basieren auf nichtlinearer Optimierung. In [JT90a] zeigen sie, welche unterschiedlichen Gestalten die mit diesen Verfahren erzeugten Dichtefunktionen haben können. Des Weiteren wird gezeigt, wie die vorgestellten Methoden genutzt werden können, um bestimmte Eigenschaften bzw. Charakteristiken der Verteilung zu beeinflussen. Diese Ideen sind grundsätzlich auch für diese Arbeit interessant. Allerdings nennen JOHNSON und TAAFFE auch Schwächen, welche die nichtlineare Optimierung mit sich bringt. Da diese u.a. sehr zeitintensiv sein kann, eignet sie sich eher für interaktive statt automatisierte Momentenanpassung. Letztere wird im Rahmen dieser Arbeit jedoch benötigt.

5.3. Beispiele

In diesem Abschnitt werden nun zwei Verfahren aus der Literatur, welche sich in ihrer Herangehensweise deutlich unterscheiden, detaillierter vorgestellt. Das zweite dieser Verfahren, d.h. insbesondere die grundlegende Modellierung, wurde von uns leicht modifiziert und schließlich für die Approximation empirischer Quellverspätungsverteilungen verwendet. Beide Verfahren eint, dass sie eher die Gestalt einer gegebenen Verteilungsfunktion approximieren als beispielsweise deren Momente.

5.3.1. Hyper-Exponentialverteilung

FELDMANN und WHITT stellen in [FW98] ein Approximationsverfahren vor, welches als grundlegende Verteilungsfamilie die Hyper-Exponentialverteilungen verwendet. Der Aufwand für die Bestimmung der Parameter der approximativen Verteilungsfunktion ist dabei, beispielsweise im Vergleich zu Momentenanpassungsverfahren, relativ gering. In Algorithmus 5.3.1 wird dieses (rekursive) Verfahren aus [FW98] vorgestellt. Einen entscheidenden Nachteil stellt der vergleichsweise kleine Bereich von Verteilungen dar, deren Verteilungsfunktionen approximiert werden können. Es wird davon ausgegangen, dass die zu approximierende Verteilungsfunktion F zu einer Long-Tail Verteilung gehört und eine monotone Dichtefunktion f besitzt. Diese Eigenschaften sind für das Approximationsverfahren selbst nicht notwendig. Sie beeinflussen aber die Güte der Approximation. Diese Anforderungen sind auch der Grund, warum sich das Verfahren nicht für unsere Zwecke einsetzen lässt. Wegen der Schlichtheit des Verfahrens, soll es an dieser Stelle trotzdem diskutiert werden. Im Folgenden wird die theoretische Grundlage des Verfahrens erläutert.

Definition 5.3.1 (Long-Tail Verteilung).

Es sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion und $F^c : \mathbb{R} \rightarrow [0, 1]$ die zugehörige komplementäre Verteilungsfunktion mit $F^c(t) := 1 - F(t)$ für alle $t \in \mathbb{R}$.

5. Approximation von Phasentypverteilungen

Dann ist F eine **Long-Tail Verteilung**, wenn ihr Schweif (engl. tail) langsamer als exponentiell fällt, d.h.

$$\forall \lambda \in \mathbb{R}_{>0} : e^{\lambda t} \cdot F^c(t) \rightarrow \infty$$

für $t \rightarrow \infty$. Im Gegensatz dazu spricht man von einer **Short-Tail Verteilung**, wenn ein $\lambda \in \mathbb{R}_{>0}$ existiert mit $e^{\lambda t} \cdot F^c(t) \rightarrow 0$ für $t \rightarrow \infty$.

Wird im Folgenden auf den Schweif einer komplementären Verteilungsfunktion F^c Bezug genommen, so ist immer der Bildbereich von F^c gemeint, in dem sich die Long-Tail bzw. Short-Tail Eigenschaft der Verteilungsfunktion F bemerkbar macht, also $F^c(I) := \{F^c(t) \mid t \in I\}$ für ein Intervall $I := (x_0, \infty)$ mit hinreichend großem $x_0 \in \mathbb{R}$.

Definition 5.3.2 (Vollständige Monotonie).

[Fel71] Eine Funktion f auf dem Intervall $(0, \infty)$ heißt **vollständig monoton**, wenn

- (i) die Ableitung $f^{(n)}$ für alle Ordnungen $n \in \mathbb{N}$ existiert und
- (ii) $(-1)^n f^{(n)}(t) \geq 0$ für alle $n \in \mathbb{N}$ und $t > 0$.

Zusätzlich wird definiert $f^{(n)}(0) := \lim_{t \rightarrow 0} f^{(n)}(t)$. Dieser Grenzwert kann sowohl endlich als auch unendlich sein.

Mit dem folgenden Theorem erläutern FELDMANN und WHITT die Fähigkeit der Familie der Hyper-Exponentialverteilungen zur Approximation anderer Verteilungen.

Satz 5.3.3.

[FW98] Jede vollständig monotone Dichtefunktion f lässt sich als Komposition von Dichten der Exponentialverteilung darstellen, d.h.

$$f(t) = \int_0^\infty \lambda e^{-\lambda t} dG(\lambda)$$

für alle $t \in \mathbb{R}_{\geq 0}$ mit geeignet gewählter Verteilungsfunktion G .

Satz 5.3.3 impliziert die Aussage des folgenden Satzes.

Satz 5.3.4.

[FW98] Ist F eine Verteilungsfunktion mit vollständig monotoner Dichtefunktion f , so existieren Verteilungsfunktionen H_n der Hyper-Exponentialverteilung, d.h.

$$H_n(t) = \sum_{m=1}^{M_n} \alpha_{n_m} (1 - e^{-\lambda_{n_m} t}) \quad \text{für alle } t \geq 0,$$

mit $\lambda_{n_m} \in \mathbb{R}_{>0} \cup \{+\infty\}$, $\sum_{m=1}^{M_n} \alpha_{n_m} = 1$, sodass $H_n \rightarrow F$ für $n \rightarrow \infty$.

Zur Verteilungsfunktion H einer Hyper-Exponentialverteilung gehört die Dichtefunktion

$$h(t) := \sum_{m=1}^M \alpha_m \cdot \lambda_m e^{-\lambda_m t} \quad \text{für alle } t \geq 0, \quad (5.9)$$

mit $\sum_{m=1}^M \alpha_m = 1$ und $\alpha_m, \lambda_m \in \mathbb{R}_{>0}$ für alle $1 \leq m \leq M$. Die Dichtefunktion h ist also eine Konvexkombination von M Dichtefunktionen der Exponentialverteilung und offensichtlich monoton. Es gelte in (5.9) o.B.d.A. $\lambda_1 < \lambda_2 < \dots < \lambda_M$. Dies hat zur Folge, dass der Schweif von Komponenten der zugehörigen Funktion $H^c := \mathbf{1}_{\mathbb{R}_{\geq 0}} - H$ mit höherem Index m entsprechend schneller fällt. Dies ermöglicht ein rekursives Vorgehen bei der Anpassung von H .

Input : Zu approximierende Verteilungsfunktion F bzw. zugehörige komplementäre Verteilungsfunktion F^c und c_1, \dots, c_M, b (wie unten erläutert)

- 1 $\lambda_1 = \frac{1}{(b-1)c_1} \ln \left(\frac{F^c(c_1)}{F^c(b \cdot c_1)} \right);$
- 2 $\alpha_1 = F^c(c_1) \cdot e^{\lambda_1 c_1};$
- 3 $F_1^c = F^c;$
- 4 **for** $2 \leq m \leq M - 1$ **do**
- 5 $F_m^c(x) = F_{m-1}^c(x) - \alpha_{m-1} \cdot e^{-\lambda_{m-1} x};$
- 6 $\lambda_m = \frac{1}{(b-1)c_m} \ln \left(\frac{F_m^c(c_m)}{F_m^c(b \cdot c_m)} \right);$
- 7 $\alpha_m = F_m^c(c_m) \cdot e^{\lambda_m c_m};$
- 8 $F_M^c(x) = F_{M-1}^c(x) - \alpha_{M-1} \cdot e^{-\lambda_{M-1} x};$
- 9 $\alpha_M = 1 - \sum_{m=1}^{M-1} \alpha_m;$
- 10 $\lambda_M = \frac{1}{c_M} \ln \left(\frac{\alpha_M}{F_M^c(c_M)} \right);$

Algorithmus 5.3.1 : Feldmann-Whitt

Bevor man mittels des Algorithmus 5.3.1 die Parameter der Hyper-Exponentialverteilung rekursiv bestimmen kann, muss die Anzahl M der Parametertupel festgelegt werden. Dabei steht M zugleich für die Anzahl der Stellen $0 < c_M < c_{M-1} < \dots < c_1$, mit $c_m \in \mathbb{R}$ für $1 \leq m \leq M$, an denen die angepasste Hyper-Exponentialverteilung mit F übereinstimmen soll. Dies geschieht so, dass die Raten c_m/c_{m+1} hinreichend groß sind. Zusätzlich wird ein $b \in \mathbb{R}$ gewählt, sodass $1 < b < c_m/c_{m+1}$. Auf diese Weise sind mit $b \cdot c_m$, $1 \leq m \leq M$, M weitere Stellen gegeben, an denen eine Übereinstimmung mit F erreicht werden soll. FELDMANN und WHITT geben als Beispiel für die Wahl der Schnittpunkte $c_m = c_1 10^{-(m-1)}$ mit $2 \leq m \leq M$

an. Für die Raten heißt das $c_m/c_{m+1} = 10$. So kann beispielsweise $b = 2$ gewählt werden.

In Algorithmus 5.3.1 wird mit der Bestimmung des Parametertupels (λ_1, α_1) begonnen. Dieses erhalten wir aus den Gleichungen $\alpha_1 e^{-\lambda_1 c_1} = F^c(c_1)$ sowie $\alpha_1 e^{-\lambda_1 b \cdot c_1} = F^c(b \cdot c_1)$. Hintergrund des weiteren Vorgehens ist die Tatsache, dass für hinreichend großes λ_2 die Summe $\sum_{m=2}^M e^{-\lambda_m t}$ für hinreichend großes t gegenüber $\alpha_1 e^{-\lambda_1 t}$ vernachlässigbar ist. Daher kann (λ_1, α_1) unabhängig von den anderen Parametern der Funktion H_M gewählt werden. Diese erste Anpassung von H bzw. H^c an die Verteilungsfunktion F bzw. die Funktion $F^c := \mathbf{1}_{\mathbb{R}_{\geq 0}} - F$ findet also anschaulich im Schweif von F^c statt. Nach Subtraktion des Terms $\alpha_1 e^{\lambda_1 t}$ sowohl von H^c also auch F^c , wird die zweite Komponente an den verbliebenen Schweif von F^c angepasst bzw. das Parametertupel (λ_2, α_2) bestimmt (s. Algorithmus 5.3.1). Dies geschieht wieder mit der Argumentation, dass für hinreichend großes λ_3 der Ausdruck $\sum_{m=3}^M e^{-\lambda_m t}$ für ebenfalls hinreichend großes t gegenüber $\alpha_2 e^{-\lambda_2 t}$ vernachlässigbar ist und somit unabhängig von den noch nicht bestimmten Parametertupeln gewählt werden kann. So wird weiter fortgefahren. Die Bestimmung der Parametertupel erfolgt dabei, mit Ausnahme des letzten Tupels (s. Algorithmus 5.3.1), immer über die Gleichungen $\alpha_m e^{-\lambda_m c_m} = F_m^c(c_m)$ sowie $\alpha_m e^{-\lambda_m b \cdot c_m} = F_m^c(b \cdot c_m)$.

Gilt $0 < \alpha_m < 1$ für alle $1 \leq m \leq M$, so ist das Resultat des Verfahrens eine Verteilungsfunktion H aus der Familie der Hyper-Exponentialverteilungen mit $H^c(c_m) > F^c(c_m)$ für alle $1 \leq m \leq M$ und $H^c(b \cdot c_m) > F^c(b \cdot c_m)$ für alle $1 \leq m \leq M - 1$. Für $m = M$ muss letztere Ungleichung nicht gelten, da $b \cdot c_M$ nicht mehr als Schnittpunkt in das Anpassungsverfahren einging. Ist F^c eine Long-Tail Verteilung, so gibt es jedoch ein $t_0 \in \mathbb{R}_{>0}$, so dass $H^c(t) < F^c(t)$ für alle $t > t_0$. Der Punkt c_1 sollte daher so gewählt werden, dass er außerhalb des Bereichs liegt, der für die Anpassung von Interesse ist.

Die oben bereits erwähnte Forderung der Monotonie der Dichtefunktionen wird von den im Anwendungsbereich dieser Arbeit vorkommenden Verteilungen (s. beispielsweise Kapitel 4) im Allgemeinen nicht erfüllt. Es wäre denkbar, dass sich das Verfahren modifizieren lässt, sodass es auf Teilintervallen von $\mathbb{R}_{\geq 0}$ eingesetzt werden kann, auf denen die jeweilige Dichte streng monoton ist. Dieser Ansatz wurde im Rahmen der vorliegenden Arbeit jedoch nicht intensiver verfolgt. BÜKER gelingt es in seiner Dissertation [Bü10], für die von ihm verwendete Verteilungsfamilie ein Approximationsverfahren zu entwickeln, mittels dessen eine gegebene Verteilungsfunktion abschnittsweise, d.h. auf Teilintervallen der reellen Zahlen, approximiert werden kann.

5.3.2. EM-Algorithmus für Hyper-Erlangverteilungen

Grundlage für das Vorgehen in der vorliegenden Arbeit ist ein Ergebnis von THÜMMLER, BUCHHOLZ und TELEK. In [TBT06] wird ein EM-Algorithmus (s. [DLR77])

vorgestellt, mit dem eine Hyper-Erlangverteilung so angepasst wird, dass sie hinreichend gut der Verteilung von gegebenen empirischen Daten entspricht. Ausgangspunkt ist das folgende Theorem, welches besagt, dass jede Verteilungsfunktion einer nicht-negativen reellen Zufallsvariablen in all ihren Stetigkeitsstellen beliebig genau durch die Verteilungsfunktion einer Hyper-Erlangverteilung approximiert werden kann.

Satz 5.3.5.

[JT88] Es sei $E_{\lambda,k}$ die Verteilungsfunktion einer k -stufigen Erlangverteilung, $k \in \mathbb{N}$, mit Parameter $\lambda \in \mathbb{R}_{>0}$ und sei F die Verteilungsfunktion einer positiven reellen Zufallsvariablen. Weiter sei die Verteilungsfunktion $F_{\lambda,M}$ definiert durch

$$F_{\lambda,M}(t) = a_{\lambda,M} \cdot \sum_{k=1}^M b_{\lambda,k} \cdot E_{\lambda,k}(t)$$

für alle $t \in \mathbb{R}_{\geq 0}$ mit $a_{\lambda,M} := F\left(\frac{M}{\lambda}\right)^{-1}$ und $b_{\lambda,k} := F\left(\frac{k}{\lambda}\right) - F\left(\frac{k-1}{\lambda}\right)$. Dann gilt: Ist die Verteilungsfunktion F in $t \in \mathbb{R}_{\geq 0}$ stetig, so folgt

$$\lim_{\lambda \rightarrow \infty} \lim_{M \rightarrow \infty} F_{\lambda,M}(t) = F(t) .$$

Es folgen nun zwei Abschnitte über zwei unterschiedliche Approximationsverfahren für Hyper-Erlangverteilungen, die beide im Rahmen dieser Arbeit ihre Anwendung finden. Im ersten der beiden Abschnitte wird erläutert, inwiefern man zu einer gegebenen, d.h. zu approximierenden, Verteilung eine Hyper-Erlangverteilung findet, deren erste drei Momente mit den Momenten der gegebenen Verteilung übereinstimmen. Diese Vorgehensweise ist, im Vergleich zum im zweiten Abschnitt diskutierten EM-Algorithmus, deutlich simpler. Allerdings lässt sich mit dem EM-Algorithmus die Approximationsgüte besser steuern. Mehr dazu folgt in Abschnitt 6.6.

Momentenanpassung

Das Problem der Momentenanpassung (*engl.* moment matching) für Hyper-Erlangverteilungen wurde in der Literatur bereits intensiv untersucht. Einen wesentlichen Teil dazu beigetragen haben JOHNSON und TAAFFE in [Joh93] sowie [JT90b, JT89]. Es wurden u.a. Bedingungen für die Lösbarkeit des Problems aufgestellt. Diese sollen hier erläutert werden.

Um die Problemstellung zu spezifizieren, sei eine Zufallsvariable X mit zugehöriger Verteilung \mathbf{P}_X gegeben. Außerdem bezeichne $\mu_k := \mathbf{E}[X^k]$, $1 \leq k \leq 3$, das k -te Moment von \mathbf{P}_X . Aufgabe der Momentenanpassung ist es, (mindestens) eine Hyper-Erlangverteilung zu finden, deren erste drei Momente mit diesen μ_k übereinstimmen. Dazu wird zunächst eine Hyper-Erlangverteilung mit zwei Zweigen, d.h. $M = 2$, mit Gewichten α_1, α_2 sowie Phasenanzahl r_1 bzw. r_2 betrachtet (s. (5.7)

auf Seite 38). Es gelte ohne Beschränkung der Allgemeinheit $r_1 \leq r_2$ und es sei $n^* \in \mathbb{N}$ die kleinste natürliche Zahl mit

$$n^* > \max \left\{ \frac{\mu_1^2}{\mu_2 - \mu_1^2}, \frac{\mu_2^2}{\mu_1 \mu_3 - \mu_2^2} - 1 \right\}. \quad (5.10)$$

Diese Definition stammt von THÜMLER ET AL. in [TBT06]. Sie ist äquivalent zur Definition in [Joh93]. Dort wird n^* über die *standardisierten Momente*

$$c := \frac{\sqrt{\mathbf{Var}(X)}}{\mathbf{E}[X]} \quad \text{und} \quad \gamma := \frac{\mathbf{E}[(X - \mathbf{E}[X])^3]}{\mathbf{Var}(X)^{3/2}}$$

definiert. In Abhängigkeit dieser Momente ist $n^* \in \mathbb{N}$ die kleinste natürliche Zahl mit $n^* > 1/c^2$ und

$$n^* > \frac{-\gamma + 1/c^3 + 1/c + 2c}{\gamma - (c - 1/c)}.$$

Dieser Zusammenhang wird unter anderem im Beweis zu Theorem 2 in [JT89] deutlich. Hierauf aufbauend werden in [Joh93] bezüglich der Lösbarkeit der Momentenanpassung vier Fälle unterschieden, gegeben $\mu_1 > 0$.

- (i) $r_1, r_2 < n^*$: Es kann keine exakte Lösung gefunden werden.
- (ii) $r_1 < n^* \leq r_2$: Es existiert (mindestens) eine Lösung.
- (iii) $n^* \leq r_1 = r_2$: Es existiert eine eindeutige Lösung.
- (iv) $n^* \leq r_1 < r_2$: Es existieren (mindestens) zwei Lösungen.

Für Fall (iii) hatten bereits JOHNSON und TAAFFE (s.o.) eine Lösung in geschlossener Form angegeben. Die Fälle (ii) und (iv) sind nur numerisch lösbar. Hierfür zeigen THÜMLER ET AL., wie man eine Lösung generieren kann. Sie geben ein Polynom fünften Grades an, dessen Nullstellen (numerisch) bestimmt werden müssen (s. [TBT06]).

Bisher wurde die Konvexkombination nur zweier Erlangverteilungen, also $M = 2$, betrachtet. Sind im Fall $M > 2$ die Parameter der anderen Zweige bekannt, so kann man das Problem der Momentenanpassung auf den Fall $M = 2$ reduzieren (s. [TBT06]). Man zieht von den ursprünglich zu treffenden Momenten den Anteil der übrigen Zweige ab. Dies führt zur Definition der *reduzierten Momente*. Die beiden anzupassenden Zweige seien m_1 und m_2 mit $1 \leq m_1 < m_2 \leq M$. Das j -te reduzierte Moment ist dann definiert als

$$\tilde{\mu}_j = \left(\mu_j - \sum_{\substack{m=1 \\ m_1 \neq m \neq m_2}}^M \alpha_m \frac{(r_m + j - 1)!}{(r_m - 1)! \cdot \lambda_m^j} \right) \frac{1}{\beta} \quad (5.11)$$

mit $\beta := \alpha_{m_1} + \alpha_{m_2}$ (vgl. (5.8)). Eine notwendige Bedingung dafür, dass diese reduzierten Momente zu einer Verteilung auf $\Omega = \mathbb{R}_{\geq 0}$ gehören, ist $\min\{\tilde{\mu}_1, \tilde{\mu}_2 - \tilde{\mu}_1^2, \tilde{\mu}_1\tilde{\mu}_3 - \tilde{\mu}_2^2\} > 0$. Ist diese Bedingung erfüllt, können analog zum Fall $M = 2$ die Parameter einer Mischung von zwei Erlangverteilungen so angepasst werden, dass die Momente mit den reduzierten Momenten $\tilde{\mu}_k$, $1 \leq k \leq 3$, übereinstimmen.

EM Algorithmus

Grundlage für diesen Abschnitt ist das folgende Modell.

$$f_{\Theta}(t) = \sum_{m=1}^M \alpha_m \cdot f_{m, \theta_m}(t) \quad \text{für alle } t \in \mathbb{R}_{\geq 0} \quad (5.12)$$

Dabei ist

- $\Theta = (\alpha_1, \dots, \alpha_M, \theta_1, \dots, \theta_M)$,
- $\alpha_1, \alpha_2, \dots, \alpha_M \in \mathbb{R}_{\geq 0}$ mit $\sum_{m=1}^M \alpha_m = 1$ sowie
- f_{m, θ_m} eine durch den Vektor $\theta_m \in \mathbb{R}^d$ (im Folgenden $d = 1$) parametrisierte Dichtefunktion einer Erlangverteilung (mit Ordnung r_m).

Es wird also eine Mischdichte betrachtet, welche eine Konvexkombination aus M Erlangdichten darstellt. Die Werte α_m , $m = 1, \dots, M$, können als Wahrscheinlichkeit interpretiert werden, dass Komponente m gewählt wird. Dieser Gedanke wird weiter unten nochmals aufgegriffen. Des Weiteren sei $\mathcal{D} = \{x_1, \dots, x_K\}$ eine von (5.12) „erzeugte“ Datenmenge (Stichprobe). Damit ist gemeint, dass die Elemente dieser Menge Realisierungen von i.i.d. verteilten Zufallsvariablen $X_k : \Omega \rightarrow \mathbb{R}_{\geq 0}$, $1 \leq k \leq K$, mit Dichte (5.12) sind. Der entsprechende log-Likelihood Term lautet dann

$$\log L(\Theta; \mathcal{D}) = \log \prod_{k=1}^K f_{\Theta}(x_k) = \sum_{k=1}^K \log \left(\sum_{m=1}^M \alpha_m \cdot f_{m, \theta_m}(x_k) \right). \quad (5.13)$$

Dieser Ausdruck und dessen Maximierung kann vereinfacht werden, wenn man die Datenmenge \mathcal{D} als *unvollständig* ansieht und die Existenz nicht beobachteter Informationen bzw. Daten $y_k \in \{1, \dots, M\}$, $1 \leq k \leq K$, annimmt. Dabei steht $y_k = m$ für die Information, dass die Realisierung x_k von der Komponente f_m aus (5.12) erzeugt wurde. Sind die Werte der y_k bekannt, vereinfacht sich Gleichung (5.13) zu

$$\log L(\Theta; \mathcal{D}, y) = \sum_{k=1}^K \log(\alpha_{y_k} \cdot f_{y_k, \theta_{y_k}}(x_k)) \quad (5.14)$$

mit $y = (y_1, \dots, y_K)$. Statt also die Mischdichte zu betrachten, kann mit der Information der y_k der Zweig (vergl. Abbildung 5.6) für den log-Likelihood Ausdruck

verwendet werden, mit dessen Dichte $f_{y_k, \theta_{y_k}}$ die Realisierung x_k erzeugt wurde. Da diese y_k aber gerade nicht bekannt sind, werden sie als Realisierungen von Zufallsvariablen Y_k , $1 \leq k \leq K$, betrachtet und ihre Zählidichte $\mathbf{P}(Y_k = y_k)$ berechnet. Dafür werden zunächst die Parameter der Dichte aus (5.12) geschätzt. Es wird also angenommen, dass $\hat{\Theta} = (\hat{\alpha}_1, \dots, \hat{\alpha}_M, \hat{\theta}_1, \dots, \hat{\theta}_M)$ die passenden Parameter sind. Mit diesem $\hat{\Theta}$ können für alle k und alle m die Komponenten $f_{m, \hat{\theta}_m}(x_k)$ aus (5.12) berechnet werden. Mit dem Satz von BAYES kann schließlich die erwähnte Zählidichte der nicht beobachteten Daten berechnet werden:

$$\begin{aligned} \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k) &= \frac{f_{\hat{\Theta}}(x_k \mid Y_k = y_k) \cdot \mathbf{P}_{\hat{\Theta}}(Y_k = y_k)}{\sum_{m=1}^M f_{\hat{\Theta}}(x_k \mid Y_k = m) \cdot \mathbf{P}_{\hat{\Theta}}(Y_k = m)} \\ &= \frac{f_{y_k, \hat{\theta}_{y_k}}(x_k) \cdot \hat{\alpha}_{y_k}}{\sum_{m=1}^M f_{m, \hat{\theta}_m}(x_k) \cdot \hat{\alpha}_m} \end{aligned} \quad (5.16)$$

Hier sei daran erinnert, dass $\hat{\alpha}_{y_k} = \mathbf{P}_{\hat{\Theta}}(Y = y_k)$ gilt, da $\hat{\alpha}_{y_k}$ gerade die Wahrscheinlichkeit ist, dass die y_k -te Komponente der Mischdichte gewählt wird.

Bemerkung 5.3.6 (Überblick über das Modell).

Das Modell sieht also wie folgt aus. Es werden i.i.d.-verteilte Zufallsvariablen $Z_k := (X_k, Y_k)$, $1 \leq k \leq K$ betrachtet. Dabei gilt

- $\mathbf{P}(Y_k = m) = \alpha_m$ für alle $1 \leq k \leq K$ und $1 \leq m \leq M$,
- $\mathbf{P}_{X_k \mid Y_k=m}$ besitzt die Dichte f_{m, θ_m} ,
- $\mathbf{P}(X_k \leq x, Y_k = m) = \alpha_m \cdot \int_0^x f_{m, \theta_m}(t) dt$ (gemeinsame Verteilung) und
- $\mathbf{P}(X_k \leq x) = \int_0^x f_{\Theta}(t) dt = \int_0^x \sum_{m=1}^M \alpha_m f_{m, \theta_m}(t) dt$ (Randverteilung).

Die Z_k besitzen obige gemeinsame Verteilung und die nach dieser Verteilung gezogene Stichprobe mit Daten $z_k := (x_k, y_k)$ vollständige Information. Beobachtbar ist allerdings nur $(X_1, \dots, X_k) = (x_1, \dots, x_k)$.

Es sei $\mathbf{P}_{\hat{\Theta}}(Y = y \mid \mathcal{D}) := \mathbf{P}_{\hat{\Theta}}(Y = y \mid (X_1, \dots, X_k) = (x_1, \dots, x_k))$. Insgesamt gilt für den Vektor $y \in \{1, \dots, M\}^K$ der unbeobachteten Daten bzw. untereinander unabhängigen Realisierungen der Zufallsvariablen $Y = (Y_1, \dots, Y_K)$

$$\mathbf{P}_{\hat{\Theta}}(Y = y \mid \mathcal{D}) = \prod_{k=1}^K \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k). \quad (5.17)$$

Dabei wurde zum einen ausgenutzt, dass die $Z_k = (X_k, Y_k)$ i.i.d.-verteilt sind und damit für alle $1 \leq k \leq K$ die Zufallsvariable Y_k unabhängig von X_l , $l \neq k$, ist. Außerdem sind die Y_k bedingt unabhängig gegebenen X_1, \dots, X_K .

Es folgt an dieser Stelle der erste Schritt des EM Algorithmus, die Berechnung des Erwartungswertes des log-Likelihood Ausdrucks mit vollständigem Datensatz (vgl. (5.14)) bezüglich der Zufallsvariablen Y , gegebenem Datensatz \mathcal{D} und aktuellem Vektor der Schätzungen für die Parameter $\hat{\Theta}$:

$$\begin{aligned} Q(\Theta, \hat{\Theta}) &:= \mathbf{E}_{\hat{\Theta}} [\log L(\Theta; \mathcal{D}, Y) \mid \mathcal{D}] \\ &= \sum_{y \in \{1, \dots, M\}^K} \log L(\Theta; \mathcal{D}, y) \cdot \mathbf{P}_{\hat{\Theta}}(Y = y \mid \mathcal{D}) \end{aligned}$$

Mit Verwendung von (5.14) und (5.17) erhält man

$$Q(\Theta, \hat{\Theta}) = \sum_{y \in \{1, \dots, M\}^K} \sum_{k=1}^K \log(\alpha_{y_k} \cdot f_{y_k, \theta_{y_k}}(x_k)) \cdot \prod_{i=1}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i). \quad (5.18)$$

THÜMMLER ET AL. beschreiben im Anhang von [TBT06] wie man (5.18) durch Umordnung der Summen und Produkte vereinfachen kann. Dafür wird in (5.18) die Indikatorfunktion $\delta_{x,y}$ eingeführt. Es gilt $\delta_{x,y} = 1$, wenn $x = y$, und $\delta_{x,y} = 0$ sonst.

$$\begin{aligned} Q(\Theta, \hat{\Theta}) &= \sum_{y \in \{1, \dots, M\}^K} \sum_{k=1}^K \sum_{m=1}^M \delta_{y_k, m} \log(\alpha_m f_{m, \theta_m}(x_k)) \cdot \prod_{i=1}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \\ &= \sum_{k=1}^K \sum_{m=1}^M \log(\alpha_m \cdot f_{m, \theta_m}(x_k)) \sum_{y \in \{1, \dots, M\}^K} \delta_{y_k, m} \prod_{i=1}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \end{aligned}$$

Im nächsten Schritt wird nur noch die Summe über die $y \in \{1, \dots, M\}^K$ betrachtet und vereinfacht.

$$\begin{aligned} &\sum_{y \in \{1, \dots, M\}^K} \delta_{y_k, m} \cdot \prod_{i=1}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \\ &= \sum_{y_1=1}^M \sum_{y_2=1}^M \dots \sum_{y_K=1}^M \delta_{y_k, m} \prod_{i=1}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \\ &= \sum_{y_k=1}^M \delta_{y_k, m} \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k) \cdot \\ &\quad \sum_{y_1=1}^M \dots \sum_{y_{k-1}=1}^M \sum_{y_{k+1}=1}^M \dots \sum_{y_K=1}^M \prod_{i=1, i \neq k}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \\ &= \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \cdot \end{aligned}$$

$$\begin{aligned}
 & \sum_{y_1=1}^M \dots \sum_{y_{k-1}=1}^M \sum_{y_{k+1}=1}^M \dots \sum_{y_K=1}^M \prod_{i=1, i \neq k}^K \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \\
 &= \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \prod_{i=1, i \neq k}^K \sum_{y_i=1}^M \mathbf{P}_{\hat{\Theta}}(Y_i = y_i \mid X_i = x_i) \\
 &= \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \prod_{i=1, i \neq k}^K 1 \\
 &= \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) .
 \end{aligned}$$

So erhält man für $Q(\Theta, \hat{\Theta})$ schließlich folgende Darstellung.

$$\begin{aligned}
 Q(\Theta, \hat{\Theta}) &= \sum_{k=1}^K \sum_{m=1}^M \log(\alpha_m \cdot f_{m, \theta_m}(x_k)) \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \\
 &= \sum_{k=1}^K \sum_{m=1}^M \log(\alpha_m) \cdot \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \\
 &\quad + \sum_{k=1}^K \sum_{m=1}^M \log(f_{m, \theta_m}(x_k)) \cdot \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k)
 \end{aligned} \tag{5.20}$$

Um im nächsten Schritt des EM Algorithmus den zuvor berechneten Erwartungswert zu maximieren, können die beiden Terme aus (5.20) getrennt voneinander maximiert werden. Um ein α_m zu finden, welches den ersten Term aus (5.20) maximiert, kann die Methode der Lagrange-Multiplikatoren (vgl. [KM97]) verwendet werden und man erhält

$$\alpha_m = \frac{1}{K} \sum_{k=1}^K \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) .$$

Für die Bestimmung des θ_m , welches den zweiten Term maximiert, benötigt man Informationen über die Komponenten f_{m, θ_m} der Mischdichte aus (5.12). In [TBT06] entsprechen diese Komponenten Erlangdichten mit r_m Phasen, es gilt also

$$f_{m, \lambda_m}(x_k) = \frac{(\lambda_m x_k)^{r_m - 1}}{(r_m - 1)!} \lambda_m e^{-\lambda_m x_k} \tag{5.21}$$

und $\Theta = (\alpha_1, \dots, \alpha_M, \lambda_1, \dots, \lambda_M)$. Mit diesem Wissen kann die partielle Ableitung nach λ_m des zweiten Term aus (5.20) berechnet und anschließend Null gesetzt werden:

$$\sum_{k=1}^K \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \frac{\partial}{\partial \lambda_m} \log(f_{m, \lambda_m}(x_k)) \stackrel{!}{=} 0$$

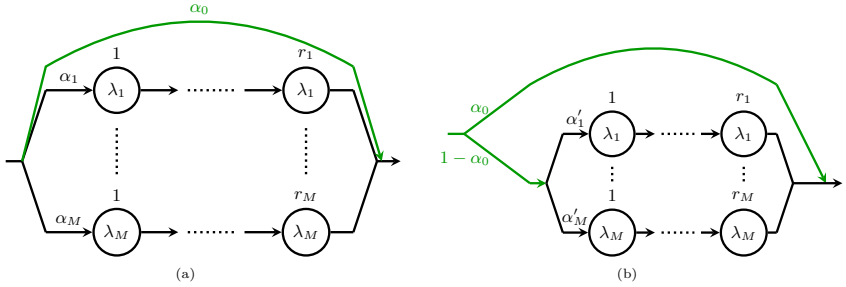


Abbildung 5.7.: Modifizierte Hyper-Erlangverteilung mit positivem Gewicht auf der Null

So erhält man

$$\lambda_m = \frac{r_m \sum_{k=1}^K \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k)}{\sum_{k=1}^K \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \cdot x_k}.$$

Im Folgenden wird dann $\hat{\Theta} := \Theta$ gesetzt und das Verfahren mit den neuen Schätzungen für die Parameter wiederholt.

5.4. Hyper-Erlangverteilung mit Gewichtung der Null

Für die Anwendung in der vorliegenden Arbeit wird ein Modell benötigt, welches dem Auftreten von Daten $x_k = 0$ in der Stichprobe \mathcal{D} eine positive Wahrscheinlichkeit zuweisen kann. Dafür wird die obige Annahme, dass die Daten Realisierungen von unabhängig identisch verteilten Zufallsvariablen einer Konvexkombination von Erlangverteilungen sind, aufgeweicht. Es handelt sich jedoch immer noch um eine Phasentypverteilung (s. Abbildung 5.7).

Im Folgenden ist $x \geq 0$. Für den Fall $x > 0$ ist die Wahrscheinlichkeitsdichte

$$f_{\Theta}(x) := \sum_{m=1}^M \alpha'_m f_{m, \theta_m}(x) \quad (5.22)$$

mit $\sum_{m=1}^M \alpha'_m = 1$ gemäß dem vorherigen Modell definiert. Das neue Modell besitzt jetzt den Parametervektor $\Theta = (\alpha_0, \alpha_1, \dots, \alpha_M, \lambda_1, \dots, \lambda_M)$. Dabei gilt für $1 \leq k \leq K$ bzw. $1 \leq m \leq M$

$$\alpha_0 := \mathbf{P}(X_k = 0) \quad \text{und} \quad \alpha_m := (1 - \alpha_0) \alpha'_m.$$

Des Weiteren gilt $\alpha_0 + \sum_{m=1}^M \alpha_m = 1$ sowie

$$\mathbf{P}(X_k \leq x) = \mathbf{P}(X_k \leq x \mid X_k = 0) \cdot \mathbf{P}(X_k = 0)$$

$$\begin{aligned}
 & + \mathbf{P}(X_k \leq x \mid X_k > 0) \cdot \mathbf{P}(X_k > 0) \\
 & = 1 \cdot \alpha_0 + \mathbf{P}(X_k \leq x \mid X_k > 0) \cdot (1 - \alpha_0) \\
 & = \alpha_0 + (1 - \alpha_0) \int_0^x f_{\Theta}(t) dt
 \end{aligned}$$

Für die im folgenden log-Likelihood Ausdruck verwendete Funktion $h : \mathcal{D} \rightarrow \mathbb{R}_{\geq 0}$ gilt

$$h_{\Theta}(x_k) = \begin{cases} \alpha_0, & \text{falls } x_k = 0, \\ \sum_{m=1}^M \alpha_m f_{m, \lambda_m}(x_k), & \text{falls } x_k > 0. \end{cases}$$

Wir erhalten

$$\begin{aligned}
 & \log L(\Theta; \mathcal{D}) \\
 & = \log \prod_{k=1}^K h_{\Theta}(x_k) \\
 & = \sum_{k=1}^K \left(\mathbf{1}_{\{0\}}(x_k) \cdot \log(\alpha_0) + \mathbf{1}_{(0, \infty)}(x_k) \cdot \log \left(\sum_{m=1}^M \alpha_m \cdot f_{m, \lambda_m}(x_k) \right) \right) \quad (5.24) \\
 & = \sum_{k=1}^K \mathbf{1}_{\{0\}}(x_k) \cdot \log(\alpha_0) + \sum_{k=1}^K \mathbf{1}_{(0, \infty)}(x_k) \cdot \log \left(\sum_{m=1}^M \alpha_m \cdot f_{m, \lambda_m}(x_k) \right).
 \end{aligned}$$

Es wird jetzt davon ausgegangen, dass die Datenmenge \mathcal{D} tatsächlich Elemente $x_k = 0$ enthält und o.B.d.A. wie folgt geordnet ist: Für $1 \leq k \leq k_0 < K$ sei $x_k > 0$, für $k_0 < k \leq K$ sei $x_k = 0$. Mit $A_0 := K - k_0$ wird so aus (5.24)

$$\log L(\Theta; \mathcal{D}) = \sum_{k=1}^{k_0} \log \left(\sum_{m=1}^M \alpha_m \cdot f_{m, \lambda_m}(x_k) \right) + A_0 \cdot \log(\alpha_0). \quad (5.25)$$

Die $x_k = 0$ können aus keiner der M Erlangdichten generiert worden sein. Also existieren für diese Werte keine unbeobachtbaren Daten. Für $1 \leq k \leq k_0$ bzw. die $x_k > 0$ existieren allerdings wieder nicht beobachtbare Daten y_k . Es gilt $y_k = m$, $y_k \in \{1, 2, \dots, M\}$, genau dann, wenn $x_k \in \mathcal{D}$ von Komponente f_m der Kombination von Erlangdichten erzeugt wurde. Mit Kenntnis der genauen Werte der y_k vereinfachte sich (5.25) zu

$$\log L(\Theta; \mathcal{D}, y) = \sum_{k=1}^{k_0} \log(\alpha_{y_k} \cdot f_{y_k, \lambda_k}(x_k)) + A_0 \cdot \log(\alpha_0)$$

mit $y \in \{1, 2, \dots, M\}^{k_0}$. Für die im vorherigen Abschnitt erklärte Zähl-dichte der nicht beobachtbaren Daten und $x_k > 0$ gilt jetzt

$$\begin{aligned} \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k) &= \frac{f_{\hat{\Theta}}(x_k \mid Y_k = y_k) \cdot \mathbf{P}_{\hat{\Theta}}(Y_k = y_k)}{\sum_{m=1}^M f_{\hat{\Theta}}(x_k \mid Y_k = m) \cdot \mathbf{P}_{\hat{\Theta}}(Y_k = m)} \\ &= \frac{f_{y_k, \hat{\theta}_{y_k}}(x_k) \cdot \hat{\alpha}_{y_k}}{\sum_{m=1}^M f_{m, \hat{\theta}_m}(x_k) \cdot \hat{\alpha}_m}. \end{aligned}$$

Wegen der bedingten Unabhängigkeit der Y_k gegebenen X_k gilt wieder $\mathbf{P}_{\hat{\Theta}}(Y = y \mid \mathcal{D}) = \prod_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k)$. Außerdem gilt

$$\begin{aligned} Q(\Theta, \hat{\Theta}) &= \mathbf{E}_{\hat{\Theta}}[\log L(\Theta; \mathcal{D}, Y) \mid \mathcal{D}] \\ &= \sum_{y \in \{1, \dots, M\}^{k_0}} \log L(\Theta; \mathcal{D}, y) \cdot \mathbf{P}_{\hat{\Theta}}(Y = y \mid \mathcal{D}) \\ &= A_0 \log(\alpha_0) \sum_{y \in \{1, \dots, M\}^{k_0}} \mathbf{P}_{\hat{\Theta}}(Y = y \mid \mathcal{D}) \\ &\quad + \sum_{y \in \{1, \dots, M\}^{k_0}} \sum_{k=1}^{k_0} \log(\alpha_{y_k} f_{y_k, \lambda_k}(x_k)) \prod_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k) \\ &= A_0 \log(\alpha_0) \\ &\quad + \sum_{y \in \{1, \dots, M\}^{k_0}} \sum_{k=1}^{k_0} \log(\alpha_{y_k} f_{y_k, \lambda_k}(x_k)) \prod_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = y_k \mid X_k = x_k). \end{aligned}$$

Die Summe über die Vektoren $y \in \{1, \dots, M\}^{k_0}$ in der letzten Zeile kann wie zuvor in [TBT06] vereinfacht werden. So erhält man

$$\begin{aligned} Q(\Theta, \hat{\Theta}) &= A_0 \log(\alpha_0) + \sum_{k=1}^{k_0} \sum_{m=1}^M \log(\alpha_m) \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \\ &\quad + \sum_{k=1}^{k_0} \sum_{m=1}^M \log(f_m(x_k \mid \lambda_m)) \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k). \end{aligned} \tag{5.27}$$

In (5.27) können so die Terme mit den α_m bzw. λ_m unabhängig voneinander maximiert werden. Für die Bestimmung der α_m wird auf einen LAGRANGE-Multiplikator zurückgegriffen. Die LAGRANGESCHE Hilfsfunktion lautet $L(\alpha, \lambda) := f(\alpha) + \lambda g(\alpha)$ mit $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_M)$, $\lambda \in \mathbb{R}$,

$$f(\alpha) = A_0 \log(\alpha_0) + \sum_{m=1}^M \sum_{k=1}^{k_0} \log(\alpha_m) \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k)$$

5. Approximation von Phasentypverteilungen

und $g(\alpha) := 1 - \sum_{m=0}^M \alpha_m = 0$. Bilden der partiellen Ableitungen und Auflösen nach α_m ergibt

$$\alpha_0 = \frac{A_0}{\lambda} \quad \text{sowie} \quad \alpha_m = \frac{1}{\lambda} \cdot \sum_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k)$$

für $1 \leq m \leq M$. Aus $1 = \alpha_0 + \sum_{m=1}^M \alpha_m = \frac{A_0 + k_0 \cdot 1}{\lambda}$ folgt $\lambda = A_0 + k_0 = K$ bzw.

$$\alpha_0 = \frac{A_0}{K} \quad \text{und} \quad \alpha_m = \frac{1}{K} \sum_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k)$$

für $1 \leq m \leq M$. Dabei entspricht also α_0 gerade der relativen Häufigkeit von Elementen $x_k = 0$. Bei der Bestimmung der λ_m aus dem entsprechenden Term in (5.27) ergibt sich wie in [TBT06]

$$\lambda_m = \frac{r_m \sum_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k)}{\sum_{k=1}^{k_0} \mathbf{P}_{\hat{\Theta}}(Y_k = m \mid X_k = x_k) \cdot x_k} .$$

6. Bestimmung der Verspätungsverteilungen

Übersicht

6.1. Theta-Exponentialpolynome	56
6.2. Einbettung der Hyper-Erlangverteilungen	59
6.3. Grundlegende Operationen	59
6.3.1. Excess Beyond	60
6.3.2. Maximum	61
6.3.3. Addition	61
6.4. Anschlussicherung	62
6.4.1. Einfache Anschlussicherung	63
6.4.2. Vorausschauende Anschlussicherung	65
6.5. Zugfolgerestriktionen	70
6.5.1. Theoretische Umsetzung	70
6.5.2. Berechnung - Subtraktion	74
6.6. Komplexitätsreduzierung	76
6.6.1. Metriken für Verteilungsfunktionen	77
6.6.2. Approximation der Gestalt einer Verteilungsfunktion	80
6.6.3. Verwendetes Verfahren	81
6.6.4. Evaluation des Verfahrens	84
6.7. Fehlerfortpflanzung	89
6.7.1. Excess-Beyond	91
6.7.2. Faltung	91
6.7.3. Produkt	92
6.8. Topologische Sortierung	92
6.9. Stochastische Abhängigkeiten	93

In diesem Kapitel stellen wir mit den Theta-Exponentialpolynomen die Klasse von Funktionen vor, welche wir für die Repräsentation von Verteilungsfunktionen verwenden, und zeigen, wie die Verteilungsfunktionen der Quellverspätungen, bzw. allgemein der Hyper-Erlangverteilungen, als Theta-Exponentialpolynom dargestellt werden. Wir erläutern, welche Operationen wir benötigen, wenn wir Verspätungen auf Ebene der Verteilungsfunktionen fortpflanzen wollen. Mit Ausnahme einer Operation, sind Theta-Exponentialpolynome gegen diese Operationen abgeschlossen. Die Parameteranzahl der aus den Operationen resultierenden Theta-Exponentialpolynome steigt in der Regel sehr stark an. Daher benötigen wir ein Verfahren, welches die Komplexität der Darstellung reduziert. Dies wird in Abschnitt

6.6 diskutiert. Hierfür wird ein approximatives Verfahren verwendet. Im Anschluss erläutern wir kurz, wie sich die im Zuge solcher Approximationen entstandenen Fehler bei Anwendung weiterer Operationen fortpflanzen können. In Abschnitt 6.8 gehen wir auf die topologische Sortierung der Ereignisse des Ereignis-Aktivitätsnetzwerkes ein, welche die Reihenfolge für die Bestimmung der Verspätungsverteilungen vorgibt. In kreisfreien Netzen lässt sich immer eine solche Sortierung finden. Abschließend gehen wir noch auf Situationen ein, in denen die stochastische Unabhängigkeit der miteinander zu verrechnenden Zufallsvariablen verletzt sein kann. Diese ist die zentrale Bedingung für Verspätungfortpflanzung auf Ebene der Verteilungsfunktionen.

6.1. Theta-Exponentialpolynome

Theta-Exponentialpolynome (s. [TG97, Fle90]) stellen eine Verallgemeinerung der azyklischen Phasentypverteilungen dar. Sie wurden bereits von FUHR in [Fuh07] für die Repräsentation und Verknüpfung von Verteilungsfunktionen verwendet. Wesentlicher Grund hierfür ist ihre Abgeschlossenheit gegenüber Multiplikation, Differentiation, Integration und Faltung. Warum wir auf diese Operationen angewiesen sind, wird in Abschnitt 6.3 erläutert. Die folgenden Ergebnisse stammen aus [TG97] von TROGEMANN und GENTE.

Definition 6.1.1 (Theta-Exponentialpolynom).

Eine Funktion $H : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$ der Art $H(t) := \sum_{i=1}^n \mathbf{1}_{[\theta_i, \infty)}(t) H_i(t)$ mit

$$H_i(t) := a_i (t - \theta_i)^{k_i} e^{-\lambda_i (t - \theta_i)} \quad (6.1)$$

und $a_i \in \mathbb{R}$, $\lambda_i \in \mathbb{R}_0^+$, $k_i \in \mathbb{N}_0$ heißt **Theta-Exponentialpolynom**. Dabei heißt

- (i) θ_i der deterministische Anteil,
- (ii) a_i der Koeffizient,
- (iii) λ_i die Rate und
- (iv) k_i die Stufenkennzahl.

Für die Darstellung in einem Phasendiagramm werden Parameter

$$c_i := \frac{a_i \cdot k_i!}{\lambda_i^{k_i+1}}$$

definiert. Diese können im Fall $c_i \in [0, 1]$ und $\sum_{i=1}^n c_i = 1$ als Verzweigungswahrscheinlichkeiten interpretiert werden. In [TG97] treten sie bei der Bestimmung der Momente einer Dichte $H(t)$ aus der Klasse der Exponentialpolynome auf.

Satz 6.1.2.

[TG97] Es sei X Zufallsvariable mit Verteilungsfunktion $F_X(t) = \int_{-\infty}^t H(s) ds$ für alle $t \in \mathbb{R}$ und H aus Definition 6.1.1. Dann gilt für alle $k \in \mathbb{N}$

$$E[X^k] = \sum_{i=1}^n \frac{a_i \cdot k_i!}{\lambda_i^{k_i+1}} \sum_{j=0}^l \binom{l}{j} \frac{\prod_{m=1}^j (k_i + m)}{\lambda_i^j} \theta_i^{l-j} .$$

Ebenfalls in [TG97] zeigen TROGEMANN und GENTE die Abgeschlossenheit der Klasse der Theta-Exponentialpolynome unter

- *Integration,*
- *Differentiation,*
- *Faltung und*
- *Multiplikation.*

Im Folgenden werden die entsprechenden Resultate jeweils ohne Beweis angegeben.

Satz 6.1.3.

[TG97] Die Klasse der Theta-Exponentialpolynome ist bezüglich der Integration abgeschlossen. Das bestimmte Integral von $H(t) = \sum_{i=1}^n \mathbf{1}_{[\theta_i, \infty)}(t) H_i(t)$ mit

$$H_i(t) := a_i(t - \theta_i)^{k_i} e^{-\lambda_i(t - \theta_i)}$$

und $\lambda_i > 0$ ist gegeben durch $\int_0^t H(t) dt = \sum_{i=1}^n \hat{H}_i(t)$ und

$$\hat{H}_i(t) = \sum_{l=0}^{k_i} \frac{a_i k_i!}{\lambda_i^{k_i+1-l} l!} (t - \theta_i)^l e^{-\lambda_i(t - \theta_i)} + a_i \frac{k_i!}{\lambda_i^{k_i+1}} .$$

Satz 6.1.4.

[TG97] Die Multiplikation zweier Theta-Exponentialpolynome h und H ist wieder ein Element der Klasse der Theta-Exponentialpolynome und gegeben durch:

$$\begin{aligned} h(t)H(t) &= \left(\sum_{i=1}^n \mathbf{1}_{[\theta_i, \infty)}(t) a_i(t - \theta_i)^{k_i} e^{-\lambda_i(t - \theta_i)} \right) \\ &\quad \cdot \left(\sum_{j=1}^m \mathbf{1}_{[\theta_j, \infty)}(t) A_j(t - \theta_j)^{K_j} e^{-\Lambda_j(t - \theta_j)} \right) \\ &= \sum_{i=1}^n \sum_{j=1}^m \mathbf{1}_{[\theta_{ij}, \infty)}(t) H_{ij}(t) \end{aligned}$$

6. Bestimmung der Verspätungsverteilungen

Dabei ist $\theta_{ij} := \max\{\theta_i, \Theta_j\}$ sowie

$$H_{ij}(t) = \begin{cases} \hat{H}_{ij}(t), & \text{falls } \theta_i > \Theta_j \\ \bar{H}_{ij}(t), & \text{falls } \theta_i < \Theta_j \\ h_{ij}(t), & \text{falls } \theta_i = \Theta_j \end{cases}$$

mit

$$\begin{aligned} \hat{H}_{ij}(t) &= \sum_{l=0}^{K_j} \binom{K_j}{l} A_j a_i (\theta_i - \Theta_j)^{K_j-l} e^{-\Lambda_j(\theta_i - \Theta_j)} \cdot (t - \theta_i)^{k_i+l} e^{-(\lambda_i + \Lambda_j)(t - \theta_i)} \\ \bar{H}_{ij}(t) &= \sum_{l=0}^{k_i} \binom{k_i}{l} a_i A_j (\Theta_j - \theta_i)^{k_i-l} e^{-\lambda_i(\Theta_j - \theta_i)} \cdot (t - \Theta_j)^{K_j+l} e^{-(\Lambda_j + \lambda_i)(t - \Theta_j)} \\ h_{ij}(t) &= A_j a_i (t - \theta_i)^{K_j+k_i} e^{-(\Lambda_j + \lambda_i)(t - \theta_i)}. \end{aligned}$$

Für die Faltungsverknüpfung gibt es das folgende Resultat.

Satz 6.1.5.

[TG97] Die Faltung zweier Theta-Exponentialpolynome h und H ist wieder ein Element der Klasse der Theta-Exponentialpolynome und gegeben durch:

$$\begin{aligned} h(t) * H(t) &= \left(\sum_{i=1}^n \mathbf{1}_{[\theta_i, \infty)}(t) a_i (t - \theta_i)^{k_i} e^{-\lambda_i(t - \theta_i)} \right) \\ &\quad * \left(\sum_{j=1}^m \mathbf{1}_{[\Theta_j, \infty)}(t) A_j (t - \Theta_j)^{K_j} e^{-\Lambda_j(t - \Theta_j)} \right) \\ &= \sum_{i=1}^n \sum_{j=1}^m \mathbf{1}_{[\theta_{ij}, \infty)}(t) H_{ij}(t) \end{aligned}$$

Dabei ist $\theta_{ij} := \theta_i + \Theta_j$ sowie

$$H_{ij}(t) = \begin{cases} \hat{H}_{ij}(t), & \text{falls } \lambda_i \neq \Lambda_j \\ \frac{A_j a_i k_i! K_j!}{(k_i + K_j + 1)!} (t - \Theta_{ij})^{k_i + K_j + 1} e^{-\lambda_i(t - \Theta_{ij})}, & \text{falls } \lambda_i = \Lambda_j \end{cases}$$

mit

$$\begin{aligned} \hat{H}_{ij}(t) &= \sum_{l=0}^{k_i} \frac{(-1)^l \binom{K_j+l}{l} a_i k_i! A_j K_j!}{(k_i - l)! (\Lambda_j - \lambda_i)^{K_j+l+1}} (t - \Theta_{ij})^{k_i-l} e^{-\lambda_i(t - \Theta_{ij})} \\ &\quad + \sum_{l=0}^{K_j} \frac{(-1)^l \binom{k_i+l}{l} a_i k_i! A_j K_j!}{(K_j - l)! (\lambda_i - \Lambda_j)^{k_i+l+1}} (t - \Theta_{ij})^{K_j-l} e^{-\Lambda_j(t - \Theta_{ij})}. \end{aligned}$$

6.2. Einbettung der Hyper-Erlangverteilungen

In diesem Abschnitt zeigen wir, wie die Familie der Hyper-Erlangverteilungen in die Klasse der Theta-Exponentialpolynome eingebettet werden kann. Haben wir beispielsweise über den in Abschnitt 5.4 erläuterten EM-Algorithmus den Parametervektor $\Theta = (\alpha_0, \alpha_1, \dots, \alpha_M, \lambda_1, \dots, \lambda_M)$ sowie die Anzahl der Zustände r_m , $1 \leq m \leq M$, in den jeweiligen Zweigen der Hyper-Erlangverteilung erhalten, repräsentieren diese Parameter die Verteilungsfunktion

$$F(t) = \alpha_0 + \int_0^t f_{\Theta}(s) ds$$

mit $f_{\Theta}(t) := \sum_{m=1}^M \alpha_m f_{m, \lambda_m}(t)$ und

$$f_{m, \lambda_m}(t) = \sum_{m=1}^M \frac{(\lambda_m t)^{r_m-1}}{(r_m-1)!} \lambda_m e^{-\lambda_m t}$$

für alle $t \in \mathbb{R}_{\geq 0}$ (vgl. Abschnitt 5.4). Hieraus lässt sich leicht eine Repräsentation der Verteilungsfunktion mittels eines Theta-Exponentialpolynoms (s. (6.1)) konstruieren. Dabei wird α_0 durch

$$\mathbf{1}_{[\theta_0, \infty)}(t) \cdot a_0 \cdot (t - \theta_0)^{k_0} \cdot e^{-\lambda_0(t - \theta_0)} \quad \text{für alle } t \in \mathbb{R}$$

mit $a_0 = \alpha_0$ und $\theta_0 = k_0 = \lambda_0 = 0$ dargestellt. Die Terme $\alpha_m \cdot f_{m, \lambda_m}(t)$ werden durch

$$\mathbf{1}_{[\theta_m, \infty)}(t) \cdot a_m \cdot (t - \theta_m)^{k_m} e^{-\lambda_m(t - \theta_m)} \quad \text{für alle } t \in \mathbb{R}$$

mit

$$a_m = \alpha_m \cdot \frac{\lambda_m^{r_m}}{(r_m-1)!},$$

$\theta_m = 0$, $k_m = r_m - 1$ und identischen Werten für λ_m für alle $1 \leq m \leq M$ dargestellt.

6.3. Grundlegende Operationen

Für die Berechnung der Verteilungsfunktionen aller Ankunfts- und Abfahrtsverspätungen (fortgepflanzte Verspätungen) im Ereignis-Aktivitätsnetzwerk benötigt man im Wesentlichen nur drei grundlegende Operationen. Wir betrachten im Folgenden jeweils sowohl die Situation auf Ebene von Zufallsvariablen als auch auf Ebene der Verteilungsfunktionen. Die Notwendigkeit dieser drei Operationen wird schon in der Literatur beschrieben (z.B. [MM07]). Sie ergibt sich außerdem direkt aus der in (3.4) und (3.5) beschriebenen Fortpflanzung von Verspätungen. Es sind

- (i) die sogenannte *Excess Beyond* Operation,
- (ii) das Maximum von Zufallsvariablen bzw. das Produkt der zugehörigen Verteilungsfunktionen,
- (iii) die Addition von Zufallsvariablen bzw. die Faltung der zugehörigen Verteilungsfunktionen.

Wie sich in den folgenden Abschnitten zeigen wird, benötigt man jedoch weitere Operationen, wenn man die Modellierung der Verspätungsfortpflanzung um Anschlussicherung (s. Abschnitt 6.4) oder Zugfolgerestriktionen (s. Abschnitt 6.5), erweitern möchte. Die Zufallsvariablen Y_1 und Y_2 sind im Folgenden stets nicht-negativ. Der Definitionsbereich der Verteilungsfunktionen ist daher stets $\mathbb{R}_{\geq 0}$.

6.3.1. Excess Beyond

Einfach umzusetzen ist die in der Literatur so genannte *Excess Beyond* Operation. Gegeben sei beispielsweise eine Zufallsvariable Y_1 für die (fortgepflanzte) Verspätung eines Zuges und der zeitliche Puffer $s \geq 0$, der für den Abbau der Verspätung zur Verfügung steht. Die Zufallsvariable

$$Y_2 := [Y_1 - s]^+ = \max\{Y_1 - s, 0\}$$

steht dann für die Verspätung nach Berücksichtigung dieses Puffers. Des Weiteren sei F_{Y_1} die Verteilungsfunktion von Y_1 . Dann gilt offensichtlich für die Verteilungsfunktion

$$F_{Y_2}(t) = F_{Y_1}(t + s) \quad \text{für alle } t \in \mathbb{R}_{\geq 0}.$$

F_{Y_2} ist die auf der t -Achse um s verschobene Verteilungsfunktion F_{Y_1} . In der Familie der Theta-Exponentialpolynome ist diese Verschiebung leicht umsetzbar. Ist hier also

$$F_{Y_1}(t) = \sum_{m=1}^M \mathbf{1}_{[\theta_m, \infty)}(t) \cdot a_m \cdot (t - \theta_m)^{k_m} \cdot e^{-\lambda_m(t - \theta_m)}$$

gegeben, erhalten wir

$$F_{Y_1}(t + s) = \sum_{m=1}^M \mathbf{1}_{[\theta'_m, \infty)}(t) \cdot a_m \cdot (t - \theta'_m)^{k_m} \cdot e^{-\lambda_m(t - \theta'_m)}$$

mit $\theta'_m := \theta_m + s$ für alle $1 \leq m \leq M$.

6.3.2. Maximum

In Umsteigestationen interessieren wir uns unter anderem für das Maximum der von Zubringern übertragenen Verspätungen (vgl. (3.5)). Im Falle zweier Zubringer bestimmen wir also die Verteilungsfunktion der Zufallsvariable $Y := \max\{Y_1, Y_2\}$. Unter der Voraussetzung, dass Y_1 und Y_2 stochastisch unabhängig sind, gilt

$$F_Y(t) = F_{Y_1} \cdot F_{Y_2}(t) \quad \text{für alle } t \in \mathbb{R}_{\geq 0}.$$

Sind die Darstellungen von F_{Y_1} und F_{Y_2} als Theta-Exponentialpolynom bekannt, kann gemäß Satz 6.1.5 die Verteilungsfunktion $F_{Y_1} \cdot F_{Y_2}$ berechnet werden.

6.3.3. Addition

Berechnung der Verteilungsfunktion

Möchte man die Verteilungsfunktion der Zufallsvariablen $Y = Y_1 + Y_2$ bestimmen, so faltet man die Verteilungsfunktionen F_{Y_1} und F_{Y_2} , d.h. $F_Y = F_{Y_1} * F_{Y_2}$. Voraussetzung hierfür ist wieder, dass Y_1 und Y_2 stochastisch unabhängig sind. Es gilt dann

$$F_Y(t) = F_{Y_1} * F_{Y_2}(t) = \int_{\mathbb{R}} F_{Y_2}(t-s) dF_{Y_1}(s) \quad \text{für alle } t \in \mathbb{R}_{\geq 0}. \quad (6.2)$$

F_{Y_1} ist, im Rahmen unserer Anwendung, im Allgemeinen keine stetige Verteilungsfunktion sondern besitzt im Punkt $t = 0$ eine Sprungstelle der Höhe $F_{Y_1}(0)$. Insbesondere existiert zu F_{Y_1} also keine Dichtefunktion. Es ist aber möglich, F_{Y_1} gemäß $F_{Y_1}(t) = F_1(t) + F_2(t)$ zu zerlegen, wobei für alle $t \in \mathbb{R}_{\geq 0}$ gilt (vgl. Definition des STIELTJES-Integrals [Neu77])

$$\begin{aligned} F_1(t) &= (F_{Y_1}(t) - F_{Y_1}(0)) \cdot \mathbf{1}_{[0, \infty)}(t) \quad \text{und} \\ F_2(t) &= F_{Y_1}(0) \cdot \mathbf{1}_{[0, \infty)}(t). \end{aligned}$$

F_1 ist jetzt offensichtlich eine stetige Funktion. F_2 ist eine Treppenfunktion mit Sprunghöhe $F_{Y_1}(0)$ in $t = 0$. Im Folgenden wird zweimal die Linearität des STIELTJES-Integrals im Integrator ausgenutzt. So folgt aus (6.2) zunächst für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} F_Y(t) &= \int_{\mathbb{R}} F_{Y_2}(t-s) dF_1(s) + \int_{\mathbb{R}} F_{Y_2}(t-s) dF_2(s) \\ &= \int_0^\infty F_{Y_2}(t-s) dF_1(s) + F_{Y_2}(t)F_{Y_1}(0). \end{aligned} \quad (6.4)$$

Im nächsten Schritt ersetzt man die Funktion F_1 . Man kann mittels der totalen Wahrscheinlichkeit leicht zeigen, dass für die bedingte Verteilungsfunktion $F_{Y_1|Y_1>0}$ im Falle $t \in \mathbb{R}_{\geq 0}$

$$F_{Y_1|Y_1>0}(t) = \frac{F_{Y_1}(t) - F_{Y_1}(0)}{1 - F_{Y_1}(0)} = \frac{F_1(t)}{1 - F_{Y_1}(0)}$$

gilt. Ersetzt man in (6.4) also F_1 durch $(1 - F_{Y_1}(0)) \cdot F_{Y_1|Y_1>0}$ und nutzt ein zweites mal die Linearität des STIELTJES-Integrals im Integrator aus, erhält man schließlich für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} F_Y(t) &= (1 - F_{Y_1}(0)) \cdot \int_0^\infty F_{Y_2}(t-s) dF_{Y_1|Y_1>0}(s) + F_{Y_2}(t)F_{Y_1}(0) \\ &= (1 - F_{Y_1}(0)) \cdot \int_0^\infty F_{Y_2}(t-s)f_{Y_1|Y_1>0}(s) ds + F_{Y_2}(t)F_{Y_1}(0). \end{aligned} \quad (6.6)$$

Dabei ist $f_{Y_1|Y_1>0}$ die zu $F_{Y_1|Y_1>0}$ gehörende Dichtefunktion. Der erste Term in der letzten Zeile in (6.6) entspricht der Faltung $F_{Y_2} * f_{Y_1|Y_1>0}(t)$ multipliziert mit der Konstanten $1 - F_{Y_1}(0)$. Die Familie der Theta-Exponentialpolynome ist unter der Faltung ebenso abgeschlossen wie unter der Multiplikation mit einer Konstanten (Konvexkombination).

Verrechnung der Momente

Es existiert noch eine Alternative zur Faltung der Verteilungsfunktionen. In Abschnitt 6.6 wird erläutert, dass wir für die Komplexitätsreduzierung eine Momentenanpassung für Hyper-Erlangverteilungen nutzen. Nach jeder Faltungsoperation auf Ebene der Theta-Exponentialpolynome ist es empfehlenswert, eine Komplexitätsreduzierung durchzuführen. Verrechnen wir stattdessen direkt die Momente der Verteilungen von Y_1 bzw. Y_2 , so führt dies zum gleichen Ergebnis und wir sparen uns die Faltungsoperation. Es gilt:

$$\begin{aligned} \mathbf{E}[Y] &= \mathbf{E}[Y_1 + Y_2] = \mathbf{E}[Y_1] + \mathbf{E}[Y_2] \\ \mathbf{E}[Y^2] &= \mathbf{E}[(Y_1 + Y_2)^2] = \mathbf{E}[Y_1^2] + \mathbf{E}[Y_2^2] + 2 \cdot \mathbf{E}[Y_1 \cdot Y_2] \\ &= \mathbf{E}[Y_1^2] + \mathbf{E}[Y_2^2] + 2 \cdot \mathbf{E}[Y_1] \cdot \mathbf{E}[Y_2] \\ \mathbf{E}[Y^3] &= \mathbf{E}[(Y_1 + Y_2)^3] = \mathbf{E}[Y_1^3] + \mathbf{E}[Y_2^3] + 3 \cdot (\mathbf{E}[Y_1^2] \cdot \mathbf{E}[Y_2] + \mathbf{E}[Y_1] \cdot \mathbf{E}[Y_2^2]) \end{aligned} \quad (6.7)$$

Diese Herangehensweise wird allerdings nur genutzt, wenn die Momente der Verteilungsfunktionen von Y_1 und Y_2 jeweils in einem Verhältnis zueinander stehen, welche sich für die Momentenanpassung besonders eignen. Mehr dazu folgt in den Abschnitten 6.6.3 und 6.6.4 (s. Abbildung 6.10). Ansonsten wird auf die oben beschriebene Faltung der Verteilungsfunktionen zurückgegriffen.

6.4. Anschlusssicherung

Anschlusssicherung ist eine Strategie, mittels der man verhindern kann, dass Reisende Anschlüsse mit wenig zeitlichem Puffer, auf Grund von Verspätungen ihrer

Zubringerzüge, verpassen. Eine Anschlusssicherungszeit von $\kappa \in \mathbb{N}_0$ Minuten bedeutet für den Abbringerzug, dass er, im Anschluss an seine fahrplanmäßige Abfahrtszeit, maximal κ Minuten auf seine planmäßigen Zubringer wartet. Könnte der Abbringer eigentlich, ohne Berücksichtigung seiner Zubringer, pünktlich abfahren, so wird seine Abfahrt durch die Einhaltung der Anschlusssicherung verspätet. Auf der anderen Seite kann man so jedoch verhindern, dass eventuelle Umsteiger eine starke Verspätung dadurch erfahren, dass ihr nächster Abbringer (fahrplanmäßig) erst einen Takt später fährt. Hierbei gehen wir davon aus, dass Reisende ihre ursprüngliche Reiseverbindung in einem solchen Fall nicht ändern. Sie weichen also nicht auf einen alternativen Abbringerzug einer anderen Linie aus.

Wir nutzen für die tiefer gehende Betrachtung die folgende Notation:

- \mathcal{Z} sei die Menge der fahrplanmäßigen Zubringer.
- $X_i \geq 0$ seien die von den Zubringern $i \in \mathcal{Z}$ auf den Abbringer übertragenen Verspätungen.
- A sei die eigene Abfahrtsverspätung des Abbringers, die er bereits besitzt, bevor durch die verwendete Anschlusssicherungsstrategie weitere, von Zubringern übertragene, Verspätung berücksichtigt wird.
- Y sei die resultierende (tatsächliche) Abfahrtsverspätung des Abbringers mit Berücksichtigung der Anschlusssicherung.

Ein extremes Beispiel ist eine Anschlusssicherungszeit $\kappa = 0$. In diesem Fall wartet der Abbringer gar nicht auf planmäßige Zubringer. Entscheidend für seine Abfahrtsverspätung ist dann nur die bisher selbst erfahrene Verspätung. Die Berechnung der Verteilungsfunktion von Y ist in diesem Falle entsprechend einfach. Unwesentlich schwieriger ist die Berechnung, falls der Abbringer beliebig lange ($\kappa = \infty$) auf seine Zubringer wartet. In dem Falle gilt $Y = \max\{A, \max\{X_i \mid i \in \mathcal{Z}\}\}$. Im Folgenden wollen wir die Anschlusssicherung für beliebiges $\kappa \in \mathbb{N}_0$ betrachten. Wir unterscheiden dabei zwei verschiedene Modellierungen einer Anschlusssicherungsstrategie.

6.4.1. Einfache Anschlusssicherung

Es sei $\kappa \in \mathbb{N}_0$. Der Abbringer wartet maximal κ Minuten auf seine Zubringer, d.h.

$$Y = \max \left\{ A, \min \left\{ \kappa, \max \{ X_i \mid i \in \mathcal{Z} \} \right\} \right\}. \quad (6.8)$$

Es sei K Zufallsvariable mit $\mathbf{P}(K \leq t) = \mathbf{1}_{[\kappa, \infty)}(t)$ für alle $t \in \mathbb{R}_{\geq 0}$, d.h. $\mathbf{P}(K = \kappa) = 1$. Es gilt für alle $t \in \mathbb{R}_{\geq 0}$

$$\mathbf{P}(Y \leq t) = \mathbf{P}(A \leq t) \cdot \mathbf{P}(\min \{ K, \max \{ X_i \mid i \in \mathcal{Z} \} \leq t)$$

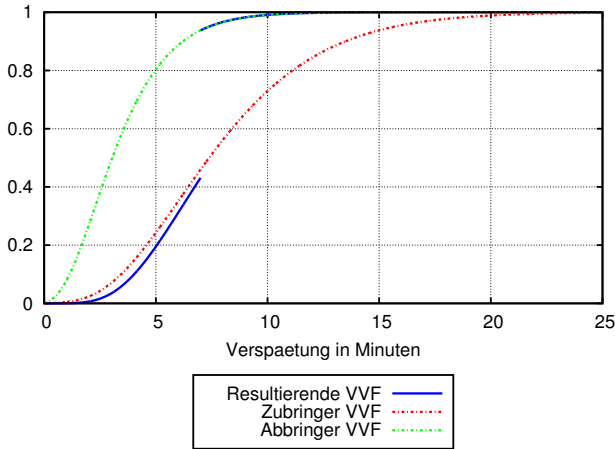


Abbildung 6.1.: Einfache Anschlussicherung, Sprungstelle der Verspätungsverteilungsfunktion (VVF) in $t = \kappa = 7$

$$\begin{aligned}
 &= \mathbf{P}(A \leq t) \cdot (1 - \mathbf{P}(K > t, \max\{X_i \mid i \in \mathcal{Z}\} > t)) \\
 &= \mathbf{P}(A \leq t) \cdot \left(1 - (1 - \mathbf{P}(K \leq t)) \cdot \left(1 - \prod_{i \in \mathcal{Z}} \mathbf{P}(X_i \leq t)\right)\right) \\
 &= \mathbf{P}(A \leq t) \cdot \left(1 - \mathbf{1}_{(-\infty, \kappa)}(t) \cdot \left(1 - \prod_{i \in \mathcal{Z}} \mathbf{P}(X_i \leq t)\right)\right) . \quad (6.9)
 \end{aligned}$$

Ergebnis der Berechnung ist im Allgemeinen eine Verteilungsfunktion mit einer Sprungstelle in $t = \kappa$. Da die Theta-Exponentialpolynome gegen alle verwendeten Operationen abgeschlossen sind, kann die Verteilungsfunktion als Theta-Exponentialpolynom repräsentiert werden. Beispielhaft wird die Situation in Abbildung 6.1 dargestellt. Durch eine folgende Faltung mit der Verteilungsfunktion einer Quellverspätung D kann die Höhe der Sprungstelle abnehmen. Im Falle $\mathbf{P}(D = 0) = 0$ wird die resultierende Verteilungsfunktion keine Sprungstelle mehr besitzen. Im Falle $\mathbf{P}(D = 0) = 1$ bleibt die Sprungstelle in voller Höhe erhalten.

In der Praxis würde für die resultierende Verteilungsfunktion eine Komplexitätsreduzierung durchgeführt. Deren Ergebnis ist eine stetige Verteilungsfunktion (s. Abschnitt 6.6.3). Daher sind die Unterschiede bezüglich der Funktionswerte in der Umgebung der Sprungstelle entsprechend groß. Die ersten drei Momente stimmen aber trotzdem überein. Da wir in der praktischen Anwendung nur bei Verteilungsfunktion von Ankunftsereignissen in Umsteigestationen an den Funktionswerten in-

teressiert sind (s. Kapitel 9), ist dies zunächst noch nicht problematisch. Wie zuvor erläutert, werden durch die bis zur Berechnung der Verteilungsfunktion eines solchen Ankunftsereignisses vorzunehmenden Faltungen, die Abweichungen bezüglich der Funktionswerte in der Regel deutlich reduziert.

6.4.2. Vorausschauende Anschlusssicherung

Die vorausschauende Anschlusssicherung ermöglicht eine etwas realistischere Vorgehensweise. Wir nehmen an, dass der Abbringer wieder maximal $\kappa \in \mathbb{N}_0$ Minuten auf verspätete Zubringer wartet. Es sind jetzt jedoch nur Zubringer $i \in \mathcal{Z}$ relevant, deren übertragene Verspätung X_i nicht größer als κ ist. Dies hat insbesondere zur Folge, dass, sollte für alle fahrplanmäßigen Zubringer $X_i > \kappa$ gelten, der Abbringer nicht warten muss. Für die resultierende Verspätung Y ist in diesem Falle folglich nur seine eigene Verspätung A relevant. Insgesamt gilt

$$Y = \max \{ A, \max \{ X_i \mid i \in \mathcal{Z}, X_i \leq \kappa \} \}. \quad (6.10)$$

Diese Verspätung tritt in der Realität genau dann auf, wenn die Dispositionentscheidung, d.h. die Entscheidung ob der Abbringer warten soll oder nicht, für alle enthaltenen Zubringer-Abbringer Beziehungen korrekt entschieden wird. Dies setzt voraus, dass der Entscheidungsträger zum Zeitpunkt seiner Entscheidung bereits weiß, mit welcher Verspätung die Zubringer die Umsteigestation erreichen. Einen alternativen Ansatz, der näher an der Realität ist, jedoch damit auch fehlerhafte Entscheidungen zulässt, wählt BÜKER in [Bü10]. Statt der tatsächlichen Ankunftsverspätung der Zubringer betrachtet er deren Verspätung zu Beginn des sogenannten „Dispositionshorizonts“, d.h. zu einem bestimmten Zeitpunkt vor der planmäßigen Abfahrt des Abbringers. Die Wahrscheinlichkeit, dass diese Verspätung dazu führt, dass der Abbringer im Rahmen der Anschlusssicherungsstrategie auf seinen Zubringer warten muss, wird als Entscheidungskriterium vorgeschlagen, ob die Verspätung des jeweiligen Zubringers in die Bestimmung der Verspätung des Abbringers einfließt. Ist diese Wahrscheinlichkeit sehr klein, so hat die Verspätung des Zubringers entsprechend wenig Einfluss auf die Abfahrtsverspätung des Abbringers.

Im folgenden Satz zeigen wir, was aus (6.10) für die zugehörige Verteilungsfunktion F_Y folgt.

Satz 6.4.1.

Es seien X_i , $i \in \mathcal{Z}$, und A stochastisch unabhängig. Außerdem sei $Y = \max \{ A, \max \{ X_i \mid i \in \mathcal{Z}, X_i \leq \kappa \} \}$ mit Verteilungsfunktion F_Y . Dann gilt für alle $t \in \mathbb{R}_{\geq 0}$

$$F_Y(t) = F_A(t) \cdot \left(\mathbf{1}_{[0, \kappa]}(t) \cdot \prod_{i \in \mathcal{Z}} (F_{X_i}(t) + (1 - F_{X_i}(\kappa)) + \mathbf{1}_{(\kappa, \infty)}(t)) \right). \quad (6.11)$$

Die resultierende Verteilungsfunktion F_Y kann als Theta-Exponentialpolynom dargestellt werden. Die Aussage des Satzes wird in mehreren Schritten bewiesen. Es sei $Z_\kappa = \{i \in \mathcal{Z} \mid X_i \leq \kappa\}$. Z_κ ist also eine Zufallsvariable. Eine vorteilhafte Eigenschaft von Satz 6.4.1 ist, dass wir für die Berechnung von F_Y nicht vorab wissen müssen, welche Zubringer $i \in \mathcal{Z}$ tatsächlich relevant sind. Das heißt, die genaue Ausprägung der Zufallsvariablen Z_κ wird nicht benötigt. Bevor wir uns mit dem ersten Lemma für Satz 6.4.1 beschäftigen, halten wir fest, dass für $\mathcal{I} \subset \mathcal{Z}$ gilt

$$Z_\kappa = \mathcal{I} \iff (\forall i \in \mathcal{I} : X_i \leq \kappa) \wedge (\forall j \in \mathcal{I}^c : X_j > \kappa).$$

Lemma 6.4.2.

Es sei $Y = \max\{A, \max\{X_i \mid i \in \mathcal{Z}, X_i \leq \kappa\}\}$ mit Verteilungsfunktion F_Y . Dann gilt für $t \in \mathbb{R}$

$$F_Y(t) = F_A(t) \cdot \left(\mathbf{1}_{[0, \kappa]}(t) \cdot \mathbf{P}(\max\{X_i \mid i \in Z_\kappa\} \leq t) + \mathbf{1}_{(\kappa, \infty)}(t) \right).$$

Beweis. Zunächst gilt wegen stochastischer Unabhängigkeit zwischen A und X_i , $i \in \mathcal{Z}$, offensichtlich

$$\begin{aligned} F_Y(t) &= \mathbf{P}(Y \leq t) = \mathbf{P}(\max\{A, \max\{X_i \mid i \in \mathcal{Z}, X_i \leq \kappa\}\} \leq t) \\ &= \mathbf{P}(A \leq t) \cdot \mathbf{P}(\max\{X_i \mid i \in \mathcal{Z}, X_i \leq \kappa\} \leq t) \\ &= F_A(t) \cdot \mathbf{P}(\max\{X_i \mid i \in \mathcal{Z}, X_i \leq \kappa\} \leq t). \end{aligned}$$

Für den zweiten Term des Produkts folgt

$$\mathbf{P}(\max\{X_i \mid i \in \mathcal{Z}, X_i \leq \kappa\} \leq t) = \begin{cases} 1 & \text{für } t > \kappa, \\ \mathbf{P}(\max\{X_i \mid i \in Z_\kappa\} \leq t) & \text{für } t \leq \kappa. \end{cases} \quad \square$$

Im Folgenden beschränken wir uns auf den Fall $t \leq \kappa$.

Lemma 6.4.3.

Es gilt für $t \leq \kappa$ und $\mathcal{I} \subset \mathcal{Z}$

$$\mathbf{P}(\max\{X_i \mid i \in Z_\kappa\} \leq t) = \prod_{i \in \mathcal{Z}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa]).$$

Beweis. Mit dem Satz der totalen Wahrscheinlichkeit erhalten wir

$$\mathbf{P}(\max\{X_i \mid i \in Z_\kappa\} \leq t) = \sum_{\mathcal{I} \subset \mathcal{Z}} \mathbf{P}(\max\{X_i \mid i \in \mathcal{I}\} \leq t \mid Z_\kappa = \mathcal{I}) \cdot \mathbf{P}(Z_\kappa = \mathcal{I}).$$

Im Folgenden beschäftigen wir uns mit der Umformung der bedingten Wahrscheinlichkeit dieser Gleichung. Für $\mathcal{I} \subset \mathcal{Z}$ gilt:

$$\mathbf{P}(\max\{X_i \mid i \in \mathcal{I}\} \leq t \mid Z_\kappa = \mathcal{I})$$

$$\begin{aligned}
 &= \mathbf{P} \left(\bigcap_{i \in \mathcal{I}} [X_i \leq t] \mid Z_\kappa = \mathcal{I} \right) \\
 &= \frac{\mathbf{P} \left(\bigcap_{i \in \mathcal{I}} [X_i \leq t] \cap \bigcap_{i \in \mathcal{I}} [X_i \leq \kappa] \cap \bigcap_{i \in \mathcal{I}^c} [X_i > \kappa] \right)}{\mathbf{P} \left(\bigcap_{i \in \mathcal{I}} [X_i \leq \kappa] \cap \bigcap_{i \in \mathcal{I}^c} [X_i > \kappa] \right)} \\
 &= \frac{\mathbf{P} \left(\bigcap_{i \in \mathcal{I}} [X_i \leq t] \cap \bigcap_{i \in \mathcal{I}^c} [X_i > \kappa] \right)}{\mathbf{P} \left(\bigcap_{i \in \mathcal{I}} [X_i \leq \kappa] \cap \bigcap_{i \in \mathcal{I}^c} [X_i > \kappa] \right)} \quad (t \leq \kappa) \\
 &= \frac{\prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq t) \cdot \prod_{i \in \mathcal{I}^c} \mathbf{P}(X_i > \kappa)}{\prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq \kappa) \cdot \prod_{i \in \mathcal{I}^c} \mathbf{P}(X_i > \kappa)} \quad (\text{stoch. unabhängig}) \\
 &= \frac{\prod_{i \in \mathcal{I}} \mathbf{P}([X_i \leq t] \cap [X_i \leq \kappa])}{\prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq \kappa)} \quad (t \leq \kappa) \\
 &= \prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq t \mid X_i \leq \kappa)
 \end{aligned}$$

Es gilt offensichtlich für alle $i \in \mathcal{Z}$ sowohl $\mathbf{P}(X_i > \kappa \mid X_i \leq \kappa) = 0$, $\mathbf{P}(X_i > \kappa \mid X_i > \kappa) = 1$ also auch $\mathbf{P}(X_i \leq t \mid X_j > \kappa) = 0$ für $t \leq \kappa$. Daraus folgt neben

$$\begin{aligned}
 \prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq t \mid X_i \leq \kappa) &= \prod_{i \in \mathcal{I}} (\mathbf{P}(X_i \leq t \mid X_i \leq \kappa) + \mathbf{P}(X_i > \kappa \mid X_i \leq \kappa)) \\
 &= \prod_{i \in \mathcal{I}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa] \mid X_i \leq \kappa)
 \end{aligned}$$

ebenfalls

$$\begin{aligned}
 1 &= \prod_{j \in \mathcal{I}^c} \mathbf{P}(X_j > \kappa \mid X_j > \kappa) \\
 &= \prod_{j \in \mathcal{I}^c} (\mathbf{P}(X_j > \kappa \mid X_j > \kappa) + \mathbf{P}(X_j \leq t \mid X_j > \kappa)) \\
 &= \prod_{j \in \mathcal{I}^c} \mathbf{P}([X_j \leq t] \cup [X_j > \kappa] \mid X_j > \kappa) .
 \end{aligned}$$

So erhalten wir

$$\begin{aligned}
 &\prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq t \mid X_i \leq \kappa) \\
 &= \prod_{i \in \mathcal{I}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa] \mid X_i \leq \kappa) \cdot \prod_{j \in \mathcal{I}^c} \mathbf{P}([X_j \leq t] \cup [X_j > \kappa] \mid X_j > \kappa)
 \end{aligned}$$

und wegen stochastischer Unabhängigkeit der X_i schließlich

$$\begin{aligned}
 & \prod_{i \in \mathcal{I}} \mathbf{P}(X_i \leq t \mid X_i \leq \kappa) \\
 &= \prod_{i \in \mathcal{I}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa] \mid X_i \leq \kappa, Z_\kappa = \mathcal{I}) \\
 &\quad \cdot \prod_{j \in \mathcal{I}^c} \mathbf{P}([X_j \leq t] \cup [X_j > \kappa] \mid X_j > \kappa, Z_\kappa = \mathcal{I}) \\
 &= \prod_{i \in \mathcal{I}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa] \mid Z_\kappa = \mathcal{I}) \\
 &\quad \cdot \prod_{j \in \mathcal{I}^c} \mathbf{P}([X_j \leq t] \cup [X_j > \kappa] \mid Z_\kappa = \mathcal{I}) \\
 &= \prod_{i \in \mathcal{Z}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa] \mid Z_\kappa = \mathcal{I}) \\
 &= \mathbf{P} \left(\bigcap_{i \in \mathcal{Z}} ([X_i \leq t] \cup [X_i > \kappa]) \mid Z_\kappa = \mathcal{I} \right).
 \end{aligned}$$

Dieses Ergebnis setzen wir jetzt in die erste Gleichung des Beweises ein. Wir erhalten

$$\begin{aligned}
 & \mathbf{P}(\max\{X_i \mid i \in \mathcal{Z}_\kappa\} \leq t) \\
 &= \sum_{\mathcal{I} \subset \mathcal{Z}} \mathbf{P} \left(\bigcap_{i \in \mathcal{Z}} ([X_i \leq t] \cup [X_i > \kappa]) \mid Z_\kappa = \mathcal{I} \right) \cdot \mathbf{P}(Z_\kappa = \mathcal{I}) \\
 &= \mathbf{P} \left(\bigcap_{i \in \mathcal{Z}} ([X_i \leq t] \cup [X_i > \kappa]) \right) \quad (\text{totale Wahrscheinlichkeit}) \\
 &= \prod_{i \in \mathcal{Z}} \mathbf{P}([X_i \leq t] \cup [X_i > \kappa]). \quad \square
 \end{aligned}$$

Korollar 6.4.4.

Für $t \leq \kappa$ gilt

$$\mathbf{P}(\max\{X_i \mid i \in \mathcal{Z}_\kappa\} \leq t) = \prod_{i \in \mathcal{Z}} (F_{X_i}(t) + (1 - F_{X_i}(\kappa))).$$

Beweis. Für $t \leq \kappa$ gilt $[X_i \leq t] \cap [X_i > \kappa] = \emptyset$. Daher folgt

$$\begin{aligned}
 \mathbf{P}([X_i \leq t] \cup [X_i > \kappa]) &= \mathbf{P}(X_i \leq t) + \mathbf{P}(X_i > \kappa) \\
 &= F_{X_i}(t) + (1 - F_{X_i}(\kappa)).
 \end{aligned}$$

Mit Lemma 6.4.3 folgt die Behauptung. □

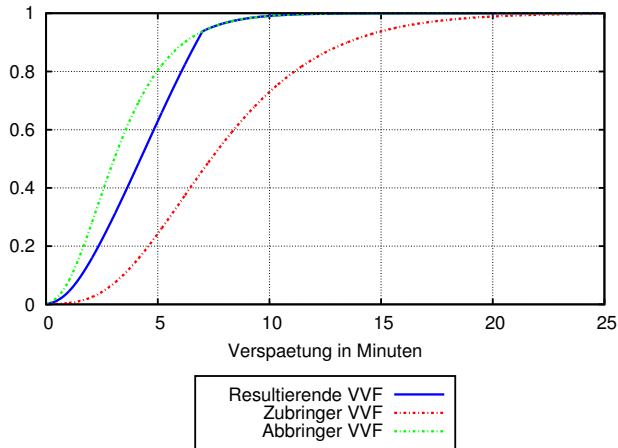


Abbildung 6.2.: Vorausschauende Anschlusssicherung, stetiger Übergang der Verspätungsverteilungsfunktion (VVF) in $t = \kappa = 7$

An dieser Stelle haben wir den Beweis von Satz 6.4.1 abgeschlossen. Abbildung 6.2 zeigt die typische Gestalt der resultierenden Verteilungsfunktion. Sie ist in $t = \kappa = 7$ stetig, besitzt an dieser Stelle jedoch einen „Knick“. Je stärker ausgeprägt dieser Knick ist, desto ungünstiger ist dies für die von uns genutzte Komplexitätsreduzierung mittels Momentenanpassung (s. Abschnitt 6.6.3). Die Approximation der Verteilungsfunktion wird um $t = \kappa$ relativ ungenau sein. Wie zuvor im Falle der einfachen Anschlusssicherung, wird diese Problematik im Allgemeinen durch folgende Faltungen mit Verteilungsfunktionen von Quellverspätungen entschärft.

Wir können allerdings auch schon während der Berechnung der Verteilungsfunktion (6.11) eine alternative Herangehensweise verwenden. Kritisch sind die Faktoren des Produktes

$$\prod_{i \in \mathcal{Z}} (F_{X_i}(t) + f_{i,\kappa})$$

mit $f_{i,\kappa} := 1 - F_{X_i}(\kappa)$, $i \in \mathcal{Z}$. Diese besitzen einen Knick in $t = \kappa$, der letztlich den Knick in der resultierenden Verteilungsfunktion erklärt. Abhängig von der Kardinalität der Menge \mathcal{Z} , erfordert obige Produktbildung unterschiedlich viele Komplexitätsreduzierungen (der Zwischenergebnisse). Da die von uns gewählte Komplexitätsreduzierung mittels Momentenanpassung einen solchen Knick nur schlecht approximiert, sind in Folge der Produktbildung der Zwischenergebnisse Abweichungen bezüglich der Funktionswerte der exakten Verteilungsfunktion zu erwarten. Wir

betrachten den Fall $\mathcal{Z} = \{i_1, i_2\}$. Dann gilt

$$\begin{aligned} & \prod_{i \in \mathcal{Z}} (F_{X_i}(t) + f_{i,\kappa}) \\ &= F_{X_{i_1}} F_{X_{i_2}}(t) + f_{i_2,\kappa} \cdot F_{X_{i_1}}(t) + f_{i_1,\kappa} \cdot F_{X_{i_2}}(t) + f_{i_1,\kappa} \cdot f_{i_2,\kappa} . \end{aligned}$$

Der Vorteil dieser Herangehensweise besteht in der Gestalt der Funktionen der einzelnen Summanden. In der Regel besitzt keine dieser Funktionen einen Knick, außer F_{X_1} bzw. F_{X_2} weisen bereits einen solchen auf. Daher können dann alle Summanden auch hinsichtlich der Gestalt der jeweiligen Funktion gut approximiert werden. Die Anzahl der durchzuführenden Operationen steigt jedoch an.

6.5. Zugfolgerestriktionen

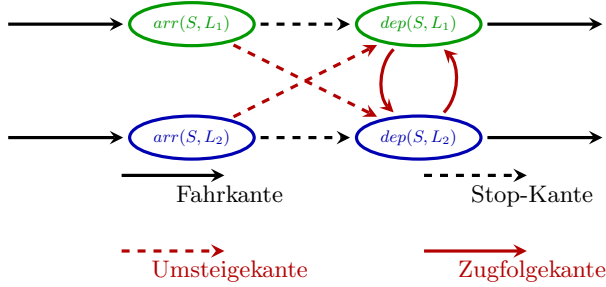
Zugfolgerestriktionen werden an dieser Stelle nur der Vollständigkeit halber betrachtet. Sie werden bisher nicht bei der Bestimmung der Verspätungsverteilungen berücksichtigt. Soll dies in Zukunft geschehen, so wird das Problem zu lösen sein, in welcher Form die resultierenden Verteilungsfunktionen repräsentiert werden können, da die Theta-Exponentialpolynome gegen die zugehörigen Operationen nicht abgeschlossen sind. Entweder wird auf eine gänzlich andere Verteilungsfamilie zurückgegriffen, oder es werden Lösungen in Betracht gezogen, wie sie zum Abschluss dieses Abschnitts erwähnt werden.

6.5.1. Theoretische Umsetzung

Zur Veranschaulichung der Berücksichtigung von Zugfolgerestriktionen nutzen wir Abbildung 6.3. Wir sehen dort einen Ausschnitt eines Ereignis-Aktivitätsnetzwerks in einer Station $S \in \mathfrak{S}$. Zwischen den beiden Linien $L_1 \in \mathfrak{L}$ und $L_2 \in \mathfrak{L}$ bestehen Umsteigebeziehungen. Außerdem existieren bezüglich der Abfahrt von Zügen der Linien Zugfolgerestriktionen. Wir nehmen an, dass von den Abfahrtsereignissen (in allen Perioden) fahrplanmäßig $dep(S, L_1)$ vor $dep(S, L_2)$ stattfindet. Als weitere Größen benötigen wir die Mindestzugfolgezeiten $head_{dep}(L_1, L_2, S)$ und $head_{dep}(L_2, L_1, S)$ sowie die zeitlichen Puffer $s(headdep, S, L_1, L_2)$ und $s(headdep, S, L_2, L_1)$. Wir nehmen an, dass Reihenfolgetausche bezüglich der Abfahrten der beiden Linien möglich sind.

Wir wollen die Verteilung der Verspätung $Y(dep, S, L_2)$ bestimmen. Es sei

$$\begin{aligned} Y_{L_1, L_2} := \max\{ & Y(arr, S, L_1) - s(change, S, L_1, L_2), \\ & \min\{Y(arr, S, L_1) + D(stop, S, L_1) - s(headdep, S, L_1, L_2), \\ & head_{dep}(S, L_1, L_2)\} \} . \end{aligned}$$

Abbildung 6.3.: Zugfolgerestriktionen bezüglich zweier Linien in Station S

die Verspätung, die von Linie L_1 auf das Abfahrtsereignis $dep(S, L_2)$ übertragen wird (s. Abbildung 6.3). Dann gilt

$$Y(dep, S, L_2) = \max\{Y(arr, S, L_2) + D(stop, S, L_2) - s(stop, S, L_2), Y_{L_1, L_2}, 0\}.$$

Linie L_1 kann im Ereignis-Aktivitätsnetzwerk auf zwei Pfaden Verspätung übertragen, über die Umsteigeaktivität sowie über die Zugfolgeaktivität. In letzterem Fall ist die übertragene Verspätung jedoch durch die Mindestzugfolgezeit begrenzt, da wir (s.o.) Reihenfolgetausche zulassen. Auf beiden Pfaden wird die Verspätung $Y(arr, S, L_1)$ fortgepflanzt. Es konnte kein Weg gefunden werden, die Verteilung der Größe Y_{L_1, L_2} zu bestimmen, ohne gegen die Voraussetzung stochastischer Unabhängigkeit (bei Faltung- oder Produktbildung) zu verstoßen. Der Ansatzpunkt war, Y_{L_1, L_2} bedingt bezüglich $Y(arr, S, L_1)$ zu berechnen. Es ist plausibel, dass die stochastische Abhängigkeit an dieser Stelle vergleichsweise stark sein kann.

Im Folgenden zeigen wir eine alternative Herangehensweise, welche auch BÜKER in seiner Dissertation [Bü10] vorstellt. Wir nehmen an, dass alle Züge gleichwertig behandelt werden. Der Zug, der theoretisch zuerst abfahren könnte, tut dies auch tatsächlich. BÜKER betrachtet und modelliert stattdessen auch mögliche Prioritäten für bestimmte Züge bzw. Linien. Wir untersuchen nun die Situation für Züge z_1 und z_2 der Linien L_1 bzw. L_2 , deren Abfahrten fahrplanmäßig aufeinander folgen. Dabei nutzen wir im Sinne der Übersichtlichkeit der Formeln

- $s_{1,2} := s(headdep, S, L_1, L_2)$, $s_{2,1} := s(headdep, S, L_2, L_1)$
- $h_{1,2} := head_{dep}(L_1, L_2, S)$, $h_{2,1} := head_{dep}(L_2, L_1, S)$.

Es gilt

$$\begin{aligned} s_{1,2} &= \pi_{z_2}(ab, S, L_2) - \pi_{z_1}(ab, S, L_1) - h_{1,2}, \\ s_{2,1} &= \pi_{z_1+1}(ab, S, L_1) - \pi_{z_2}(ab, S, L_2) - h_{2,1}. \end{aligned}$$

Verläuft alles fahrplanmäßig, wird natürlich keine Verspätung übertragen. Ob und wie viel Verspätung durch die Zugfolgerestriktionen übertragen wird hängt, neben Mindestzugfolge- und Pufferzeiten, von dem Verhältnis der Verspätungen der Züge zueinander ab. Es seien $Y^*(dep, S, L_1)$ und $Y^*(dep, S, L_2)$ die Abfahrtsverspätungen der Linien L_1 und L_2 , falls keine Zugfolgerestriktionen berücksichtigt werden. Wir interessieren uns für die Zufallsvariable

$$Y^* := Y^*(dep, S, L_1) - Y^*(dep, S, L_2)$$

und unterscheiden dabei mehrere Fälle. Zunächst betrachten wir die Situation, in der Zug z_1 Verspätung auf Zug z_2 überträgt. Fände die Abfahrt von z_2 , ohne Einfluss der Zugfolgerestriktionen, fahrplanmäßig statt, so überträgt z_1 erst dann eine Verspätung auf z_2 , wenn seine eigene Verspätung größer als der vorhandene Puffer $s_{1,2}$ ist. Ab einer bestimmten Höhe seiner Verspätung überträgt er schließlich keine Verspätung mehr. Allgemeiner beschreiben wir diese Situation durch die beiden Ereignisse

$$E_1 := [Y^* \leq s_{1,2}] , \tag{6.12}$$

$$E_2 := [s_{1,2} < Y^* \leq s_{1,2} + h_{1,2}] . \tag{6.13}$$

Tritt E_1 ein, so wird keine Verspätung übertragen. Theoretisch wäre es möglich, dass in diesem Fall der Zug z_2 der Linie L_2 so verspätet ist, dass er Verspätung auf den Zug $z_1 + 1$ der Linie L_1 in der nächsten Periode überträgt. Dieser Umstand wird an dieser Stelle aber ausgeschlossen. Tritt das Ereignis E_2 ein, so überträgt z_1 eine Verspätung der Höhe $Y^* - s_{1,2}$ auf z_2 . Die fahrplanmäßige Reihenfolge für die Abfahrt der Züge wird eingehalten. Im folgenden Fall ist dem nicht mehr so.

$$E_3 := [s_{1,2} + h_{1,2} < Y^* \leq s_{1,2} + h_{1,2} + h_{2,1}] \tag{6.14}$$

Tritt Ereignis E_3 ein, so ist der Zug z_1 der Linie L_1 , relativ zur Verspätung des Zuges z_2 der Linie L_2 , so stark verspätet, dass die fahrplanmäßige Reihenfolge für Abfahrten nicht eingehalten werden kann. Zug z_2 fährt nun zuerst ab und überträgt eine Verspätung der Höhe $Y^* - s_{1,2} - h_{1,2}$ auf Zug z_1 . Letztlich kann die Differenz der Abfahrtsverspätungen offensichtlich so groß werden, dass es zwar zu einem Reihenfolgetausch der Abfahrten kommt, aber keine Verspätung mehr zwischen den Zügen der beiden Linien übertragen wird. Dieser Fall wird durch das Ereignis E_4 beschrieben.

$$E_4 := [s_{1,2} + h_{1,2} + h_{2,1} < Y^*] \tag{6.15}$$

Nun bleibt zu untersuchen, wie die resultierenden Verteilungsfunktionen bei Eintritt eines der Ereignisse aussehen.

Wir untersuchen im Folgenden die Verspätung der Linie L_2 . Es gilt

$$\mathbf{P}(Y(dep, S, L_2) \leq t) = \sum_{i=1}^4 \mathbf{P}(Y(dep, S, L_2) \leq t \mid E_i) \cdot \mathbf{P}(E_i)$$

für alle $t \in \mathbb{R}_{\geq 0}$. Im Folgenden sei zur Vereinfachung der Darstellung $Y_{L_1}^* := Y^*(dep, S, L_1)$ sowie $Y_{L_2}^* := Y^*(dep, S, L_2)$. Wir setzen voraus, dass die Zufallsvariablen $Y_{L_1}^*$ und $Y_{L_2}^*$ stochastisch unabhängig sind. Ereignis E_1 ist identisch mit dem Ereignis

$$E_1' = [Y^*(dep, S, L_2) - Y^*(dep, S, L_1) \geq -s_{1,2}]$$

und daher

$$\mathbf{P}(Y(dep, S, L_2) \leq t \mid E_1) = \mathbf{P}(Y_{L_2}^* \leq t \mid E_1') = \frac{1}{\mathbf{P}(E_1')} \cdot \mathbf{P}([Y_{L_2}^* \leq t] \cap E_1')$$

für alle $t \in \mathbb{R}_{\geq 0}$. Denn wie oben erläutert, wird im Falle des Ereignisses E_1 keine (zusätzliche) Verspätung übertragen. Mit

$$\begin{aligned} \mathbf{P}(E_1') &= 1 - \mathbf{P}(Y^*(dep, S, L_2) - Y^*(dep, S, L_1) < -s_{1,2}) \\ &= \int_{\mathbb{R}} 1 \, dF_{Y_{L_2}^*}(u) - \int_{\mathbb{R}} F_{-Y_{L_1}^*}(-s_{1,2} - u) \, dF_{Y_{L_2}^*}(u) \end{aligned}$$

folgt für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} \mathbf{P}(Y(dep, S, L_2) \leq t \mid E_1) \\ = \frac{1}{\mathbf{P}(E_1')} \cdot \left(F_{Y_{L_2}^*}(t) - \int_{-\infty}^t F_{-Y_{L_1}^*}(-s_{1,2} - u) \, dF_{Y_{L_2}^*}(u) \right). \end{aligned}$$

Im Falle des Ereignisses E_2 gilt für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} \mathbf{P}(Y(dep, S, L_2) \leq t \mid E_2) &= \mathbf{P}(Y_{L_2}^* + Y^* - s_{1,2} \leq t \mid E_2) \\ &= \frac{1}{\mathbf{P}(E_2)} \cdot \mathbf{P}([Y_{L_1}^* - s_{1,2} \leq t] \cap E_2) \\ &= \frac{1}{\mathbf{P}(E_2)} \cdot \mathbf{P}([Y_{L_1}^* - s_{1,2} \leq t] \\ &\quad \cap [0 \leq Y_{L_1}^* - s_{1,2} - Y_{L_2}^* \leq h_{1,2}]) . \end{aligned}$$

Mit $Y_{L_1,s}^* := Y_{L_1}^* - s_{1,2}$ folgt für alle $t \in \mathbb{R}_{\geq 0}$

$$\mathbf{P}(Y(dep, S, L_2) \leq t \mid E_2) = \frac{1}{\mathbf{P}(E_2)} \cdot \left(\int_{\infty}^t F_{-Y_{L_2}^*}(h_{1,2} - u) \, dF_{Y_{L_1,s}^*}(u) \right)$$

$$- \int_{-\infty}^t F_{-Y_{L_2}^*}(-u) dF_{Y_{L_1,s}^*}(u) \Big) .$$

Es seien $c_1 := -s_{1,2} - h_{1,2}$ und $c_2 := -s_{1,2} - h_{1,2} - h_{2,1}$. Im Falle des Ereignisses E_3 gilt für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} \mathbf{P}(Y(dep, S, L_2) \leq t \mid E_3) &= \mathbf{P}(Y_{L_2}^* \leq t \mid E_3) = \frac{1}{\mathbf{P}(E_3)} \cdot \mathbf{P}([Y_{L_2}^* \leq t] \cap E_3) \\ &= \frac{1}{\mathbf{P}(E_3)} \cdot \left(\int_{-\infty}^t F_{-Y_{L_1}^*}(c_1 - u) dF_{Y_{L_2}^*}(u) - \int_{-\infty}^t F_{-Y_{L_1}^*}(c_2 - u) dF_{Y_{L_2}^*}(u) \right) . \end{aligned}$$

Schließlich gilt im Falle des Ereignisses E_4 für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} \mathbf{P}(Y(dep, S, L_2) \leq t \mid E_4) &= \mathbf{P}(Y_{L_2}^* \leq t \mid E_4) \\ &= \frac{1}{\mathbf{P}(E_4)} \cdot \mathbf{P}([Y_{L_2}^* \leq t] \cap E_4) \\ &= \frac{1}{\mathbf{P}(E_4)} \cdot \int_{-\infty}^t F_{-Y_{L_1}^*}(c_2 - u) dF_{Y_{L_2}^*}(u) . \end{aligned}$$

Analog erhält man die Verspätung der Linie L_1 bei Berücksichtigung von Zugfolgerestriktionen. Auch hier gilt

$$\mathbf{P}(Y(dep, S, L_1) \leq t) = \sum_{i=1}^4 \mathbf{P}(Y(dep, S, L_1) \leq t \mid E_i) \cdot \mathbf{P}(E_i) .$$

Einschränkend muss an dieser Stelle erwähnt werden, dass die Repräsentation von Verteilungsfunktionen mittels Theta-Exponentialpolynomen im Rahmen der Verspätungsfortpflanzung unter Berücksichtigung von Zugfolgerestriktionen an seine Grenzen stößt. In den hergeleiteten Formeln für die Ereignisse E_i , $1 \leq i \leq 4$, werden die Verteilungsfunktionen $F_{-Y_{L_1}^*}$ und $F_{-Y_{L_2}^*}$ verwendet. Diese Verteilungsfunktionen können wir nur für einzelne $t \in \mathbb{R}$ auswerten (s. Abschnitt 6.5.2). Sie können jedoch nicht als Theta-Exponentialpolynom dargestellt werden.

6.5.2. Berechnung - Subtraktion

Berechnung der Verteilungsfunktion

Es seien Y_1, Y_2 nicht-negative stochastisch unabhängige Zufallsvariablen. Wie im vorherigen Abschnitt gesehen, wird bei der Betrachtung von Zugfolgezeiten die Verteilungsfunktion der Differenz $Y := Y_1 - Y_2$ solcher Zufallsvariablen benötigt. Während solche Verteilungsfunktionen nicht als Theta-Exponentialpolynom dargestellt werden können, gibt es zumindest die Möglichkeit, Werte $F_Y(t)$ für alle $t \in \mathbb{R}$ zu berechnen.

Es gilt für alle $t \in \mathbb{R}$

$$\begin{aligned} F_{-Y_2}(t) &= \mathbf{P}(-Y_2 \leq t) = \mathbf{P}(Y_2 \geq -t) = 1 - \mathbf{P}(Y_2 < -t) \\ &= \mathbf{1}_{(-\infty, 0)}(t) \cdot (1 - F_{Y_2}(-t)) + \mathbf{1}_{[0, \infty)}(t) . \end{aligned}$$

F_{Y_2} und F_{-Y_2} sind im Rahmen unserer Anwendung im Allgemeinen keine stetigen Verteilungsfunktionen. In $t = 0$ besitzt F_{-Y_2} eine Sprungstelle der Höhe $F_{Y_2}(0)$. Es ist möglich, F_{-Y_2} durch $F_{-Y_2} := F_1 + F_2$ mit $F_1 := F_{1,1} + F_{1,2}$,

$$F_{1,1}(t) = \mathbf{1}_{(-\infty, 0)}(t) \cdot (1 - F_{Y_2}(-t)) , \text{ für alle } t \in \mathbb{R} ,$$

$$F_{1,2}(t) = \mathbf{1}_{[0, \infty)}(t) \cdot \lim_{s \rightarrow -0} F_{1,1}(s) = \mathbf{1}_{[0, \infty)}(t) \cdot (1 - F_{Y_2}(0)) , \text{ für alle } t \in \mathbb{R} ,$$

und

$$F_2(t) = \mathbf{1}_{[0, \infty)}(t) \cdot F_{Y_2}(0)$$

für alle $t \in \mathbb{R}$ zu beschreiben. F_1 ist eine stetige Funktion und auf $\mathbb{R} \setminus \{0\}$ differenzierbar. F_2 ist eine Treppenfunktion mit Sprunghöhe $F_{Y_2}(0)$ in $t = 0$. Für die Verteilungsfunktion von Y folgern wir

$$\begin{aligned} F_Y(t) &= \int_{\mathbb{R}} F_{Y_1}(t-s) dF_{-Y_2}(s) = \int_{\mathbb{R}} F_{Y_1}(t-s) dF_1(s) + \int_{\mathbb{R}} F_{Y_1}(t-s) dF_2(s) \\ &= \int_{(-\infty, 0)} F_{Y_1}(t-s) dF_1(s) + \int_{[0, \infty)} F_{Y_1}(t-s) dF_1(s) + F_{Y_1}(t) \cdot F_{Y_2}(0) \\ &= \int_{(-\infty, 0)} F_{Y_1}(t-s) dF_{1,1}(s) + \int_{[0, \infty)} F_{Y_1}(t-s) dF_{1,2}(s) + F_{Y_1}(t) \cdot F_{Y_2}(0) \\ &= \int_{(-\infty, 0)} F_{Y_1}(t-s) \cdot f_{1,1}(s) ds + F_{Y_1}(t) \cdot F_{Y_2}(0) \end{aligned}$$

mit $f_{1,1}(t) := \frac{d}{dt} F_{1,1}(t)$, $\mathbb{R} \setminus \{0\}$. Für $t < 0$ gilt

$$f_{1,1}(t) = f_{Y_2}(-t) ,$$

für $t > 0$ gilt $f_{1,1}(t) = 0$. So erhalten wir

$$\begin{aligned} F_Y(t) &= \int_{-\infty}^0 F_{Y_1}(t-s) \cdot f_{Y_2}(-s) ds + F_{Y_1}(t) \cdot F_{Y_2}(0) \\ &= \int_0^{\infty} F_{Y_1}(t+s) \cdot f_{Y_2}(s) ds + F_{Y_1}(t) \cdot F_{Y_2}(0) . \end{aligned}$$

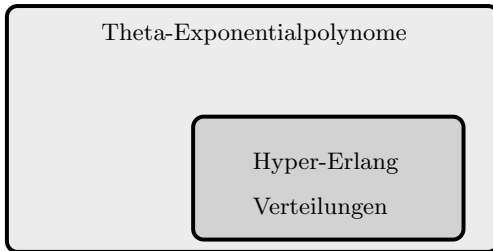


Abbildung 6.4.: Beziehung zwischen Theta-Exponentialpolynomen und Hyper-Erlangverteilungen

Simulation der ersten drei Momente

Eine Möglichkeit, die Familie der Theta-Exponentialpolynome auch bei Berücksichtigung von Zugfolgerestriktionen zu verwenden wäre, ein insgesamt „hybrides“ Verfahren zu verwenden. Dabei könnten, speziell für die Bestimmung der Verteilungsfunktion von Differenzen von Zufallsvariablen, die ersten drei Momente simuliert werden. Im Anschluss könnte eine Hyper-Erlangverteilung bestimmt werden, deren erste drei Momente mit diesen übereinstimmen. Umgesetzt wird diese Herangehensweise in dieser Arbeit nicht mehr.

6.6. Komplexitätsreduzierung

Wie zuvor erwähnt, hat die Repräsentation von Verteilungsfunktionen durch Theta-Exponentialpolynome das Problem, dass im Zuge der fortdauernden Verknüpfung verschiedener Verteilungsfunktionen die Anzahl M der Parametertupel $(\theta_m, a_m, k_m, \lambda_m)$ rapide ansteigt. Dies führt während der Berechnung auf einem Rechner schnell zu Speicherplatz- und numerischen Problemen (s. [Fuh07]). Um die Bestimmung von Verspätungsverteilungen ganzer Netzwerke praktikabel zu halten, muss ein Weg gefunden werden, die Komplexität der Theta-Exponentialpolynome zu reduzieren. Dabei soll möglichst wenig Information der Verteilung verloren gehen. Abbildung 6.4 veranschaulicht die Beziehung zwischen Theta-Exponentialpolynomen und Hyper-Erlangverteilungen.

In [Fuh07] wurde die direkte Manipulation der Theta-Exponentialpolynome bereits erprobt. Dies geschah insbesondere vor dem Hintergrund, die Koeffizienten der Summanden des Theta-Exponentialpolynoms weder zu groß noch zu klein werden zu lassen, und so die Berechnungen numerisch stabil zu halten. So war eine naheliegende Idee, dass man die Summanden mit sehr kleinen Koeffizienten aus der Darstellung streicht. Hierdurch verlässt man jedoch den Raum der Verteilungs-

funktionen. Um zurück in die Schnittmenge der Theta-Exponentialpolynome und Wahrscheinlichkeitsverteilungen zu kommen, muss das Theta-Exponentialpolynom wieder zu einer Verteilungsfunktion normiert werden. Dieser Ansatz wird in [Fuh07] als wenig erfolgreich und zudem ungenau beurteilt. Es dürfte auch schwierig sein, den Wert allgemein festzulegen, unter dem Koeffizienten liegen müssen, damit der zugehörige Summand gestrichen wird. Ansonsten wäre dieser Ansatz auch für die Reduzierung der Anzahl der Parametertupel erwägenswert gewesen.

Als Ursache für sehr große Koeffizienten wurde in [Fuh07] die Faltungsoperation (s. Satz 6.1.5) ausgemacht, sofern hier Polynome mit nur unwesentlichen Unterschieden bei Paaren von λ_m -Werten gefaltet werden sollen. Hier wurde versucht, bei nahezu identischen Paaren, diese gleichzusetzen und bei der Faltungsoperation den zu berechnenden Koeffizienten bezüglich der entsprechenden Vorschrift zu bilden. Ein Nebeneffekt wäre, dass bei der Faltungsoperation ein starkes Anwachsen der Anzahl der Parametertupel vermieden werden könnte. Aber wie in [Fuh07] festgestellt wird, ist es auch in diesem Fall schwer einen Wert festzulegen, ab dem dieses Gleichsetzen durchgeführt werden soll.

Schließt man die Reduzierung der Komplexität mittels direkter Manipulation der Parametertupel der Theta-Exponentialpolynome aus, bleibt unter anderem die Möglichkeit einer Approximation der Verteilungsfunktion auf Grundlage bestimmter Kenngrößen der Verteilung. Das Ergebnis der Approximation muss wieder als Theta-Exponentialpolynom dargestellt werden können. Über die Approximation von allgemeinen Verteilungen durch Phasentypverteilungen wurde bereits in Kapitel 5 einiges gesagt. Wesentlich war, dass es bei der Approximation verschiedene Abstandsbegriffe für den Vergleich von Verteilungsfunktionen gibt. In Kapitel 9 wird deutlich, dass sowohl die Momente der Verteilung als auch die Gestalt der Verteilungsfunktion selbst möglichst gut approximiert werden sollte. Darüber hinaus ist für unsere Anwendung entscheidend, dass das Verfahren zur Komplexitätsreduzierung möglichst simpel und wenig zeitintensiv ist.

6.6.1. Metriken für Verteilungsfunktionen

Um Unterschiede in der Gestalt von Verteilungsfunktionen zu beschreiben, kann man sich Metriken für Verteilungen bedienen. Dabei werden die Verteilungen durch ihre Verteilungsfunktionen repräsentiert. Zwei Beispiele sind die folgenden Metriken (s. [Rac91]):

$$\rho(F, G) := \sup\{|F(t) - G(t)| \mid t \in \mathbb{R}_{\geq 0}\} \quad (\text{KOLMOGOROV}) \quad (6.16)$$

$$\kappa(F, G) := \int_0^\infty |F(t) - G(t)| dt \quad (\text{KANTOROVICH}) \quad (6.17)$$

Diese Metriken sind für beliebige Verteilungsfunktionen F und G auf $\mathbb{R}_{\geq 0}$ definiert. Sie werden auch von MEESTER und MUNS in [MM07] verwendet, um die

Fortpflanzung von Fehlern nach Operationen aus Abschnitt 6.3 abzuschätzen (s. auch Abschnitt 6.7).

Wollen wir Metriken verwenden, müssen wir zuvor die Frage klären, ob unsere Art der Repräsentation von Verteilungsfunktionen deren Berechnung ermöglicht. Im Falle von Theta-Exponentialpolynomen können beide Metriken grundsätzlich analytisch berechnet werden. Da die Familie der Theta-Exponentialpolynome bezüglich der Betragsbildung nicht abgeschlossen ist, müssten in beiden Fällen zunächst die Nullstellen der Funktion $H := F - G$ bestimmt werden. Dies muss im Allgemeinen numerisch, beispielsweise mittels des NEWTON-Verfahrens, gelöst werden. Im Anschluss folgt, auf Teilintervallen des $\mathbb{R}_{\geq 0}$, eine Extremwertbestimmung für ρ bzw. Integration für κ . Die Familie der Theta-Exponentialpolynome ermöglicht ein solches Vorgehen zwar grundsätzlich, letztlich erweist sich aber insbesondere die Nullstellenberechnung als zeitlich relativ aufwendig.

Eine Alternative zur analytischen ist die näherungsweise Berechnung der beiden Metriken. Diese erfordert im Falle ρ die Auswertung zweier Verteilungsfunktionen an nur endlich vielen Stellen. Präzisiert wird dieses Vorgehen in Algorithmus 6.6.1. Wie häufig die Verteilungsfunktionen F und G ausgewertet werden müssen, hängt bei gegebenem c von der Wahl des Δt ab. Für $\Delta t \rightarrow 0$ erhalten wir den exakten Wert von ρ in (6.16). Allerdings wird sich dies auch negativ auf die Berechnungsdauer auswirken.

Input : Verteilungsfunktionen F und G

```

1  $\rho(F, G) := 0$ ;  $t := 0$ ;
2 while  $t \in [0, c]$  do
3    $\rho(F, G) = \max\{\rho(F, G), |F(t) - G(t)|\}$ ;
4    $t = t + \Delta t$ ;
5 return  $\rho(F, G)$ ;
```

Algorithmus 6.6.1 : Näherungsweise Berechnung der Metrik ρ

Metrik κ kann in Form von (6.17) nicht berechnet werden. Die Familie der Theta-Exponentialpolynome ist gegen Betragsbildung nicht abgeschlossen. Für die näherungsweise Berechnung der Metrik κ soll Algorithmus 6.6.2 verwendet werden. Im Folgenden soll der Zusammenhang zwischen der eigentlichen Metrik und dem Algorithmus gezeigt werden. Wir betrachten die Pseudometrik

$$\kappa^*(F, G) := \left| \int_0^\infty (F(t) - G(t)) dt \right| \leq \kappa(F, G) .$$

Eine Pseudometrik unterscheidet sich von einer Metrik darin, dass aus $\kappa^*(F, G) = 0$

nicht $F \equiv G$ folgen muss. Dass es sich bei κ^* um eine Pseudometrik handelt, wird unter anderem aus dem folgenden Zusammenhang deutlich. Sind X und Y Zufallsvariablen auf $\mathbb{R}_{\geq 0}$ mit Verteilungsfunktionen F bzw. G , so gilt $|\mathbf{E}[X] - \mathbf{E}[Y]| = \kappa^*(F, G)$. Erwartungswerte allein legen eine Verteilung bekanntlich nicht eindeutig fest. Abgesehen davon hat κ^* aber den Vorteil, dass man diese Pseudometrik bei Verwendung von Theta-Exponentialpolynomen berechnen kann, da diese gegen Integration abgeschlossen sind. Für κ^* muss zunächst der Integrand berechnet werden. Das entstehende Theta-Exponentialpolynom wird in Algorithmus dann an den Stellen c_i , $1 \leq i \leq n$, ausgewertet.

Mit Kenntnis der Nullstellen könnte κ auch exakt berechnet werden. Es sei im Folgenden weiterhin $H = F - G$. Repräsentieren die c_i mit $1 \leq i \leq n - 1$ die Nullstellen der Funktion H und ist $c_0 := 0$ sowie $c_n = \infty$, so gilt

$$\kappa(F, G) := \int_0^\infty |F(t) - G(t)| dt = \sum_{i=1}^n \left| \int_{c_{i-1}}^{c_i} (F(t) - G(t)) dt \right|.$$

Input : Verteilungsfunktionen F und G

```

1  $\kappa(F, G) := 0;$ 
2 for  $i = 1, i \leq n$  do
   //  $H(c_i, c_j) := \left| \int_{c_i}^{c_j} F(t) - G(t) dt \right|$ , mit  $c_i, c_j \in \mathbb{R}_{>0}$ 
3    $\kappa(F, G) = \kappa(F, G) + H(c_{i-1}, c_i);$ 
4    $i \leftarrow i + 1;$ 
5 return  $\kappa(F, G);$ 

```

Algorithmus 6.6.2 : Näherungsweise Berechnung der Metrik κ

Allerdings wird diese Berechnung, wie schon oben beschrieben, durch die Nullstellenberechnung unpraktikabel. Algorithmus 6.6.2 soll ein Kompromiss zwischen κ^* und κ sein. Die Wahl der c_i hat hierbei entscheidenden Einfluss auf das Ergebnis. Letztlich berechnet Algorithmus 6.6.2 also

$$\sum_{i=1}^n \left| \int_{c_{i-1}}^{c_i} (F(t) - G(t)) dt \right|$$

mit $c_0 = 0$ sowie $c_n < \infty$.

6.6.2. Approximation der Gestalt einer Verteilungsfunktion

Im Rahmen dieser Arbeit wurden drei Verfahren untersucht, welche bei der Approximation einer gegebenen Verteilungsfunktion die Gestalt berücksichtigen. Das erste Verfahren ist das in Abschnitt 5.3.1 detaillierter betrachtete von FELDMANN und WHITT.

Hyper-Exponentialverteilungen

Ein Argument für das Verfahren aus [FW98] ist dessen Einfachheit, welche sich in einer niedrigen Rechenzeit niederschlägt. Da das Verfahren für die Approximation Hyper-Exponentialverteilungen verwendet, beschränkt sich der natürliche Anwendungsbereich des Verfahrens auf Verteilungen mit monotoner Dichtefunktion. Die von uns betrachteten Verteilungen der Quellverspätungen sowie die während der Verspätungsfortpflanzung entstehenden Verteilungen besitzen zum Teil gar keine Dichtefunktion (Sprungstelle in $t = 0$). Besitzen sie eine, muss diese nicht monoton sein. Daher ließe sich das Verfahren aus [FW98] für unsere Zwecke lediglich auf Einzelfälle anwenden.

Simulation und EM-Algorithmus

Eine weitere Idee war, den EM-Algorithmus aus [TBT06] nicht nur für die Approximation empirischer Verteilungsfunktionen zu nutzen. Möchte man die Komplexität einer Verteilungsfunktionsdarstellung reduzieren, so wäre ein erster Schritt die Simulation der Verteilung. Aus der erhaltenen Stichprobe kann dann mittels des EM-Algorithmus von THUEMMLER ET AL. eine weniger komplexe Darstellung einer Hyper-Erlangverteilung gewonnen werden. Die benötigte Rechenzeit ist von mehreren Faktoren abhängig. Dies ist zunächst die Ausführungszeit des EM-Algorithmus, welche von der Stichprobengröße und der zugelassenen Variation der in Frage kommenden Hyper-Erlangverteilungen abhängt. Die Ausführungszeit der Simulation selbst hängt, neben der Komplexität der zu simulierenden Verteilung, auch vom speziellen Simulationsverfahren ab.

Letztlich erwies sich die Kombination aus Simulation und EM-Algorithmus zwar als fähig, die Komplexität zu reduzieren. Die Verwendung des EM-Algorithmus ist jedoch zu zeitaufwendig. Daher wurde der Ansatz nicht weiter verfolgt.

Momentenanpassung und Optimierung

Wie schon erwähnt, gibt es umfangreiche Literatur zur Approximation von Verteilungen mittels Momentenanpassungsverfahren. Beiträge, wie dabei auch die Gestalt der Verteilungsfunktionen berücksichtigt werden kann, lieferten zum Beispiel JOHNSON und TAAFFE (s.o.). Die verwendeten Ansätze mittels nichtlinearer Optimierung können sich aber ebenfalls als zeitintensiv erweisen.

THUEMLER ET AL. geben darüber hinaus in [TBT06] Lösungen für den Fall an, dass man sich nicht auf Hyper-Erlangverteilungen mit Komponenten gleicher Ordnung beschränken möchte. Sie verwenden die Momentenanpassung allerdings nicht direkt für die Approximation einer empirischen Verteilungsfunktion. Dies geschieht zunächst mittels ihres EM-Algorithmus. Im Anschluss existiert die Möglichkeit, zwei Zweige der approximierten Hyper-Erlangverteilung für die Momentenanpassung zu verwenden (s. reduzierte Momente in Abschnitt 5.3.2). Sie machten dabei die Erfahrung, dass die mit dem EM-Algorithmus geschätzten Parameter durch die Momentenanpassung nur unwesentlich verändert werden. Wenn wir die Verteilungen jedoch nicht simulieren wollen, ist dieses Vorgehen nicht möglich.

Da Momentenanpassungsverfahren alleine bezüglich der Gestalt einer Verteilungsfunktion zu ungenau und das Lösen von Optimierungsproblemen zu zeitintensiv sein können, versuchten wir einen Kompromiss zu finden und schon zur Verfügung stehende Werkzeuge zu verwenden. Wir nutzen für die Approximation bzw. Momentenanpassung Konvexkombinationen aus nur zwei Erlangverteilungen. Die Ordnung der beiden Erlangverteilungen wird durch r_1 bzw. r_2 festgelegt (s. Seite 46). Der erste Schritt war immer, die Momentenanpassung für den Fall gleicher Ordnung zu lösen. Dieser Fall hat den Vorteil, dass es eine eindeutige Lösung in geschlossener Form gibt. Da die anderen beiden Fälle (s. Seite 46) numerische Methoden benötigen, sollten sie nur gelöst werden, wenn die Approximation für den Fall gleicher Ordnung ($r_1 = r_2 = n^*$) nicht exakt genug war. Die Genauigkeit wurde über eine der beiden oben beschriebenen Metriken bzw. deren approximative Berechnungen gemessen. Um den Zeitaufwand gering zu halten, wurden hier nur wenige Variationen zugelassen wie beispielsweise $r_1 + 1 = n^* = r_2$ bzw. $r_1 = n^* = r_2 - 1$. Der Gewinn bezüglich der Genauigkeit stand jedoch in keinem Verhältnis zum zusätzlichen Zeitaufwand.

Insgesamt lässt sich festhalten, dass wir auch im Falle gleicher Ordnung bereits überzeugende Ergebnisse erhalten. Deshalb wurde diese Herangehensweise Grundlage für das im nächsten Abschnitt beschriebene Verfahren.

6.6.3. Verwendetes Verfahren

Wegen der im vorherigen Abschnitt beschriebenen Erfahrungen, wurde die Idee der Kombination von Momentenanpassungsverfahren und Optimierung hinsichtlich der Genauigkeit der Gestalt der zu approximierenden Verteilungsfunktion nicht weiter verfolgt. Solange keine realistischen Daten zu Quellverspätungen zur Verfügung stehen, scheint ein aussagekräftiger Vergleich zwischen zusätzlichem Zeitaufwand und tatsächlich gewonnener Genauigkeit nicht möglich. Daher wurde die zeitlich sehr effiziente Momentenanpassung für die Konvexkombination zweier Erlangverteilungen gleicher Ordnung ($r_1 = r_2$) implementiert. Die Genauigkeit wird somit ausschließlich bezüglich der ersten drei Momente einer Verteilung bemessen.

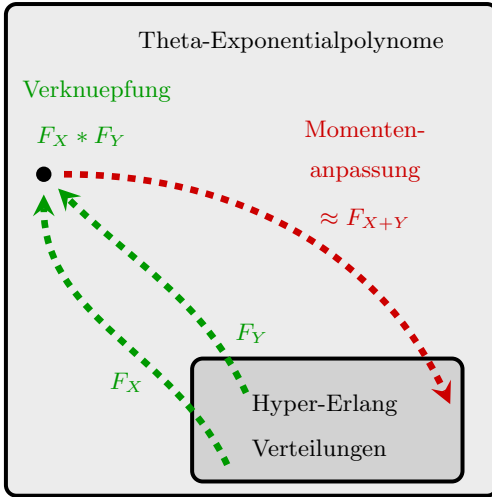


Abbildung 6.5.: Komplexitätsreduzierung

Ein Vorteil dieser Vorgehensweise ist es, dass man (relativ aufwendige) Faltungsoperationen auf Ebene der Theta-Exponentialpolynome durch die direkte Verrechnung der Momente ersetzen kann. Bezüglich der Genauigkeit in den Momenten werden hier keine (zusätzlichen) Ungenauigkeiten bzw. Fehler produziert. Ein Problem ergab sich jedoch noch, wenngleich dieses in der Praxis selten auftritt. Die Größe n^* (s. Seite 46) geht (indirekt) als Fakultät in die Berechnung der Momente, der Dichte- sowie der Verteilungsfunktion ein (s. Seite 38). Dies sorgt bei großen Werten von n^* zu numerischen Problemen bzw. dazu, dass die Berechnung auf einem Rechner nicht mehr durchgeführt werden kann.

In der Literatur [JT89] findet man bereits Hinweise auf dieses Problem sowie den Vorschlag, beispielsweise das dritte Moment zu verfälschen, um den Wert von n^* zu beeinflussen. Spielen für das Verhalten eines Warteschlangensystems etwa nur die ersten beiden Momente eine Rolle, hätte dies keine gravierenden Auswirkungen auf die Ergebnisse. Es sei $N_{max} \in \mathbb{N}$ die größte natürliche Zahl, für die $(N_{max} - 1)!$ noch auf dem Rechner berechnet werden kann. Weiter sei $\mu_k := \mathbf{E}[X^k]$ für $1 \leq k \leq 3$. Dann gilt Folgendes:

$$\frac{\mu_1^2}{\mu_2 - \mu_1^2} < N_{max} \iff \mu_2 > \mu_1^2 \cdot \left(1 + \frac{1}{N_{max}}\right) \quad (6.18)$$

$$\frac{\mu_2^2}{\mu_1 \mu_3 - \mu_2^2} < N_{max} \iff \mu_3 > \frac{\mu_2^2}{\mu_1} \cdot \left(1 + \frac{1}{N_{max}}\right) \quad (6.19)$$

An den beiden Gleichungen wird deutlich, dass die ersten drei Momente einer zu approximierenden Verteilung in einem bestimmten Verhältnis zueinander stehen müssen, damit diese Verteilung als Kombination zweier Erlangverteilungen gleicher Ordnung dargestellt werden kann. Die Varianz einer Zufallsvariable ist natürlich nie negativ. Es gilt also immer $\mu_2 \geq \mu_1^2$. In (6.18) sieht man, um wie viel μ_2 , in Abhängigkeit von N_{max} , echt größer sein muss als μ_1^2 . In (6.19) sehen wir einen ähnlichen Zusammenhang zwischen dem dritten Moment und den ersten beiden Momenten.

Eine Situation aus der Anwendung, in der eine solche Modifikation der Momente grundsätzlich notwendig wäre, wird gesondert behandelt. So können im Zuge der Berechnung Verspätungsverteilungen mit großem Erwartungswert und relativ kleiner Varianz auftreten. Es sei in diesem Falle X die Zufallsvariable der Verspätungsverteilung. Die grundlegende Idee ist, die Momentenanpassung nicht für die Verteilung von X , sondern für die Verteilung einer Zufallsvariablen $Y := X - a$ mit

$$a := \sup\{t \in \mathbb{R}_{\geq 0} \mid \mathbf{P}(X \leq t) = 0\} \quad (6.20)$$

durchzuführen. So erhalten wir die approximierte Verteilung der Zufallsvariablen \tilde{Y} . Im Anschluss wird die Verschiebung wieder rückgängig gemacht, wir erhalten also die Verteilung der Zufallsvariablen $Z := \tilde{Y} + a$. Es lässt sich leicht zeigen, dass $\mathbf{E}[X^k] = \mathbf{E}[Z^k]$ für $1 \leq k \leq 3$. Es gilt

$$\begin{aligned} \mathbf{E}[Y] &= \mathbf{E}[X] - a \\ \mathbf{E}[Y^2] &= \mathbf{E}[X^2] - 2 \cdot a \cdot \mathbf{E}[X] + a^2 \\ \mathbf{E}[Y^3] &= \mathbf{E}[X^3] - 3 \cdot a \cdot \mathbf{E}[X^2] + 3 \cdot a^2 \cdot \mathbf{E}[X] - a^3 \end{aligned}$$

und

$$\begin{aligned} \mathbf{E}[Z] &= \mathbf{E}[Y] + a = \mathbf{E}[X] \\ \mathbf{E}[Z^2] &= \mathbf{E}[Y^2] + 2 \cdot a \cdot \mathbf{E}[Y] + a^2 \\ &= \mathbf{E}[X^2] - 2 \cdot a \cdot \mathbf{E}[X] + a^2 + 2 \cdot a \cdot (\mathbf{E}[X] - a) + a^2 = \mathbf{E}[X^2] \\ \mathbf{E}[Z^3] &= \mathbf{E}[Y^3] + 3 \cdot a \cdot \mathbf{E}[Y^2] + 3 \cdot a^2 \cdot \mathbf{E}[Y] + a^3 \\ &= \mathbf{E}[X^3] - 3 \cdot a \cdot \mathbf{E}[X^2] + 3 \cdot a^2 \cdot \mathbf{E}[X] - a^3 \\ &\quad + 3 \cdot a \cdot (\mathbf{E}[X^2] - 2 \cdot a \cdot \mathbf{E}[X] + a^2) \\ &\quad + 3 \cdot a^2 \cdot (\mathbf{E}[X] - a) + a^3 = \mathbf{E}[X^3] . \end{aligned}$$

In der Praxis wird man den Wert a aus (6.20) natürlich näherungsweise bestimmen. Im folgenden Abschnitt zeigt Abbildung 6.10 ein Beispiel für die Vorteile einer solchen Verschiebung der Verteilung.

Eine weitere Vereinfachung findet statt, wenn für den Erwartungswert der Zufallsvariablen X der zu approximierenden Verteilung $\mathbf{E}[X] \approx 0$ gilt. In diesem Fall wird die Approximation nicht als Hyper-Erlangverteilung sondern als Indikatorfunktion $\mathbf{1}_{\mathbb{R}_{\geq 0}}$ in Form eines Theta-Exponentialpolynoms repräsentiert. In Algorithmus C.2.1 (s. Anhang) wird das Vorgehen zusammengefasst.

6.6.4. Evaluation des Verfahrens

In diesem Abschnitt wollen wir das zuvor beschriebene Verfahren zur Komplexitätsreduzierung evaluieren. Dabei wird die Güte mittels der Metrik (6.16) bzw. (näherungsweise) Algorithmus 6.6.1 gemessen. Wir interessieren uns also für die Gestalt der Verteilungsfunktion.

In Kapitel 4 haben wir bereits Hinweise erhalten, wie die Verteilungsfunktionen realistischer Quellverspätungen aussehen könnten. MEESTER und MUNS verwenden in [MM07] für die Modellierung von Quellverspätungen Weibullverteilungen, welche in ihrer Gestalt der (exakten) Verteilungsfunktion in Abbildung 6.6 ähneln. Dies widerspricht auch nicht den Eindrücken, die wir in Kapitel 4 erhalten haben. Abbildung 6.6 zeigt ein repräsentatives Beispiel von Verteilungsfunktionen, deren Gestalt besonders gut durch eine Momentenanpassung mittels Hyper-Erlangverteilungen gleicher Ordnung approximiert werden können. Der gemessene maximale Abstand zwischen exakter und approximierter Verteilungsfunktion liegt hier bei $\rho = 0.00617428$.

Weniger gut lassen sich Verteilungen approximieren, deren Verteilungsfunktionen in ihrer Gestalt denen von Exponentialverteilungen ähneln. Zwei Beispiele hierfür sind in den Abbildungen 6.7 sowie 6.8 zu sehen. Dort wurden Exponentialverteilungen mit Parametern 0.5 bzw. 0.1 durch Hyper-Erlangverteilungen approximiert. In solchen Fällen kann der maximale Abstand zwischen exakter und approximierter Verteilungsfunktion Werte im Bereich $\rho = 0.05$ annehmen.

Im Falle von Verspätungsverteilungen von Zufallsvariablen X mit $\mathbf{P}(X = 0) > 0$ ist die Approximation hinsichtlich der Metrik ρ in der Regel besser als im zuvor beschriebenen Fall. Als repräsentatives Beispiel ist in Abbildung 6.9 eine Verspätungsverteilung gegeben, welche nach Abzug eines Puffers (s. Abschnitt 6.3.1) aus der in Abbildung 6.8 gezeigten Verspätungsverteilung hervorgegangen ist. Der gemessene maximale Abstand zwischen exakter und approximierter Verteilungsfunktion liegt hier nur noch bei $\rho = 0.0192604$.

Im vorherigen Abschnitt wurde der Sonderfall thematisiert, wenn, bei einem ungünstigem Verhältnis zwischen Momenten einer Verteilung, die Momente einer zu approximierenden Verteilung zum Teil modifiziert werden müssen. Ist der Erwartungswert im Vergleich zur Varianz relativ groß, so wird geprüft, ob die Verteilung vor der Momentenanpassung verschoben werden kann. Der Vorteil dieses Vorgehens wird in Abbildung 6.10 deutlich (s. Seite 90). In Abbildung 6.10(a) ist das Ergebnis einer Momentenanpassung ohne vorherige Verschiebung zu sehen. Sowohl das

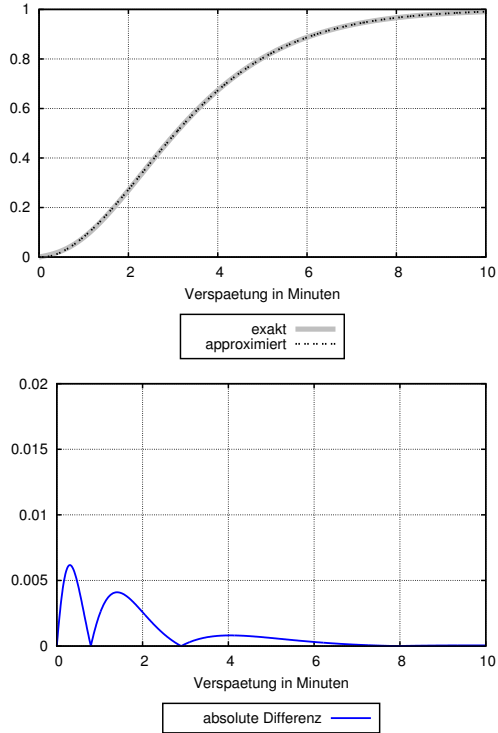


Abbildung 6.6.: $\mathbf{E}[X] = 3.42685$, $\mathbf{E}[X^2] = 16.055$, $\mathbf{E}[X^3] = 94.5939$,
 $\rho = 0.00617428$ an Stelle $t = 0.306$

6. Bestimmung der Verspätungsverteilungen

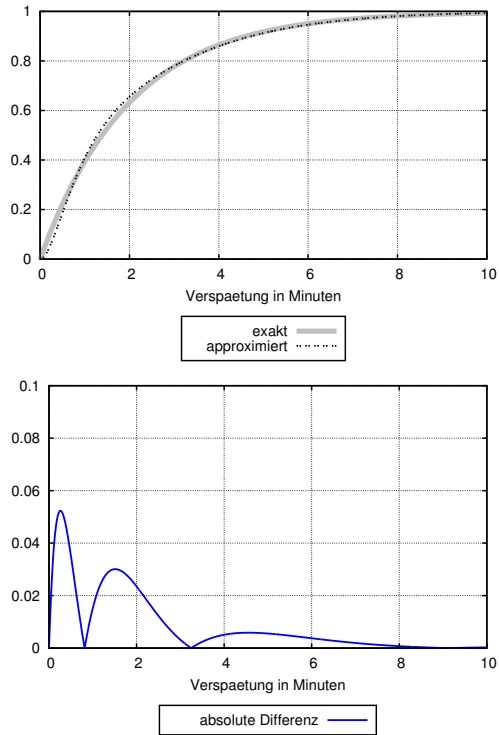


Abbildung 6.7.: $\mathbf{E}[X] = 2.0$, $\mathbf{E}[X^2] = 8.0$, $\mathbf{E}[X^3] = 48.0$,
 $\rho = 0.052355$ an Stelle $t = 0.261$

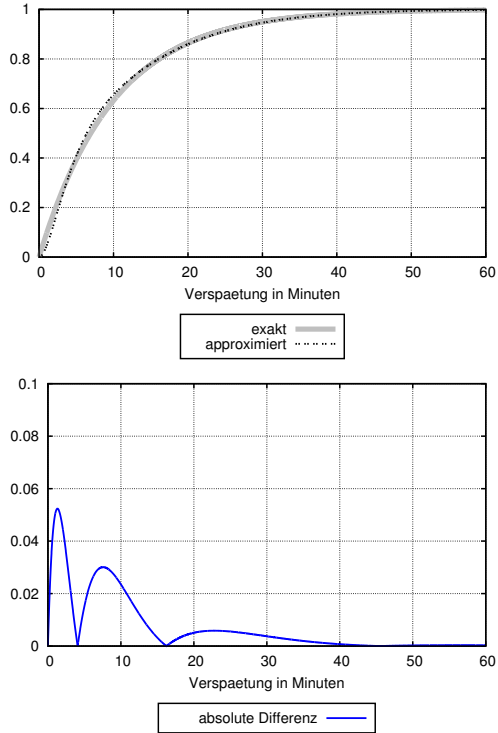


Abbildung 6.8.: $\mathbf{E}[X] = 10.0$, $\mathbf{E}[X^2] = 200.0$, $\mathbf{E}[X^3] = 6000.0$,
 $\rho = 0.0523551$ an Stelle $t = 1.307$

6. Bestimmung der Verspätungsverteilungen

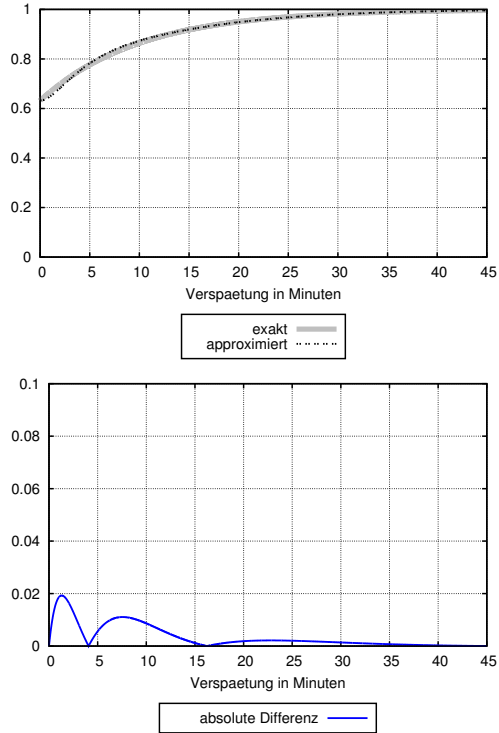


Abbildung 6.9.: $\mathbf{E}[X] = 3.67879$, $\mathbf{E}[X^2] = 73.5759$, $\mathbf{E}[X^3] = 2207.28$,
 $\rho = 0.0192604$ an Stelle $t = 1.307$

zweite als auch das dritte Moment müssen modifiziert werden, um die Verteilung auf dem Rechner darstellen zu können. In diesem Beispiel gilt $\mathbf{E}[\tilde{X}^2] = 4.6662$ sowie $\mathbf{E}[\tilde{X}^3] = 10.9706$. Die gemessene maximale Abweichung zwischen exakter und approximierter Verteilungsfunktion ist mit $\rho = 0.452332$ sehr hoch. In Abbildung 6.10(b) ist Situation dargestellt, in der vor der Momentenanpassung eine Verschiebung der zu approximierenden Verteilungsfunktion durchgeführt wurde. Das zweite und dritte Moment mussten nicht modifiziert werden. Die gemessene maximale Abweichung liegt schließlich bei nur noch $\rho = 0.0523541$.

Im folgenden Abschnitt wird diskutiert, wie die im Zuge der Approximationen erfahrenen Fehler sich, bei Anwendung weiterer Operationen, fortpflanzen können. Ein Nachweis, dass sich die Fehler nicht unkontrolliert aufschaukeln, konnte durch Vergleich mit den Ergebnissen des in Kapitel 8 vorgestellten Simulationsverfahrens geführt werden. Mittels dieses Verfahrens vergleichen wir die in der Fahrplanoptimierung (s. Kapitel 9) interessierenden Kostenfunktionswerte. Ein hiervon unabhängiger Vergleich, direkt auf Ebene der resultierenden Verteilungsfunktionen, fand in [HKKS14] statt. Als Vergleichsmodell diente ein szenarienbasierter Ansatz der Arbeitsgruppe „Optimierung“ der Georg-August-Universität Göttingen. Die Ergebnisse stimmten in hohem Maße überein.

6.7. Fehlerfortpflanzung

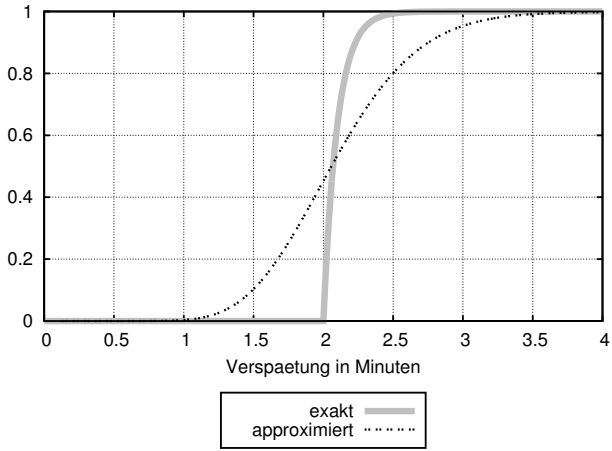
Es seien X, Y stochastisch unabhängige, positive Zufallsvariablen mit Verteilungsfunktionen F_X und F_Y . Weiter seien \tilde{X}, \tilde{Y} Zufallsvariablen mit Verteilungsfunktionen $F_{\tilde{X}}$ und $F_{\tilde{Y}}$, die durch Momentenanpassung (Approximation) aus F_X bzw. F_Y hervorgegangen sind. Wir wollen in diesem Abschnitt untersuchen, wie sich durch die Approximation ergebene Ungenauigkeiten durch weitere Verknüpfungen der Verteilungsfunktionen fortpflanzen. Bei der Auswertung (Kapitel 9) der Verteilungsfunktionen interessieren wir uns für

1. die mittleren Verspätungen, d.h. die zugehörigen Erwartungswerte, sowie
2. die Funktionswerte der jeweiligen Verteilungsfunktion.

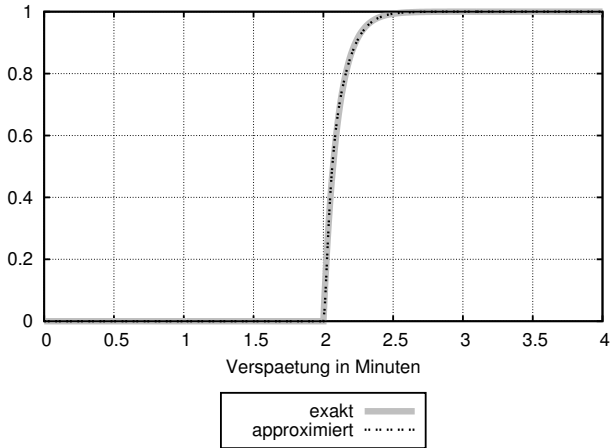
Grundlage für die nötigen Approximationen sind immer die jeweils ersten drei Momente. Nach einer Approximation bleiben wir diesbezüglich in der Regel exakt. Lediglich das zweite und dritte Moment können in Ausnahmefällen modifiziert worden sein. Es muss vermieden werden, dass sich Fehler im Zuge weiterer Operationen unkontrolliert aufsummieren und wir ungenaue Ergebnisse erhalten.

Wir beziehen uns im Folgenden auf Aussagen aus [MM07] und betrachten auch an dieser Stelle die dort verwendeten Metriken für Verteilungsfunktionen ρ und κ aus (6.16) bzw. (6.17) auf Seite 77.

6. Bestimmung der Verspätungsverteilungen



(a) $\rho = 0.452332$ an Stelle $t = 2.0$



(b) $\rho = 0.0523541$ an Stelle $t = 2.013$

Abbildung 6.10.: $\mathbf{E}[X] = 2.1$, $\mathbf{E}[X^2] = 4.42$, $\mathbf{E}[X^3] = 9.326$

6.7.1. Excess-Beyond

Es sei $Z = [X - s]^+$ sowie $\tilde{Z} = [\tilde{X} - s]^+$. Wir betrachten zunächst die Fehlerfortpflanzung hinsichtlich des Erwartungswertes. Wegen der Approximation durch Momentenanpassung gilt immer $|\mathbf{E}[X] - \mathbf{E}[\tilde{X}]| = 0$. Es sei für alle $t \in \mathbb{R}_{\geq 0}$

$$\delta_{X, \tilde{X}}(t) := \int_0^t (F_X(u) - F_{\tilde{X}}(u)) \, du .$$

Dann gilt

$$|\mathbf{E}[Z] - \mathbf{E}[\tilde{Z}]| = |\delta_{X, \tilde{X}}(s)| .$$

Anschaulich hängt die Fehlerfortpflanzung davon ab, wie in dem Teil der Verteilungsfunktionen, der durch die Excess-Beyond Operation gewissermaßen abgeschnitten wird, die Differenz der beiden Verteilungsfunktionen aussieht. Es existiert auch eine Schranke für mögliche Fehlerzuwächse. Denn es gilt offensichtlich $|\mathbf{E}[X] - \mathbf{E}[\tilde{X}]| = 0 \leq \kappa(F_X, F_{\tilde{X}})$ sowie (s. [MM07])

$$\kappa(F_Z, F_{\tilde{Z}}) = \int_0^\infty |F_X(t+s) - F_{\tilde{X}}(t+s)| \, dt \leq \kappa(F_X, F_{\tilde{X}}) .$$

Bezüglich der Metrik κ kann also der Fehler nach einer Anwendung der Excess-Beyond Operation nie größer werden. Gleiches gilt, wenn wir die Fehlerfortpflanzung hinsichtlich der Gestalt der Verteilungsfunktionen betrachten. Es gilt offensichtlich (s. [MM07])

$$\rho(F_Z, F_{\tilde{Z}}) = \sup\{|F_X(t+s) - F_{\tilde{X}}(t+s)| \mid t \in \mathbb{R}_{\geq 0}\} \leq \rho(F_X, F_{\tilde{X}}) .$$

6.7.2. Faltung

Es sei $Z = X + Y$ sowie $\tilde{Z} = \tilde{X} + \tilde{Y}$. Es gilt $F_Z = F_X * F_Y$ sowie $F_{\tilde{Z}} = F_{\tilde{X}} * F_{\tilde{Y}}$. Statt die Faltungsoperation zu verwenden, machen wir jedoch von der Linearität des Erwartungswertes Gebrauch, indem wir die jeweils ersten drei Momente direkt miteinander verrechnen (s. (6.7)). Da \tilde{X} und \tilde{Y} nach Annahme jeweils mittels Momentenanpassung aus X bzw. Y hervorgegangen sind, verursacht dieses Vorgehen keinen Fehler bezüglich der Erwartungswerte, d.h.

$$|\mathbf{E}[Z] - \mathbf{E}[\tilde{Z}]| = 0 .$$

Über den Fehler bezüglich der Gestalt der Verteilungsfunktion, d.h. $\rho(F_Z, F_{\tilde{Z}})$, lässt sich nur schwer eine allgemeine Aussage treffen. Dieser kann sowohl zunehmen als auch abnehmen. Er hängt von der Güte der Momentenanpassung hinsichtlich dieses Kriteriums ab. Er wird in Folge aufeinander folgender Faltungsoperationen nicht

beliebig anwachsen sondern ist erfahrungsgemäß in der Regel durch $\rho \approx 0.05$ nach oben beschränkt (s. Abschnitt 6.6.4).

Werden die Momente nicht direkt miteinander verrechnet, sondern wird stattdessen auf die tatsächliche Faltungsoperation zurückgegriffen, so können die Fehler, vergleichbar zum im folgenden Abschnitt betrachteten Produkt zweier Verteilungsfunktionen, maximal linear anwachsen.

6.7.3. Produkt

Es sei $Z = \max\{X, Y\}$ sowie $\tilde{Z} = \max\{\tilde{X}, \tilde{Y}\}$. Hier gilt $F_Z = F_X \cdot F_Y$ sowie $F_{\tilde{Z}} = F_{\tilde{X}} \cdot F_{\tilde{Y}}$. Das Produkt wird, im Gegensatz zur Faltungsoperation im vorherigen Abschnitt, immer mittels der entsprechenden (exakten) Operation auf Ebene der Theta-Exponentialpolynome berechnet. Es werden also keine Momente miteinander verrechnet. Für den Fehlerzuwachs existieren die folgenden Schranken (s. [MM07]). Es gilt $|\mathbf{E}[Z] - \mathbf{E}[\tilde{Z}]| \leq \kappa(F_Z, F_{\tilde{Z}})$ und

$$\begin{aligned} \kappa(F_Z, F_{\tilde{Z}}) &\leq \kappa(F_X, F_{\tilde{X}}) + \kappa(F_Y, F_{\tilde{Y}}) , \\ \rho(F_Z, F_{\tilde{Z}}) &\leq \rho(F_X, F_{\tilde{X}}) + \rho(F_Y, F_{\tilde{Y}}) . \end{aligned}$$

6.8. Topologische Sortierung

MEESTER, MUNS weisen in [MM07] darauf hin, dass man für die Berechnung der Verspätungen bzw. der Verspätungsverteilungen zunächst eine zulässige Berechnungsreihenfolge benötigt. Ob es eine solche gibt, hängt im Rahmen unserer Anwendung, d.h. ohne Berücksichtigung von Zugfolgerestriktionen, von den *Umsteigebeziehungen* zwischen den Linien ab. Wollen wir die Verspätungsverteilung eines Abfahrtsereignisses einer Linie in einer Umsteigestation bestimmen, so müssen wir zuvor bereits die Verspätungsverteilungen der Ankunftsereignisse aller Zubringer bestimmt haben. So muss in allen Umsteigestationen verfahren werden. Wir erhalten eine zulässige Sortierung aller Ereignisse genau dann, wenn kein Kreis bezüglich der Umsteigebeziehungen bzw. kein Kreis im Anschlussgraphen existiert.

Wir betrachten das Streckennetz aus Abbildung 6.11(a). In der graphischen Darstellung des Streckennetzes ist ein Kreis erkennbar, der die Stationen S_2, S_3, S_5 und S_6 enthält. Dies ist jedoch nicht gleichbedeutend mit der Existenz eines Kreises im Anschlussgraphen. Entscheidend ist hier, ob in den Stationen S_2 und S_5 Umsteigebeziehungen zwischen den Linien L_1 und L_2 existieren. Hierfür spielen die kürzesten Reisewege für Passagiere aus Kapitel 9 eine Rolle. Denn nur entlang dieser kürzesten Reisewege existieren Umsteigebeziehungen.

Wir nehmen nun an, dass zuvor erwähnte Umsteigebeziehungen existieren. Dann zeigt 6.11(b) den zum Streckennetz aus Abbildung 6.11(a) gehörenden Kreis im Anschlussgraphen. Der Anschlussgraph komprimiert Informationen des Ereignis-Aktivitätsnetzwerkes. Er liefert für die Suche einer zulässigen Berechnungsreihenfolge,

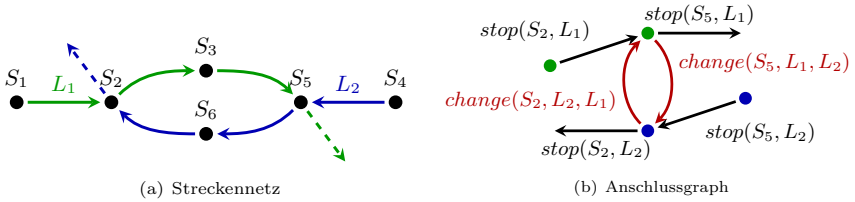


Abbildung 6.11.: Bei wechselseitigen Zubringerbeziehungen der Linien $L_1, L_2 \in \mathcal{L}$ ist keine topologische Sortierung möglich.

d.h. einer zulässigen Reihenfolge aller Ereignisse, die passende Datenstruktur. Es existiert eine zulässige Berechnungsreihenfolge genau dann, wenn der Anschlussgraph keinen Kreis enthält. Aus der Graphentheorie ist bekannt, dass die Kreisfreiheit äquivalent zur Existenz einer topologischen Sortierung der Knoten des Graphen ist. Die Suche einer topologischen Sortierung kann mittels eines Tiefensuchealgorithmus durchgeführt werden.

Wie die Sortierung für die Bestimmung der Verspätungsverteilungen verwendet wird, und insbesondere wie wir mit Kreisen im Anschlussgraphen umgehen, wird in Kapitel 7 behandelt.

6.9. Stochastische Abhängigkeiten

Bei den in Abschnitt 6.3 diskutierten Verspätungsoperationen wurde häufig die stochastische Unabhängigkeit der miteinander zu verrechnenden Zufallsvariablen vorausgesetzt. In der Literatur zu ganzzahliger linearer Programmierung findet man eine analoge Forderung, die *never-meet* Annahme (s. [Sch07]). Sie wird dort für die Korrektheit der Linearisierung im Rahmen des Verspätungsmanagement Problems vorausgesetzt. BÜKER hat sich in [Bü10] ebenfalls intensiver mit der Verletzung stochastischer Unabhängigkeit beschäftigt. Er hält dazu abschließend fest, dass man im Allgemeinen, durch die Verrechnung der Verspätungen mit Zeitpuffern, keine signifikant fehlerhaften Ergebnisse erhält.

Wir erläutern im Folgenden, wann stochastische Unabhängigkeit in der Praxis

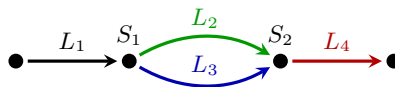


Abbildung 6.12.: Ungerichteter Kreis

nicht vorausgesetzt werden kann. Für diesen Fall zeigen wir, wann und wie sich die Verletzung der stochastischen Unabhängigkeit während der Berechnung der Verspätungsverteilungen vermeiden lässt. Wir betrachten nun den ungerichteten Kreis im Streckennetz in Abbildung 6.12. Dieser führt im Falle existierender Umsteigebeziehungen in den Stationen S_1 und S_2 auch zu einem ungerichteten Kreis im Anschlussgraphen. Existieren nur gerichtete Kreise, stellt dies bezüglich der geforderten stochastischen Unabhängigkeit kein Problem dar. Dieses Problem tritt nur dann auf, wenn zwei in ein Maximum eingehende Zufallsgrößen hinsichtlich der Verspätungsförpflanzung einen gemeinsamen Ursprung besitzen.

Anschaulich ist die Abhängigkeit der Ankunftsverspätungen der Linien L_2 und L_3 in Station S_2 schnell offensichtlich: Ist die Ankunft Linie L_1 in Station S_1 hinreichend stark verspätet, so verspätet sich dadurch sowohl die Abfahrt von Linie L_2 als auch die von Linie L_3 . Enthält der Fahrplan nun keine entsprechend großen Puffer für die jeweiligen Fahrten von S_1 nach S_2 , so wird sich die von Linie L_1 auf die Linien L_2 und L_3 übertragene Verspätungen bis zur Station S_2 fortpflanzen.

Im Folgenden betrachten wir die zugehörige formale Darstellung. Für die Abfahrtsereignisse der Linien L_2, L_3 in S_1 gilt

$$Y(dep, S_1, L_2) = [Y(arr, S_1, L_1) - s(change, S_1, L_1, L_2)]^+ , \quad (6.21)$$

$$Y(dep, S_1, L_3) = [Y(arr, S_1, L_1) - s(change, S_1, L_1, L_3)]^+ . \quad (6.22)$$

Es folgen die Fahrten von S_1 nach S_2 .

$$Y(arr, S_2, L_2) = [Y(dep, S_1, L_2) + D(drive, S_1, S_2, L_2) - s(drive, S_1, S_2, L_2)]^+ \quad (6.23)$$

$$Y(arr, S_2, L_3) = [Y(dep, S_1, L_3) + D(drive, S_1, S_2, L_3) - s(drive, S_1, S_2, L_3)]^+ \quad (6.24)$$

Für die Abfahrt der Linie L_4 in S_2 gilt

$$Y(dep, S_2, L_4) = \max\{Y(arr, S_2, L_2) - s(change, S_2, L_2, L_4), \\ Y(arr, S_2, L_3) - s(change, S_2, L_3, L_4), 0\} \quad (6.25)$$

Zusammengesetzt erhält man

$$Y(dep, S_2, L_4) = \max \{ [Y(dep, S_1, L_2) + D(drive, S_1, S_2, L_2) \\ - s(drive, S_1, S_2, L_2)]^+ - s(change, S_2, L_2, L_4), \\ [Y(dep, S_1, L_3) + D(drive, S_1, S_2, L_3) \\ - s(drive, S_1, S_2, L_3)]^+ - s(change, S_2, L_3, L_4), 0 \} \\ = \max \{ \max\{Y(dep, S_1, L_2) + D(drive, S_1, S_2, L_2) \\ - s(drive, S_1, S_2, L_2), 0\} - s(change, S_2, L_2, L_4),$$

$$\begin{aligned}
 & \max\{Y(dep, S_1, L_3) + D(drive, S_1, S_2, L_3) \\
 & \quad - s(drive, S_1, S_2, L_3), 0\} - s(change, S_2, L_3, L_4), 0\} \\
 = & \max \left\{ \max\{[Y(arr, S_1, L_1) - s(change, S_1, L_1, L_2)]^+ \right. \quad (6.26) \\
 & \quad + D(drive, S_1, S_2, L_2) \\
 & \quad - s(drive, S_1, S_2, L_2), 0\} - s(change, S_2, L_2, L_4), \\
 & \quad \max\{[Y(arr, S_1, L_1) - s(change, S_1, L_1, L_3)]^+ \\
 & \quad + D(drive, S_1, S_2, L_3) \\
 & \quad - s(drive, S_1, S_2, L_3), 0\} - s(change, S_2, L_3, L_4), 0\} \\
 = & \max \left\{ \max \left\{ \max\{Y(arr, S_1, L_1) - s(change, S_1, L_1, L_2), 0\} \right. \right. \\
 & \quad + D(drive, S_1, S_2, L_2) \\
 & \quad - s(drive, S_1, S_2, L_2), 0\} - s(change, S_2, L_2, L_4), \\
 & \quad \max \left\{ \max\{Y(arr, S_1, L_1) - s(change, S_1, L_1, L_3), 0\} \right. \\
 & \quad + D(drive, S_1, S_2, L_3) \\
 & \quad \left. \left. - s(drive, S_1, S_2, L_3), 0\} - s(change, S_2, L_3, L_4), 0 \right\} \right\}.
 \end{aligned}$$

Bei der Berechnung Verteilungsfunktion von $Y(dep, S_2, L_4)$ kann man schrittweise, in der Reihenfolge der Gleichungen (6.21)-(6.25), vorgehen. In diesem Fall wird man jedoch einen Fehler machen, wenn man das Produkt der Verteilungsfunktionen der stochastisch abhängigen Zufallsvariablen

$$\begin{aligned}
 & \max \left\{ \max \left\{ Y(arr, S_1, L_1) - s(change, S_1, L_1, L_2), 0 \right\} + D(drive, S_1, S_2, L_2) \right. \\
 & \quad \left. - s(drive, S_1, S_2, L_2), 0 \right\} - s(change, S_2, L_2, L_4) \quad (6.27)
 \end{aligned}$$

und

$$\begin{aligned}
 & \max \left\{ \max \left\{ Y(arr, S_1, L_1) - s(change, S_1, L_1, L_3), 0 \right\} + D(drive, S_1, S_2, L_3) \right. \\
 & \quad \left. - s(drive, S_1, S_2, L_3), 0 \right\} - s(change, S_2, L_3, L_4) \quad (6.28)
 \end{aligned}$$

aus (6.26) bildet. In beiden Ausdrücken ist die Verspätung $Y(arr, S_1, L_1)$ enthalten. Die Auswirkungen machen wir uns an einem noch weiter vereinfachten Beispiel deutlich. Wir gehen davon aus, dass es im Beispiel aus Abbildung 6.12 weder Puffer noch positive Quellverspätungen gibt, d.h. $Y(arr, S_2, L_2) = Y(arr, S_2, L_3) = Y(arr, S_1, L_1)$ und

$$Y(dep, S_2, L_4) = \max\{Y(arr, S_2, L_2), Y(arr, S_2, L_3)\} \quad (6.29)$$

(vgl. Gleichung (6.25)).

6. Bestimmung der Verspätungsverteilungen

Wir berechnen die zugehörige Verteilungsfunktion wie oben beschrieben. Es wäre korrekt, wenn wir aus (6.29) zunächst

$$Y(dep, S_2, L_4) = \max\{Y(arr, S_1, L_1), Y(arr, S_1, L_1)\} = Y(arr, S_1, L_1)$$

und somit

$$F_{Y(dep, S_2, L_4)} \equiv F_{Y(arr, S_1, L_1)} \quad (6.30)$$

folgerten. Stattdessen betrachten wir zwei i.i.d.-verteilte Zufallsvariablen

$$Y(arr, S_1, L_1)_1 \quad \text{und} \quad Y(arr, S_1, L_1)_2$$

mit Verteilungsfunktion $F_{Y(arr, S_1, L_1)}$ und

$$Y(arr, S_2, L_2) = Y(arr, S_1, L_1)_1, \quad Y(arr, S_2, L_3) = Y(arr, S_1, L_1)_2.$$

Es gilt dann für alle $t \in \mathbb{R}$

$$F_{Y(dep, S_2, L_4)}(t) = F_{Y(arr, S_2, L_2)}(t) \cdot F_{Y(arr, S_2, L_3)}(t) = F_{Y(arr, S_1, L_1)}(t)^2. \quad (6.31)$$

Offensichtlich folgt (6.30) \geq (6.31). Der absolute Fehler in (6.31) ist für alle $t \in \mathbb{R}$ daher $F_{Y(arr, S_1, L_1)}(t) - F_{Y_2(arr, S_1, L_1)}(t)^2$.

Wir gehen im Folgenden wieder zurück zu den Verhältnissen in Gleichung (6.26). Wollen wir verhindern, bei der Berechnung der Verteilungsfunktion einen Fehler zu machen, müssen wir eine Situation wie in (6.27) und (6.28) vermeiden. $Y(arr, S_1, L_1)$ darf nicht gleichzeitig in Zufallsausdrücken enthalten sein, von deren Maximum die Verteilungsfunktion berechnet werden soll (vgl. vereinfachtes Beispiel oben). Wir formen (6.27) und (6.28) um in

$$\begin{aligned} & \max\{Y(arr, S_1, L_1) - s(change, S_1, L_1, L_2) + D(drive, S_1, S_2, L_2) \\ & \quad - s(drive, S_1, S_2, L_2) - s(change, S_2, L_2, L_4), \\ & D(drive, S_1, S_2, L_2) - s(drive, S_1, S_2, L_2) - s(change, S_2, L_2, L_4), \\ & \quad - s(change, S_2, L_2, L_4)\} \end{aligned}$$

bzw.

$$\begin{aligned} & \max\{Y(arr, S_1, L_1) - s(change, S_1, L_1, L_3) + D(drive, S_1, S_2, L_3) \\ & \quad - s(drive, S_1, S_2, L_3) - s(change, S_2, L_3, L_4), \\ & D(drive, S_1, S_2, L_3) - s(drive, S_1, S_2, L_3), \\ & \quad - s(change, S_2, L_3, L_4)\}. \end{aligned}$$

Für $Y(dep, S_2, L_4)$ gilt schließlich

$$Y(dep, S_2, L_4) =$$

$$\begin{aligned} \max \left\{ Y(arr, S_1, L_1) + \max \left\{ D(drive, S_1, S_2, L_2) \right. \right. \\ \left. \left. - s(drive, S_1, S_2, L_2) - s(change, S_2, L_2, L_4), \right. \right. \\ \left. \left. D(drive, S_1, S_2, L_3) - s(drive, S_1, S_2, L_3) - s(change, S_2, L_3, L_4) \right\}, \right. \\ \left. D(drive, S_1, S_2, L_2) - s(drive, S_1, S_2, L_2) - s(change, S_2, L_2, L_4), \right. \\ \left. D(drive, S_1, S_2, L_3) - s(drive, S_1, S_2, L_3) - s(change, S_2, L_3, L_4), 0 \right\}. \end{aligned}$$

Die Verteilungsfunktion solcher Zufallsausdrücke kann mit Hilfe der Theta-Exponentialpolynome (Produkt von Verteilungsfunktionen) berechnet werden. Die Berechnung wird jedoch aufwendiger. Außerdem müsste der Anschlussgraph hinsichtlich vorkommender ungerichteter Kreise untersucht werden. Sobald wir durch die Hinzunahme von Anschlusssicherungsstrategien (s. Abschnitt 6.4) in unsere Modellierung zusätzlich auch das Minimum vom Zufallsvariablen berechnen müssen, ist das zuvor beschriebene Vorgehen nicht mehr möglich.

7. Verspätungsfortpflanzung in kreisartigen Strukturen

Übersicht

7.1. Modellierung	99
7.1.1. Kreisartige Strukturen	99
7.1.2. Verspätungsfortpflanzung	104
7.2. Verteilungskonvergenz	114
7.2.1. Bedingungen für Verteilungskonvergenz	115
7.2.2. Einfluss unterschiedlicher Sortierungen	120
7.3. Endlichkeit der Verspätungen	123
7.3.1. Übertragbare Ergebnisse aus der Warteschlangentheorie	123
7.3.2. Stabilität in einem Kreis der Kreisvereinigung	126
7.3.3. Stabilität in der Kreisvereinigung	135
7.4. Pseudo-topologische Sortierung	138
7.5. Empirische Resultate - Konvergenzgeschwindigkeit .	138

In diesem Kapitel wird die langfristige Entwicklung der Verspätungsfortpflanzung in kreisartigen Strukturen des Anschlussgraphen untersucht. Grundlage für die Beobachtung der Verspätungsfortpflanzung bzw. -entwicklung ist die Idee einer iterativen Bestimmung der Verteilungsfunktionen von Abfahrts- und Ankunftsereignissen in solchen Strukturen. Es wird der Frage nachgegangen, unter welchen Bedingungen die Verteilungsfunktionen konvergieren. Zunächst wird hierfür nur ein Teil der Struktur betrachtet. Im Anschluss werden die so gewonnenen Ergebnisse verallgemeinert und auf die Situation in der kompletten Struktur übertragen. Nachdem der Einfluss unterschiedlicher Sortierungen empirisch untersucht wurde, wird ein Verfahren vorgestellt, das eine sinnvolle Sortierung der Ereignisse für das iterative Vorgehen liefert.

7.1. Modellierung

7.1.1. Kreisartige Strukturen

Zu Beginn legen wir die grundlegende Notation fest und definieren Kreise im Anschlussgraphen.

Definition 7.1.1 (Kreis).

Im Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ bildet der durch $K = (V_1, V_2, \dots, V_n)$ symbolisierte Teilgraph einen (gerichteten) **Kreis**, falls

- (i) $V_i \in \mathfrak{V}$ für $1 \leq i \leq n$,
- (ii) $V_i \neq V_j$ für $1 \leq i < j \leq n$ und
- (iii) $(V_i, V_{i+1}) \in \mathfrak{E}$ für $1 \leq i \leq n$ mit $V_{n+1} := V_1$.

Mit $\mathcal{K}_{\mathfrak{G}}$ wird die Menge aller Kreise des Anschlussgraphen \mathfrak{G} bezeichnet. Des Weiteren ist \mathcal{V}_K die Menge der Knoten des Kreises $K \in \mathcal{K}_{\mathfrak{G}}$.

Die Relation $R \subset \mathcal{K}_{\mathfrak{G}} \times \mathcal{K}_{\mathfrak{G}}$ ist definiert durch $R = \{(K_i, K_j) \in \mathcal{K}_{\mathfrak{G}} \times \mathcal{K}_{\mathfrak{G}} \mid \mathcal{V}_{K_i} \cap \mathcal{V}_{K_j} \neq \emptyset\}$. R^+ sei die zugehörige transitive Hülle der Relation.

Definition 7.1.2 (Kreisvereinigung).

Eine **Kreisvereinigung** ist eine Mengen $\mathcal{M} \subset \mathcal{K}_{\mathfrak{G}}$ von Kreisen des Anschlussgraphen \mathfrak{G} für die gilt

- (i) $\forall K_1, K_2 \in \mathcal{M} : K_1 R^+ K_2$,
- (ii) $\forall K_1 \notin \mathcal{M} \forall K_2 \in \mathcal{M} : \neg(K_1 R^+ K_2)$.

$\mathcal{V}_{\mathcal{M}}$ ist die Menge aller in den Kreisen der Kreisvereinigung \mathcal{M} enthaltenen Knoten. Die Menge aller Kreisvereinigungen des Anschlussgraphen wird mit $\mathcal{M}_{\mathfrak{G}}$ bezeichnet.

Gemäß dieser Definition enthält eine Kreisvereinigung nur Kreise. Sie repräsentiert einen maximal zusammenhängenden Teilgraphen des Anschlussgraphen. In der Graphentheorie nennt man diese Teilgraphen *starke Zusammenhangskomponenten*.

Definition 7.1.3 (Singuläre Knoten).

Ein Knoten $V \in \mathfrak{V}$ des Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ heißt **singulärer Knoten** genau dann, wenn

$$\forall \mathcal{M} \in \mathcal{M}_{\mathfrak{G}} : \neg(\exists K \in \mathcal{M} : V \in \mathcal{V}_K) .$$

Jeder singuläre Knoten ist also in keiner der Kreisvereinigungen des Anschlussgraphen enthalten. Die Menge der singulären Knoten des Anschlussgraphen \mathfrak{G} wird mit $\mathcal{I}_{\mathfrak{G}}$ bezeichnet.

Im Folgenden bezeichnet $\mathcal{O}_{\mathfrak{G}} := \mathcal{M}_{\mathfrak{G}} \cup \mathcal{I}_{\mathfrak{G}}$ die Menge aller Kreisvereinigungen und singulären Knoten. Diese Menge bildet später die Knotenmenge eines neuen Graphentyps. Zuvor definieren wir noch die Nachfolger bzw. Vorgänger von Kreisvereinigungen im Anschlussgraphen. Diese sind selbst entweder eine Kreisvereinigung oder ein singulärer Knoten.

Definition 7.1.4 (Kreisvereinigungsgraph).

Es sei $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ ein Anschlussgraph. Der **Kreisvereinigungsgraph** $\mathfrak{M} = (\mathcal{O}_{\mathfrak{G}}, \mathcal{E}_{\mathfrak{G}})$ besitzt als Knotenmenge $\mathcal{O}_{\mathfrak{G}} = \mathcal{M}_{\mathfrak{G}} \cup \mathcal{I}_{\mathfrak{G}}$. Die Kantenmenge $\mathcal{E}_{\mathfrak{G}}$ ist anschaulich eine Teilmenge der Kantenmenge des Anschlussgraphen. Für $O_i, O_j \in \mathcal{O}_{\mathfrak{G}}$ gilt $e := (O_i, O_j) \in \mathcal{E}_{\mathfrak{G}}$, falls O_i Vorgänger von O_j bzw. O_j Nachfolger von O_i im Sinne der Definition 7.1.5, Definition 7.1.6 bzw. Definition 3.4.2 ist.

Definition 7.1.5 (Nachfolger von Kreisvereinigungen).

Nachfolger von Kreisvereinigungen $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ im Kreisvereinigungsgraphen \mathfrak{M} sind Objekte aus $\mathcal{O}_{\mathfrak{G}}$. Ein singulärer Knoten V^* ist Nachfolger einer Kreisvereinigung \mathcal{M} , falls er Nachfolger eines Knoten V aus \mathcal{M} ist. Eine Kreisvereinigung \mathcal{M}^* ist Nachfolger einer anderen Kreisvereinigung \mathcal{M} , falls sie einen Knoten V^* enthält, der Nachfolger eines Knoten V aus \mathcal{M} ist.

(i) $V^* \in \mathcal{I}_{\mathfrak{G}}$ **Nachfolger** von \mathcal{M} , $V^* \in \text{Suc}_{\mathfrak{M}}(\mathcal{M})$, genau dann, wenn

$$\exists K \in \mathcal{M} \exists V \in \mathcal{V}_K : V^* \in \text{Suc}_{\mathfrak{G}}(V),$$

(ii) $\mathcal{M}^* \in \mathcal{M}_{\mathfrak{G}}$ **Nachfolger** von \mathcal{M} , $\mathcal{M}^* \in \text{Suc}_{\mathfrak{M}}(\mathcal{M})$, genau dann, wenn

$$\exists K \in \mathcal{M} \exists K^* \in \mathcal{M}^* \exists V \in K \exists V^* \in K^* : V^* \in \text{Suc}_{\mathfrak{G}}(V) .$$

Definition 7.1.6 (Vorgänger von Kreisvereinigungen).

Vorgänger von Kreisvereinigungen $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ im Kreisvereinigungsgraphen \mathfrak{M} sind Objekte aus $\mathcal{O}_{\mathfrak{G}}$. Ein singulärer Knoten V^* ist Vorgänger einer Kreisvereinigung \mathcal{M} , falls er Vorgänger eines Knoten V aus \mathcal{M} ist. Eine Kreisvereinigung \mathcal{M}^* ist Vorgänger einer anderen Kreisvereinigung \mathcal{M} , falls sie einen Knoten V^* enthält, der Vorgänger eines Knoten V aus \mathcal{M} ist.

(i) $V^* \in \mathcal{I}_{\mathfrak{G}}$ **Vorgänger** von \mathcal{M} , $V^* \in \text{Pre}_{\mathfrak{M}}(\mathcal{M})$, genau dann, wenn

$$\exists K \in \mathcal{M} \exists V \in \mathcal{V}_K : V^* \in \text{Pre}_{\mathfrak{G}}(V),$$

(ii) $\mathcal{M}^* \in \mathcal{M}_{\mathfrak{G}}$ **Vorgänger** von \mathcal{M} , $\mathcal{M}^* \in \text{Pre}_{\mathfrak{M}}(\mathcal{M})$, genau dann, wenn

$$\exists K \in \mathcal{M} \exists K^* \in \mathcal{M}^* \exists V \in K \exists V^* \in K^* : V^* \in \text{Pre}_{\mathfrak{G}}(V) .$$

Wie erinnern uns, dass für die korrekte Berechnung der Verteilungsfunktionen in den Knoten des Anschlussgraphen eigentlich eine topologische Sortierung *aller* Knoten vorliegen muss. Eine topologische Sortierung liegt genau dann vor, wenn jeder Knoten erst dann besucht wird, wenn zuvor bereits *alle* Vorgänger des Knoten besucht wurden. Bezüglich der zu bestimmenden Verteilungen heißt das anschaulich, dass alle Informationen über Verspätungen vorhanden sein müssen, welche die aktuell zu bestimmende Verspätungsverteilung beeinflussen. Existieren jedoch

Kreise, werden wir für Knoten innerhalb eines Kreises zunächst nur unvollständige Information über aus Vorgängerknoten übertragene und somit in die Berechnung einfließende Verspätungen haben. Wir zeigen nun, dass für die Knoten des Kreisvereinigungsgraphen stets eine topologische Sortierung existiert. Der Kreisvereinigungsgraph ist also immer kreisfrei. Ist die Sortierung erstellt, müssen wir nur noch eine sinnvolle Sortierung der Knoten innerhalb der Kreisvereinigungen finden.

Satz 7.1.7.

Es sei $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ ein gegebener Anschlussgraph und $\mathfrak{M} = (\mathcal{O}_{\mathfrak{G}}, \mathcal{E}_{\mathfrak{G}})$ der zugehörige Kreisvereinigungsgraph. Dann existiert für Elemente aus $\mathcal{O}_{\mathfrak{G}}$ immer eine topologische Sortierung.

Beweis. Die Existenz einer topologischen Sortierung für die im Anschlussgraphen enthaltenen Kreisvereinigungen und singulären Knoten ist äquivalent zur Kreisfreiheit im Kreisvereinigungsgraph. Die Aussage des Satzes ist trivial. Isolierte Knoten können gemäß ihrer Definition nie Teil eines Kreises aus \mathfrak{G} sein. Daher können sie auch in \mathfrak{M} nicht Element eines Kreises sein. Wären zwei Kreisvereinigungen Teil eines Kreises in \mathfrak{M} , so bildeten sie zusammen eine Kreisvereinigung (starke Zusammenhangskomponente), die mehr Kreise umfasst als beide einzeln. Dies widerspricht ihrer Definition. □

Grundlage für das weitere Vorgehen in diesem Kapitel ist die Idee, die Verspätungsverteilungen iterativ zu bestimmen. Für die iterative Bestimmung der Verspätungsverteilungen innerhalb einer Kreisvereinigung suchen wir eine Sortierung aller enthaltenen Knoten. Ein Schritt der Iteration steht dann für die einmalige Berechnung aller Knoten der Kreisvereinigung gemäß dieser Sortierung. Dabei gehen für jeden Knoten sowohl die von Vorgängerknoten übertragenen Verspätungen als auch die im aktuell betrachteten Knoten auftretenden Quellverspätungen in die Berechnung ein. Im ersten Iterationsschritt gibt es immer mindestens einen Knoten,

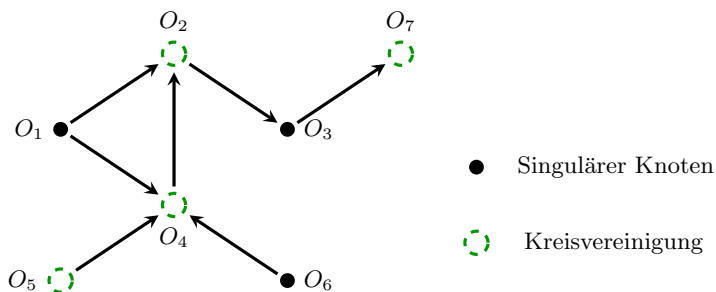


Abbildung 7.1.: Beispiel eines Kreisvereinigungsgraphen

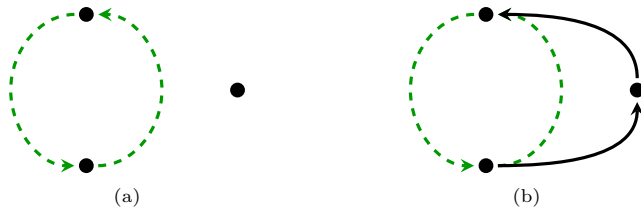


Abbildung 7.2.: Kreisvereinigung und singulärer Knoten

der einen Vorgänger besitzt, dessen Verspätungsverteilungen noch nicht bestimmt wurden. Die von diesen Vorgängern übertragenen Verspätungen werden daher im ersten Iterationsschritt als identisch 0 angesehen. In allen weiteren Iterationsschritten sind die Verspätungsverteilungen in allen Knoten mindestens einmal bestimmt worden. Es wird dann der jeweils aktuelle Stand der Verspätungsverteilungen der Vorgängerknoten verwendet. Die Iteration bricht ab, wenn die Verspätungsverteilungen in allen Knoten hinreichend gut approximiert wurden. Dies könnte in der praktischen Anwendung beispielsweise der Fall sein, wenn die Veränderung der Erwartungswerte der Verspätungsverteilungen bezüglich zweier aufeinander folgender Iterationsschritte unter eine vorgegebene Schranke fällt.

Input : Sortierter Vektor S aller Knoten der Kreisvereinigung

```

1 for  $i = 0; i < S.size(); i ++$  do
2    $S[i].SetzeVerspaetungenIdentischNull();$ 
3 while Genauigkeit der Approximation der Verspätungsverteilungen nicht
   ausreichend do
4   for  $i = 0; i < S.size(); i ++$  do
5      $S[i].BestimmeVerspaetungsverteilungen(Mit\ allen$ 
        $VorgaengerKnoten);$ 

```

Algorithmus 7.1.1 : Iterative Bestimmung der Verteilungsfunktionen in Kreisvereinigungen

Bei einer solchen Vorgehensweise sind insbesondere die folgenden beiden Fragen zu beantworten:

1. Welche Voraussetzungen müssen für die Konvergenz der Verspätungsverteilungen erfüllt sein?

2. Welchen Einfluss hat die der Iteration zu Grunde liegende Sortierung der Knoten der Kreisvereinigung auf die Ergebnisse?

Zunächst betrachten wir die Situation in Abbildung 7.4. Hier ist eine vereinfachte Illustration für die allgemeine Situation in einem Kreis gegeben, der Element einer Kreisvereinigung ist. Für diesen Fall wird erläutert, wie sich Verspätungen fortpflanzen (s. Abschnitt 7.1.2). Hierauf aufbauend werden anschließend in Abschnitt 7.2 Bedingungen für die Konvergenz der Verteilungsfunktionen hergeleitet und der Einfluss unterschiedlicher Sortierungen diskutiert. In Abschnitt 7.3 zeigen wir, unter welchen Voraussetzungen die Grenzfunktionen Verteilungsfunktionen endlicher Zufallsvariablen sind. Auch die Eindeutigkeit der Grenzverteilungsfunktionen wird hier untersucht. Der Einfluss unterschiedlicher Sortierung Einfluss auf die Konvergenzgeschwindigkeit wird in Abschnitt 7.5 experimentell nachgewiesen. Zuletzt wird in Abschnitt 7.4 eine sinnvolle Sortierung vorgestellt, die im Allgemeinen zu guten Ergebnissen führt.

7.1.2. Verspätungsfortpflanzung

Im Rahmen unserer Untersuchungen stützen wir uns unter anderem auf den Artikel [Loy62] von LOYNES. Dort wird die Verteilung der Wartezeit W_n des n -ten Kunden in einem Bediensystem mit einem Bediener untersucht. In Abbildung 7.3(a) wird die Situation illustriert. Das betrachtete System enthält nur einen Bediener. Ankommende Kunden werden erst bedient, wenn alle zuvor angekommenen Kunden bereits bedient wurden. Die Wartezeit des $n+1$ -ten Kunden im System hängt daher von folgenden Zufallsgrößen ab:

- Die Wartezeit des n -ten Kunden (des Vorgängers): W_n
- Die Bedienzeit des n -ten Kunden: S_n
- Die Zeit, die zwischen seiner Ankunft und der Ankunft des n -ten Kunden vergeht: T_n

Es wird angenommen, dass für alle $n \in \mathbb{N}$ sowohl $S_n < \infty$ (\mathbf{P} -f.s.) als auch $T_n < \infty$ (\mathbf{P} -f.s.) gilt. Kommt ein Kunde zu einem Zeitpunkt ins Bediensystem, zu dem der zuvor angekommene Kunde noch warten muss, so muss er die Wartezeit seines Vorgängers, abzüglich der Zeit, die er noch nicht im System war, selbst abwarten. Zusätzlich muss er bis zu seiner eigenen Bedienung noch die Bedienzeit seines Vorgängers abwarten. In [Loy62] wird angenommen, dass der erste Kunde im System sofort bedient wird, d.h. $W_1 \equiv 0$. Mit $U_n := S_n - T_n$ gilt für alle $n \in \mathbb{N}$ somit $W_{n+1} = f(W_n, U_n) := [W_n + U_n]^+ := \max\{W_n + U_n, 0\}$.

In Abbildung 7.3(b) wird dargestellt, wie die Situation im beschriebenen Bediensystem auf die Situation von Verspätungen im Bahnverkehr übertragen werden kann. Statt Wartezeiten betrachten wir dort Abfahrtsverspätungen Y_n in Station n .

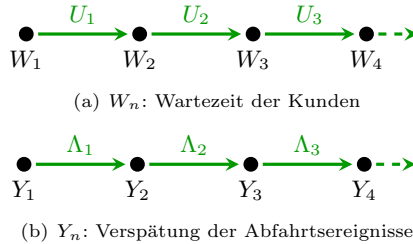


Abbildung 7.3.: Analogie zwischen Wartezeit im Bediensystem und Verspätung im Bahnverkehr

Die Bedienzeiten werden ersetzt durch Quellverspätungen D_n . Statt Zwischenankunftszeiten benutzen wir deterministische Puffer s_n und definieren $\Lambda_n := D_n - s_n$ für alle $n \in \mathbb{N}$. Wir starten in Station S_1 mit Verspätung $Y_1 \equiv 0$ und betrachten nur Quellverspätungen und zeitliche Puffer auf Strecken. Dann ergeben sich die Abfahrtsverspätungen in den folgenden Stationen gerade gemäß $Y_{n+1} = f(Y_n, \Lambda_n) = [Y_n + \Lambda_n]^+$. Wegen dieser Analogie zwischen Wartezeiten und Verspätungen, können wir einen Teil der Resultate aus [Loy62] auf das Problem der Verspätungsfortpflanzung anwenden.

Der Grund, dass wir manche Ergebnisse aus [Loy62] nicht direkt anwenden können, liegt in der Annahme von LOYNES, dass die Folge $(U_n)_{n \in \mathbb{N}}$ stationär ist, d.h. insbesondere identisch verteilte Folgenglieder hat. Diese Bedingung ist im Bahnverkehr sicherlich unrealistisch, da hier analog die Quellverspätungen $(D_n)_{n \in \mathbb{N}}$ auf den Strecken stationär sein müssten. Diese werden sich jedoch im Allgemeinen hinsichtlich ihrer Verteilung unterscheiden. Die für unsere Untersuchungen verwendete (realistischere) Modellierung unterscheidet sich jedoch auch von der bisher erläuterten. Denn statt Züge entlang einer (unendlich langen) Folge verschiedener Strecken fahren zu lassen, betrachten wir Züge, die sich innerhalb von Kreisvereinigungen bewegen. In Abbildung 7.4(a) wird ein einzelner Kreis der Kreisvereinigung betrachtet. Die Situation auf Ebene des Streckennetzes wird in Abbildung 7.4(b) dargestellt. Diese stark vereinfachte Modellierung der allgemeinen Situation ist für unsere Fragestellungen immer noch ausreichend. Die Vereinfachung besteht unter anderem darin, dass jeder Knoten des Kreises genau ein Abfahrts- und ein Ankunftsereignis enthält. Der Fahrt zwischen den beiden zugehörigen Stationen wird eine Quellverspätungsverteilung sowie ein deterministischer zeitlicher Puffer zugewiesen. Die Knoten sind entweder durch eine Halte- oder eine Umsteigeaktivität miteinander verbunden. Es werden zur Vereinfachung der Modellierung keine zu Halte-Aktivitäten gehörenden Quellverspätungen und Puffer berücksichtigt, denn diese sind bezüglich der Verspätungsfortpflanzung gleichwertig zu zusätzli-

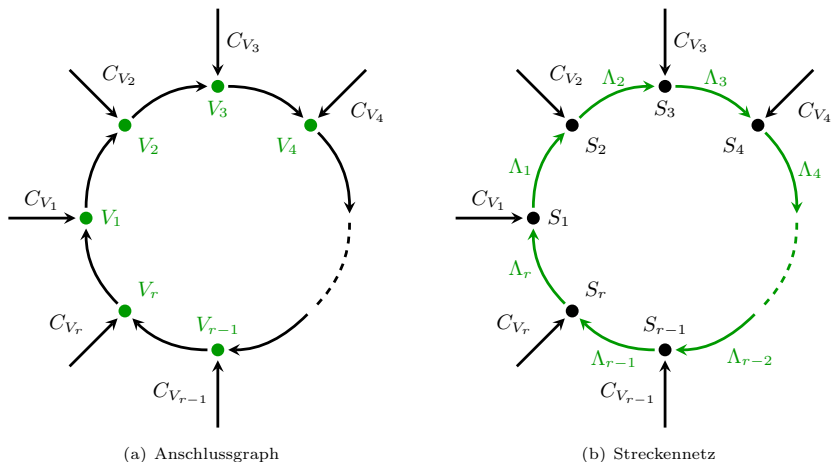


Abbildung 7.4.: Vereinfachte allgemeine Situation für Verspätungsfortpflanzung in einem Kreis

chen Quellverspätungen und Puffern der jeweils vorangehenden Fahraktivität. Mit analoger Begründung werden auch keine Puffer von Umsteigekanten betrachtet. Für die Beantwortung der Frage, ob sich die Verspätungen stabilisieren, erweist es sich als gleichwertig, diesen Puffer als zusätzlichen Puffer für die vorangehenden Fahraktivität zu behandeln.

Die Verspätungen bzw. deren Verteilungsfunktionen werden von uns iterativ, mittels mehrmaligen Durchlaufens der Knoten des Kreises, bestimmt. In den folgenden Abschnitten interessieren wir uns im Rahmen der Untersuchung der Verspätungsfortpflanzung insbesondere für die Grenzverteilungsfunktion des jeweiligen Abfahrtsereignisses eines jeden Knotens. Die Verspätungsfortpflanzung innerhalb der Knoten wird in jedem Iterationsschritt von, bezüglich der Iterationsschritte, identischen (Zeitpuffer) bzw. identisch verteilten (Quellverspätungen) Größen bestimmt. Wir bestimmen keine gemeinsame Verteilung sondern Randverteilungen.

Definition 7.1.8.

Wir betrachten die Knoten eines Kreises $K \in \mathcal{M}$ der Kreisvereinigung \mathcal{M} . Für $1 \leq i \leq r$, $r = |\mathcal{V}_K|$, sei für alle $n \in \mathbb{N}$ im n -ten Iterationsschritt

- (i) $\Lambda_{V_i}^{(n)} = D_{V_i}^{(n)} - s_i$ die Differenz aus Quellverspätung und Puffer der Fahraktivität in Knoten V_i bzw. auf Strecke (S_i, S_{i+1}) ,
- (ii) $C_{V_i}^{(n)}$ die über Umsteigeaktivitäten auf das Abfahrtsereignis des Knotens V_i

bzw. in Station S_i übertragene Verspätung von Zubringern, welche die Station von außerhalb des Kreises erreichen, sowie

- (iii) $\mathbf{Y}_{V_i}^{(n)}$ die Verspätung des ersten Abfahrtsereignisses des Knotens V_i in Station S_i .

Gemäß dieser Definition ist für festes $i \in \{1, 2, \dots, r\}$

- $(D_{V_i}^{(n)})_{n \in \mathbb{N}}$ eine Folge identisch verteilter, stochastisch unabhängiger sowie nicht-negativer Zufallsvariablen,
- $(\Lambda_{V_i}^{(n)})_{n \in \mathbb{N}}$ eine Folge identisch verteilter und stochastisch unabhängiger Zufallsvariablen und
- $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ ein Folge nicht-negativer Zufallsvariablen.

Eine genauere Analyse, welche Verspätungen die Zufallsvariablen $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ repräsentieren, erfolgt weiter unten in Bemerkung 7.1.10. Wir setzen nun voraus, dass für die Knoten des Kreisvereinigungsgraphen (s. Abbildung 7.1) eine topologische Sortierung erstellt wurde. Nach Satz 7.1.7 ist dies immer möglich. Verteilungsfunktionen von singulären Knoten oder Kreisvereinigungen, die in der topologischen Sortierung vor der aktuell betrachteten Kreisvereinigung stehen, wurden schon bestimmt. Dies entspricht dem ersten Schritt in Abbildung 7.5. In jedem Iterationsschritt $n \in \mathbb{N}$ wird die Entwicklung der Verspätungen innerhalb des betrachteten Kreises der Kreisvereinigung von, bezüglich der Iterationsschritte, identisch verteilten Verspätungen $(\Lambda_{V_i}^{(n)})_{n \in \mathbb{N}}$ und Verspätungen $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ von Knoten innerhalb bzw. außerhalb des Kreises beeinflusst. In Abbildung 7.1 werden beispielsweise die Verspätungsverteilungen der Kreisvereinigung O_2 von den Verteilungen der vom singulären Knoten O_1 und von der Kreisvereinigung O_4 übertragenen Verspätungen beeinflusst.

Bemerkung 7.1.9 (Unbegrenzte Verspätungsübertragung).

Für das weitere Vorgehen ist die Annahme entscheidend, dass Verspätungen immer in vollem Umfang übertragen werden, d.h. wir betrachten zunächst keine Anschlussstrategien (s. Abschnitt 6.4), welche die Verspätungsübertragung begrenzen könnten. Dies gilt, solange wir diese Annahme nicht explizit fallen lassen (s. Abschnitt 7.3.2).

Bemerkung 7.1.10 (Verspätungen $C_{V_i}^{(n)}$).

An dieser Stelle soll näher darauf eingegangen werden, welche Verspätungen die Zufallsvariablen $C_{V_i}^{(n)}$, $1 \leq i \leq r$ und $n \in \mathbb{N}$, repräsentieren. Dies hängt unter anderem von den Vorgängern des jeweiligen Knotens V_i im Anschlussgraphen \mathfrak{G} ab. Grundsätzlich gilt für alle $1 \leq i \leq r$ und $n \in \mathbb{N}$

$$C_{V_i}^{(n)} = \max \left\{ [Y_{V_j}^{(n)} + \Lambda_{V_j}^{(n)}]^+ \mid V_j \in \mathfrak{Pre}_{\mathfrak{G}}(V_i) \right\}. \quad (7.1)$$

Wir erinnern uns dabei daran, dass wir zur Vereinfachung der Modellierung keine (zusätzlichen) Puffer für Umsteigeaktivitäten betrachten wollen. Die Zufallsvariablen $Y_{V_j}^{(n)}$ und damit $[Y_{V_j}^{(n)} + \Lambda_{V_j}^{(n)}]^+$ sind für alle $n \in \mathbb{N}$ identisch verteilt, falls $V_j \notin \text{Pre}_{\mathfrak{G}}(V_i) \cap \mathcal{V}_{\mathcal{M}}$, $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$, gilt. Der Vorgängerknoten V_j ist dann also nicht selbst Teil der Kreisvereinigung. Daher findet die Annahme Anwendung, dass die Verteilungsfunktion der Zufallsvariablen Y_{V_j} zuvor bereits abschließend bestimmt wurde (s.o.). Im Falle $V_j \in \text{Pre}_{\mathfrak{G}}(V_i) \cap \mathcal{V}_{\mathcal{M}}$ sind die Zufallsvariablen $Y_{V_j}^{(n)}$ nicht für alle $n \in \mathbb{N}$ identisch verteilt, da sie selbst der Verspätungsfortpflanzung innerhalb der Kreisvereinigung unterliegen. Die zugehörige Verteilungsfunktion wird iterativ approximiert.

Wir wollen nun für einen bestimmten Knoten $V \in \mathcal{V}_K$ des Kreises $K \in \mathcal{M}$ die Grenzverteilung der Abfahrtsverspätung $(Y_V^{(n)})_{n \in \mathbb{N}}$ bestimmen. Dafür führen wir eine Sortierung ein, die uns eine vorteilhafte Beschreibung der iterativen Herangehensweise ermöglicht. So wird im Anschluss an die folgende Bemerkung die Verspätungsfortpflanzung während eines Iterationsschrittes mittels einer Transformation beschrieben. Es wird sowohl gezeigt, welche Eigenschaften als auch welche grundsätzliche Struktur diese Transformation im Rahmen unserer Anwendung besitzt. Die allgemeine Beschreibung mittels einer solchen Transformation hat den Vorteil, dass die Ergebnisse leichter auf andere Fragestellungen, wie etwa mögliche Modellerweiterungen, übertragen werden können. Wir orientieren uns dabei an Lemma 7.2.1 aus [Loy62]. Hier wurde die Entwicklung der Wartezeit der Kunden ebenfalls in einem allgemeineren Rahmen mit Hilfe einer Transformation untersucht, die gewisse Eigenschaften besitzen muss.

Bemerkung 7.1.11 (Knotensortierung $\sigma_K(V)$).

Wir nehmen zunächst an, dass wir eine Sortierung der Knoten $V' \in \mathcal{V}_K$ eines Kreises $K \in \mathcal{M}$ der Kreisvereinigung \mathcal{M} nutzen, die mit jenem Knoten V beginnt, dessen Grenzverteilungen wir bestimmen wollen (s.o.). Davon ausgehend soll die Sortierung der restlichen Knoten der natürlichen Sortierung im Kreis entsprechen. In Abbildung 7.4(a) hieße das beispielsweise, dass die Sortierung mit Knoten $V_1 \in \mathcal{V}_K$ beginnt und dann die Knoten $V_2, V_3, \dots, V_r \in \mathcal{V}_K$ folgen. Eine solche Sortierung stellt sicher, dass Verspätung $Y_V^{(n)}$ den Kreis nach einem Iterationsschritt vollständig durchlaufen hat. Das heißt, dass Verspätung $Y_V^{(n)}$ im n -ten Iterationsschritt in die Berechnung der Verspätungen aller übrigen Knoten des Kreises eingeht. Letztlich beeinflusst $Y_V^{(n)}$ sogar direkt $Y_V^{(n+1)}$ im folgenden Iterationsschritt. Mehr dazu folgt im Anschluss an diese Bemerkung (s. Satz 7.1.12). Wir symbolisieren diese Art der Sortierung der Knoten des Kreises K , beginnend mit Knoten V , durch $\sigma_K(V)$.

Nutzen wir für die Iteration Sortierung $\sigma_K(V)$, so können wir die Verspätungen $Y_V^{(n)}$ für alle $n \in \mathbb{N}$ mittels einer von der Menge der Fahrzeitzuschläge (Puffer)

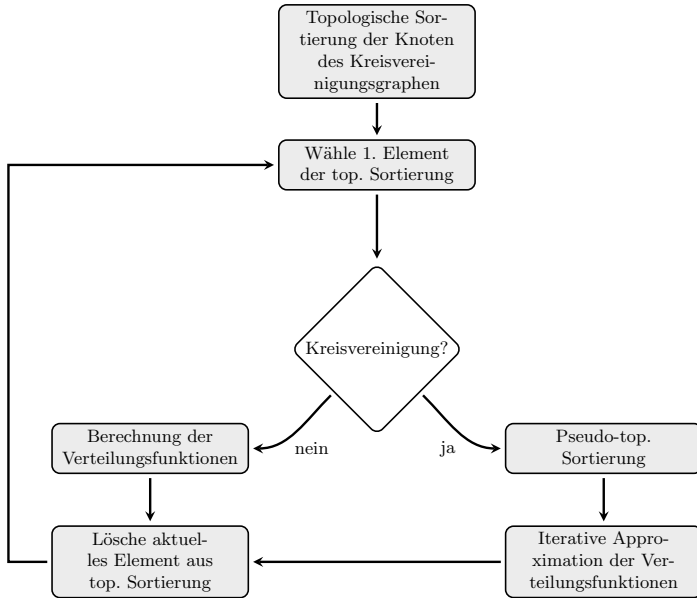


Abbildung 7.5.: Umgang mit Kreisen im Anschlussgraphen

$\mathcal{S} = \{s(a) \mid a \in \mathfrak{A}^{drive}\}$ abhängigen Transformation $f_{\mathcal{S}}^*$ gemäß

$$Y_V^{(n+1)} = f_{\mathcal{S},V}^*(Y_V^{(n)}, D^{(n)}, C^{(n)}) \quad (7.2)$$

mit $D^{(n)} = (D_{V_1}^{(n)}, D_{V_2}^{(n)}, \dots, D_{V_r}^{(n)})$ und $C^{(n)} = (C_{V_1}^{(n)}, C_{V_2}^{(n)}, \dots, C_{V_r}^{(n)})$, $r = |\mathcal{V}_K|$, berechnen (s. Satz 7.1.12). Bemerkenswert ist an dieser Stelle, dass wir die Verspätung $Y_V^{(n+1)}$ im $n+1$ -ten Iterationsschritt (für alle $n \in \mathbb{N}$) direkt aus der Verspätung $Y_V^{(n)}$ im vorherigen Iterationsschritt und den Größen $D^{(n)}$ sowie $C^{(n)}$ bestimmen können. Die genaue Gestalt der Transformation hängt von der Struktur des Kreises K ab. Der folgende Satz beschreibt die Struktur der Transformation im allgemeinen, in Abbildung 7.4 illustrierten, Fall.

Satz 7.1.12 (Allgemeine Struktur der Transformation $f_{\mathcal{S}}^*$).

Berechnen wir die Verspätungen $(Y_V^{(n)})_{n \in \mathbb{N}}$ der Knoten $V \in \mathcal{V}_K$ iterativ mittels der Sortierung $\sigma_K(V_1)$ mit Startknoten $V_1 \in \mathcal{V}_K$ (s. Bemerkung 7.1.11), so gilt für alle $n \in \mathbb{N}$ mit $n \geq 2$ und $\mathcal{S} = \{s(a) \mid a \in \mathfrak{A}^{drive}\}$

$$Y_{V_1}^{(n)} = f_{\mathcal{S},V_1}^*(Y_{V_1}^{(n-1)}, D^{(n-1)}, C^{(n-1)})$$

$$\begin{aligned}
 & := \max \left\{ Y_{V_1}^{(n-1)} + K_{1,r}^{(n-1)}, \right. \\
 & \quad C_{V_2}^{(n-1)} + K_{2,r}^{(n-1)}, C_{V_3}^{(n-1)} + K_{3,r}^{(n-1)}, \dots, \\
 & \quad \left. C_{V_{r-1}}^{(n-1)} + K_{r-1,r}^{(n-1)}, C_{V_r}^{(n-1)} + K_{r,r}^{(n-1)}, C_{V_1}^{(n-1)} \right\} \\
 & = \max \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)}, \right. \\
 & \quad X_2^{(1)} + \sum_{l=2}^{n-1} K_{1,r}^{(l)}, \dots, X_r^{(1)} + \sum_{l=2}^{n-1} K_{1,r}^{(l)}, C_{V_1}^{(1)} + \sum_{l=2}^{n-1} K_{1,r}^{(l)}, \\
 & \quad X_2^{(2)} + \sum_{l=3}^{n-1} K_{1,r}^{(l)}, \dots, X_r^{(2)} + \sum_{l=3}^{n-1} K_{1,r}^{(l)}, C_{V_1}^{(2)} + \sum_{l=3}^{n-1} K_{1,r}^{(l)}, \\
 & \quad \vdots \\
 & \quad X_2^{(n-2)} + \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, \dots, \\
 & \quad X_r^{(n-2)} + \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, C_{V_1}^{(n-2)} + \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, \\
 & \quad \left. X_2^{(n-1)}, \dots, X_r^{(n-1)}, C_{V_1}^{(n-1)} \right\}. \tag{7.4}
 \end{aligned}$$

Dabei gilt $K_{j,l}^{(n)} := \sum_{k=j}^l \Lambda_{V_k}^{(n)}$ für alle $n \in \mathbb{N}$, $j, l \in \mathbb{N}$, und $X_i^{(n)} := C_{V_i}^{(n)} + K_{i,r}^{(n)}$ für alle $n \in \mathbb{N}$ und $1 \leq i \leq r$.

Beweis. Im ersten Iterationsschritt gilt im Falle $Y_{V_1}^{(1)} \equiv 0$

$$Y_{V_2}^{(1)} = \max \left\{ [Y_{V_1}^{(1)} + \Lambda_{V_1}^{(1)}]^+, C_{V_2}^{(1)} \right\} = \max \left\{ [\Lambda_{V_1}^{(1)}]^+, C_{V_2}^{(1)} \right\}.$$

Weiter folgt $Y_{V_3}^{(1)} = \max \left\{ [Y_{V_2}^{(1)} + \Lambda_{V_2}^{(1)}]^+, C_{V_3}^{(1)} \right\}$ sowie nach Abschluss des ersten Iterationsschrittes schließlich

$$Y_{V_r}^{(2)} = \max \left\{ [Y_{V_r}^{(1)} + \Lambda_{V_r}^{(1)}]^+, C_{V_1}^{(1)} \right\}.$$

Im Folgenden wollen wir $Y_{V_1}^{(2)}$ in Abhängigkeit von $Y_{V_1}^{(1)}$ darstellen. Wir erhalten durch sukzessives Einsetzen:

$$Y_{V_2}^{(1)} = \max \left\{ \max \left\{ Y_{V_1}^{(1)} + \Lambda_{V_1}^{(1)}, 0 \right\}, C_{V_2}^{(1)} \right\}$$

$$\begin{aligned}
&= \max \left\{ Y_{V_1}^{(1)} + \Lambda_{V_1}^{(1)}, C_{V_2}^{(1)} \right\} \\
Y_{V_3}^{(1)} &= \max \left\{ \max \left\{ Y_{V_2}^{(1)} + \Lambda_{V_2}^{(1)}, 0 \right\}, C_{V_3}^{(1)} \right\} \\
&= \max \left\{ Y_{V_1}^{(1)} + \sum_{k=1}^2 \Lambda_{V_k}^{(1)}, C_{V_2}^{(1)} + \Lambda_{V_2}^{(1)}, C_{V_3}^{(1)} \right\} \\
Y_{V_4}^{(1)} &= \max \left\{ \max \left\{ Y_{V_3}^{(1)} + \Lambda_{V_3}^{(1)}, 0 \right\}, C_{V_4}^{(1)} \right\} \\
&= \max \left\{ Y_{V_1}^{(1)} + \sum_{k=1}^3 \Lambda_{V_k}^{(1)}, C_{V_2}^{(1)} + \sum_{k=2}^3 \Lambda_{V_k}^{(1)}, C_{V_3}^{(1)} + \Lambda_{V_3}^{(1)}, C_{V_4}^{(1)} \right\}
\end{aligned}$$

Induktiv erhalten wir so für $1 \leq i \leq r$

$$\begin{aligned}
Y_{V_i}^{(1)} &= \max \left\{ Y_{V_1}^{(1)} + \sum_{k=1}^{i-1} \Lambda_{V_k}^{(1)}, C_{V_2}^{(1)} + \sum_{k=2}^{i-1} \Lambda_{V_k}^{(1)}, \right. \\
&\quad \left. C_{V_3}^{(1)} + \sum_{k=3}^{i-1} \Lambda_{V_k}^{(1)}, \dots, C_{V_{i-1}}^{(1)} + \Lambda_{V_{i-1}}^{(1)}, C_{V_i}^{(1)} \right\}
\end{aligned}$$

und

$$\begin{aligned}
Y_{V_1}^{(2)} &= \max \left\{ Y_{V_1}^{(1)} + \sum_{k=1}^r \Lambda_{V_k}^{(1)}, C_{V_2}^{(1)} + \sum_{k=2}^r \Lambda_{V_k}^{(1)}, C_{V_3}^{(1)} + \sum_{k=3}^r \Lambda_{V_k}^{(1)}, \right. \\
&\quad \left. \dots, C_{V_{r-1}}^{(1)} + \sum_{k=r-1}^r \Lambda_{V_k}^{(1)}, C_{V_r}^{(1)} + \Lambda_{V_r}^{(1)}, C_{V_1}^{(1)} \right\}.
\end{aligned}$$

Zur Vereinfachung der Darstellung nutzen wir nun identisch verteilte und stochastisch unabhängige Zufallsfolgen $(K_{j,l}^{(n)})_{n \in \mathbb{N}}$ (Definition s. oben). So erhalten wir

$$\begin{aligned}
Y_{V_1}^{(2)} &= \max \left\{ Y_{V_1}^{(1)} + K_{1,r}^{(1)}, \right. \\
&\quad \left. C_{V_2}^{(1)} + K_{2,r}^{(1)}, C_{V_3}^{(1)} + K_{3,r}^{(1)}, \dots, C_{V_{r-1}}^{(1)} + K_{r-1,r}^{(1)}, C_{V_r}^{(1)} + K_{r,r}^{(1)}, C_{V_1}^{(1)} \right\}
\end{aligned}$$

und für die folgenden Iterationsschritte

$$\begin{aligned}
Y_{V_1}^{(3)} &= \max \left\{ Y_{V_1}^{(2)} + K_{1,r}^{(2)}, \right. \\
&\quad \left. C_{V_2}^{(2)} + K_{2,r}^{(2)}, C_{V_3}^{(2)} + K_{3,r}^{(2)}, \dots, C_{V_{r-1}}^{(2)} + K_{r-1,r}^{(2)}, C_{V_r}^{(2)} + K_{r,r}^{(2)}, C_{V_1}^{(2)} \right\} \\
&= \max \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^2 K_{1,r}^{(l)}, \right.
\end{aligned}$$

$$\left. \begin{aligned} &C_{V_2}^{(1)} + K_{2,r}^{(1)} + K_{1,r}^{(2)}, \dots, C_{V_r}^{(1)} + K_{r,r}^{(1)} + K_{1,r}^{(2)}, C_{V_1}^{(1)} + K_{1,r}^{(2)}, \\ &C_{V_2}^{(2)} + K_{2,r}^{(2)}, \dots, C_{V_r}^{(2)} + K_{r,r}^{(2)}, C_{V_1}^{(2)} \end{aligned} \right\}$$

und

$$\begin{aligned} Y_{V_1}^{(4)} &= \max \left\{ Y_{V_1}^{(3)} + K_{1,r}^{(3)}, \right. \\ &\quad \left. C_{V_2}^{(3)} + K_{2,r}^{(3)}, C_{V_3}^{(3)} + K_{3,r}^{(3)}, \dots, C_{V_{r-1}}^{(3)} + K_{r-1,r}^{(3)}, C_{V_r}^{(3)} + K_{r,r}^{(3)}, C_{V_1}^{(3)} \right\} \\ &= \max \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^3 K_{1,r}^{(l)}, \right. \\ &\quad \left. C_{V_2}^{(1)} + K_{2,r}^{(1)} + \sum_{l=2}^3 K_{1,r}^{(l)}, \dots, C_{V_r}^{(1)} + K_{r,r}^{(1)} + \sum_{l=2}^3 K_{1,r}^{(l)}, C_{V_1}^{(1)} + \sum_{l=2}^3 K_{1,r}^{(l)}, \right. \\ &\quad \left. C_{V_2}^{(2)} + K_{2,r}^{(2)} + K_{1,r}^{(3)}, \dots, C_{V_r}^{(2)} + K_{r,r}^{(2)} + K_{1,r}^{(3)}, C_{V_1}^{(2)} + K_{1,r}^{(3)} \right. \\ &\quad \left. C_{V_2}^{(3)} + K_{2,r}^{(3)}, \dots, C_{V_r}^{(3)} + K_{r,r}^{(3)}, C_{V_1}^{(3)} \right\} \end{aligned}$$

Wiederum induktiv erhalten wir schließlich für alle $n \in \mathbb{N}$ mit $n > 1$

$$\begin{aligned} Y_{V_1}^{(n)} &= \max \left\{ Y_{V_1}^{(n-1)} + K_{1,r}^{(n-1)}, \right. \\ &\quad \left. C_{V_2}^{(n-1)} + K_{2,r}^{(n-1)}, C_{V_3}^{(n-1)} + K_{3,r}^{(n-1)}, \dots, \right. \\ &\quad \left. C_{V_{r-1}}^{(n-1)} + K_{r-1,r}^{(n-1)}, C_{V_r}^{(n-1)} + K_{r,r}^{(n-1)}, C_{V_1}^{(n-1)} \right\} \\ &= \max \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)}, \right. \\ &\quad \left. C_{V_2}^{(1)} + K_{2,r}^{(1)} + \sum_{l=2}^{n-1} K_{1,r}^{(l)}, \dots, C_{V_r}^{(1)} + K_{r,r}^{(1)} + \sum_{l=2}^{n-1} K_{1,r}^{(l)}, C_{V_1}^{(1)} + \sum_{l=2}^{n-1} K_{1,r}^{(l)}, \right. \\ &\quad \left. C_{V_2}^{(2)} + K_{2,r}^{(2)} + \sum_{l=3}^{n-1} K_{1,r}^{(l)}, \dots, C_{V_r}^{(2)} + K_{r,r}^{(2)} + \sum_{l=3}^{n-1} K_{1,r}^{(l)}, C_{V_1}^{(2)} + \sum_{l=3}^{n-1} K_{1,r}^{(l)}, \right. \\ &\quad \vdots \\ &\quad \left. C_{V_2}^{(n-2)} + K_{2,r}^{(n-2)} + \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, \dots, \right. \end{aligned}$$

$$\begin{aligned}
& C_{V_r}^{(n-2)} + K_{r,r}^{(n-2)} + \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, C_{V_1}^{(n-2)} + \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, \\
& \left. C_{V_2}^{(n-1)} + K_{2,r}^{(n-1)}, \dots, C_{V_r}^{(n-1)} + K_{r,r}^{(n-1)}, C_{V_1}^{(n-1)} \right\}.
\end{aligned}$$

□

Lemma 7.1.13 (Eigenschaften der Transformation $f_{\mathcal{S}}^*$).

Eine jede durch (7.2) für einen Knoten $V \in \mathcal{M}$ und fahrplanabhängige Menge $\mathcal{S} = \{s(a) \mid a \in \mathfrak{A}^{\text{drive}}\}$ beschriebene Transformation $f_{\mathcal{S},V}^*$ ist

- (a) nicht-negativ,
- (b) linksseitig stetig in der ersten Komponente (auch in ∞) sowie
- (c) monoton wachsend in allen Komponenten.

Beweis. Im Folgenden seien $x_1, x_2, s \in \mathbb{R}_{\geq 0}$ und $s \in \mathcal{S}$. Die Transformation $f_{\mathcal{S},V}^*$ besteht immer aus einer (mehrfachen) Verknüpfung der folgenden drei (Verspätungs-) Operationen:

$$\begin{aligned}
f_1(x_1, x_2) &:= x_1 + x_2 \\
f_2(x_1, x_2) &:= \max\{x_1, x_2\} \\
f_3(x_1) &:= \max\{x_1 - s, 0\}
\end{aligned}$$

Beispielsweise wäre

$$f_{\mathcal{S}}^*(y, d, c) := \max\{\max\{y + d - s, 0\}, c\} = f_2(f_3(f_1(y, d)), c).$$

1. Es ist unmittelbar klar, dass $z_1 := f_1(x_1, x_2) \geq 0$, $z_2 := f_2(x_1, x_2) \geq 0$ sowie $z_3 := f_3(x_1) \geq 0$. Jede weitere Anwendung einer der Operationen f_1 , f_2 bzw. f_3 auf die bisherigen Ergebnisse führt zu einer nicht-negativen reellen Zahl. Daraus folgt, dass auch das Ergebnis der Transformation $f_{\mathcal{S}}^*$ immer eine nicht-negative reelle Zahl ist.
2. Im Folgenden sollen Monotonie und linksseitige Stetigkeit der Transformation $f_{\mathcal{S}}^*$ (mindestens) in der ersten Komponente nachgewiesen werden. Es folgt für $x_1 \leq x_2$ und $y \in \mathbb{R}$:

- $f_1(x_1, y) = x_1 + y \leq x_2 + y = f_1(x_2, y)$
- $f_1(y, x_1) = y + x_1 \leq y + x_2 = f_1(y, x_2)$
- $f_2(x_1, y) = \max\{x_1, y\} \leq \max\{x_2, y\} = f_2(x_2, y)$

- $f_2(y, x_1) = \max\{y, x_1\} \leq \max\{y, x_2\} = f_2(y, x_2)$
- $f_3(x_1) = \max\{x_1 - s, 0\} \leq \max\{x_2 - s, 0\} = f_3(x_2)$

Offensichtlich sind alle drei Operationen monoton wachsend und stetig in der ersten Komponente. Für f_1 und f_2 gilt dies auch für die zweite Komponente. Da für beliebige stetige Funktionen $h_1, h_2 : \mathbb{R} \rightarrow \mathbb{R}$ auch die Konkatenation $h_1 \circ h_2$ stetig ist, folgt, dass die Transformation f_S^* insgesamt in der ersten Komponente stetig ist. Analog schließen wir auf die Monotonie von f_S^* in der ersten Komponente. \square

Es sei noch angemerkt, dass Verspätungsoperation f_3 in (7.4) gar nicht benötigt wird, da die Nicht-Negativität durch die Maximumsbildung mit der übertragenen nicht-negativen Verspätung $C_{V_1}^{(n-1)}$ garantiert wird. Für die Bildung dieser Verspätung brauchen wir Operation f_3 (s. Bemerkung 7.1.10).

7.2. Verteilungskonvergenz

In [Loy62] ist, neben dem später von uns angewendeten Resultat zur Endlichkeit der Verspätungen bzw. Eindeutigkeit der Verteilungsfunktionen, ebenfalls ein Resultat zur Verteilungskonvergenz enthalten. Die Modellierung ist vergleichbar zu unserer Modellierung in (7.2). Die im folgenden Lemma verwendete Transformation f besitzt nach Voraussetzung dieselben Eigenschaften wie die in Abschnitt 7.1.2 beschriebene Transformation f_S^* . Allerdings sorgt die Forderung der Stationarität der Folge $(U_n)_{n \in \mathbb{Z}}$ dafür, dass wir dieses Ergebnis nicht direkt auf unser Verspätungsfortpflanzungsproblem anwenden können. Daher weisen wir zunächst Verteilungskonvergenz mittels eines von diesem Lemma unabhängigen Resultats nach (s. Satz 7.2.6). Im Anschluss zeigen wir noch, welche Annahmen bzw. Voraussetzungen geändert werden müssten, um ein vergleichbares Resultat für die Anwendung auf unsere Verspätungsmodellierung zu erhalten (s. Korollar 7.2.7). Dieses zum Teil allgemeinere Resultat könnte sich dafür eignen, die Auswirkungen möglicher Modellerweiterungen auf die Verteilungskonvergenz zu untersuchen.

Lemma 7.2.1 (Lemma von Loynes).

[Loy62] Seien $(W_n)_{n \in \mathbb{N}}$ Zufallsvariablen mit

$$W_{n+1} = f(W_n, U_n) \tag{7.5}$$

und $(U_n)_{n \in \mathbb{Z}}$ als stationäre Folge von Zufallsvariablen. Außerdem sei

- (i) die Transformation $f(x, y)$, $x, y \in \mathbb{R}$, nicht-negativ sowie monoton wachsend und linksseitig stetig in x (auch in ∞),
- (ii) $W_1 \equiv 0$.

Dann existiert eine stationäre Folge von Zufallsvariablen $(M_n)_{n \in \mathbb{Z}}$, mit

$$M_{n+1} = f(M_n, U_n)$$

für alle $n \in \mathbb{Z}$, für welche im Falle $\mathbf{P}(M_0 < \infty) = 1$ folgt, dass $W_n \xrightarrow{D} M_0$ ($n \rightarrow \infty$). Insbesondere gilt für alle Folgen $(A_n)_{n \in \mathbb{Z}}$, die (7.5) für alle $n \in \mathbb{Z}$ erfüllen, $A_n \geq M_n$ (\mathbf{P} -f.s.).

Der Beweis ist im Anhang auf Seite 167 dargestellt.

7.2.1. Bedingungen für Verteilungskonvergenz

Wir erinnern uns, dass der Ausgangspunkt für die Bestimmung der Verspätungsverteilungen eine bestimmte Sortierung für *alle* Knoten der Kreisvereinigung war. Dieser Sortierung folgend, werden die Verspätungsverteilungen iterativ bestimmt. Wir betrachten zur Illustration im Folgenden Abbildung 7.6. Wir wollen die Verteilung der Verspätung $Y_{V_1} := \lim_{n \rightarrow \infty} Y_{V_1}^{(n)}$ bestimmen. Wir zeigen zunächst anhand eines Beispiels, dass es im Allgemeinen nicht möglich ist, dass $Y_{V_1}^{(n)}$ in jedem Iterationsschritt $n \in \mathbb{N}$ nur mit Elementen stationärer Zufallsfolgen verknüpft wird. Dies hat zur Folge, dass Lemma 7.2.1 nicht angewendet werden kann bzw. modifiziert werden muss. Wenn wir im Folgenden davon reden, dass eine Verspätung die Knoten eines Kreises *durchläuft*, dann ist damit immer eine Verspätungsfortpflanzung im Sinne von Satz 7.1.12 gemeint.

Im Falle der Kreisvereinigung in Abbildung 7.6(a) ist eine Verknüpfung mit Elementen ausschließlich stationärer Zufallsfolgen zunächst noch möglich. Wir betrachten hierfür Sortierung $\sigma = (V_1, V_2, V_3, V_4, V_5)$. In jedem Iterationsschritt durchläuft die vom Knoten V_1 übertragene Verspätung die Kreise $K_1 := (V_1, V_2, V_3)$ und $K_2 := (V_1, V_4, V_5)$. In jedem dieser Kreise wirken als Verspätungen nur Elemente stationärer Zufallsfolgen. $Y_{V_1}^{(n+1)}$ hängt also, neben diesen Elementen, nur noch von $Y_{V_1}^{(n)}$ ab.

Für die Kreisvereinigung in Abbildung 7.6(b) gilt dies jedoch nicht. Wir betrachten die Sortierung $\sigma = (V_1, V_2, V_3, V_4, V_5, V_6, V_7)$. Wieder durchläuft die vom Knoten V_1 übertragene Verspätung die beiden Kreise K_1 und K_2 . Dieses Mal gibt es jedoch in Knoten V_4 eine weitere Verzweigung. $Y_{V_1}^{(n)}$ durchläuft natürlich auch diesen Teil der Kreisvereinigung, den Kreis $K_3 := (V_4, V_6, V_7)$. Der Unterschied zum Fall in Abbildung 7.6(a) besteht darin, dass die (fortgepflanzte) Verspätung $Y_{V_1}^{(n)}$, sobald sie diesen Kreis erreicht hat, in allen folgenden Iterationsschritten die Verspätungen in diesem Kreis und damit auch die Verspätungen $Y_{V_1}^{(m)}$ mit $m > n + 1$ beeinflusst. Im folgenden Beispiel präzisieren wir dies.

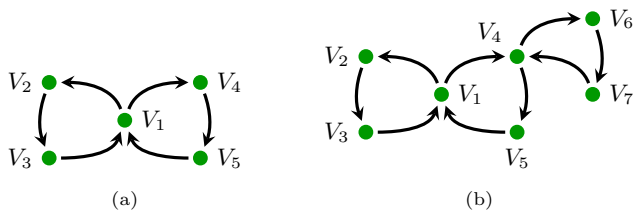


Abbildung 7.6.: Beispiele für Kreisvereinigungen

Beispiel 7.2.2.

Wir betrachten weiter die Situation in Abbildung 7.6(b). Es sei $Y_{V_1}^{(1)} \equiv 0$. Wir verwenden Sortierung $\sigma = (V_1, V_2, V_3, V_4, V_5, V_6, V_7)$. Dann folgt (mit Satz 7.1.12)

$$\begin{aligned}
 Y_{V_1}^{(2)} &= \max \left\{ Y_{V_3}^{(1)} + \Lambda_{V_3}^{(1)}, Y_{V_5}^{(1)} + \Lambda_{V_5}^{(1)}, 0 \right\} \\
 &= \max \left\{ Y_{V_1}^{(1)} + \max \left\{ \sum_{l=1,2,3} \Lambda_{V_l}^{(1)}, \sum_{l=1,4,5} \Lambda_{V_l}^{(1)} \right\}, \right. \\
 &\quad \left. \sum_{l=2,3} \Lambda_{V_l}^{(1)}, \sum_{l=4,5} \Lambda_{V_l}^{(1)}, \Lambda_{V_3}^{(1)}, \Lambda_{V_5}^{(1)}, 0 \right\} \\
 Y_{V_1}^{(3)} &= \max \left\{ Y_{V_3}^{(2)} + \Lambda_{V_3}^{(2)}, Y_{V_5}^{(2)} + \Lambda_{V_5}^{(2)}, 0 \right\} \\
 &= \max \left\{ Y_{V_1}^{(2)} + \sum_{l=1,2,3} \Lambda_{V_l}^{(2)}, [\dots], \right. \\
 &\quad Y_{V_1}^{(2)} + \sum_{l=1,4,5} \Lambda_{V_l}^{(2)}, [\dots], \\
 &\quad \left. Y_{V_1}^{(1)} + \sum_{l=1,4,6,7} \Lambda_{V_l}^{(1)} + \sum_{l=4,5} \Lambda_{V_l}^{(2)}, [\dots], 0 \right\}
 \end{aligned}$$

$Y_{V_1}^{(n)}$ ist für alle $n \geq 3$ unter anderem von der Zufallsvariable

$$Z^{(n)} := Y_{V_1}^{(1)} + X^{(n)}$$

abhängig. Im Falle $n = 3$ wäre (s.o.) beispielsweise

$$X^{(3)} = \sum_{l=1,4,6,7} \Lambda_{V_l}^{(1)} + \sum_{l=4,5} \Lambda_{V_l}^{(2)} .$$

Die Verteilung der $X^{(n)}$ wird im Allgemeinen nicht für alle $n \geq 3$ identisch verteilt sein. Gleiches gilt für die Verteilung der $Z^{(n)}$. Also ist die Folge $(Z^{(n)})_{n \in \mathbb{N}}$ insbesondere nicht stationär.

Wir suchen nun, unabhängig von Lemma 7.2.1, nach Voraussetzungen, welche die Verteilungskonvergenz implizieren. In einem ersten Schritt beschäftigen wir uns mit Bedingungen für die stochastische Monotonie der Zufallsvariablen der Verspätungen.

Lemma 7.2.3.

Wir nehmen an,

- (i) dass wir die Verspätungen in den Knoten einer Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ des Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ iterativ mittels einer beliebigen Sortierung $\sigma(V_1)$ bestimmen, die mit Knoten $V_1 \in \mathcal{V}_{\mathcal{M}}$ beginnt,
- (ii) und dass $Y_{V_1}^{(1)} \equiv 0$.

Dann sind für alle Knoten $V \in \mathcal{V}_{\mathcal{M}}$ der Kreisvereinigung die Glieder der Zufallsfolge $(Y_V^{(n)})_{n \in \mathbb{N}}$ (monoton) stochastisch wachsend.

Beweis. Wesentliche Voraussetzung ist $Y_{V_1}^{(1)} \equiv 0$. Die Tatsache, dass die Sortierung mit Knoten $V_1 \in \mathcal{V}_{\mathcal{M}}$ beginnt hat zur Folge, dass jeder Iterationsschritt $n \in \mathbb{N}$ mit der Berechnung der Verspätung $Y_{V_1}^{(n)}$ startet. Wir nehmen im Folgenden wieder an, dass die Verspätungen möglicher Vorgängerknoten $V_p \in \mathfrak{Pre}_{\mathfrak{G}}(V) \setminus \mathfrak{Pre}_{\mathcal{M}}(V)$ eines Knotens $V \in \mathcal{V}_{\mathcal{M}}$ in allen Iterationsschritten stochastisch unabhängig und identisch verteilt sind. Denn deren Verteilungsfunktionen wurden zuvor bereits bestimmt bzw. (beliebig exakt) approximiert.

Wir betrachten einen bestimmten Knoten $V \in \mathcal{V}_{\mathcal{M}}$ und Vorgängerknoten $V_p \in \mathfrak{Pre}_{\mathfrak{G}}(V)$. Sei $Z_{V_p}^{(n)}(V)$ die Verspätung, die V_p im n -ten Iterationsschritt, $n \in \mathbb{N}$, auf V überträgt. Es gilt dann für alle $V \in \mathcal{V}_{\mathcal{M}}$

$$Y_V^{(n)} = \max \left\{ Z_{V_p}^{(n)}(V) \mid V_p \in \mathfrak{Pre}_{\mathfrak{G}}(V) \right\} . \quad (7.6)$$

Im Folgenden betrachten wir nur noch Vorgänger $V_p \in \mathfrak{Pre}_{\mathcal{M}}(V)$. Die Verspätung $Z_{V_p}^{(n)}(V)$ hängt von der Position des Knotens V_p in der Sortierung $\sigma(V_1)$ ab, relativ zur Position von V . Steht V_p in $\sigma(V_1)$ vor V , so wurden die Verspätungen in V_p im n -ten Iterationsschritt bereits berechnet, bevor die Berechnung der Verspätungen in V im n -ten Iterationsschritt begonnen wird. Es gilt dann

$$Z_{V_p}^{(n)}(V) = [Y_{V_p}^{(n)} + \Lambda_{V_p}^{(n)}]^+ . \quad (7.7)$$

Steht jedoch V_p in $\sigma(V_1)$ hinter V , so setzen wir

$$Z_{V_p}^{(n)}(V) = [Y_{V_p}^{(n-1)} + \Lambda_{V_p}^{(n-1)}]^+ . \quad (7.8)$$

Besonders behandelt wird der erste Iterationsschritt, d.h. $n = 1$. Hier setzen wir im letzteren Falle $Z_{V_p}^{(1)}(V) \equiv 0$.

Für alle $t \in \mathbb{R}_{\geq 0}$ und $n = 2$ gilt offensichtlich

$$\mathbf{P} \left(Y_{V_1}^{(n)} \leq t \right) \leq \mathbf{P} \left(Y_{V_1}^{(n-1)} \leq t \right) = 1 . \quad (7.9)$$

Wir nehmen nun an, dass (7.9) für $2 \leq n \leq m \in \mathbb{N}$ gilt. Es sei $m \in \mathbb{N}$ der erste Iterationsschritt und $V \in \mathcal{V}_{\mathcal{M}}$, $V \neq V_1$, in diesem Fall der erste Knoten in der Sortierung $\sigma(V_1)$ mit

$$\mathbf{P} \left(Y_V^{(m)} \leq t \right) > \mathbf{P} \left(Y_V^{(m-1)} \leq t \right) , \quad (7.10)$$

für einige $t \in \mathbb{R}_{\geq 0}$. Im m -ten Iterationsschritt tritt im Knoten V also zum ersten Mal der Fall ein, dass die Verspätung $Y_V^{(m)}$ nicht stochastisch größer ist als im Iterationsschritt zuvor. Wegen der Monotonie der Maximumsoperation und (7.6) folgt, dass (7.10) auch für die übertragene Verspätung $Z_{V_p}^{(n)}(V)$ eines Vorgängerknotens $V_p \in \mathfrak{Pre}_{\mathcal{M}}(V)$ gelten muss. Hierfür wird Lemma B.1.2 (s. Anhang auf Seite 169) angewandt.

Wir unterscheiden wieder zwei Fälle. Steht Knoten V_p in der Sortierung vor Knoten V , so gilt (7.7) mit $n = m$. Wegen der Monotonie von Summe und Excess-Beyond Operation (vgl. Lemma 7.1.13), müsste (7.10) dann auch für V_p bzw. $Y_{V_p}^{(m)}$ gelten. Dies widerspricht jedoch der Annahme, dass V der erste Knoten in der Sortierung ist, der (7.10) erfüllt. Steht Knoten V_p in der Sortierung hinter Knoten V , so gilt (7.8) mit $n = m$. Analog zum Fall zuvor folgern wir, dass nun (7.10) auch für Knoten V_p bzw. Verspätung $Y_{V_p}^{(m-1)}$ gelten muss. Diese ist aber nach Induktionsannahme stochastisch größer geworden, d.h. es gilt $\mathbf{P}(Y_V^{(m-1)} \leq t) \leq \mathbf{P}(Y_V^{(m-2)} \leq t)$ für alle $t \in \mathbb{R}_{\geq 0}$. Ungleichung (7.10) ist daher verletzt. Im m -ten Iterationsschritt kann es somit insgesamt keinen Knoten geben, für den (7.10) erfüllt ist. Da die Verspätungen $Y_V^{(m)}$ also für alle Knoten V der Sortierung stochastisch größer geworden sind, gilt (7.9) auch für $n = m + 1$. \square

Satz 7.2.4 (Satz von Helly).

[Bil86] *Es sei $(F_n)_{n \in \mathbb{N}}$ eine Folge von Verteilungsfunktionen. Dann existiert eine Teilfolge $(F_{n_k})_{k \in \mathbb{N}}$ sowie eine monoton wachsende und rechtsseitig stetige Funktion F , sodass*

$$\lim_{k \rightarrow \infty} F_{n_k}(t) = F(t) \quad (7.11)$$

für alle Stellen $t \in \mathbb{R}$ gilt, an denen F stetig ist.

Korollar 7.2.5.

Gilt für eine Folge reeller Zufallsvariablen $(X_n)_{n \in \mathbb{N}}$ für alle $t \in \mathbb{R}$

$$\mathbf{P}(X_{n+1} \leq t) \leq \mathbf{P}(X_n \leq t) ,$$

d.h. die Folgenglieder sind stochastisch wachsend, sowie $F_n(t) := \mathbf{P}(X_n \leq t)$, so existiert eine monoton wachsende und rechtsseitig stetige Funktion $F : \mathbb{R} \rightarrow [0, 1]$ mit

$$\lim_{n \rightarrow \infty} F_n(t) = F(t)$$

für alle Stellen $t \in \mathbb{R}$ gilt, an denen F stetig ist. F muss jedoch keine Verteilungsfunktion sein (z.B. im Falle $F \equiv 0$).

Satz 7.2.6 (Verteilungskonvergenz).

Wir nehmen an,

- (i) dass wir die Verspätungen in den Knoten einer Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ des Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$ iterativ mittels einer beliebigen Sortierung $\sigma(V_1)$ bestimmen, die mit Knoten $V_1 \in \mathcal{V}_{\mathcal{M}}$ beginnt,
- (ii) und dass $Y_{V_1}^{(1)} \equiv 0$.

Dann konvergieren die Verspätungen aller Knoten $V \in \mathcal{V}_{\mathcal{M}}$ in Verteilung, falls für alle $V \in \mathcal{V}_{\mathcal{M}}$ gilt, dass $\lim_{n \rightarrow \infty} Y_V^{(n)} < \infty$ (\mathbf{P} -f.s.).

Beweis. Die Aussage folgt unmittelbar aus Lemma 7.2.3 und Korollar 7.2.5. □

Zum Abschluss dieses Abschnittes wollen wir noch festhalten, wie aus einem von Lemma 7.2.1 abgeleiteten Resultat auf die Verteilungskonvergenz geschlossen werden kann. Die Modellierung ist wieder passend zur Situation in Abbildung 7.4, in der nur ein einzelner Kreis einer Kreisvereinigung betrachtet wird. Die Aussage des Korollars ist bezüglich der Transformation $f_{\mathcal{S}, V}$ allgemeiner gefasst als Satz 7.2.6. Die Transformation muss nicht der in Satz 7.1.12 beschriebenen Transformation $f_{\mathcal{S}, V}^*$ entsprechen. Sie kann beispielsweise andere bzw. weitere Operationen umfassen (vgl. Lemma 7.1.13). Sie muss Eigenschaft (i) aus Korollar 7.2.7 erfüllen. Bei der Untersuchung der Auswirkungen möglicher Modellerweiterungen auf die Verteilungskonvergenz ist die Überprüfung dieser Eigenschaften ein möglicher Ansatzpunkt.

Korollar 7.2.7.

Es gelte für die Zufallsfolge $(Y_V^{(n)})_{n \in \mathbb{N}}$ des Knotens $V \in K$ des Kreises $K \in \mathcal{M}$ der Kreisvereinigung \mathcal{M}

$$Y_V^{(n+1)} = f_{\mathcal{S}, V}(Y_V^{(n)}, D^{(n)}, C^{(n)})$$

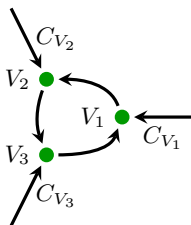


Abbildung 7.7.: Aus nur einem Kreis bestehende Kreisvereinigung

mit für größer werdende $n \in \mathbb{N}$ elementweise stochastisch wachsenden Zufallsvektoren $D^{(n)} = (D_{V_1}^{(n)}, D_{V_2}^{(n)}, \dots, D_{V_r}^{(n)})$ und $C^{(n)} = (C_{V_1}^{(n)}, C_{V_2}^{(n)}, \dots, C_{V_r}^{(n)})$, $r = |\mathcal{V}_K|$. Es seien $Y_V^{(n)}$, $D^{(n)}$ und $C^{(n)}$ für alle $n \in \mathbb{N}$ stochastisch unabhängig. Außerdem sei

(i) die Transformation $f_{S,V}$ nicht-negativ sowie linksseitig stetig in der ersten Komponente (auch in ∞) und monoton wachsend in allen Komponenten,

(ii) $Y_V^{(1)} \equiv 0$.

Dann existiert für alle $V \in K$ eine Zufallsvariable M_V , für die im Falle $\lim_{n \rightarrow \infty} Y_V^{(n)} < \infty$ (\mathbf{P} -f.s.) folgt, dass $Y_V^{(n)} \xrightarrow{\mathcal{D}} M$ ($n \rightarrow \infty$).

Der Beweis zu Korollar 7.2.7 findet sich im Anhang auf Seite 169.

7.2.2. Einfluss unterschiedlicher Sortierungen

In diesem Abschnitt zeigen wir, dass die Konvergenz der Verspätungsverteilungen in einem einzelnen Kreis K einer Kreisvereinigung von der Wahl der Sortierung $\sigma_K(V)$ unabhängig ist. Wir betrachten dazu zunächst ein Beispiel mit zwei unterschiedlichen Sortierungen, die jedoch mit demselben Knoten beginnen.

Beispiel 7.2.8.

Für die Knoten des Kreises K in Abbildung 7.7 gilt $\sigma_K(V_1) = (V_1, V_2, V_3)$ sowie

$$\begin{aligned} Y_{V_1}^{(n+1)} &= f_{S,V_1}^* \left(Y_{V_1}^{(n)}, (D_{V_1}^{(n)}, D_{V_2}^{(n)}, D_{V_3}^{(n)}), (C_{V_1}^{(n)}, C_{V_2}^{(n)}, C_{V_3}^{(n)}) \right) \\ &= \max \left\{ Y_{V_1}^{(n)} + \sum_{l=1}^3 \Lambda_{V_1}^{(n)}, C_{V_2}^{(n)} + \sum_{l=2}^3 \Lambda_{V_1}^{(n)}, C_{V_3}^{(n)} + \Lambda_{V_3}^{(n)}, C_{V_1}^{(n)} \right\} \end{aligned} \quad (7.12)$$

für $n \in \mathbb{N}$ mit $n > 1$. Für alle anderen Sortierungen, die mit Knoten V_1 beginnen, existieren Teilfolgen, für die die Verspätungsentwicklung ebenfalls mittels der

Transformation $f_{\mathcal{S}, V_1}^*$ beschrieben werden kann. Dazu untersuchen wir beispielhaft die Auswirkung der Sortierung $\sigma := (V_1, V_3, V_2)$. In diesem Falle gilt

$$\begin{aligned} Y_{V_1}^{(2)} &= \max \left\{ C_{V_3}^{(1)} + \Lambda_{V_3}^{(1)}, C_{V_1}^{(1)} \right\} \text{ und} \\ Y_{V_1}^{(3)} &= \max \left\{ Y_{V_1}^{(1)} + \Lambda_{V_1}^{(1)} + \Lambda_{V_2}^{(1)} + \Lambda_{V_3}^{(2)}, \right. \\ &\quad \left. C_{V_2}^{(1)} + \Lambda_{V_2}^{(1)} + \Lambda_{V_3}^{(2)}, C_{V_3}^{(2)} + \Lambda_{V_3}^{(2)}, C_{V_1}^{(2)} \right\}. \end{aligned}$$

Allgemein gilt im Falle dieser Sortierung für alle $n \in \mathbb{N}_0$

$$\begin{aligned} &Y_{V_1}^{(2(n+1)+1)} \\ &= \max \left\{ Y_{V_1}^{(2n+1)} + \Lambda_{V_1}^{(2n+1)} + \Lambda_{V_2}^{(2n+1)} + \Lambda_{V_3}^{(2n+2)}, \right. \\ &\quad \left. C_{V_2}^{(2n+1)} + \Lambda_{V_2}^{(2n+1)} + \Lambda_{V_3}^{(2n+2)}, C_{V_3}^{(2n+2)} + \Lambda_{V_3}^{(2n+2)}, C_{V_1}^{(2n+2)} \right\} \\ &= f_{\mathcal{S}, V_1}^* \left(Y_{V_1}^{(2n+1)}, (D_{V_1}^{(2n+1)}, D_{V_2}^{(2n+1)}, D_{V_3}^{(2n+2)}), (C_{V_1}^{(2n+2)}, C_{V_2}^{(2n+1)}, C_{V_3}^{(2n+2)}) \right). \end{aligned} \quad (7.13)$$

Es dauert bei von $\sigma_K(V_1)$ abweichenden, aber mit Knoten V_1 beginnenden, Sortierungen länger, bis sich die Verspätung $Y_{V_1}^{(n)}$ einmal durch den Kreis fortgepflanzt hat.

Beispiel 7.2.8 legt nahe, dass sich zwar die Konvergenzgeschwindigkeit bezüglich der beiden Sortierungen unterscheidet, nicht aber die Grenzverteilungen der Verspätungen $Y_{V_1}^{(n)}$. Dass die Grenzverteilungen unter den Gegebenheiten dieses Beispiels tatsächlich übereinstimmen müssen, zeigt der folgende Satz. Die Eindeutigkeit der Grenzverteilungsfunktionen im allgemeinen Fall, wird in Abschnitt 7.3 behandelt. Dass die Wahl der Sortierung keinen Einfluss auf die Frage hat, ob die Verspätungsverteilungen konvergieren, haben wir bereits in Satz 7.2.6 gezeigt.

Satz 7.2.9 (Einfluss unterschiedlicher Sortierungen).

Es sei Knoten $V_1 \in \mathcal{V}_K$ eines Kreises $K \in \mathcal{K}_{\mathfrak{G}}$ des Anschlussgraphen \mathfrak{G} gegeben. Wir nehmen an,

- (i) dass wir die Verspätung Y_{V_1} iterativ mittels einer beliebigen Knotensortierung bestimmen, die mit Knoten V_1 beginnt,
- (ii) dass für alle $1 \leq i \leq |\mathcal{V}_K|$ alle Elemente der Zufallsfolge $(\Lambda_{V_i}^{(n)})_{n \in \mathbb{N}}$ stochastisch unabhängig und identisch verteilt sind,
- (iii) dass für alle $1 \leq j \leq |\mathcal{V}_K|$ alle Elemente der Zufallsfolge $(C_{V_j}^{(n)})_{n \in \mathbb{N}}$ stochastisch unabhängig und identisch verteilt sind,

7. Verspätungsfortpflanzung in kreisartigen Strukturen

- (iv) dass für alle $1 \leq i \leq |\mathcal{V}_K|$ die Zufallsfolge $(\Lambda_{V_i}^{(n)})_{n \in \mathbb{N}}$ für alle $1 \leq j \leq |\mathcal{V}_K|$ stochastisch unabhängig zur Zufallsfolge $(C_{V_j}^{(n)})_{n \in \mathbb{N}}$ und zur Zufallsfolge $(\Lambda_{V_j}^{(n)})_{n \in \mathbb{N}}$ ($i \neq j$) ist,
- (v) dass für alle $1 \leq i \leq |\mathcal{V}_K|$ die Zufallsfolge $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ für alle $1 \leq j \leq |\mathcal{V}_K|$ stochastisch unabhängig zur Zufallsfolge $(C_{V_j}^{(n)})_{n \in \mathbb{N}}$ und zur Zufallsfolge $(\Lambda_{V_j}^{(n)})_{n \in \mathbb{N}}$ ($i \neq j$) ist,
- (vi) und dass $Y_{V_1}^{(1)} \equiv 0$.

Dann folgt, dass die resultierende Grenzverteilung eindeutig ist, d.h. unabhängig von der speziellen Wahl der Sortierung.

Beweis. Entscheidend ist, dass in jedem Iterationsschritt alle Knoten $V_i \in \mathcal{V}_K$ des Kreises, d.h. genauer alle vorkommenden Quellverspätungen D_{V_i} und alle von Knoten außerhalb des Kreises einwirkende Verspätungen C_{V_i} , genau einmal betrachtet bzw. verwendet werden. Wir untersuchen zunächst die Verwendung der Sortierung $\sigma_K(V)$. Diese sorgt dafür, dass die Verspätung $Y_V^{(n)}$, $n \in \mathbb{N}$, in einem Iterationsschritt durch den kompletten Kreis im Sinne von Satz 7.1.12 fortgepflanzt wird. Das heißt, dass sich die Verspätungen $Y_V^{(n)}$ gemäß (7.2) bestimmen lassen. Die Sortierung folgt dabei gewissermaßen der natürlichen, durch die Struktur des Kreises vorgegebenen Verspätungsfortpflanzung im Kreis. Jede von $\sigma_K(V)$ abweichende, mit Knoten V beginnende Knotensortierung unterbricht diese Fortpflanzung lediglich. So benötigt es mehr Iterationsschritte, bis die Verspätung $Y_V^{(n)}$, $n \in \mathbb{N}$, einmal durch den kompletten Kreis fortgepflanzt wurde und wieder Knoten V erreicht. Es werden dabei, im Vergleich zu Sortierung $\sigma_K(V)$, Quellverspätungen und von außerhalb des Kreises einwirkende Verspätungen anderer Iterationsschritte miteinander verrechnet (vgl. Beispiel 7.2.8). Die resultierenden Verspätungen sind aber wegen der Voraussetzungen (ii) - (v) immer identisch verteilt. So existiert für jede mögliche mit Knoten V beginnende Sortierung eine Teilfolge der Verspätungen (vgl. Beispiel 7.2.8), deren Glieder identisch verteilt sind zu den Gliedern der Folge, deren Grundlage die Sortierung $\sigma_K(V)$ ist. Die Grenzverteilung dieser Folge stimmt daher mit der Grenzverteilung der jeweiligen Teilfolge überein.

Mit Lemma 7.2.3 erhalten wir, dass die Glieder der Folge $(Y_V^{(n)})_{n \in \mathbb{N}}$ für jede Sortierung, die mit Knoten V beginnt, stochastisch wachsen, falls $Y_V^{(1)} \equiv 0$. Daraus folgt mit Korollar 7.2.5 für jede solche Sortierung, dass sämtliche zu ihr gehörenden Teilfolgen gegen dieselbe Grenzverteilung konvergieren. Mit den obigen Überlegungen ergibt sich daraus, dass die Grenzverteilung der Zufallsvariable $\lim_{n \rightarrow \infty} Y_V^{(n)}$ unabhängig von der verwendeten Sortierung ist. \square

Mit den bisherigen Ergebnissen wissen wir nun zum einen, unter welchen Bedingungen die Verspätungsverteilungen in starken Zusammenhangskomponenten des

Ereignis-Aktivitätsnetzwerke konvergieren. Insbesondere wissen wir, dass die Frage der Verteilungskonvergenz unter diesen Bedingungen unabhängig von der speziellen Wahl der Knotensortierung ist. In Satz 7.2.9 wird bereits auf die Eindeutigkeit der Grenzverteilungen eingegangen. Allgemeinere Aussagen dazu folgen in Abschnitt 7.3.

7.3. Endlichkeit der Verspätungen

In Satz 7.2.6 wird darauf hingewiesen, dass die Zufallsvariablen $\lim_{n \rightarrow \infty} Y_V^{(n)}$ nicht endlich sein müssen. Unter welchen Bedingungen dies der Fall ist, bleibt also noch zu untersuchen. Wir betrachten die Stabilität der Verspätungen aller Ankunfts- und Abfahrtsereignisse innerhalb einer Kreisvereinigung in zwei Schritten. Zunächst beschäftigen wir uns mit der Situation für einen einzelnen Kreis der Kreisvereinigung. Im Anschluss schließen wir darauf aufbauend auf Bedingungen für die Stabilität der Verspätungen in der gesamten Kreisvereinigung.

7.3.1. Übertragbare Ergebnisse aus der Warteschlangentheorie

In diesem Abschnitt beschäftigen wir uns mit einem Resultat aus [Loy62] zur Stabilität von Wartezeiten im dort betrachteten Warteschlangensystem, welches wir auch im Rahmen unserer Modellierung der Verspätungsförderung anwenden können.

Definition 7.3.1 (Stabilität).

[Loy62] Eine Folge von Zufallsvariablen $(Y_n)_{n \in \mathbb{N}}$ mit $\mathbf{P}(Y_n < \infty) = 1$ und Verteilungsfunktionen F_n für alle $n \in \mathbb{N}$ wird als stabil bezeichnet, falls eine Zufallsvariable Y existiert mit

- (i) $Y_n \xrightarrow{D} Y$ und
- (ii) $\mathbf{P}(Y < \infty) = 1$.

Die Folge wird als substabil bezeichnet, wenn jede ihrer Teilfolgen selbst eine stabile Teilfolge enthält. Ist sie nicht substabil, so wird sie instabil genannt.

Definition 7.3.2 (Terminale σ -Algebra).

[Kle08] Es sei \mathcal{I} eine abzählbar unendliche Indexmenge und $(\mathfrak{A}_i)_{i \in \mathcal{I}}$ eine Familie von σ -Algebren. Dann heißt

$$\mathfrak{T}((\mathfrak{A}_i)_{i \in \mathcal{I}}) := \bigcap_{\mathcal{J} \subset \mathcal{I}, |\mathcal{J}| < \infty} \sigma \left(\bigcup_{j \in \mathcal{I} \setminus \mathcal{J}} \mathfrak{A}_j \right)$$

die terminale σ -Algebra von $(\mathfrak{A}_i)_{i \in \mathcal{I}}$. Ist $(X_i)_{i \in \mathcal{I}}$ eine Familie von Zufallsvariablen, so setzen wir

$$\mathfrak{T}((X_i)_{i \in \mathcal{I}}) := \mathfrak{T}((\sigma(X_i))_{i \in \mathcal{I}}) .$$

In der terminalen σ -Algebra sind alle Ereignisse A enthalten, deren Eintreten von jeder festen endlichen Teilfamilie der X_i nicht abhängt.

Satz 7.3.3 (Kolomogorov'sches 0-1 Gesetz).

[Kle08] Sei \mathcal{I} eine abzählbar unendliche Indexmenge, und sei $(\mathfrak{A}_i)_{i \in \mathcal{I}}$ eine unabhängige Familie von σ -Algebren. Dann gilt

$$\mathbf{P}(A) \in \{0, 1\}$$

für alle $A \in \mathfrak{T}((\mathfrak{A}_i)_{i \in \mathcal{I}})$.

Satz 7.3.4.

[Loy62] Es sei die Folge $(U_n)_{n \in \mathbb{Z}}$ stochastisch unabhängig und identisch verteilt und $W_1 \equiv 0$. Dann ist die Folge $(W_n)_{n \in \mathbb{N}}$ mit $W_{n+1} = [W_n + U_n]^+$ für alle $n \in \mathbb{N}$ genau dann stabil oder substabil, falls

$$\mathbf{P} \left(\limsup_{n \rightarrow \infty} \sum_{l=1}^n U_l < \infty \right) = 1 .$$

Ansonsten ist sie instabil.

Beweis. Im Beweis zu Lemma 7.2.1 (s. Anhang auf Seite 167) wird für alle $m \in \mathbb{Z}$ die Folge $(W_n^m)_{n \in \mathbb{Z}}$ definiert. Für diese Folge gilt

$$\begin{aligned} W_1^0 &= 0, \\ W_2^0 &= \sup\{W_1^0 + U_1, 0\} = \sup\{U_1, 0\} \\ W_3^0 &= \sup\{W_2^0 + U_2, 0\} = \sup \left\{ \sum_{l=1}^2 U_l, \sum_{l=2}^2 U_l, 0 \right\} . \end{aligned}$$

Es folgt induktiv $W_n^0 = [\sup_{1 \leq m \leq n-1} \sum_{l=m}^{n-1} U_l]^+$. Außerdem erhalten wir mit

$$\begin{aligned} W_0^1 &= 0, \\ W_0^2 &= \sup\{W_{-1}^2 + U_{-1}, 0\} = \sup\{U_{-1}, 0\}, \\ W_0^3 &= \sup\{W_{-1}^3 + U_{-1}\} = \sup \left\{ \sup\{W_{-2}^3 + U_{-2}, 0\} + U_{-1}, 0 \right\} \\ &= \sup \left\{ \sup\{U_{-2}, 0\} + U_{-1}, 0 \right\} = \sup \left\{ \sum_{l=1}^2 U_{-l}, \sum_{l=1}^1 U_{-l}, 0 \right\} \end{aligned}$$

induktiv $W_0^n = [\sup_{1 \leq m \leq n-1} \sum_{l=1}^m U_{-l}]^+$. Im Beweis zu Lemma 7.2.1 wird gezeigt, dass W_n^0 und W_0^n für alle $n \in \mathbb{N}$ identisch verteilt sind. Außerdem wird dort gezeigt, dass W_0^n für $n \rightarrow \infty$ in Verteilung gegen M_0 konvergiert. Also folgt

$$\mathbf{P}(M_0 \leq t) = \mathbf{P} \left(\sup_{m \in \mathbb{N}} \sum_{l=1}^m U_{-l} \leq t \right) = \mathbf{P} \left(\sup_{m \in \mathbb{N}} \sum_{l=1}^m U_l \leq t \right) .$$

Nach Voraussetzung gilt $U_n < \infty$ (\mathbf{P} -f.s.) für alle $n \in \mathbb{N}$ (s. Seite 104). Daher ist die Zufallsvariable $\sup_{m \in \mathbb{N}} \sum_{l=1}^m U_l$ genau dann endlich, wenn auch die Zufallsvariable

$$W := \limsup_{m \rightarrow \infty} \sum_{l=1}^m U_l$$

endlich ist. Das Ereignis $[W < \infty]$ ist terminal. Es gilt daher entweder $\mathbf{P}(W < \infty) = 1$ oder $\mathbf{P}(W < \infty) = 0$. \square

Korollar 7.3.5.

[Loy62] *Es sei die Folge $(U_n)_{n \in \mathbb{Z}}$ stochastisch unabhängig und identisch verteilt. Falls $\mathbf{E}[U_1] < 0$, so ist $(W_n)_{n \in \mathbb{N}}$ für alle Startwerte der Zufallsvariablen W_1 stabil. $(W_n)_{n \in \mathbb{N}}$ besitzt insbesondere für alle Startverteilungen (von W_1) dieselbe (eindeutige) Grenzverteilung.*

Beweis. Mit dem starken Gesetz der großen Zahlen gilt (s. [Kle08])

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n U_l = \mathbf{E}[U_1] \quad (\mathbf{P}\text{-f.s.}) .$$

Daher folgt aus $\mathbf{E}[U_1] < 0$ und $W_1 \equiv 0$, dass

$$\mathbf{P}(M_0 < \infty) = \mathbf{P}\left(\limsup_{n \rightarrow \infty} \sum_{l=1}^n U_l < \infty\right) = 1 .$$

Nun wird eine beliebige Folge von nicht-negativen Zufallsvariablen $(A_n)_{n \in \mathbb{N}}$ mit $\mathbf{P}(A_1 < \infty) = 1$ betrachtet, welche Gleichung (7.5) (s. Seite 114) erfüllt. Es muss also insbesondere nicht $A_1 \equiv 0$ gelten. Im Beweis von Lemma 7.2.1 wurde die Minimalität der W_n^0 , gezeigt. Es gilt also für alle $n \in \mathbb{N}$, dass $A_n \geq W_n^0$ (\mathbf{P} -f.s.). Sobald für ein $n^* \in \mathbb{N}$ das Ereignis $[A_{n^*} = 0]$ eintritt, tritt auch das Ereignis $[W_{n^*}^0 = 0]$ ein. Da beide Folgen Gleichung (7.5) erfüllen, gilt dann $A_n = W_n^0$ (\mathbf{P} -f.s.) für alle $n > n^*$. Ein solches n^* existiert (\mathbf{P} -f.s.). Anderenfalls wäre für alle $n \in \mathbb{N}$ immer (\mathbf{P} -f.s.)

$$A_{n+1} = A_1 + \sum_{l=1}^n U_l > 0 .$$

Dies kann nur dann der Fall sein, wenn $A_1 = +\infty$ oder $\sum_{l=1}^n U_l$ nicht gegen $-\infty$ strebt. Beide Ereignisse sind wegen der Voraussetzungen Nullmengen. Also existiert für alle Werte der Zufallsvariablen A_1 (\mathbf{P} -f.s.) ein $n \in \mathbb{N}$, ab dem die Realisierungen der Folgen W_n^0 und A_n sowie daher auch deren Verteilungen übereinstimmen. Die Folge A_n besitzt eine eindeutige Grenzverteilung, unabhängig von der Startverteilung des Elementes A_1 . \square

Zufallsvariable	Definition
$\Lambda_{V_i}^{(n)}$	$D_{V_i}^{(n)} - s_{V_i}$
$K_{j,l}^{(n)}$	$\sum_{k=j}^l \Lambda_{V_k}^{(n)}$

Tabelle 7.1.: Übersicht verwendeter Zufallsvariablen

Korollar 7.3.5 liefert Aussagen zur Stabilität der Zufallsfolge $(W_n)_{n \in \mathbb{N}}$ und zur Eindeutigkeit der zugehörigen Grenzverteilung im Falle $\mathbf{E}[U_1] < 0$. Für den Fall $\mathbf{E}[U_1] > 0$ kann entsprechend gezeigt werden, dass sich die Wartezeiten $(W_n)_{n \in \mathbb{N}}$ nicht stabilisieren. Die Zufallsvariable der Grenzverteilung wird nicht endlich sein. Schwieriger ist es, allgemeine Aussagen für den Fall $\mathbf{E}[U_1] = 0$ zu erhalten, also für den Fall, wenn die erwartete Bedienzeit $\mathbf{E}[S_n]$ und die erwartete Zwischenankunftszeit $\mathbf{E}[T_n]$ übereinstimmen.

7.3.2. Stabilität in einem Kreis der Kreisvereinigung

Die Modellierung dieses Abschnitts wird in Abbildung 7.4 auf Seite 106 illustriert und in Abschnitt 7.1.2 näher beschrieben. Wir betrachten also einen Kreis $K \in \mathcal{M}$ der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$. Für die Untersuchung der Endlichkeit der Zufallsvariable $Y := \lim_{n \rightarrow \infty} Y_{V_1}^{(n)}$ nehmen wir zunächst an, dass für alle $1 \leq i \leq r$, $r = |\mathcal{V}_K|$, die Zufallsfolgen $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ stochastisch wachsen. Außerdem soll für alle $1 \leq i \leq r$ gelten, dass

- (i) $\mathbf{P}(K_{i,r}^{(1)} < \infty) = 1$, und dass
- (ii) eine Konstante $c_i \in \mathbb{R}_{\geq 0}$ existiert mit $\mathbf{P}(C_{V_i}^{(n)} \in [0, c_i]) = 1$ für alle $n \in \mathbb{N}$.

Annahme (i) ist sicherlich sinnvoll. Andernfalls gäbe es mit positiver Wahrscheinlichkeit eine unendlich große (positive) Quellverspätung. Annahme (ii) ist weniger intuitiv. Sie bedeutet im Wesentlichen, dass von Knoten außerhalb des gerade betrachteten Kreises übertragene Verspätungen beschränkt sind. Dies ist insbesondere dann nicht klar, wenn diese Knoten ebenfalls Teil der Kreisvereinigung sind, d.h. in der Menge $\mathcal{V}_{\mathcal{M}} \setminus \mathcal{V}_K$ liegen. In der Praxis kann dies mittels der bereits diskutierten Anschlussicherungsstrategien (s. Abschnitt 6.4) erreicht werden.

Zunächst wenden wir ein zuvor bereits in Abschnitt 7.3.1 diskutiertes Ergebnis aus [Loy62] auf unsere Modellierung der Verspätungsfortpflanzung an.

Korollar 7.3.6.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$. Es gelte $\mathbf{P}(K_{1,r}^{(1)} < \infty)$,

$r = |\mathcal{V}_K|$. Dann gilt

$$\limsup_{n \rightarrow \infty} \sum_{l=1}^n K_{1,r}^{(l)} < \infty \text{ (P-f.s.) ,}$$

falls $\mathbf{E}[K_{1,r}^{(1)}] < 0$.

Beweis. Wir wenden Satz 7.3.4 an mit $U_n := K_{1,r}^{(n)}$ für alle $n \in \mathbb{N}$. Zusammen mit Korollar 7.3.5 folgt die Behauptung. \square

Korollar 7.3.7.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$. Es gelte $\mathbf{P}(K_{1,r}^{(1)} < \infty)$, $r = |\mathcal{V}_K|$. Dann gilt

$$\limsup_{n \rightarrow \infty} \sum_{l=1}^n K_{1,r}^{(l)} < \infty \text{ (P-f.s.)} \iff \limsup_{n \rightarrow \infty} \sup_{1 \leq m \leq n} \sum_{l=m}^n K_{1,r}^{(l)} < \infty \text{ (P-f.s.) .}$$

Beweis. Folgt unmittelbar aus dem Beweis zu Satz 7.3.4 mit $U_n := K_{1,r}^{(n)}$ für alle $n \in \mathbb{N}$. \square

Zusätzlich benötigen wir noch die Aussage in Lemma 7.3.10. Dessen Beweis wird vorab mittels mehrerer Lemmata geführt.

Lemma 7.3.8.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$. Es gelte $\mathbf{P}(K_{1,r}^{(1)} < \infty)$, $r = |\mathcal{V}_K|$. Falls (P-f.s.) für alle $i \in \{2, 3, \dots, r\}$ gilt

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \left(K_{i,r}^{(n+1)} + \sum_{l=2}^n K_{1,r}^{(l)} \right) < 0 , \quad (7.14)$$

so folgt für alle $i \in \{2, 3, \dots, r\}$

$$\limsup_{n \rightarrow \infty} \max_{2 \leq m \leq n} \left\{ K_{i,r}^{(n+1)} + \sum_{l=2}^m K_{1,r}^{(l)} \right\} < \infty \text{ (P-f.s.) .}$$

Beweis. Aus (7.14) folgt, dass (P-f.s.) für alle $i \in \{2, 3, \dots, r\}$ gilt

$$\exists \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N}, n \geq n_0 : \frac{1}{n} \left(K_{i,r}^{(n)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)} \right) \leq -\varepsilon < 0 ,$$

d.h. es gilt (P-f.s.) für alle $i \in \{2, 3, \dots, r\}$

$$\exists \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \in \mathbb{N}, n \geq n_0 : K_{i,r}^{(n)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)} \leq -n \cdot \varepsilon < 0 .$$

Daher gilt für alle $i \in \{2, 3, \dots, r\}$ und oben beschriebenes $n_0 \in \mathbb{N}$

$$\begin{aligned}
 & \left[\limsup_{n \rightarrow \infty} \max_{2 \leq m \leq n} \left\{ K_{i,r}^{(n+1)} + \sum_{l=2}^m K_{1,r}^{(l)} \right\} \right]^+ \\
 &= \left[\limsup_{n \rightarrow \infty} \max \left\{ \max_{2 \leq m < n_0} \left\{ K_{i,r}^{(m+1)} + \sum_{l=2}^m K_{1,r}^{(l)} \right\}, \right. \right. \\
 & \quad \left. \left. \max_{n_0 \leq m \leq n} \left\{ K_{i,r}^{(m+1)} + \sum_{l=2}^m K_{1,r}^{(l)} \right\} \right\} \right]^+ \\
 &= \left[\max_{2 \leq m < n_0} \left\{ K_{i,r}^{(m+1)} + \sum_{l=2}^m K_{1,r}^{(l)} \right\} \right]^+ < \infty. \quad \square
 \end{aligned}$$

Lemma 7.3.9.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\Theta}$. Es gelte $\mathbf{P}(K_{1,r}^{(1)} < \infty)$, $r = |\mathcal{V}_K|$. Dann gilt für alle $i \in \{2, 3, \dots, r\}$ (**P-f.s.**)

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \left(K_{i,r}^{(n+1)} + \sum_{l=2}^n K_{1,r}^{(l)} \right) < 0,$$

falls $\mathbf{E}[K_{1,r}^{(1)}] < 0$.

Beweis. Nach Voraussetzung ist $\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{l=2}^n K_{1,r}^{(l)} < 0$. Die Folge $(K_{i,r}^{(n)})_{n \in \mathbb{N}}$ ist für alle $i \in \{2, 3, \dots, r\}$ stochastisch unabhängig und identisch verteilt mit $\mathbf{E}[K_{i,r}^{(1)}] < \infty$. Es gilt (**P-f.s.**)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{l=1}^n K_{i,r}^{(l)} = \mathbf{E}[K_{i,r}^{(1)}].$$

Daraus folgt $\lim_{n \rightarrow \infty} \frac{1}{n+1} K_{i,r}^{(n+1)} = 0$ (**P-f.s.**). Anderenfalls gäbe es eine Teilfolge $(n_k)_{k \in \mathbb{N}}$ mit

$$\exists a \in \mathbb{R}_{>0} \forall k \in \mathbb{N} : \frac{1}{n_k} K_{i,r}^{(n_k)} > a.$$

Daraus folgte

$$\begin{aligned}
 \frac{1}{n_k} \sum_{l=1}^{n_k} K_{i,r}^{(l)} &> \frac{1}{n_k} \sum_{l=1}^{n_k-1} K_{i,r}^{(l)} + a \\
 &= \frac{n_k - 1}{n_k} \cdot \frac{1}{n_k - 1} \cdot \sum_{l=1}^{n_k-1} K_{i,r}^{(l)} + a \xrightarrow{k \rightarrow \infty} \mathbf{E}[K_{i,r}^{(1)}] + a.
 \end{aligned}$$

Dies steht im Widerspruch zu $\lim_{k \rightarrow \infty} \frac{1}{n_k} \sum_{l=1}^{n_k} K_{i,r}^{(l)} = \mathbf{E}[K_{i,r}^{(1)}]$. Analog kann für $a \in \mathbb{R}_{<0}$ geschlossen werden. Also erhalten wir

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \left(K_{i,r}^{(n+1)} + \sum_{l=2}^n K_{1,r}^{(l)} \right) &\leq \limsup_{n \rightarrow \infty} \frac{1}{n} K_{i,r}^{(n+1)} + \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{l=2}^n K_{1,r}^{(l)} \\ &= \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{l=2}^n K_{1,r}^{(l)} < 0. \quad \square \end{aligned}$$

Lemma 7.3.10.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$. Es gelte $\mathbf{P}(K_{1,r}^{(1)} < \infty)$, $r = |\mathcal{V}_K|$. Dann gilt für alle $i \in \{2, 3, \dots, r\}$

$$\limsup_{n \rightarrow \infty} \max_{2 \leq m \leq n} \left\{ K_{i,r}^{(m-1)} + \sum_{l=m}^n K_{1,r}^{(l)} \right\} < \infty \quad (\mathbf{P}\text{-f.s.}), \quad (7.15)$$

falls $\mathbf{E}[K_{1,r}^{(1)}] < 0$.

Beweis. $(K_{1,r}^{(l)})_{l \in \mathbb{N}}$ sowie $(K_{i,r}^{(m)})_{m \in \mathbb{N}}$ für alle $i \in \{2, 3, \dots, r\}$ sind stochastisch unabhängig und identisch verteilt. Die Zufallsvariablen

$$K_{i,r}^{(m)} \quad \text{und} \quad \sum_{l=m+1}^n K_{1,r}^{(l)}$$

sind ebenfalls für alle $m \in \mathbb{N}$ stochastisch unabhängig. Daher gilt (7.15) genau dann, wenn

$$\limsup_{n \rightarrow \infty} \max_{2 \leq m \leq n} \left\{ K_{i,r}^{(n+1)} + \sum_{l=2}^m K_{1,r}^{(l)} \right\} < \infty \quad (\mathbf{P}\text{-f.s.}).$$

Mit Lemma 7.3.9 und Lemma 7.3.8 folgt die Behauptung. □

Wir untersuchen im Folgenden zunächst nur die Situation in einem einzelnen Kreis der Kreisvereinigung. Stabilität in der kompletten Kreisvereinigung wird darauf aufbauend in Abschnitt 7.3.3 betrachtet.

Satz 7.3.11.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ und $V_1 \in \mathcal{V}_K$ ein beliebiger Knoten dieses Kreises. Es gelte

- (a) $Y_{V_1}^{(1)} < \infty$ (\mathbf{P} -f.s.),
- (b) $\mathbf{P}(K_{i,r}^{(1)} < \infty) = 1$ für alle $1 \leq i \leq r$, und

7. Verspätungsfortpflanzung in kreisartigen Strukturen

(c) es existiere eine Konstante $c_i \in \mathbb{R}_{\geq 0}$ mit $\mathbf{P} \left(C_{V_i}^{(n)} \in [0, c_i] \right) = 1$ für alle $n \in \mathbb{N}$.

Dann folgt

$$\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} < \infty \quad (\mathbf{P}\text{-f.s.}), \quad (7.16)$$

falls $\mathbf{E}[K_{1,r}^{(1)}] < 0$.

Beweis. Grundlage für diesen Beweis ist Gleichung (7.4) aus Satz 7.1.12. Es gilt für alle $n \in \mathbb{N}$ mit $n > 1$

$$\begin{aligned} Y_{V_1}^{(n)} = \max & \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)}, \right. \\ & \max_{2 \leq m \leq n-1, 2 \leq i \leq r} \left\{ C_{V_i}^{(m-1)} + K_{i,r}^{(m-1)} + \sum_{l=m}^{n-1} K_{1,r}^{(l)} \right\}, \\ & \max_{2 \leq m \leq n-1} \left\{ C_{V_1}^{(m-1)} + \sum_{l=m}^{n-1} K_{1,r}^{(l)} \right\}, \\ & \left. \max_{2 \leq i \leq r} \left\{ C_{V_i}^{(n-1)} + K_{i,r}^{(n-1)} \right\}, C_{V_1}^{(n-1)} \right\}. \end{aligned} \quad (7.17)$$

Wir folgern

$$\begin{aligned} Y_{V_1}^{(n)} \leq & \max_{1 \leq i \leq r, 1 \leq m \leq n-1} C_{V_i}^{(m)} + \\ & \max \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)}, \max_{2 \leq i \leq r, 2 \leq m \leq n-1} \left\{ K_{i,r}^{(m-1)} + \sum_{l=m}^{n-1} K_{1,r}^{(l)} \right\}, \right. \\ & \left. \max_{2 \leq m \leq n-1} \sum_{l=m}^{n-1} K_{1,r}^{(l)}, \max_{2 \leq i \leq r} K_{i,r}^{(n-1)}, 0 \right\}. \end{aligned} \quad (7.18)$$

Mit Annahme (c) gilt, dass

$$\limsup_{n \rightarrow \infty} \max_{1 \leq i \leq r, 1 \leq m \leq n-1} C_{V_i}^{(m)}$$

beschränkt und daher insbesondere (\mathbf{P} -f.s.) endlich ist. Wir untersuchen nun das Grenzverhalten der einzelnen Terme des Maximumsausdruckes in (7.18). Nach Voraussetzung gilt mit Korollar 7.3.6 und Korollar 7.3.7

$$\limsup_{n \rightarrow \infty} \sup_{1 \leq m \leq n} \sum_{l=m}^n K_{1,r}^{(l)} < \infty \quad (\mathbf{P}\text{-f.s.}).$$

Daher folgt für

$$B_1^{(n-1)} := \max_{2 \leq m \leq n-1} \sum_{l=m}^{n-1} K_{1,r}^{(l)}$$

unmittelbar $\limsup_{n \rightarrow \infty} B_1^{(n-1)} < \infty$ (\mathbf{P} -f.s.). Mit $Y_{V_1}^{(1)} < \infty$ (\mathbf{P} -f.s.) folgt für

$$B_2^{(n-1)} := Y_{V_1}^{(1)} + \sum_{l=1}^{n-1} K_{1,r}^{(l)}$$

offensichtlich ebenso $\limsup_{n \rightarrow \infty} B_2^{(n-1)} < \infty$ (\mathbf{P} -f.s.). Da $K_{i,r}^{(m)}$ nach Voraussetzung für alle $1 \leq i \leq r$ und $m \in \mathbb{N}$ (\mathbf{P} -f.s.) endlich ist, folgern wir für das über einen endlichen Bereich definierte Maximum

$$B_3^{(n-1)} := \max_{2 \leq i \leq r} K_{i,r}^{(n-1)},$$

dass $\limsup_{n \rightarrow \infty} B_3^{(n-1)}$ (\mathbf{P} -f.s.) endlich ist. Es bleibt abschließend der Ausdruck

$$\begin{aligned} B_4^{(n-1)} &:= \max_{2 \leq i \leq r, 2 \leq m \leq n-1} \left\{ K_{i,r}^{(m-1)} + \sum_{l=m}^{n-1} K_{1,r}^{(l)} \right\} \\ &= \max_{2 \leq i \leq r} \left\{ \max_{2 \leq m \leq n-1} \left\{ K_{i,r}^{(m-1)} + \sum_{l=m}^{n-1} K_{1,r}^{(l)} \right\} \right\} \end{aligned}$$

zu untersuchen. Mit Lemma 7.3.10 folgt unmittelbar

$$\limsup_{n \rightarrow \infty} \max_{2 \leq m \leq n-1} \left\{ K_{i,r}^{(m)} + \sum_{l=m+1}^n K_{1,r}^{(l)} \right\} < \infty \text{ (\mathbf{P} -f.s.)},$$

für alle $2 \leq i \leq r$. Nach Vertauschung von Grenzwerts- und Maximumsbildung erhalten wir damit $\limsup_{n \rightarrow \infty} B_4^{(n-1)} < \infty$ (\mathbf{P} -f.s.). Insgesamt gilt

$$\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} \leq \max_{1 \leq k \leq 4} \left\{ \limsup_{n \rightarrow \infty} B_k^{(n)}, 0 \right\} < \infty \text{ (\mathbf{P} -f.s.)}. \quad \square$$

Korollar 7.3.12 (Kriterium für Endlichkeit).

Unter den Bedingungen von Satz 7.3.11 ist $\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} < \infty$ (\mathbf{P} -f.s.), falls

$$\sum_{k=1}^r \mathbf{E} \left[D_k^{(1)} \right] < \sum_{k=1}^r s_k. \quad (7.19)$$

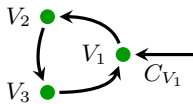


Abbildung 7.8.: Aus nur einem Kreis bestehende Kreisvereinigung

Beweis. Es gilt

$$\begin{aligned} \mathbf{E} \left[K_{1,r}^{(1)} \right] < 0 &\iff \mathbf{E} \left[\sum_{k=1}^r D_k^{(1)} - s_k \right] < 0 \\ &\iff \sum_{k=1}^r \mathbf{E} \left[D_k^{(1)} \right] < \sum_{k=1}^r s_k . \end{aligned} \quad \square$$

Der für die Aussage $\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} < \infty$ (**P**-f.s.) kritische Ausdruck im Beweis zu Satz 7.3.11 ist

$$\limsup_{n \rightarrow \infty} \max_{1 \leq i \leq r, 1 \leq m \leq n-1} C_{V_i}^{(m)} . \quad (7.20)$$

Dieser muss endlich sein. In Satz 7.3.11 wurde dies durch Voraussetzung (c) erreicht. Ein andere Herangehensweise wäre, zu untersuchen, ob es reicht, wenn an Stelle der Beschränktheit nur endliche Erwartungswerte gefordert werden. Das folgende Beispiel liefert hierfür eine Grundlage.

Beispiel 7.3.13.

Wir betrachten die Situation in Abbildung 7.8. Es existiert ein $K = (V_1, V_2, V_3)$. Es sei

- (i) $Y_{V_1}^{(1)} = C_{V_1}^{(1)}$,
- (ii) $(C_{V_1}^{(n)})_{n \in \mathbb{N}}$ sei stochastisch unabhängig und identisch verteilt sowie
- (iii) $\mathbf{E}[C_{V_1}^{(1)}] < \infty$.

Wir nehmen an, dass es innerhalb des Kreises keine positive Quellverspätung gibt. Es sei einzig ein Zeitpuffer $s \in \mathbb{R}_{>0}$ enthalten, sodass

$$Y_{V_1}^{(n+1)} = \max \left\{ Y_{V_1}^{(n)} - s, C_{V_1}^{(n+1)} \right\}$$

für alle $n \in \mathbb{N}$. $\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} < \infty$ ist ein terminales Ereignis. Es gilt daher

$$\mathbf{P} \left(\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} < \infty \right) \in \{0, 1\} .$$

Für die Verteilungsfunktionen der Folge $(Y_{V_1}^{(n)})_{n \in \mathbb{N}}$ folgt für alle $t \in \mathbb{R}_{\geq 0}$

$$\begin{aligned} F_{Y_1^{(1)}}(t) &= F_{C_{V_1}^{(1)}}(t) , \\ F_{Y_1^{(2)}}(t) &= F_{C_{V_1}^{(1)}}(t) \cdot F_{C_{V_1}^{(1)}}(t+s) , \\ F_{Y_1^{(3)}}(t) &= F_{C_{V_1}^{(1)}}(t) \cdot F_{C_{V_1}^{(1)}}(t+s) \cdot F_{C_{V_1}^{(1)}}(t+2s) , \\ &\vdots \\ F_{Y_1^{(n)}}(t) &= \prod_{l=1}^n F_{C_{V_1}^{(1)}}(t+(l-1)s) . \end{aligned}$$

Es gilt weiter für alle $t \in \mathbb{R}_{\geq 0}$

$$\sum_{l=1}^{\infty} \left(1 - F_{C_{V_1}^{(1)}}(t+(l-1)s) \right) < \int_0^{\infty} \left(1 - F_{C_{V_1}^{(1)}}(s) \right) ds = \mathbf{E} \left[C_{V_1}^{(1)} \right] < \infty .$$

Daraus schließen wir, mittels eines grundlegenden Resultats zum Zusammenhang zwischen der Konvergenz von Reihen und endlichen Produkten, für alle $t \in \mathbb{R}_{\geq 0}$

$$\prod_{l=1}^{\infty} F_{C_{V_1}^{(1)}}(t+(l-1)s) > 0 .$$

Daraus folgt jedoch noch nicht $\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} < \infty$ (\mathbf{P} -f.s.).

Eine hierauf aufbauende allgemeinere Untersuchung, beispielsweise auch für stochastisch wachsende $(C_{V_1}^{(n)})_{n \in \mathbb{N}}$, wird im Rahmen dieser Arbeit nicht mehr durchgeführt.

Satz 7.3.14.

Es sei $K \in \mathcal{M}$ ein Kreis der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ und $V_1 \in \mathcal{V}_K$ ein beliebiger Knoten dieses Kreises. Es gelte

- (a) $Y_{V_1}^{(1)} < \infty$ (\mathbf{P} -f.s.),
- (b) $\mathbf{P}(K_{i,r}^{(1)} < \infty) = 1$ für alle $1 \leq i \leq r$, und
- (c) es existiere eine Konstante $c_i \in \mathbb{R}_{\geq 0}$ mit $\mathbf{P}(C_{V_i}^{(n)} \in [0, c_i]) = 1$ für alle $n \in \mathbb{N}$.

Dann folgt

$$\limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} = \infty \text{ (\mathbf{P} -f.s.) ,}$$

falls $\mathbf{E}[K_{1,r}^{(1)}] > 0$.

Beweis. Aus $\mathbf{E}[K_{1,r}^{(1)}] > 0$ folgt mit dem starken Gesetz der großen Zahlen (vgl. Beweis zu Korollar 7.3.5), dass $\limsup_{n \rightarrow \infty} \sum_{l=1}^n K_{1,r}^{(l)} = \infty$ (\mathbf{P} -f.s.). Mit Korollar 7.3.7 ist das äquivalent zur Aussage $\limsup_{n \rightarrow \infty} S_n = \infty$ (\mathbf{P} -f.s.) mit

$$S_n := \sup_{1 \leq m \leq n} \sum_{l=m}^n K_{1,r}^{(l)}.$$

für alle $n \in \mathbb{N}$. Wir definieren für alle $n \in \mathbb{N}$ mit $n > 1$

$$A^{(n-1)} := \min \left\{ Y_{V_1}^{(1)}, \min_{2 \leq m \leq n-1, 2 \leq i \leq r} \left\{ C_{V_i}^{(m-1)} + K_{i,r}^{(m-1)} \right\}, \right. \\ \left. \min_{2 \leq m \leq n-1} C_{V_1}^{(m-1)}, \min_{2 \leq i \leq r} \left\{ C_{V_i}^{(n-1)} + K_{i,r}^{(n-1)} \right\}, C_{V_1}^{(n-1)} \right\}.$$

Aus den Voraussetzungen (b) und (c) folgt $\mathbf{P}(\limsup_{n \rightarrow \infty} |A^{(n)}| < \infty) = 1$. Außerdem erhalten wir aus (7.17) für alle $n \in \mathbb{N}$ mit $n > 1$

$$Y_{V_1}^{(n)} \geq A^{(n-1)} + \max \left\{ \sum_{l=1}^{n-1} K_{1,r}^{(l)}, \dots, \sum_{l=n-1}^{n-1} K_{1,r}^{(l)}, 0 \right\} = A^{(n-1)} + [S_{n-1}]^+.$$

Hieraus schließen wir

$$\limsup_{n \rightarrow \infty} S_n = \infty \text{ (\mathbf{P} -f.s.)} \implies \limsup_{n \rightarrow \infty} Y_{V_1}^{(n)} = \infty \text{ (\mathbf{P} -f.s.)}. \quad \square$$

Korollar 7.3.15 (Eindeutigkeit).

Wir berechnen die Verspätungen $(Y_V^{(n)})_{n \in \mathbb{N}}$ in einem Kreis $K \in \mathcal{M}$ der Kreisvereinigung \mathcal{M} gemäß (7.3). Es gelte Folgendes:

- (a) $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ ist für alle $i \in \{1, \dots, r\}$ ein stochastischer Prozess stochastisch wachsender Zufallsvariablen,
- (b) $(C_{V_i}^{(n)})_{n \in \mathbb{N}}$ ist für alle $i \in \{1, \dots, r\}$ stochastisch unabhängig von $Y_{V_1}^{(1)}$ und
- (c) Bedingung (7.19) ist erfüllt.

Dann ist die Grenzverteilung der Zufallsvariablen $\lim_{n \rightarrow \infty} Y_V^{(n)}$ eindeutig, d.h. unabhängig von der speziellen Startverteilung der Größe $Y_V^{(1)}$, falls $\lim_{n \rightarrow \infty} Y_V^{(n)} < \infty$ (\mathbf{P} -f.s.).

Beweis. Im Folgenden wird der oben genannte Knoten V mit V_1 bezeichnet. Es sei $Z_n := [Y_{V_1}^{(1)} + \sum_{l=1}^n K_{1,r}^{(l)}]^+$ für alle $n \in \mathbb{N}$. Analog zum Beweis zu Korollar 7.3.5 gilt hier, dass die Grenzverteilung der Zufallsgröße $\lim_{n \rightarrow \infty} Z_n$ für $n \rightarrow \infty$ eindeutig ist. Falls $Y_{V_1}^{(1)}$ (\mathbf{P} -f.s.) endlich ist, folgt aus $\lim_{n \rightarrow \infty} \sum_{l=1}^n K_{1,r}^{(l)} = -\infty$ (\mathbf{P} -f.s.), dass

$\lim_{n \rightarrow \infty} Z_n = 0$ (**P**-f.s.). In (7.4) ist zu sehen, dass Z_n auch die einzige der in die Berechnung von $Y_{V_1}^{(n)}$ einfließenden Zufallsgrößen ist, die von $Y_{V_1}^{(1)}$ abhängt. Daher ist auch die Grenzverteilung $\lim_{n \rightarrow \infty} Y_{V_1}^{(n)}$ eindeutig, d.h. unabhängig von der Startverteilung von $Y_V^{(1)}$, falls die Bedingung in (7.19) erfüllt ist. \square

7.3.3. Stabilität in der Kreisvereinigung

In diesem Abschnitt untersuchen wir den allgemeinen Fall einer aus mehreren Kreisen bestehenden Kreisvereinigung. Der erste Ansatz ist immer, die zuvor erhaltenen Ergebnisse zu übertragen, indem wir die jeweiligen Bedingungen für jeden Kreis der Kreisvereinigung fordern.

Satz 7.3.16 (Stabilität, notwendige Bedingung).

Sei $V \in \mathcal{V}_K \subset \mathfrak{V}$ ein beliebiger Knoten in einem Kreis $K \in \mathcal{M}$ einer Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ im Anschlussgraphen $\mathfrak{G} = (\mathfrak{V}, \mathfrak{E})$. Außerdem existiere in der Kreisvereinigung ein Knoten $V_1 \in \mathcal{V}_{\mathcal{M}}$ mit $Y_{V_1}^{(1)} \equiv 0$. Mit diesem Knoten werde die Berechnung der Verspätungen in jedem Iterationsschritt $n \in \mathbb{N}$ begonnen. Ist die Verspätungsübertragung zwischen den Kreisen unbeschränkt, so kann die Zufallsvariable der Grenzverteilung der Zufallsfolge $(Y_V^{(n)})_{n \in \mathbb{N}}$ nur dann stabil sein, wenn Bedingung (7.19) in jedem Kreis $K' \in \mathcal{M}$ erfüllt ist.

Beweis. Sind die Zufallsvariablen $\lim_{n \rightarrow \infty} Y_V^{(n)}$ für alle $V \in \mathcal{V}_{\mathcal{M}}$ (**P**-f.s.) endlich, so ist die Verteilungskonvergenz dieser Verspätungen unter den gegebenen Voraussetzungen bereits mit Satz 7.2.6 nachgewiesen. Es bleibt die Endlichkeit der Zufallsvariablen der Grenzverteilungen zu untersuchen.

Nehmen wir an, dass in der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ ein Kreis $K \in \mathcal{M}$ existiert, in dem Bedingung (7.19) nicht erfüllt ist. Für die Knoten $V \in \mathcal{V}_K$ dieses Kreises gilt, dass sich die Verspätungen $(Y_V^{(n)})_{n \in \mathbb{N}}$ nicht stabil entwickeln. Die jeweilige Zufallsvariable der Grenzverteilung ist (**P**-f.s.) nicht endlich. Dann folgt, dass auch für alle anderen Knoten der Kreisvereinigung die Zufallsvariable der jeweiligen Grenzverteilung nicht endlich sein kann. Wir erinnern uns, dass eine Kreisvereinigung eine starke Zusammenhangskomponente des Graphen darstellt. Jeder Knoten ist von jedem anderen Knoten aus erreichbar. Verspätungsübertragungen von einem Kreis in einen anderen finden mittels der Maximums-Operation statt. Die Zufallsgrößen $C_{V_i}^{(n)}$, $1 \leq i \leq \mathcal{V}_K$ mit $K \in \mathcal{M}$, wurden in den Voraussetzungen dieses Satzes nicht beschränkt. \square

Dieser Satzes liefert also eine notwendige Bedingung für die Endlichkeit der Zufallsvariablen aller Grenzverspätungsverteilungen in der betrachteten Kreisvereinigung. Aus Satz 7.3.11 schließen wir auf hinreichende Bedingungen.

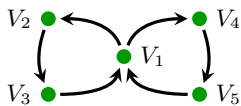


Abbildung 7.9.: Beispiel für eine Kreisvereinigung

Satz 7.3.17.

Sind die Bedingungen (a), (c) aus Satz 7.3.11 sowie Bedingung (7.19) für alle Kreise $K \in \mathcal{M}$ der Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ erfüllt, so gilt

$$\lim_{n \rightarrow \infty} Y_V^{(n)} < \infty \text{ (P-f.s.)}$$

für alle Knoten $V \in \mathcal{V}_{\mathcal{M}}$.

Mit Satz 7.3.17 erhalten wir die Stabilität der Grenzverteilungen, falls wir beispielsweise im ersten Iterationsschritt mit einer (P-f.s.) nicht positiven Verspätung starten und zusätzlich zu (7.19) eine (einfache) Anschlussicherungsstrategie (s. Abschnitt 6.4.1) mit endlicher maximaler Wartezeit verwenden. Letztere wird dann nur für Umsteigebeziehungen benötigt, bei denen Verspätung von Ereignissen außerhalb des jeweils betrachteten Kreises übertragen wird.

Satz 7.3.18 (Eindeutigkeit).

Sind die Grenzverteilungen der Verspätungen $(Y_V^{(n)})_{n \in \mathbb{N}}$ für alle Knoten $V \in \mathcal{V}_{\mathcal{M}}$ einer Kreisvereinigung $\mathcal{M} \in \mathcal{M}_{\mathfrak{G}}$ stabil, so sind sie außerdem eindeutig, falls in jedem Kreis $K \in \mathcal{M}$ die Bedingungen aus Korollar 7.3.15 erfüllt sind. Ist $V_1 \in \mathcal{V}_{\mathcal{M}}$ der Knoten, mit dem die Verspätungsberechnung in jedem Iterationsschritt beginnt, so sind die Grenzverteilungen unabhängig von der Startverteilung der Zufallsgröße $Y_{V_1}^{(1)}$.

Tatsächlich wird die Forderung der stochastischen Unabhängigkeit in Bedingung (b) aus Korollar 7.3.15 im Allgemeinen nicht immer bzw. nur näherungsweise erfüllt sein. Sie ist die zentrale Bedingung für Verspätungsfortpflanzung auf Ebene der Verteilungsfunktionen. Ist sie nicht erfüllt, können die Resultate fehlerhaft sein. Wir zeigen im Folgenden für ein stark vereinfachtes Beispiel, wie wir auf Eindeutigkeit der Grenzverteilung schließen können, auch wenn diese Bedingung nicht erfüllt ist. Ein allgemeiner theoretischer Beweis wird in dieser Arbeit nicht mehr geführt.

Beispiel 7.3.19.

Wir betrachten im Folgenden die Situation in Abbildung 7.9 und untersuchen die Fortpflanzung der Verspätung $Y_{V_1}^{(1)}$. Die Kreisvereinigung enthält die beiden Kreise

$$K_1 = (V_1, V_2, V_3) \quad \text{und} \quad K_2 = (V_1, V_4, V_5) .$$

Als Knotensortierung verwenden wir $\sigma(V_1) = (V_1, V_2, V_3, V_4, V_5)$. Lediglich für Knoten V_1 fließen in die benötigte Maximumsoperation eventuell stochastisch abhängige Zufallsgrößen ein. Es sei für alle $n \in \mathbb{N}$ und einen Kreis $K \in \mathcal{M} = \{K_1, K_2\}$

$$B_K^{(n)} := \sum_{V \in \mathcal{V}_K} \Lambda_V^{(n)} \quad \text{und} \quad B^{(n)} := \max \left\{ B_{K_1}^{(n)}, B_{K_2}^{(n)} \right\}.$$

$(B^{(n)})_{n \in \mathbb{N}}$ ist eine stochastisch unabhängige und identische verteilte Zufallsfolge. Es gilt für alle Iterationsschritte $n \in \mathbb{N}$

$$Y_{V_1}^{(n+1)} = \max \left\{ Y_{V_1}^{(n)} + B_{K_1}^{(n)}, Y_{V_1}^{(n)} + B_{K_2}^{(n)}, U^{(n)} \right\} = \max \left\{ Y_{V_1}^{(n)} + B^{(n)}, U^{(n)} \right\}$$

mit

$$U^{(n)} = \max \left\{ \sum_{i=2}^3 \Lambda_{V_i}^{(n)}, \Lambda_{V_3}^{(n)}, \sum_{i=4}^5 \Lambda_{V_i}^{(n)}, \Lambda_{V_5}^{(n)}, 0 \right\}.$$

$U^{(n)}$ repräsentiert die von $Y_{V_1}^{(n)}$ unabhängigen Terme. Zusätzlich beschränken wir im Sinne der Stabilität die Verspätungsübertragung in Knoten V_1 , d.h.

$$Y_{V_1}^{(n+1)} = \min \left\{ c, \max \left\{ Y_{V_1}^{(n)} + B^{(n)}, U^{(n)} \right\} \right\}$$

für ein $c \in \mathbb{R}_{>0}$. Durch rekursives Einsetzen erhalten wir

$$\begin{aligned} Y_{V_1}^{(n+1)} &= \min \left\{ c, \max \left\{ \min \left\{ c + B^{(n)}, \max \left\{ [\dots], \right. \right. \right. \right. \\ &\quad \min \left\{ c + \sum_{l=2}^n B^{(l)}, \right. \\ &\quad \left. \left. \max \left\{ Y_{V_1}^{(1)} + \sum_{l=1}^n B^{(l)}, U^{(1)} + \sum_{l=2}^n B^{(l)} \right\} \right\}, [\dots] \right\} \right\}. \end{aligned}$$

Für obiges Beispiel nehmen wir wie in Satz 7.3.18 an, dass $\mathbf{E}[B_{K_1}^{(1)}] < 0$ und $\mathbf{E}[B_{K_2}^{(1)}] < 0$ gilt. Gilt außerdem $\mathbf{E}[B^{(1)}] < 0$, so existiert (\mathbf{P} -f.s.) ein $n_0 \in \mathbb{N}$, sodass für alle $n \geq n_0$ gilt

$$Y_{V_1}^{(1)} + \sum_{l=1}^n B^{(l)} < 0.$$

Damit wäre $\lim_{n \rightarrow \infty} Y_{V_1}^{(n)}$, wegen der Maximumsbildung mit 0, unabhängig von der Startverteilung. Gilt jedoch $\mathbf{E}[B^{(1)}] > 0$, so existiert (\mathbf{P} -f.s.) ein $n_0 \in \mathbb{N}$, sodass für alle $n \geq n_0$ gilt

$$Y_{V_1}^{(1)} + \sum_{l=1}^n B^{(l)} > c.$$

In diesem Falle wäre $\lim_{n \rightarrow \infty} Y_{V_1}^{(n)}$, wegen der Minimumsbildung mit c , ebenfalls unabhängig von der Startverteilung. Ein hierauf aufbauendes allgemeineres Resultat wird in dieser Arbeit nicht mehr angegeben.

7.4. Pseudo-topologische Sortierung

Es sei $\mathcal{S} \subset \mathcal{V}_{\mathcal{M}}$ die Menge der bereits sortierten Knoten der Kreisvereinigung \mathcal{M} , d.h. die Menge der Knoten, für die bereits eine Reihenfolge existiert. $\mathcal{U} = \mathcal{V}_{\mathcal{M}} \setminus \mathcal{S}$ ist entsprechend die Menge der unsortierten Knoten. Die Idee der im Folgenden vorgestellten Sortierung ist es, einer topologischen Sortierung möglichst nahe zu kommen. Ein Knoten soll erst dann einsortiert werden, wenn bereits *möglichst viele* seine Vorgänger zuvor einsortiert wurden. Um dieses Vorgehen zu präzisieren definieren wir mit

$$\alpha_V = \frac{\#\{V \in \text{Pre}(V) \mid V \in \mathcal{U}\}}{\#\text{Pre}(V)} \in [0, 1], \tag{7.21}$$

den Anteil noch nicht sortierter Vorgänger eines Knoten $V \in \mathcal{V}_{\mathcal{M}}$ im Anschlussgraphen. Suchen wir den als nächstes einzusortierenden Knoten, wählen wir denjenigen mit minimalem α_V . Je geringer die Werte für α_V dabei sind, desto mehr ähnelt die sich ergebene Sortierung einer topologischen Sortierung. Falls eine topologische Sortierung existiert, besitzen die gewählten Knoten bei diesem Vorgehen immer die Eigenschaft $\alpha_V = 0$. Algorithmus C.4.1 beschreibt die einzelnen Schritte genauer.

7.5. Empirische Resultate - Konvergenzgeschwindigkeit

Knoten	Linie	Stationen	Knoten	Linie	Stationen
V_1	L_1	S_1, S_8, S_2	V_2	L_2	S_2, S_9, S_3
V_3	L_3	S_3, S_{10}, S_4	V_4	L_4	S_4, S_{11}, S_5
V_5	L_5	S_5, S_{12}, S_6	V_6	L_6	S_6, S_{13}, S_1
V_7	L_7	S_2, S_7, S_1	V_8	L_8	S_3, S_7, S_1
V_9	L_9	S_4, S_7, S_1	V_{10}	L_{10}	S_5, S_7, S_1
V_{11}	L_{11}	S_{15}, S_{14}, S_4	V_{12}	L_{12}	S_4, S_{14}, S_{15}

Tabelle 7.2.: Übersicht der Beziehung zwischen Knoten und Linien

In den vorangegangenen Abschnitten dieses Kapitels wurde die Verspätungsfortpflanzung mittels eines theoretischen Ansatzes untersucht. Ursprünglich ging dem

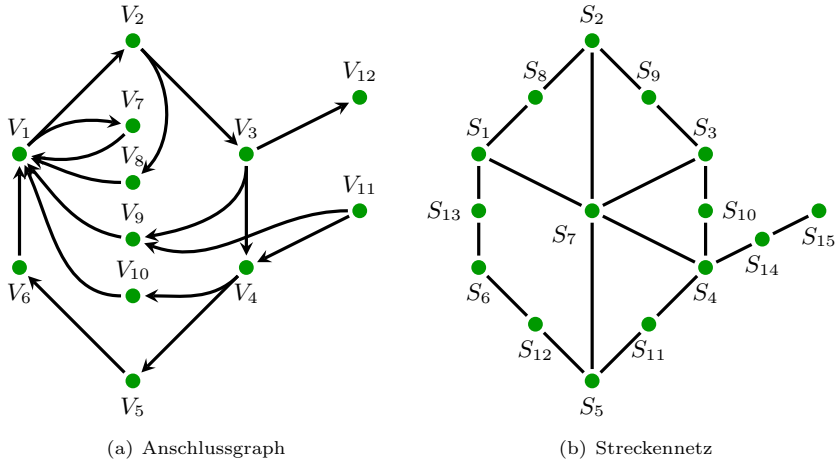


Abbildung 7.10.: Beispiel für eine Kreisvereinigung

jedoch eine empirische Analyse des Einflusses unterschiedlicher Knotensortierungen des Anschlussgraphen voraus. Der theoretische Modellierungsansatz sollte die empirisch gewonnenen Erkenntnisse bestätigen bzw. erklären.

Grundlage für die empirischen Untersuchungen war der Anschlussgraph in Abbildung 7.10(a). In Abbildung 7.10(b) ist das zugehörige Streckennetz zu sehen. Insgesamt 12 Linien verkehren innerhalb dieses Netzes. Der Zusammenhang zwischen Knoten des Anschlussgraphen und Linien im Streckennetz wird in Tabelle 7.2 angegeben. Sämtliche Kanten des Anschlussgraphen repräsentieren Umsteigeaktivitäten des zugehörigen Ereignis-Aktivitätsnetzwerkes. Für diese Umsteigeaktivitäten gilt jeweils Folgendes:

- Die Mindestumsteigezeit beträgt 4 Minuten.
- Der Zeitpuffer beträgt 1 Minute.

Alle weiteren Fahrplaninformationen werden in Tabelle 7.5 angegeben. Nur Fahraktivitäten wird eine (einheitliche) Quellverspätungsverteilung zugewiesen. Halteaktivitäten werden also nicht verspätet. Die verwendete Hyper-Erlangverteilung ist durch die folgenden Parameter ihrer Dichtefunktion gegeben (vgl. Abschnitt 5.4):

- $\alpha_0 = 0$, $M = 3$,
- $\alpha_1 = 0.060439562579102$, $r_1 = 1$, $\lambda_1 = 0.55853096523127$,

- $\alpha_2 = 0.0061855243197251$, $r_2 = 1$, $\lambda_2 = 0.55853096686656$,
- $\alpha_3 = 0.93337491310117$, $r_3 = 3$, $\lambda_3 = 0.84658212153167$.

Die zugehörige Verteilungsfunktion ist in Abbildung 6.6 (s. Seite 85) zu sehen. Der Erwartungswert beträgt 3.42685.

Das Ziel bei der Erzeugung dieses Anwendungsbeispiels war zum einen, bezüglich der Sortierung der Knoten des Anschlussgraphen verschiedene Ansätze vergleichen zu können. Zunächst betrachten wir die topologische Sortierung der Knoten auf Ebene des Kreisvereinigungsgraphens. Es existieren nur zwei singuläre Knoten: V_{11} und V_{12} . Die übrigen Knoten bilden die Kreisvereinigung \mathcal{M} . Die topologische Sortierung lautet $(V_{11}, \mathcal{M}, V_{12})$. Lediglich die Wahl der Sortierung für Knoten der Kreisvereinigung ist frei wählbar. Im Folgenden werden die im Rahmen der empirischen Untersuchungen verwendeten Sortierungen (s. beispielsweise Tabelle 7.3) kurz erläutert. Die pseudo-topologische Sortierung wurde bereits im vorherigen Abschnitt diskutiert.

- **Vorgänger:** Knoten V_1 ist bezüglich der Anzahl der Vorgängerknoten der wichtigste bzw. größte Knoten des Anschlussgraphen. Gemäß dieser Sortierung sollen in jedem Iterationsschritt zuerst die Verspätungen der Ereignisse der Vorgängerknoten bestimmt werden, bevor die Verspätungen der Ereignisse des Knotens V_1 bestimmt werden. Die übrigen Knoten werden der natürlichen Abfolge in den Kreisen folgend sortiert.
- **Größe:** Diese Sortierung verfolgt eine, im Vergleich zur vorherigen Sortierung, gegensätzliche Idee. Es wird in jedem Iterationsschritt mit der Bestimmung der Verspätungen der Ereignisse des Knotens V_1 begonnen. Die übrigen Knoten werden wieder der natürlichen Abfolge in den Kreisen folgend sortiert.
- **sprunghaft:** Statt der natürlichen Abfolge der Knoten in den Kreisen zu folgen, werden in dieser Sortierung einzelne Knoten übersprungen. Die Vorstellung ist, dass die direkte Verspätungsfortpflanzung so unterbrochen wird.
- **anti-topologisch:** Diese Sortierung stellt gewissermaßen das Gegenteil einer topologischen Sortierung dar. Eine direkte, d.h. natürliche, Verspätungsfortpflanzung wird weitestgehend unterbunden.

Ein weiteres Ziel war, für einen aussagekräftigen Vergleich der Knotensortierungen eine möglichst „gleichmäßige“ Verspätungsfortpflanzung innerhalb und zwischen den Knoten zu garantieren. In allen Knoten sollten daher identische Quellverspätungsverteilungen und Zeitpuffer verwendet werden. Ebenso sollten die Zeitpuffer für Umsteigeaktivitäten übereinstimmen. Für alle Sortierungen wurde ein einheitliches Kriterium für den Abbruch der Iteration verwendet. Für vorgegebenes

$\alpha \in (0, 1)$ ist

$$n := \min \left\{ m \in \mathbb{N} \setminus \{1\} \mid \forall V \in \mathcal{V}_{\mathcal{M}} : \left| \mathbf{E}[Y_V^{(m)}] - \mathbf{E}[Y_V^{(m-1)}] \right| \leq \mathbf{E}[Y_V^{(m)}] \cdot \alpha \right\}$$

die Anzahl der durchgeführten Iterationsschritte. Tabelle 7.3 zeigt die benötigte Anzahl an Iterationsschritten in Abhängigkeit von der Wahl der Sortierung und unterschiedlichen Werten des Parameters α . Je kleiner der Wert von α ist, desto kleiner sind die zugelassenen Veränderungen der Erwartungswerte im Vergleich zum Iterationsschritt zuvor. Daher ist es plausibel, dass die Anzahl der Iterationsschritte für alle Sortierungen größer wird, wenn der Wert des Parameters α kleiner gewählt wird. Für $\alpha = 0.1$ sieht die Situation für alle Sortierungen vergleichbar aus. Es werden zwischen 4 und 6 Iterationsschritte durchgeführt. Für $\alpha = 0.01$ weicht nur die pseudo-topologische Sortierung etwas von den anderen ab. Für $\alpha = 0.0001$ ist die Anzahl an Iterationsschritten die ersten drei Sortierungen vergleichbar. Die „sprunghafte“ Sortierung benötigt zwischen 6 und 7 Schritten mehr, die „anti-topologische“ benötigt sogar zwischen 11 und 12 Schritten mehr.

Das allein sagt über die Güte der Sortierungen noch nicht viel aus. Wichtig ist außerdem, wie nahe die Erwartungswerte zum Zeitpunkt des Abbruchs der Iteration an ihren (empirischen) Grenzwerten liegen. Wir betrachten dazu Tabelle 7.4. Für $\alpha = 0.0001$ sehen wir, dass die Erwartungswerte bei den drei betrachteten Sortierungen empirisch, trotz der zuvor diskutierten Unterschiede in den Sortierungen, in allen Knoten gegen die gleichen Grenzwerte streben. Dies geschieht allerdings mit unterschiedlicher Geschwindigkeit. Tabelle 7.3 entnehmen wir, dass sich die Sortierungen bezüglich der Anzahl der Iterationsschritte für vorgegebenes α unterscheiden. Obwohl bei Wahl der „sprunghaften“ bzw. der „anti-topologischen“ Sortierung im Falle $\alpha = 0.1$ ein bzw. zwei Iterationsschritte mehr durchgeführt werden als bei Wahl der „pseudo-topologischen“ Sortierung, ist die Konvergenz der Erwartungswerte bei letzterer in fast allen Knoten weiter vorangeschritten (s. Tabelle 7.4).

	Sortierung	Anz. Iterationsschritte		
		$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.0001$
pseudo-topologisch	$V_9, V_4, V_{10}, V_5, V_6, V_1, V_7, V_2, V_8, V_3$	4	8	22
Vorgänger	$V_6, V_{10}, V_9, V_8, V_7, V_1, V_2, V_3, V_4, V_5$	5	12	23
Größe	$V_1, V_2, V_3, V_4, V_5, V_6, V_{10}, V_9, V_8, V_7$	4	11	23
sprunghaft	$V_2, V_4, V_6, V_9, V_7, V_3, V_5, V_{10}, V_8, V_1$	5	13	29
anti-topologisch	$V_1, V_6, V_5, V_4, V_3, V_2, V_{10}, V_9, V_8, V_7$	6	12	34

Tabelle 7.3.: Benötigte Anzahl an Iterationsschritten

Dabei unterstellen wird, dass die in Tabelle 7.4 für $\alpha = 0.0001$ gemessenen Werte näherungsweise den wahren Grenzwerten entsprechen.

Als Ergebnis dieses Experiments können wir zusammenfassend festhalten, dass die Sortierungen die Geschwindigkeit der Konvergenz der Erwartungswerte beeinflussen. Die jeweiligen Grenzwerte stimmen empirisch überein. Des Weiteren sprechen die Ergebnisse hinsichtlich der Konvergenzgeschwindigkeit für die Wahl der im vorherigen Abschnitt vorgestellten „pseudo-topologischen“ Sortierung.

Liniestation	pseudo-top.	sprunghaft	anti-top.	pseudo-top.	sprunghaft	anti-top.
	$\alpha = 0.1$			$\alpha = 0.0001$		
(L_1, S_8)	2.85065	2.85976	2.81675	3.12977	3.12981	3.12962
(L_1, S_2)	2.14881	2.15848	2.12466	2.40107	2.40111	2.40095
(L_2, S_9)	1.81395	1.7627	1.79744	2.01018	2.01011	2.01009
(L_2, S_3)	1.46144	1.4221	1.44903	1.61677	1.61671	1.6167
(L_3, S_{10})	1.39572	1.36725	1.35626	1.54467	1.54463	1.54456
(L_3, S_4)	1.169	1.14689	1.13842	1.23751	1.23747	1.23741
(L_4, S_{11})	1.37364	1.35969	1.35631	1.44574	1.44573	1.44568
(L_4, S_5)	1.12501	1.11369	1.11096	1.19079	1.19078	1.19074
(L_5, S_{12})	1.15622	1.14774	1.12466	1.21221	1.2122	1.21213
(L_5, S_6)	0.984385	0.977643	0.959435	1.03239	1.03238	1.03233
(L_6, S_{13})	1.06956	1.05088	1.03433	1.10712	1.1071	1.10705
(L_6, S_1)	0.922273	0.907508	0.894593	0.95329	0.953267	0.953227
(L_7, S_7)	1.81395	1.76278	1.79744	2.01018	2.01011	2.01009
(L_7, S_1)	1.46144	1.4221	1.44903	1.61677	1.61671	1.6167
(L_8, S_7)	1.39572	1.36725	1.38692	1.54467	1.54463	1.54462
(L_8, S_1)	1.169	1.14689	1.16225	1.23751	1.23747	1.23746
(L_9, S_7)	1.37364	1.35969	1.38023	1.44574	1.44573	1.44573
(L_9, S_1)	1.12501	1.11369	1.13042	1.19079	1.19078	1.19078
(L_{10}, S_7)	1.15622	1.14774	1.14571	1.21221	1.2122	1.21217
(L_{10}, S_1)	0.984385	0.977643	0.976035	1.03239	1.03238	1.03236
(L_{11}, S_{14})	0.591177	0.591177	0.591177	0.591177	0.591177	0.591177
(L_{11}, S_4)	0.600676	0.600676	0.600676	0.600676	0.600676	0.600676
(L_{12}, S_{14})	1.20307	1.18689	1.18073	1.24878	1.24876	1.24871
(L_{12}, S_{15})	1.02774	1.01489	1.01002	1.0707	1.07068	1.07064

Tabelle 7.4.: Erwartete Ankunftsverspätungen in Abhängigkeit der genutzten Sortierung

7. Verspätungsfortpflanzung in kreisartigen Strukturen

Linienstation	Ankunft	Abfahrt	Mindesthaltezeit	Mindestfahrzeit	Puffer Station	Puffer Strecke
(L ₁ , S ₁)	-	0	-	7	-	4
(L ₁ , S ₈)	11	13	2	7	0	5
(L ₁ , S ₂)	25	-	-	-	-	-
(L ₂ , S ₂)	-	30	-	7	-	4
(L ₂ , S ₉)	41	43	2	7	0	5
(L ₂ , S ₃)	55	-	-	-	-	-
(L ₃ , S ₃)	-	0	-	7	-	4
(L ₃ , S ₁₀)	11	13	2	7	0	5
(L ₃ , S ₄)	25	-	-	-	-	-
(L ₄ , S ₄)	-	30	-	7	-	4
(L ₄ , S ₁₁)	41	43	2	7	0	5
(L ₄ , S ₅)	55	-	-	-	-	-
(L ₅ , S ₅)	-	0	-	7	-	4
(L ₅ , S ₁₂)	11	13	2	7	0	5
(L ₅ , S ₆)	25	-	-	-	-	-
(L ₆ , S ₆)	-	30	-	7	-	4
(L ₆ , S ₁₃)	41	43	2	7	0	5
(L ₆ , S ₁)	55	-	-	-	-	-
(L ₇ , S ₂)	-	30	-	7	-	4
(L ₇ , S ₇)	41	43	2	7	0	5
(L ₇ , S ₁)	55	-	-	-	-	-
(L ₈ , S ₃)	-	0	-	22	-	4
(L ₈ , S ₇)	26	28	2	22	0	5
(L ₈ , S ₁)	55	-	-	-	-	-
(L ₉ , S ₄)	-	30	-	7	-	4
(L ₉ , S ₇)	41	43	2	7	0	5
(L ₉ , S ₁)	55	-	-	-	-	-
(L ₁₀ , S ₅)	-	0	-	22	-	4
(L ₁₀ , S ₇)	26	28	2	22	0	5
(L ₁₀ , S ₁)	55	-	-	-	-	-
(L ₁₁ , S ₁₅)	-	0	-	7	-	4
(L ₁₁ , S ₁₄)	11	13	2	7	0	5
(L ₁₁ , S ₄)	25	-	-	-	-	-
(L ₁₂ , S ₄)	-	30	-	7	-	4
(L ₁₂ , S ₁₄)	41	43	2	7	0	5
(L ₁₂ , S ₁₅)	55	-	-	-	-	-

Tabelle 7.5.: Fahrplan, Zeitpuffer und Mindestzeiten

8. Ereignisgesteuerte stochastische Simulation

Übersicht

8.1. Elemente der Simulation	147
8.2. Der Simulationsalgorithmus	151

Die verwendete ereignisgesteuerte stochastische Simulation basiert auf einer studentischen Arbeit von FLORIAN ARNOLD. Das Verfahren sowie die Implementation wurden im Rahmen dieser Arbeit wesentlich weiterentwickelt. Ziel war es, neben einem Performance-Vergleich zwischen Simulation und analytischer Berechnung auch eine Möglichkeit zur Validierung der Ergebnisse der analytischen Berechnung zu erhalten. Die Kernidee der Arbeit von FLORIAN ARNOLD war die spezielle Berücksichtigung von Zubringer-Abbringer-Beziehungen zwischen Zügen verschiedener Linien. Es musste ein möglichst effizienter Weg gefunden werden, um Abbringern in den jeweiligen Stationen die Information zu geben, ob und auf welche Zubringer sie warten müssen.

Die wesentlichen Phasen der Simulation sind in Abbildung 8.1 zu sehen. In der ersten Phase, wird anschaulich das Streckennetz mit Zügen der jeweiligen Linien gefüllt. Tatsächlich werden (mindestens) so lange verspätungsfreie Ereignisse erzeugt, bis jede Linie mit einem Zug an der Endstation angekommen ist. Der weitere Ablauf findet in gewissen Zyklen statt. Jedem dieser Zyklen geht eine Phase voran, in der Verspätungen erzeugt aber noch nicht ausgewertet werden. Genau genommen werden Quellverspätungen simuliert. Deren Fortpflanzung wird durch verspätete Ereignisse registriert. In dieser Phase sind wir noch nicht an der Größe der Verspätungen interessiert. Vielmehr soll zunächst nur sichergestellt werden, dass die simulierten Quellverspätungen bis zu den relevanten Stellen (Ereignissen) im Netz fortgepflanzt wurden, bevor wir die Verspätung an diesen Stellen messen. Wir warten also den Aufbau der Verspätungen ab. Ohne diese Entwicklungsphase für Verspätungen beobachteten wir zunächst tendenziell zu geringe Verspätungen.

Die nun folgende Phase stellt den Ablauf eines der erwähnten Zyklen dar. Er unterscheidet sich von der vorangegangenen Phase nur darin, dass wir die für uns (hinsichtlich unserer Kostenfunktion) relevanten Verspätungen notieren. Grob gesagt wird bei jeder Ankunft eines Zuges (für jedes Ankunftsereignis) die entsprechende Ankunftsverspätung festgehalten. Zusätzlich wird entschieden, ob Passagiere, die in diesem Zug saßen, einen ihrer Anschlüsse verpassen. Die dadurch resultierende

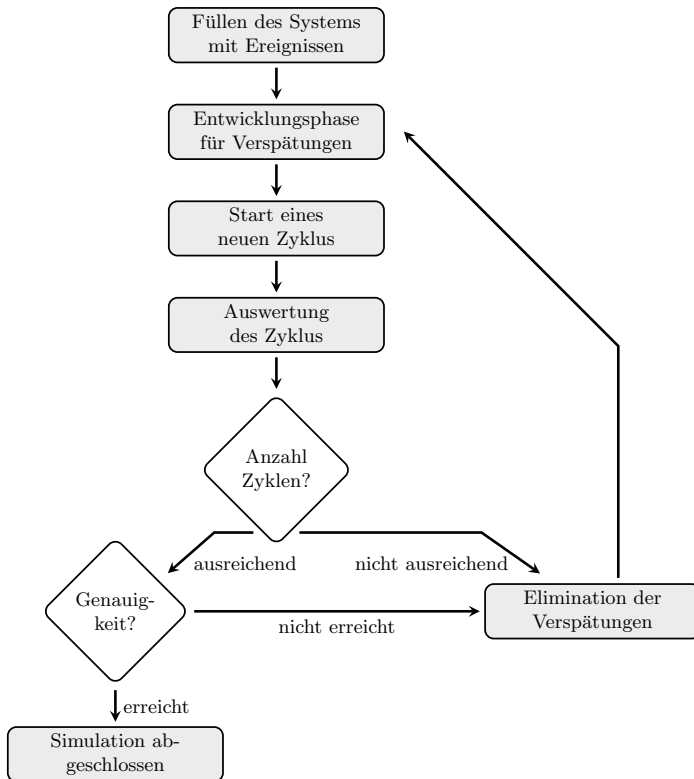


Abbildung 8.1.: Grobablauf der ereignisgesteuerten stochastischen Simulation

Verspätung wird ebenfalls festgehalten. Eine präzisere Beschreibung der Vorgehensweise folgt zu einem späteren Zeitpunkt. Die Länge eines Zyklus wird in gewissem Maße dynamisch, abhängig vom Netz und den verkehrenden Linien, festgelegt. Am Ende eines jeden Zyklus erfolgt die Erstellung eines Stichprobenwertes. Die zuvor notierten Verspätungen werden hierfür gemittelt. Auch hier sei gesagt, dass das genaue Vorgehen im Folgenden noch detaillierter dargestellt wird.

In Abbildung 8.1 ist zu sehen, dass im nächsten Schritt geprüft wird, ob die Mindestanzahl der bisher ausgewerteten Zyklen erreicht ist. Diese ist gleichbedeutend mit der Mindestanzahl der erzeugten Stichprobenwerte. Hintergrund ist, dass die Entscheidung, ob die geforderte Signifikanz der Stichprobe erfüllt ist, mit Hilfe eines Konfidenzintervalls getroffen wird. Zu Beginn wird der Mittelwert der Stich-

probenwerte stark schwanken. Mit der Forderung einer Mindestgröße der Stichprobe vermeiden wir ein zu frühes Abbrechen der Simulation. Des Weiteren können wir so die Normalverteilungsannahme der Summe der Stichprobenwerte rechtfertigen. Ist die geforderte Signifikanz erreicht, so wird die Simulation abgeschlossen und das Ergebnis (Kostenfunktionswert) zurückgegeben.

Ist die geforderte Signifikanz jedoch noch nicht erreicht, so muss ein weiterer Stichprobenwert erzeugt bzw. ein neuer Zyklus gestartet werden. Damit die im vorherigen Zyklus gemessenen Verspätungen keinen Einfluss auf die im folgenden Zyklus zu messenden Verspätungen haben, führen wir eine Art Regeneration der Simulation durch, in dem wir die zu diesem Zeitpunkt im System bzw. den Ereignissen enthaltenen Verspätungen entfernen. Es stehen im Anschluss dieselben Ereignisse in der Ereignisliste. Allerdings sind sie dann verspätungsfrei und haben eventuell ihre Position in der (sortierten) Ereignisliste verändert. Anschaulich betrachten wir immer noch die gleichen Züge wie zum Ende des vorangegangenen Zyklus, allerdings sind diese jetzt wieder fahrplanmäßig unterwegs. Bevor wir erneut Verspätungen messen, d.h. einen Zyklus starten, werden zunächst, wie oben bereits beschrieben, in einer vorangehenden Phase Verspätungen wieder aufgebaut.

8.1. Elemente der Simulation

Den zentralen Bestandteil der Simulation bilden natürlich die Ereignisse. Wir verwenden drei verschiedene Arten von Ereignissen:

1. **Abfahrt / Ankunft eines Zuges:** Abfahrten und Ankünfte von Zügen in Stationen ihrer jeweiligen Linien werden durch entsprechende Abfahrts- bzw. Ankunftsereignisse repräsentiert.
2. **Start eines Zuges:** Der Start eines Zuges stellt einen Spezialfall des Abfahrtsereignisses dar. Dieses Ereignis beschreibt die jeweilige Abfahrt eines Zuges in der Startstation der zugehörigen Linie.

Definition 8.1.1 (Ereignis).

Ereignisse sind vom Typ Start, Abfahrt oder Ankunft und repräsentieren die Züge in der Simulation. Als zusätzliche Information enthalten sie

- (i) den Zeitpunkt des Ereignisses,
- (ii) die Nummer der Station,
- (iii) die Nummer der Linie,
- (iv) die Nummer des Zuges der Linie sowie
- (v) die Verspätung des Zuges.

Die Ereignisse werden in einer, bezüglich ihrer jeweiligen Zeitpunkte, sortierten Liste gehalten. Aus dieser Liste wird immer das erste Element entnommen und verarbeitet. Genauer wird dieses Vorgehen in Algorithmus C.1.6 dargestellt. Zunächst stellen wir aber zwei weitere zentrale Bestandteile der Simulation vor. In der Simulation gibt es für jede Linie einen ersten Zug, der gestartet wird. Wir weisen diesem Zug die Nummer 0 zu. Die Zeit ist in Minuten gegeben und beginnt zum Zeitpunkt 0. Wir verwenden im Folgenden eine besondere Form des Fahrplans π aus Definition 3.2.2. Der Fahrplan π^* enthält die Ankunfts- und Abfahrtszeiten des jeweils ersten Zuges einer jeden Linie in absoluter Zeit, d.h. nicht modulo der Taktzeit τ . Bevor die eigentliche Simulation starten kann, müssen zunächst die unten definierten relativen Zugbeziehungen erstellt werden. Es müssen also für jede Umsteigebeziehung, relativ zum jeweiligen Zug der Abbringerlinie, die Zugnummern der Zubringerzüge bestimmt werden. Algorithmus 8.1.1 beschreibt die Vorgehensweise.

Definition 8.1.2 (Relative Zugbeziehung).

Für eine Umsteigebeziehung der Linien $L_i, L_j \in \mathcal{L}$, $L_i \neq L_j$, in Station $S \in \mathcal{S}$ gibt die **relative Zugbeziehung** $z_{S,L_i,L_j} \in \mathbb{Z}$ die Abweichung der Zugnummer $z_i \in \mathbb{N}$ der Abbringerlinie L_i von der Zugnummer $z_j \in \mathbb{N}$ der Abbringerlinie L_j an. Es gilt also

$$z_{S,L_i,L_j} = z_i - z_j .$$

Input : Fahrplan π^* und Taktzeit τ

// Mindestumsteigezeit $c(S, L_i, L_j)$

1 $s = \pi^*(ab, S, L_j) - \pi^*(an, S, L_i) - c(S, L_i, L_j)$;

2 **return** $\lfloor \frac{s}{\tau} \rfloor$;

Algorithmus 8.1.1 : Erzeugung der relativen Zugbeziehungen

Definition 8.1.3 (Zubringer).

Zubringer stellen in der Simulation Züge dar. Für jede Umsteigebeziehung in einer Station existieren Zubringer-Abbringer-Paare. Da wir ein periodisches System betrachten, existiert für jede Periode (Takt) genau ein solches Paar. Als Kennzeichnung enthält der jeweilige Zubringer die Information über

- (i) die Nummer seiner Linie,
- (ii) die zu seinem Abbringer relative Zugbeziehung.

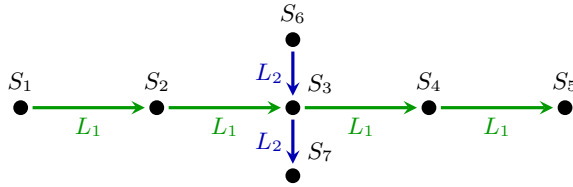


Abbildung 8.2.

Als nächstes wollen wir die sogenannten *tatsächlichen Ankünfte* erläutern. Sie setzen die zuvor erwähnte Idee aus der Arbeit von FLORIAN ARNOLD um und ermöglichen in den meisten Fällen, die Frage, ob ein Anschluss erreicht wird, korrekt zu beantworten.

Definition 8.1.4 (Tatsächliche Ankunft).

Eine *tatsächliche Ankunft* gehört immer zu einer bestimmten Linie. Sie gibt Auskunft darüber, zur welcher Zeit und in welcher Station ein bestimmter Zug dieser Linie ankommen wird oder angekommen ist. Sie enthält daher als wesentliche Information

- (i) die Nummer des Zuges,
- (ii) die Nummer der Station,
- (iii) die Ankunftszeit in dieser Station.

Für jede Linie gibt es während der Simulation eine Liste ihrer aktuellen tatsächlichen Ankünfte. Sie werden in der Simulation während der Erstellung eines Ankunftsereignisses erzeugt. Zunächst hat diese tatsächliche Ankunft, im Vergleich zum Ankunftsereignis, keinen Mehrwert bezüglich der enthaltenen Information. Grundsätzlich werden sie für etwaige Abbringer, also Abfahrtsereignisse, in der jeweiligen Station benötigt. Es gilt hier, verschiedene Situationen zu unterscheiden. Veranschaulichen können wir uns diese mittels Abbildung 8.2. Wir betrachten im Folgenden immer die Anschlussbeziehung zwischen Abbringerzug $z_2 \in \mathbb{Z}$ von Linie $L_2 \in \mathfrak{L}$ und Zubringerzug $z_1 \in \mathbb{Z}$ von Linie $L_1 \in \mathfrak{L}$ in Station $S_3 \in \mathfrak{S}$. Die Entscheidung, ob der Anschluss gehalten werden kann, wird immer zum Zeitpunkt der Ankunft des Abbringers getroffen, d.h. wenn das aktuelle Ereignis gerade die Ankunft des Zuges z_2 in S_3 repräsentiert.

1. Die Abfahrt des Zuges z_1 der Linie L_1 in der vorherigen Station S_2 hat zuvor bereits stattgefunden. Der Zug z_1 befindet sich noch auf dem Weg zur Station S_3 . In diesem Fall bräuchte es eigentlich keine tatsächliche Ankunft, um die Information zu bekommen, wann Zubringerzug z_1 ankommt. Alternativ

könnte man in der Ereignisliste das entsprechende Ankunftsereignis suchen. Trotzdem wird die Einführung der tatsächlichen Ankünfte hier in der Regel einen Vorteil haben. Denn hier suchen wir in der im Allgemeinen wesentlich kürzeren Liste der aktuellen Ankünfte der Linie L_1 . In der Ereignisliste stehen stattdessen Ereignisse verschiedenen Typs und verschiedener Linien.

- Die Abfahrt des Zuges z_1 der Linie L_1 in der vorherigen Station S_2 hat zuvor bereits stattgefunden. Der Zug z_1 ist allerdings auch an Station S_3 bereits wieder abgefahren. Es wäre beispielsweise denkbar, dass Zug z_2 sehr verspätet in Station S_3 ankommt. In dieser Situation ist die Ereignisliste als Informationsquelle unbrauchbar. Das Ankunftsereignis des Zuges z_1 in Station S_3 wurde bereits abgearbeitet und im Anschluss aus der Ereignisliste entfernt. Ob die Information in der Liste der tatsächlichen Ankünfte von Linie L_1 enthalten ist, hängt davon ab, zu welchem Zeitpunkt man die Elemente dieser Liste wieder entfernt. Je später man dies macht, desto größer bzw. länger wird die Liste werden. Je stärker das Verhalten der Züge vom Fahrplan abweicht, desto größer wird die benötigte Historie der Liste sein müssen. Man benötigt dann die schon stattgefundenen Ankünfte vorangegangener Takte. Allerdings nimmt auch der Speicherplatzbedarf sowie die Zugriffszeit bei der Suche der jeweiligen tatsächlichen Ankunft in der Liste zu. Es muss also ein Kompromiss zwischen Laufzeit und der Möglichkeit, stark vom Fahrplan abweichendes Verhalten simulieren zu können, gewählt werden.

Im Rahmen der Optimierung werden gerade solche Fahrpläne nicht lange betrachtet, die zu stark vom Fahrplan abweichendem Verhalten des Systems führen (s. Abschnitt 9.2). Entweder man setzt die Simulation erst ein, wenn diese Fahrpläne bereits aussortiert wurden, oder man findet eine gesonderte Bewertung. Dies setzt voraus, dass der Algorithmus erkennt, wenn ein Fahrplan nicht korrekt simuliert werden kann.

- Die Abfahrt des Zuges z_1 der Linie L_1 in der vorherigen Station S_2 hat noch nicht stattgefunden. Dies kann wiederum mehrere Gründe haben. Verschärft wird diese Situation grundsätzlich dadurch, dass die Simulation bereits bei Ankunft des Abbringerzuges (hier also z_2) prüft, wo sich der Zubringerzug befindet bzw. in welcher Station er mit welcher Verspätung als nächstes eintrifft. Hat Abbringerzug z_2 fahrplanmäßig eine lange Aufenthaltszeit in Station S_3 , d.h. es ist in dieser Station ein großer zeitlicher Puffer vorgesehen, so kann Zubringerzug z_1 bei Ankunft von Zug z_2 in S_3 fahrplanmäßig noch weit entfernt sein, ohne die Anschlussbeziehung zwischen diesen beiden Zügen zu verletzen.
 - Eine Möglichkeit wäre, dass die Strecke $(S_2, S_3) \in \mathfrak{B}$ sehr kurz ist. In ICE-Netzen, wird dies eher selten der Fall sein.

- Die Situation tritt auch häufiger auf, wenn bereits die erste angefahrene Station einer Zubringerlinie eine Umsteigestation ist. In Abbildung 8.2 wäre dies also der Fall, wenn es Station S_1 nicht gäbe und stattdessen Station S_2 Startstation für Linie L_1 wäre. Ist die Strecke $(S_2, S_3) \in \mathfrak{B}$ wieder hinreichend kurz, so ist Zug z_1 bei Ankunft von Zug z_2 in Station S_3 noch nicht in S_2 losgefahren. Im Unterschied zum vorherigen Fall, existiert z_2 zu diesem Zeitpunkt auch noch gar nicht im System. Entschärft wird das Problem im Falle von im Vergleich zu den Fahrzeiten relativ kleinen Taktzeiten.

Es ist in allen Fällen nicht möglich, die genaue (simulierte) Verspätung des Zubringerzuges z_1 vorab zu bestimmen. Erst wenn über die Bearbeitung der Ereignisse das Ankunftsereignis für z_1 in S_3 erstellt wurde, steht die Verspätung schließlich fest. Eine erste Idee war, die bis zum Zeitpunkt der Entscheidungsfindung, ob der Anschluss erreicht wird, für den Zubringer z_1 registrierte Verspätung auch für die Ankunft in S_3 anzunehmen. Diese Art der Verspätungsprognose ist in vielen Fällen sicher sehr grob. Eine andere Möglichkeit ist, die Information über die Erwartungswerte der noch nicht simulierten Quellverspätungen sowie die noch folgenden zeitlichen Puffer zu nutzen.

8.2. Der Simulationsalgorithmus

Algorithmus C.1.6 (s. Anhang) zeigt den bereits in Abbildung 8.1 anschaulich dargestellten Ablauf der Simulation in Pseudocode. Ein zentraler Punkt ist in Zeile 16 der Test, ob die Simulationszeit soweit vorangeschritten ist, dass ein neuer *Messpunkt* erreicht wurde. Das Ergebnis dieser Abfrage hängt, insbesondere im Falle des ersten Messpunktes, von mehreren Faktoren ab. Im Falle des ersten Messpunktes, muss zunächst mindestens so viel Zeit vergangen sein, dass die Phase des Füllens des Systems (s. Abbildung 8.1) abgeschlossen ist. Wichtig ist, dass sich „Zeit“ hier auf die Zeitpunkte der Ereignisse bezieht. Im Anschluss muss zusätzlich mindestens die Zeit vergangen sein, die für die Entwicklung der Verspätungen vorgesehen ist. Erst dann beginnt der Zeitraum, in dem die Verspätungen tatsächlich gemessen werden. Wie lange diese Messung erfolgt, hängt von der vorgegebenen Länge eines Messzyklus ab. Für alle weiteren Messpunkte gilt dann, dass sie dann erreicht werden, wenn nach dem vorherigen Messpunkt die für die Entwicklungsphase der Verspätungen benötigte Zeit zuzüglich einer Zykluslänge vergangen ist.

Wurde noch kein neuer Messpunkt erreicht, so wird die Simulation direkt mit der Bearbeitung des aktuellen Ereignisses fortgesetzt. Je nach Ereignistyp wird natürlich unterschiedlich verfahren. Im Falle eines *Startereignisses* müssen zwei Aufgaben erledigt werden. Zunächst muss sichergestellt werden, dass einen Takt später erneut ein Zug dieser Linie startet. Es wird also ein neues Startereignis in

derselben Station erzeugt, wie in Zeile 1 des Algorithmus C.1.1 zu sehen ist. Dieses Startereignis ist verspätungsfrei. Die zugehörige Zugnummer wird, im Vergleich zur Zugnummer des vorherigen (übergebenen) Startereignisses, um 1 erhöht. Als nächstes muss die tatsächliche Abfahrt des zum aktuellen Startereignis gehörenden Zuges bestimmt werden, damit ein Abfahrtsereignis erzeugt werden kann. Wir gehen davon aus, dass Züge, wenn möglich, pünktlich starten. Diese Annahme kann auch leicht verworfen werden. Sie ist aber Grundlage für das Vorgehen in Zeile 6. So kann die tatsächliche Abfahrt des Zuges nur durch eventuelle Zubringer verzögert werden. Der zum Startereignis gehörende Zug erfährt selbst keine Quellverspätung in seiner Startstation. Wie hoch die übertragene Verspätung ist, hängt neben der Verspätung der Zubringer auch von der Anschlusssicherungsstrategie bzw. der Anschlusssicherungskonstanten ab. Letztere gibt an, wie lange auf verspätete Zubringer maximal gewartet werden soll, um den Anschluss zu ermöglichen (s. Abschnitt 6.4). In Zeile 8 wird ein neues Abfahrtsereignis angelegt. Wie zuvor bei der Erstellung des neuen Startereignisses, wird die Zugnummer bzgl. der Zugnummer des Ereignisses E um 1 erhöht. Die Stationen stimmen natürlich noch überein.

Ein wenig aufwendiger gestaltet sich die Bearbeitung eines Abfahrtsereignisses (s. Algorithmus C.1.2). Neben der Erzeugung eines neuen Ankunftsereignisses muss die Liste der tatsächlichen Ankünfte überarbeitet werden. Bevor in Zeile 4 ein neues Ankunftsereignis angelegt wird, muss zuvor die Quellverspätung für die Fahrt auf der vor der Ankunft zurückzulegenden Strecke simuliert werden. Im Anschluss kann berechnet werden, mit welcher Verspätung der Zug in der nächsten Station ankommt. Weiter oben sind wir bereits auf die Problematik eingegangen, dass wir in verschiedenen Situationen die tatsächlichen Ankünfte eines größeren Zeitraums benötigen. Je größer dieser Zeitraum ist, desto größer ist auch die Komplexität bzgl. Speicherbedarf und Zugriff auf die Liste der tatsächlichen Ankünfte. Wir suchen in der Liste den jeweiligen Zug bezüglich der zugehörigen Zugnummer. Kommt eine Zugnummer öfters vor, muss zusätzlich die Station der tatsächlichen Ankunft abgefragt werden. Wir haben uns für folgendes Vorgehen entschieden. Haben wir ein Abfahrtsereignis wie oben bearbeitet, so löschen wir die tatsächliche Ankunft des zugehörigen Zuges in der vorherigen Station. Dieser Schritt ist in Zeile 6 des Algorithmus C.1.2 zu finden. Die tatsächliche Ankunft dieses Zuges in der zum Abfahrtsereignis gehörenden Station bleibt erhalten. Zusätzlich wird für den Zug natürlich eine neue tatsächliche Ankunft angelegt, welche sich die benötigten Informationen mit dem zuvor erzeugten Ankunftsereignis teilt (s. Zeile 8).

Die Bearbeitung von Ankunftsereignissen erfordert den meisten Aufwand (siehe Algorithmus C.1.3). Hier werden Entscheidungen bzgl. der Erreichbarkeit möglicher Anschlüsse getroffen. Außerdem werden Ankunftsverspätungen und durch verpasste Anschlüsse verursachte Verspätungen registriert. Hat der jeweilige Zug seine Endstation noch nicht erreicht, so muss ein neues Abfahrtsereignis in derselben Station erzeugt werden. Dafür ist insbesondere die Simulation der Quellverspätung, die der

Zug beim Halt in dieser Station erfährt, notwendig. In Zeile 6 wird damit begonnen, die tatsächliche, eventuell vom Fahrplan abweichende, Aufenthaltszeit des Zuges in der Station zu bestimmen. Im Anschluss lässt sich so mit Kenntnis der aktuellen Ankunftszeit die Abfahrtszeit des Zuges sowie seine Verspätung bestimmen. Es kann ein neues Abfahrtsereignis angelegt werden. Die tatsächliche Aufenthaltszeit hängt offensichtlich unter anderem vom Abbringerzug selbst ab. Der Zug darf frühestens zum im Fahrplan vorgesehenen Zeitpunkt abfahren. Die Abfahrt findet verspätet statt, wenn dieser Zeitpunkt durch Einhaltung der Mindesthaltezeit und erfahrener Quellverspätung überschritten wird (s. Zeile 7). Ebenso kann die Abfahrt durch Zubringerzüge verspätet werden. Die hängt insbesondere von der Größe der Anschlusssicherungskonstanten ab. In Zeile 8 wird der Zeitpunkt festgelegt, zu dem Passagiere aus Zubringerzügen spätestens zum Einstieg in den Anschlusszug bereitstehen müssen. Kann dieser Zeitpunkt nicht eingehalten werden, so wird auf diese Passagiere nicht gewartet. Die Abfahrt des Abbringerzuges wird von diesem Zubringer dann nicht beeinflusst. Wie dies im einzelnen aussieht, wird in Zeile 9 bzw. Algorithmus C.1.4 betrachtet. Dort wird insbesondere die durch die, eventuell verspäteten, Ankünfte der Zubringerzüge bedingte Aufenthaltszeit des Abbringerzuges berechnet. Die endgültige Aufenthaltszeit erhält man durch Maximumsbildung (siehe Zeile 10). Im Anschluss wird, wie schon oben beschrieben, ein neues Abfahrtsereignis erzeugt.

Zentraler Bestandteil der Simulation ist somit Algorithmus C.1.4. Es wird unter anderem das Ankunftsereignis des Abbringers übergeben. Hier enthaltene Informationen sind unter anderem Zugnummer und Linie des Abbringers. Für jeden Zubringer (siehe Zeile 2) wird dessen aktuelle Liste der tatsächlichen Ankünfte durchsucht (siehe Zeile 1). Im Optimalfall finden wir in dieser Liste eine Ankunft, die in der aktuellen Station des Abbringers stattfindet. Mit Kenntnis der Ankunftszeit des Zuges sowie der mindestens benötigten Umsteigezeit der Passagiere können wir in Zeile 6 entscheiden, ob der Anschluss noch erreicht wird. Verpassen die Passagiere ihren Anschluss, so wird die zugehörige Verspätung für die Linienstation des Zubringers vermerkt. Andernfalls wird registriert, ob und wie lange der Abbringer in der Anschlussstation verweilen muss, um einen Umstieg der Passagiere zu gewährleisten (siehe Zeile 9). Wird in Zeile 6 festgestellt, dass in der Anschlussstation keine tatsächliche Ankunft des gerade betrachteten Zubringers vermerkt ist, so kann das zunächst verschiedene Gründe haben. Der Zug kann an der Anschlussstation bereits wieder abgefahren und seine tatsächliche Ankunft in dieser Station aus der Liste gelöscht worden sein. Dies wird allerdings nicht vorkommen, wenn die tatsächlichen Ankünfte hinreichend lange in der Liste verwahrt bleiben. Auch kann es passieren, dass der Zubringerzug an der vorherigen Station noch nicht abgefahren und daher noch keine tatsächliche Ankunft in der Anschlussstation erstellt worden ist. Falls die vorherige Station die Startstation der Zubringerlinie ist, so ist es wie bereits erwähnt sogar möglich, dass der Zubringerzug noch gar nicht im System existiert.

In beiden Fällen ist also noch nicht bekannt, mit welcher Verspätung der Zubringerzug in der Anschlussstation ankommen wird. Die Entscheidung, ob der Anschluss erreicht wird, kann nicht getroffen werden. Wir können die Ankunftsverspätung allerdings in der Regel *hochrechnen* (siehe Zeile 7). Entweder wir simulieren die Verspätungen vorab, d.h. bevor die Reihenfolge der Ereignisse in der Ereignisliste dies vorsieht, oder wir verwenden die Erwartungswerte der Verspätungen. In beiden Fällen würden die Verspätungen mit den zeitlichen Puffern verrechnet.

Wurde ein neuer Messpunkt erreicht, d.h. der Zeitpunkt des aktuellen Ereignisses hat den hierfür berechneten Zeitpunkt erreicht oder überschritten, so wird aus den im gerade abgeschlossenen Zyklus registrierten Verspätungen ein Stichprobenwert erzeugt. Wurden die Mindestanzahl der Stichprobenwerte erzeugt, wird auf Grundlage von Konfidenzintervallen entschieden, ob die Schätzung für den Kostenfunktionswert hinreichend genau ist. Mehr Details zum Kostenfunktionswert folgen in Kapitel 9.

9. Anwendung der Methoden im Rahmen der Optimierung

Übersicht

9.1. Generierung der Reisendenströme	155
9.2. Bewertung eines Fahrplans	156
9.2.1. Mittlere fahrplanmäßige Reisezeit	157
9.2.2. Mittlere Ankunftsverspätung	157
9.3. Fahrplanoptimierung in HiTT	158

In diesem Kapitel beschäftigen wir uns damit, in welchem Rahmen die in den vorherigen Kapiteln beschriebene Verspätungsfortpflanzung Anwendung in der Optimierung von Fahrplänen findet. Dabei sollen die beiden verschiedenen Ansätze, d.h. Verknüpfung bzw. Modifikation von Verteilungsfunktionen oder eine ereignisgesteuerte stochastische Simulation, anhand eines Beispielnetzes miteinander verglichen werden.

Es werden Bewertungskriterien für Fahrpläne eingeführt, die berücksichtigen, welche Wege die Reisenden auf ihrem Weg von der jeweiligen Startstation zur Zielstation nutzen. Die Erzeugung solcher Reisendenströme wird im ersten Abschnitt dieses Kapitels erläutert.

9.1. Generierung der Reisendenströme

Zur Generierung der Reisendenströme suchen wir im erweiterten Ereignis-Aktivitätsnetzwerk (s. Definition 3.2.5) für jede (mögliche) Abfahrtsstation *alle* kürzesten Wege zu allen erreichbaren Ankunftsstationen (kürzeste Wege zwischen Metaknoten). Es kann durchaus mehr als einen kürzesten Weg zwischen zwei Metaknoten geben. Dies ist beispielsweise der Fall, wenn Linien eine Zeit lang parallel zueinander verlaufen. Für die Bestimmung kürzester Wege müssen den Kanten bzw. Aktivitäten Gewichte zugeordnet werden. Ziel der Suche kürzester Wege ist es, die vermutlichen Reisendenströme in einem Netzwerk sichtbar zu machen. Die Gewichte müssen daher so bestimmt werden, dass der kürzeste Weg zwischen zwei Stationen eine plausible Route für die Fahrt der Reisenden zu ihrer Zielstation darstellt. Wir wählen die Gewichte wie folgt:

- $a \in \mathfrak{A}^{drive}$: Fahraktivitäten bekommen ihre Mindestdauer zugewiesen.
- $a \in \mathfrak{A}^{change}$: Umsteigeaktivitäten bekommen das Maximum des Gewichts aller Fahraktivitäten zugewiesen.
- $a \in \mathfrak{A}^{Ein} \cup \mathfrak{A}^{Aus}$: Ein- und Aussteigeaktivitäten bekommen das mit einer Konstanten multiplizierte Gewicht der Umsteigeaktivitäten zugewiesen.
- $a \in \mathfrak{A}^{stop}$: Halteaktivitäten bekommen alle das Gewicht 0 zugewiesen.

Das im Vergleich zu den Fahraktivitäten höhere Gewicht der Umsteigeaktivitäten lässt sich damit erklären, dass Wege bevorzugt werden sollen, die möglichst wenige Umstiege beinhalten. Damit wird das Entscheidungsverhalten von Reisenden bei Wahl ihrer Verbindung nachgebildet. Auf dem Weg zu einer Zielstation benötigt ein Reisender neben Aktivitäten $a \in \mathfrak{A}^{drive} \cup \mathfrak{A}^{change} \cup \mathfrak{A}^{stop}$ genau eine Einsteige- sowie Aussteigeaktivität. Deshalb dürfen keine kürzesten Wege gefunden werden, die mehr als zwei Metaknoten beinhalten. Das hieße, dass die Reisenden in einem Metaknoten, der weder Start- noch Zielstation ist, aus- und wieder einsteigen. Erreicht werden kann dies, indem Einsteige- und Aussteigeaktivitäten ein entsprechend hohes Gewicht zugewiesen bekommen.

Für die Suche nach kürzesten Wegen wird mit Algorithmus C.3.4 ein leicht modifizierter DIJKSTRA-Algorithmus (s. [Dij59]) verwendet. Wesentlicher Unterschied ist der Block von Zeile 24 bis Zeile 29. Hierdurch wird die Speicherung von Wegen ermöglicht, welche dieselbe Länge besitzen wie der aktuell kürzeste Weg. Wird beim weiteren Durchlaufen der Schleife mit Beginn in Zeile 10 ein kürzerer Weg zu der betrachteten Ecke gefunden, werden alle bis dahin als kürzeste Wege gespeicherten Pfade wieder gelöscht. Neben der Suche nach kürzesten Wegen berechnet Algorithmus C.3.4 zusätzlich für jeden Knoten des Graphen die Anzahl kürzester Wege vom jeweiligen Startknoten aus. Mittels dieser Informationen ist es möglich, die Gewichte aus einer OD-Matrix auf die Umsteige- und Aussteigeaktivitäten umzulegen.

Algorithmus C.3.6 beschreibt die Vorgehensweise. Dabei wird der Graph von jedem möglichen Metaknoten rückwärts bis zum Startknoten durchlaufen. Dies geschieht rekursiv und gemäß einer Tiefensuche, falls mehrere kürzeste Wege existieren.

9.2. Bewertung eines Fahrplans

CAREY vergleicht in [Car99] eine Vielzahl existierender Bewertungskriterien für die Verlässlichkeit bzw. Funktionsfähigkeit von Fahrplänen. Die Schwierigkeit bei der Festlegung von Bewertungskriterien für Fahrpläne ist, dass wir dabei in einem System zwei verschiedene Größen betrachten, die sich dort ganz unterschiedlich bewegen. Dies sind zum einen *Züge*, die sich gemäß ihrer zugehörigen Linie bewegen.

Die andere Größe stellen *Passagiere* dar, welche sich mittels verschiedener Linien bzw. Züge auf dem Weg zu ihrem Ziel durch das System bewegen. Möchte man einen Fahrplan bewerten, muss man sich zunächst entscheiden, bezüglich welcher der beiden Größen dies geschehen soll. Mittels der zu findenden Kriterien soll beurteilt werden, inwieweit ein Fahrplan eine bestimmte (wünschenswerte) Eigenschaft besitzt. Diese Eigenschaften konkurrieren oft miteinander:

- „Anschlussicherung“ vs. „Kurze Aufenthaltszeiten in der Stationen“
- „Verspätungsresistenz“ vs. „Kurze Reisezeiten“

Für Züge könnte ein Fahrplan beispielsweise optimiert werden, indem die mittlere Ankunftsverspätung der Züge in den Stationen ihrer Linie minimiert wird. Gewichten wir dabei jede Linie und insbesondere jede Ankunft gleich, so wird dies im Allgemeinen einen Fahrplan ergeben, der für die Passagiere nicht optimal ist. Denn hier bleibt unberücksichtigt, wie stark welche Linien von den Passagieren genutzt werden. Die Verspätung eines jeden Passagiers direkt zu berechnen, bedeutete unverhältnismäßig viel Aufwand. Wir suchen daher einen Weg, die Verspätungen der Züge auf Passagiere umzulegen. Wir haben uns schließlich dafür entschieden, eine bikriterielle Bewertung für Fahrpläne zu verwenden.

9.2.1. Mittlere fahrplanmäßige Reisezeit

Die mittlere fahrplanmäßige Reisezeit eines Passagiers hängt, neben den im Fahrplan π enthaltenen Informationen, von seiner persönlichen Route ab. Sie ist jedoch im Gegensatz zur Ankunftsverspätung eine deterministische Größe. In Abschnitt 9.1 haben wir erläutert, wie wir die unterschiedlichen Routen der Passagiere berücksichtigen. Für die mittlere fahrplanmäßige Reisezeit $r(\pi)$ aller Reisenden wird das Produkt aus dem in Algorithmus C.3.6 berechneten Reisengewicht γ_a und der jeweiligen fahrplanmäßigen Dauer über alle Aktivitäten $a \in \mathfrak{A}$ des Ereignis-Aktivitätsnetzwerkes $\mathbb{N} = (\mathfrak{E}, \mathfrak{A})$ aufsummiert, d.h.

$$r(\pi) := \sum_{a \in \mathfrak{A}} \gamma_a \cdot \pi(a) .$$

Dabei sei $\pi(a)$ die fahrplanmäßige Dauer der Aktivität $a \in \mathfrak{A}$.

9.2.2. Mittlere Ankunftsverspätung

Grundlage für das zweite Bewertungskriterium ist die mittlere Verspätung der Passagiere bei Ankunft in ihrer Zielstation. Dies ist offensichtlich eine stochastische Größe. Sie lässt sich mittels der Information über die Verspätung der genutzten Züge bestimmen. Jeder verpasste Anschluss wird mit der vollen (einheitlichen) Taktzeit τ als Verspätung gewertet. Die endgültige Ankunftsverspätung in der Zielstation

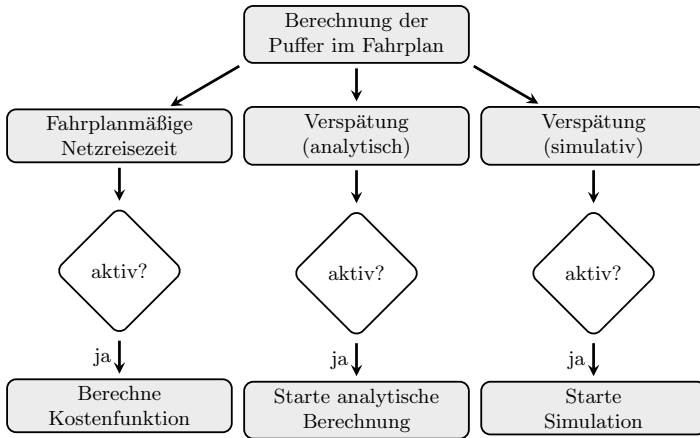


Abbildung 9.1.: Bewertung eines Individuums

ergibt sich dann durch Addition der Verspätung des zuletzt genutzten Zuges. Die Algorithmen C.3.1 und C.3.2 zeigen, wie diese Verspätung, über alle Passagiere gemittelt, im Rahmen der ereignisgesteuerten stochastischen Simulation (s. Kapitel 8) bestimmt wird. Algorithmus C.3.3 zeigt die Bestimmung unter Verwendung zuvor berechneter Verspätungsverteilungsfunktionen.

9.3. Fahrplanoptimierung in HiTT

Im Softwareprojekt HiTT werden Fahrpläne mittels eines genetischen Algorithmus optimiert. Die Optimierung selbst war nicht Gegenstand dieser Arbeit. Fahrpläne stellen die Individuen einer Population dar. Ihre Bewertung findet in einem 3-dimensionalen Kostenraum statt. Eine Dimension stellt die mittlere fahrplanmäßige Reisezeit (in Minuten) dar, die anderen beiden Dimensionen entsprechen der simulierten bzw. mittels der Verteilungsfunktionen bestimmten mittleren Ankunftsverspätung (in Minuten) aus dem vorherigen Abschnitt. In Abbildung 9.1 wird die Fahrplanbewertung veranschaulicht. Zunächst werden die im Fahrplan enthaltenen Zeitpuffer bestimmt. Dies ist für die Berechnung der Verteilungsfunktionen notwendig. Gleichzeitig wird in diesem Schritt der Fahrplan in die für die Simulation benötigte Struktur transformiert. Vor Beginn der Optimierung wird festgelegt, welche der drei Kostenfunktionen berechnet werden sollen (bis zu drei). Für die Optimierung sind maximal zwei Kostenfunktionen auswählbar. Soll hinsichtlich der mittleren Ankunftsverspätung optimiert werden, entscheiden wir uns entweder für die Simulation oder die Bestimmung der Verteilungsfunktionen.

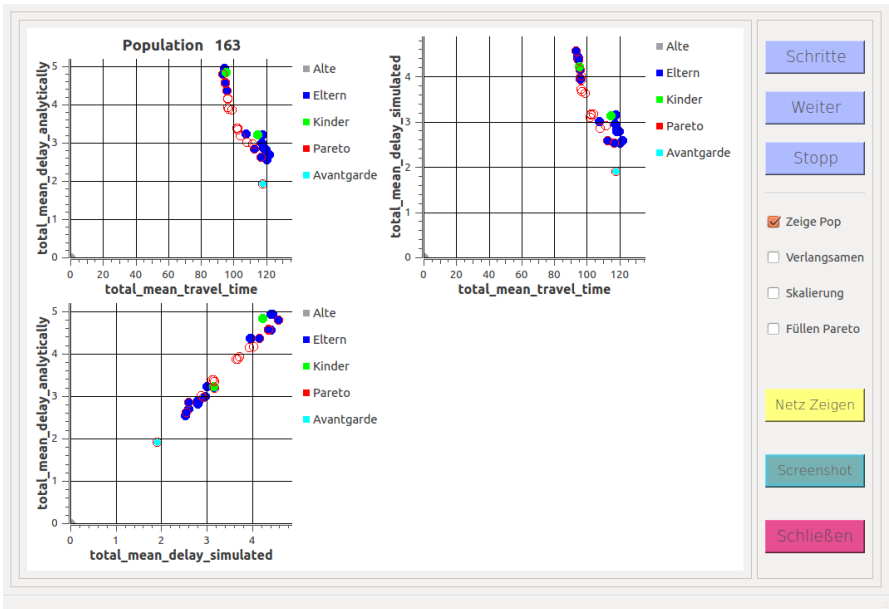


Abbildung 9.2.: Graphische Oberfläche des Softwareprojektes HiTT

In Abbildung 9.2 ist die graphische Oberfläche von HiTT mit der Visualisierung des 3-dimensionalen Kostenraumes zu sehen. In zwei der drei Fenster wird die mittlere fahrplanmäßige Reisezeit zusammen mit der mittleren Ankunftsverspätung (simulativ und unter Verwendung der Verteilungsfunktionen) in einem Streudiagramm dargestellt. Das dritte Fenster zeigt das Streudiagramm für die beiden Methoden zur Bestimmung der mittleren Ankunftsverspätung. Je näher die Punkte hier an der Ursprungsgeraden liegen, desto besser stimmen Simulation und Bestimmung der Verteilungsfunktionen für die jeweiligen Fahrpläne überein. Die Genauigkeit der Simulation lässt sich über die Parameter der Konfidenzintervalle, der vorgegeben Breite sowie des Signifikanzniveaus, steuern. Die gute Übereinstimmung der Kostenfunktionswerte spricht, wegen der deutlichen Vorteile hinsichtlich der Rechenzeit, für die Verwendung der Verspätungspropagation auf Ebene der Verteilungsfunktionen.

In Abbildung 9.3 ist eine typische Entwicklung einer Population zu sehen, wenn sowohl die mittlere Reisezeit als auch die mittlere Ankunftsverspätung als Optimierungskriterium verwendet werden. Wir erhalten eine Menge pareto-optimaler Fahrpläne (rote Markierung), wobei die mittlere Verspätung für kleiner werdende

9. Anwendung der Methoden im Rahmen der Optimierung

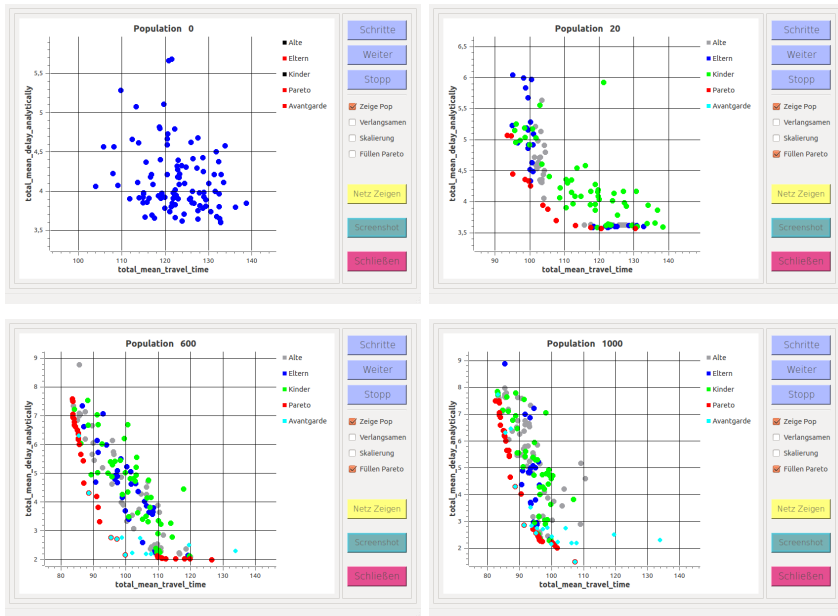


Abbildung 9.3.: Populationsentwicklung während der Optimierung

Reisezeiten zunehmend stark anwächst.

10. Zusammenfassung

Die Erstellung von Fahrplänen im Bahnverkehr unterliegt verschiedenen Herausforderungen. Zufällige Einflüsse können Betriebsabläufe stören und letztlich zu Abweichungen vom Fahrplan, d.h. Verspätungen, führen. Dies können beispielsweise nicht schließende bzw. öffnende Türen während des Aufenthalts eines Zuges in einer Station sein. Ein anderes Beispiel sind außergewöhnliche Wetterumstände während der Fahrt eines Zuges. Durch externe Störungen verursachte Verzögerungen werden Quellverspätungen (s. Kapitel 3) genannt. Sowohl aus Sicht des Dienstleisters als auch aus Sicht der Bahnreisenden ist es wünschenswert, wenn ein Fahrplan das tatsächliche, reale Fahrgeschehen der Züge möglichst genau widerspiegelt. Dies bedeutet, dass er gegen moderate Störungen robust sein muss. In solchen Fahrplänen steht für kritische Betriebsabläufe, zusätzlich zu deren Mindestzeiten, ein Zeitpuffer zur Verfügung.

Die Grundlage der Modellierung in dieser Dissertation stellt das in Kapitel 3 erklärte Ereignis-Aktivitätsnetzwerk dar. Den Aktivitäten werden Quellverspätungen zugeordnet. Darüber hinaus betrachten wir auch fortgepflanzte Verspätungen, die, direkt oder indirekt, durch Quellverspätungen verursacht werden. Diese beziehen sich immer auf ein Ereignis. Verspätungen werden grundsätzlich durch Zufallsvariablen bzw. die zugehörigen Verteilungsfunktionen repräsentiert. Ziel ist es, für alle Ankunfts- und Abfahrtsereignisse, d.h. bezüglich aller Kombinationen von Linien und zugehörigen Stationen, die Verteilungsfunktion der auftretenden Verspätungen zu bestimmen. Neben den Quellverspätungen fließen hier die im Fahrplan enthaltenen Zeitpuffer sowie Regeln für die Wartezeit von Abbringerzügen auf die jeweiligen Zubringer (s. Kapitel 6) ein. Mittels der resultierenden Verspätungsverteilungen können Kenngrößen für die Robustheit eines Fahrplans berechnet werden, welche im Rahmen der Optimierung von Fahrplänen Anwendung finden (s. Kapitel 9).

Es wird mit der Familie der erweiterten Hyper-Erlangverteilungen (s. Kapitel 5) eine Verteilungsfamilie verwendet, die es ermöglicht, die empirischen Verteilungen der Quellverspätungen hinreichend genau zu approximieren. Dazu wurde ein aus der Literatur bekanntes Verfahren modifiziert, welches für die Anpassung empirischer Verteilungen einen EM-Algorithmus verwendet. Die erhaltenen Verteilungsfunktionen werden als Theta-Exponentialpolynom (s. Kapitel 6) dargestellt. Diese Polynome sind abgeschlossen gegen die für die Propagation der Verspätungsverteilungsfunktionen benötigten Operationen. Im Zuge der Berechnung der Verteilungs-

funktionen werden diese miteinander verknüpft bzw. modifiziert. Die Komplexität der Repräsentation der resultierenden Verteilungsfunktionen steigt dabei im Allgemeinen stark an. Daher ist es nötig, ein Verfahren zur Komplexitätsreduzierung zu nutzen, welches die Verteilungsfunktion hinreichend gut approximiert. Hierbei wird eine (erweiterte) Hyper-Erlangverteilung bestimmt, die in den ersten drei Momenten mit der zu approximierenden Verteilung übereinstimmt (s. Kapitel 6). Das Ergebnis kann wieder als Theta-Exponentialpolynom dargestellt werden.

Den Kern dieser Arbeit stellt die Untersuchung der Verspätungsfortpflanzung in starken Zusammenhangskomponenten des Ereignis-Aktivitätsnetzwerks dar (s. Kapitel 7). In diesen Teilen des Netzwerkes ist es nicht möglich, eine topologische Sortierung der Ereignisse herzustellen. Zunächst wurde ein iteratives Verfahren entwickelt, welches es ermöglicht, die Verteilungsfunktionen der Verspätungen aller Ereignisse approximativ zu bestimmen. Es wurden verschiedene Sortierungen für die Ereignisse verwendet. Dabei zeigte sich, dass die spezielle Wahl einer Sortierung empirisch keinen Einfluss auf die Konvergenz der Verteilungen sowie insbesondere auf die Grenzverteilungen hat. Lediglich die Konvergenzgeschwindigkeit variierte in Abhängigkeit der gewählten Sortierung. In der vorliegenden Arbeit wird eine spezielle Sortierung vorgeschlagen, die diesbezüglich im Allgemeinen zu guten Ergebnissen führt.

Darauf aufbauend wurden die empirisch erhaltenen Ergebnisse mittels eines theoretischen Modellierungsansatzes untersucht (s. Kapitel 7). Es gelang, die empirischen Ergebnisse zu erklären bzw. zu konkretisieren. Dabei konnten zwei zentrale Bedingungen für die Durchführbarkeit bzw. Korrektheit des iterativen Verfahrens aufgestellt werden.

- (i) **Unabhängigkeit:** Die Zufallsvariablen aller miteinander zu verrechnender Verteilungsfunktionen sind stets stochastisch unabhängig.
- (ii) **Stabilität:** Der Erwartungswert der aufsummierten Quellverspätungen ist in jedem Kreis einer Kreisvereinigung stets kleiner als die Summe der dort enthaltenen Zeitpuffer.

Für die Stabilität der Verspätungen in einer starken Zusammenhangskomponente ist unter anderem Bedingung (ii) wesentlich. Ist zusätzlich die Verspätungsübertragung zwischen den Kreisen der starken Zusammenhangskomponente beschränkt, so entwickeln sich die Verspätungen stabil. Ihre Grenzverteilungsfunktion ist unabhängig von der jeweiligen Startverteilung zu Beginn der Iteration, falls Bedingung (i) erfüllt ist. Bedingung (i) ist dabei zugleich Grundlage für die gesamte Verspätungsfortpflanzung auf Ebene der Verteilungsfunktion, nicht nur in starken Zusammenhangskomponenten. Ist diese Bedingung verletzt, werden wir fehlerhafte Resultate erhalten.

Weitere potentielle Fehlerquellen liegen in den an drei verschiedenen Stellen durchgeführten Approximationen. Im Rahmen des iterativen Verfahrens werden

die Grenzverteilungsfunktionen approximiert. Während der Verrechnung der Verteilungsfunktionen, approximieren wird diese zwecks Komplexitätsreduzierung mittels eines Momentenanpassungsverfahrens. Auch die empirischen Verteilungen der Quellverspätungen werden approximiert. Dass die dabei gemachten Fehler nicht unkontrolliert anwachsen, und damit die Zulässigkeit des Verfahrens, wurde durch empirische Vergleiche gezeigt. So wurde im Rahmen dieser Arbeit zur Validierung der Ergebnisse eine ereignisgesteuerte stochastische Simulation entwickelt (s. Kapitel 8). Mittels dieser lassen sich Verspätungen relativ einfach propagieren. Insbesondere Bedingung (i) ist hier unkritisch, da statt Verteilungsfunktionen die reellen Verspätungswerte miteinander verrechnet werden. Jedoch ist die Simulation im Vergleich zur Verspätungsfortpflanzung auf Ebene der Verteilungsfunktionen sehr zeitaufwendig. Eine weitere Validierung fand im Rahmen eines Vergleiches mit einem szenarienbasierten Modellierungsansatz statt. Dieser wurde unabhängig von dieser Arbeit an der Georg-August-Universität Göttingen entwickelt und im Rahmen des Simulationswissenschaftlichen Zentrums mit dem in dieser Arbeit entwickelten Ansatz zusammengeführt. In [HKKS14] vergleichen wir die mit dem szenarienbasierten Ansatz resultierenden empirischen Verteilungsfunktionen der Verspätungen in einem Ereignis-Aktivitätsnetzwerk mit den entsprechenden Verteilungsfunktionen, die mit den in dieser Arbeit beschriebenen Methoden bestimmt wurden. Die Ergebnisse zeigen eine große Übereinstimmung.

A. Bezeichnungen

$[x]^+$	$\max\{x, 0\}$
$\sigma_K(V)$	spezielle Sortierung der Knoten des Kreises K (s. Bemerkung 7.1.11)
\mathbb{N}	die Menge der natürlichen Zahlen
\mathbb{R}	die Menge der reellen Zahlen
\mathfrak{S}	die Menge der Stationen eines Streckennetzes
\mathfrak{B}	die Menge der Strecken eines Streckennetzes
\mathfrak{L}	die Menge der Linien eines Streckennetzes
\mathcal{S}_L	die Menge der Linienstationen der Linie L
\mathcal{B}_L	die Menge der Linienstrecken der Linie L
\mathfrak{N}	das Ereignis-Aktivitätsnetzwerk
\mathfrak{N}^*	das erweiterte Ereignis-Aktivitätsnetzwerk
\mathfrak{G}	der Anschlussgraph
$\mathcal{M}_{\mathfrak{G}}$	die Menge der Kreisvereinigungen des Anschlussgraphen \mathfrak{G}
$\mathcal{I}_{\mathfrak{G}}$	die Menge der singulären Knoten des Anschlussgraphen \mathfrak{G}
$\mathfrak{Pre}_G(\cdot)$	die Menge der Vorgänger eines Knoten in Graph bzw. Netzwerk G
$\mathfrak{Suc}_G(\cdot)$	die Menge der Nachfolger eines Objektes in Graph bzw. Netzwerk G
\mathcal{V}_K	die Menge der Knoten des Kreises K
$\mathcal{V}_{\mathcal{M}}$	die Menge der Knoten aller Kreise der Kreisvereinigung \mathcal{M}

B. Beweise

B.1. Loynes

Lemma B.1.1 (Lemma von Loynes).

[Loy62] Seien $(W_n)_{n \in \mathbb{N}}$ Zufallsvariablen mit

$$W_{n+1} = f(W_n, U_n) \tag{B.1}$$

und $(U_n)_{n \in \mathbb{Z}}$ als stationäre Folge von Zufallsvariablen. Außerdem sei

- (i) die Transformation $f(x, y)$, $x, y \in \mathbb{R}$, nicht-negativ sowie monoton wachsend und linksseitig stetig in x (auch in ∞),
- (ii) $W_1 \equiv 0$.

Dann existiert eine stationäre Folge von Zufallsvariablen $(M_n)_{n \in \mathbb{Z}}$, mit

$$M_{n+1} = f(M_n, U_n)$$

für alle $n \in \mathbb{Z}$, für welche im Falle $\mathbf{P}(M_0 < \infty) = 1$ folgt, dass $W_n \xrightarrow{D} M_0$ ($n \rightarrow \infty$). Insbesondere gilt für alle Folgen $(A_n)_{n \in \mathbb{Z}}$, die (B.1) für alle $n \in \mathbb{Z}$ erfüllen, $A_n \geq M_n$ (\mathbf{P} -f.s.).

Beweis. Für den Beweis der Aussage wird in [Loy62] für alle $m \in \mathbb{Z}$ die Folge $(W_n^m)_{n \in \mathbb{Z}}$ durch

$$W_n^m \equiv 0, \quad \text{falls } n \leq -m + 1, \tag{B.2}$$

$$W_n^m = f(W_{n-1}^m, U_{n-1}), \quad \text{falls } n > -m + 1 \tag{B.3}$$

definiert. Es wird nun für festes $n \in \mathbb{Z}$ die Folge $(W_n^m)_{m \in \mathbb{Z}}$ betrachtet. Der Zusammenhang zur in Gleichung (B.1) erklärten Folge $(W_n)_{n \in \mathbb{N}}$ ergibt sich wie folgt. W_n stimmt offensichtlich für alle $n \in \mathbb{N}$ mit W_n^0 überein. Im Folgenden wird gezeigt, dass W_n^0 und W_0^n identisch verteilt sind, und W_0^n in Verteilung monoton gegen M_0 konvergiert. Damit ist dann auch die Konvergenz von W_n gegen M_0 gezeigt. Es gelten für W_n^m die folgenden beiden Eigenschaften:

1. Für festes $r \in \mathbb{Z}$ ist die Folge $(W_n^{r-n})_{n \in \mathbb{Z}}$ stationär.

Beweis: Aus $r \leq 1$ folgt

$$m := r - n \leq 1 - n \iff n \leq -m + 1$$

für alle $n \in \mathbb{Z}$. Dann gilt mit (B.2) $W_n^{r-n} \equiv 0$ für alle $n \in \mathbb{Z}$. Für $r \leq 1$ ist die Aussage damit gezeigt. Aus $r > 1$ folgt analog zu oben $n > -m + 1$. Es gilt also mit (B.3) $W_n^{r-n} = f(W_{n-1}^{r-n}, U_{n-1})$. Die Rekursion lässt sich mit W_{n-1}^{r-n} fortsetzen, bis wir wieder die Situation in (B.2) erhalten. Für festes r ist die Anzahl der Rekursionsschritte bis dies eintritt, unabhängig von n , stets $r - 1$. Es sei nun f^m , $m \in \mathbb{N}$, die m -fache Anwendung der Transformation f , sodass gilt

$$\begin{aligned} W_n^{r-n} &= f^{r-1}(W_{n-(r-1)}^{r-n}, U_{n-1}, U_{n-2}, \dots, U_{n-(r-1)}) \\ &= f^{r-1}(0, U_{n-1}, U_{n-2}, \dots, U_{n-(r-1)}) . \end{aligned}$$

Aus der Stationarität der Folge $(U_n)_{n \in \mathbb{Z}}$ folgt

$$\begin{aligned} W_n^{r-n} &\sim f^{r-1}(0, U_{(n+1)-1}, U_{(n+1)-2}, \dots, U_{(n+1)-(r-1)}) \\ &= f^{r-1}(W_{(n+1)-(r-1)}^{r-(n+1)}, U_{n-1}, U_{n-2}, \dots, U_{n-(r-1)}) = W_{n+1}^{r-(n+1)} . \end{aligned}$$

Für $(W_n^{r-n}, W_{n+1}^{r-(n+1)}, \dots, W_{n+k}^{r-(n+k)})$ mit $k > 1$ ist die Vorgehensweise analog.

2. Für festes $n \in \mathbb{Z}$ gilt für alle $m \in \mathbb{Z}$ immer $W_n^m \leq W_n^{m+1}$ (**P**-f.s.).

Beweis: Sei zunächst $m \leq -n + 1$. Dann gilt mit (B.2) $W_n^m \equiv 0$. Wegen der Nicht-Negativitätseigenschaft der Transformation f und (B.3) folgt im Falle $m = -n + 1$, dass

$$W_n^{m+1} = f(W_{n-1}^{m+1}, U_{n-1}) = f(0, U_{n-1}) \geq 0 \equiv W_n^m \quad (\mathbf{P}\text{-f.s.}) . \quad (\text{B.4})$$

Es sei nun $r \in \mathbb{N}$ und $m = r - n + 1$, d.h. $m > -n + 1$. Dann gilt $n = r - m + 1 \iff n - r = -m + 1$ und daher mit (B.4), d.h. $W_{n-r}^{m+1} \geq W_{n-r}^m$ (**P**-f.s.),

$$\begin{aligned} W_n^{m+1} &= f(W_{n-1}^{m+1}, U_{n-1}) = f^r(W_{n-r}^{m+1}, U_{n-1}, U_{n-2}, \dots, U_{n-r}) \\ &\geq f^r(W_{n-r}^m, U_{n-1}, U_{n-2}, \dots, U_{n-r}) = W_n^m \quad (\mathbf{P}\text{-f.s.}) . \end{aligned}$$

Es folgt, dass für festes $n \in \mathbb{Z}$ für alle $m \in \mathbb{Z}$ und $t \in \mathbb{R}_{\geq 0}$ immer $\mathbf{P}(W_n^m \leq t) \geq \mathbf{P}(W_n^{m+1} \leq t)$ gilt. Nach Korollar 7.2.5 konvergiert daher die Verteilungsfunktion $F_{n,m}(t) := \mathbf{P}(W_n^m \leq t)$, $t \in \mathbb{R}_{\geq 0}$, punktweise für $m \rightarrow \infty$. Die Grenzfunktion muss allerdings keine Verteilungsfunktion sein (z.B. falls $\lim_{m \rightarrow \infty} F_{n,m} \equiv 0$).

Es sei nun $M_n := \lim_{m \rightarrow \infty} W_n^m$ für alle $n \in \mathbb{Z}$. Dann gilt

$$M_{n+1} = \lim_{m \rightarrow \infty} W_{n+1}^m = \lim_{m \rightarrow \infty} f(W_n^m, U_n) = f\left(\lim_{m \rightarrow \infty} W_n^m, U_n\right) = f(M_n, U_n) .$$

Dabei wurde die Monotonie der Folge $(W_n^m)_{m \in \mathbb{Z}}$ sowie die linksseitige Stetigkeit der Transformation f ausgenutzt. Die Stationarität der Folge $(M_n)_{n \in \mathbb{Z}}$ folgt aus der Stationarität der Folge $(W_n^{r-n})_{n \in \mathbb{Z}}$ für festes $r \in \mathbb{Z}$. Es gilt für alle $k, l \in \mathbb{N}$

$$\begin{aligned} & (M_n, M_{n+1}, \dots, M_{n+k}) \\ &= \left(\lim_{m \rightarrow \infty} W_n^m, \lim_{m \rightarrow \infty} W_{n+1}^m, \dots, \lim_{m \rightarrow \infty} W_{n+k}^m \right) \\ &= \lim_{m \rightarrow \infty} (W_n^m, W_{n+1}^m, \dots, W_{n+k}^m) \sim \lim_{m \rightarrow \infty} (W_{n+l}^{m-l}, W_{n+l+1}^{m-l}, \dots, W_{n+l+k}^{m-l}) \\ &= (M_{n+l}, M_{n+l+1}, \dots, M_{n+l+k}) . \end{aligned}$$

Es bleibt noch die Minimalität der Folge $(M_n)_{n \in \mathbb{Z}}$ zu zeigen. Wir nehmen an, dass es eine weitere Folge $(A_n)_{n \in \mathbb{Z}}$ gibt mit $A_{n+1} = f(A_n, U_n)$ für alle $n \in \mathbb{Z}$. Es sei zunächst $m \leq -n + 1$. Dann gilt offensichtlich $W_n^m \equiv 0 \leq A_n$ (**P**-f.s.). Für alle $m > -n + 1$ folgt

$$\begin{aligned} W_n^m &= f(W_{n-1}^m, U_{n-1}) \\ &= f^{n-1+m}(W_{n-(n-1+m)}^m, U_{n-1}, U_{n-2}, \dots, U_{n-(n-1+m)}) \\ &= f^{n-1+m}(0, U_{n-1}, U_{n-2}, \dots, U_{1-m}) \\ &\leq f^{n-1+m}(A_{1-m}, U_{n-1}, U_{n-2}, \dots, U_{1-m}) = A_n \text{ (**P**-f.s.)} . \end{aligned}$$

Es gilt also für alle $m \in \mathbb{Z}$, dass $W_n^m \leq A_n$ (**P**-f.s.). Daher gilt auch $M_n = \lim_{m \rightarrow \infty} W_n^m \leq A_n$ (**P**-f.s.). \square

Lemma B.1.2.

[SS07] Es seien A_1, A_2, \dots, A_n , $n \in \mathbb{N}$, stochastisch unabhängige reellwertige Zufallsvariablen. Ebenso seien B_1, B_2, \dots, B_n stochastisch unabhängige reellwertige Zufallsvariablen. Es gelte

$$\mathbf{P}(A_i \leq t) \geq \mathbf{P}(B_i \leq t)$$

für alle $1 \leq i \leq n$ und $t \in \mathbb{R}$. Ist dann $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine monoton wachsende Funktion, so folgt

$$\mathbf{P}(f(A_1, A_2, \dots, A_n) \leq t) \geq \mathbf{P}(f(B_1, B_2, \dots, B_n) \leq t)$$

für alle $t \in \mathbb{R}$.

Korollar B.1.3.

Es gelte für die Zufallsfolge $(Y_V^{(n)})_{n \in \mathbb{N}}$ des Knotens $V \in K$ des Kreises $K \in \mathcal{M}$ der Kreisvereinigung \mathcal{M}

$$Y_V^{(n+1)} = f_{S,V}(Y_V^{(n)}, D^{(n)}, C^{(n)})$$

mit für größer werdende $n \in \mathbb{N}$ elementweise stochastisch wachsenden Zufallsvektoren $D^{(n)} = (D_{V_1}^{(n)}, D_{V_2}^{(n)}, \dots, D_{V_r}^{(n)})$ und $C^{(n)} = (C_{V_1}^{(n)}, C_{V_2}^{(n)}, \dots, C_{V_r}^{(n)})$, $r = |\mathcal{V}_K|$. Es seien $Y_V^{(n)}$, $D^{(n)}$ und $C^{(n)}$ für alle $n \in \mathbb{N}$ stochastisch unabhängig. Außerdem sei

(i) die Transformation $f_{S,V}$ nicht-negativ sowie monoton wachsend und linksseitig stetig in der ersten Komponente (auch in ∞),

(ii) $Y_V^{(1)} \equiv 0$.

Dann existiert für alle $V \in K$ eine Zufallsvariable M_V , für die im Falle $\lim_{n \rightarrow \infty} Y_V^{(n)} < \infty$ (\mathbf{P} -f.s.) folgt, dass $Y_V^{(n)} \xrightarrow{\mathcal{D}} M$ ($n \rightarrow \infty$).

Beweis. Aus der Nichtnegativität der Transformation f folgt aus

$$Y_V^{(2)} = f_{S,V}(Y_V^{(1)}, D^{(1)}, C^{(1)})$$

offensichtlich für alle $t \in \mathbb{R}$

$$\mathbf{P}\left(Y_V^{(1)} \leq t\right) = 1 \geq \mathbf{P}\left(f_{S,V}(Y_V^{(1)}, D^{(1)}, C^{(1)}) \leq t\right) = \mathbf{P}\left(Y_V^{(2)} \leq t\right).$$

Mit der Annahme $\mathbf{P}(Y_V^{(n-1)} \leq t) \geq \mathbf{P}(Y_V^{(n)} \leq t)$ und nach Voraussetzung

$$\mathbf{P}\left(D_{V_i}^{(n-1)} \leq t\right) \geq \mathbf{P}\left(D_{V_i}^{(n)} \leq t\right) \quad \text{sowie} \quad \mathbf{P}\left(C_{V_i}^{(n-1)} \leq t\right) \geq \mathbf{P}\left(C_{V_i}^{(n)} \leq t\right)$$

für alle $1 \leq i \leq r$ und $t \in \mathbb{R}$ folgern wir mit Lemma B.1.2 und der Monotonie der Transformation f induktiv

$$\begin{aligned} \mathbf{P}\left(Y_V^{(n)} \leq t\right) &= \mathbf{P}\left(f(Y_V^{(n-1)}, D^{(n-1)}, C^{(n-1)}) \leq t\right) \\ &\geq \mathbf{P}\left(f(Y_V^{(n)}, D^{(n)}, C^{(n)}) \leq t\right) = \mathbf{P}\left(Y_V^{(n+1)} \leq t\right) \end{aligned}$$

für alle $t \in \mathbb{R}$. Mit Korollar 7.2.5 ist die Verteilungskonvergenz der Folge $(Y_V^{(n)})_{n \in \mathbb{N}}$ gezeigt. \square

C. Algorithmen

C.1. Ereignisgesteuerte stochastische Simulation

```
Input : Ereignisliste  $\mathcal{E}$  und aktuelles Ereignis  $E$ 

// Fuer den naechsten Takt wird ein neues Startereignis erzeugt:
1  $E' =$ 
  ErzeugeNeuesStartereignis( $E$ .Zeitpunkt, $E$ .Zugnummer, $E$ .Linie, $E$ .Station);
2  $\mathcal{E}$ .SortiereEin( $E'$ );
  // Fahrplanmaessige und spaeteste Abfahrt (Anschlusssicherung)
3  $t_{Ab,FP} = E$ .Zeitpunkt -  $E$ .Verspaetung;
  // Anschlusssicherungskonstante  $\kappa$ 
4  $t_{Ab,spact} = t_{Ab,FP} + \kappa$ ;
  // Wartezeit auf spaetesten Zubringer innerhalb der
  // Anschlusssicherungsstrategie ab Ankunft des Abbringers:
5  $t_{Zubringer} =$  BerechneSpaetestenAnschluss( $t_{Ab,spact}$ );
  // Tatsaechliche Abfahrtszeit:
6  $t_{AB} = \max\{ E$ .Zeitpunkt +  $t_{Zubringer}, t_{Ab,FP} \}$ ;
7  $AktuelleVerspaetung = t_{AB} - t_{Ab,FP}$ ;
  // Erstelle ein Abfahrtsereignis in derselben Station:
8  $E' =$  ErzeugeNeuesAbfahrtsereignis( $t_{AB}$ , $AktuelleVerspaetung$ ,
   $E$ .Zugnummer, $E$ .Linie, $E$ .Station);
9  $\mathcal{E}$ .SortiereEin( $E'$ );
```

Algorithmus C.1.1 : Bearbeitung eines Startereignisses

Input : Liste der tatsächlichen Ankuenfte \mathcal{A} , Ereignisliste \mathcal{E} und aktuelles Ereignis E

```
// Quellverspaetung auf der folgenden Strecke:
1  $d_{QV} = \text{NaechsteLinienStrecke.SimuliereQuellverspaetung}();$ 
// Tatsaechliche Ankunftszeit
// (implizit Verrechnung des Puffers):
2  $t_{AN} = E.\text{Zeitpunkt} + \text{Mindestfahrzeit} + d_{QV};$ 
// Tatsaechliche - fahrplanmaessige Ankunftszeit:
3  $\text{AktuelleVerspaetung} = \max\{t_{AN} - t_{An,FP}, 0\};$ 
// Erstelle ein Ankunftsereignis in
// der naechsten Station der Linie:
4  $E' =$ 
  ErzeugeNeuesAnkunftsereignis( $t_{AN}, \text{AktuelleVerspaetung}, E.\text{Zugnummer},$ 
   $E.\text{Linie}, E.\text{Station}$ );
5  $\mathcal{E}.\text{SortiereEin}(E')$ ;
// Update der Liste tatsaechlichen Ankuenfte  $\mathcal{A}$ :
6  $p = \text{GibPositionDerTatsaechlichenAnkunftIn}(\mathcal{A}, E.\text{Station}, E.\text{Zugnummer});$ 
7  $\mathcal{A}.\text{LoeschePosition}(p);$ 
8  $\mathcal{A} = \text{ErstelleNeueTatsaechlicheAnkunft}(E')$ ;
9  $\mathcal{A}.\text{FuegeHinzu}(A);$ 
```

Algorithmus C.1.2 : Bearbeitung eines Abfahrtsereignisses

```

Input : Ereignisliste  $\mathcal{E}$  und aktuelles Ereignis  $E$ 

// Pruefe, ob der aktuelle Zug noch
// eine weitere Station anfaehrt:
1 if  $E$ .StationIstEndstationDerLinie then
2 |   RegistriereAnkunftsversaetung( $E$ );
3 else
4 |   RegistriereAnkunftsversaetung( $E$ );
5 |   // Quellversaetung in der Station:
6 |    $d_{QV} = E$ .LinienStation.SimuliereQuellversaetung();
7 |   // Tatsaechliche Aufenthaltsdauer (Abfahrt
8 |   // nicht vor fahrplanmaessiger Abfahrt):
9 |    $t_{halt} = E$ .LinienStation.Mindesthaltezeit +  $d_{QV}$ ;
10 |   $t_{halt} = \max\{t_{halt}, t_{Ab,FP} - E$ .Zeitpunkt};
11 |  // Beruecksichtigung von (evtl. versaeteten) Zubringern
12 |  // mit Anschlussicherungskonstanten  $\kappa$ :
13 |   $t_{Ab,spact} = t_{Ab,FP} + \kappa$ ;
14 |   $t_{Anschluss} =$  Algorithmus C.1.4( $E, E$ .LinienStation.Zubringervektor,
15 |   $t_{Ab,spact}$ );
16 |   $t_{halt} = \max\{t_{halt}, t_{Anschluss}\}$ ;
17 |  // Tatsaechliche Abfahrtszeit:
18 |   $t_{AB} = E$ .Zeitpunkt +  $t_{halt}$ ;
19 |   $AktuelleVersaetung = t_{AB} - t_{Ab,FP}$ ;
20 |  // Erstelle ein neues Abfahrtsereignis:
21 |   $E' =$  ErzeugeNeuesAbfahrtsereignis( $t_{AB}, AktuelleVersaetung,$ 
22 |   $E$ .Zugnummer,  $E$ .Linie,  $E$ .Station);
23 |   $\mathcal{E}$ .SortiereEin( $E'$ );

```

Algorithmus C.1.3 : Bearbeitung eines Ankunftsereignisses

Input : Ereignis E , Zubringervektor \mathcal{Z} , Array \mathfrak{A} der tatsaechlichen Ankuenfte einer jeden Linie und die durch die Anschluss-sicherungsstrategie festgelegte spaeteste Abfahrtszeit des Abbringers $t_{Ab,spael}$

```
1  $t_{Anschluss} = 0;$ 
   // Zubringervektor  $\mathcal{Z}$ 
2 forall the  $z \in \mathcal{Z}$  do
3    $ZubringerzugGefunden = false;$ 
   // Liste der tatsaechlichen Ankuenfte der Zubringerlinie:
4    $\mathcal{A}_z = \mathfrak{A}[z.Linie];$ 
5    $t_{Anschluss} = \text{Algorithmus C.1.5}(z, \mathcal{A}_z, t_{Anschluss});$ 
6   if  $!ZubringerzugGefunden$  then
       // Der Zubringerzug ist noch nicht an der vorherigen
       // Station abgefahren:
7      $t_z = \text{VerspaetungsHochrechnung}(z, \mathcal{A}_z);$ 
8     if  $t_z > t_{Ab,spael}$  then
9       |  $\text{RegistriereVerpasstenAnschluss}(E.Linienstation);$ 
10    else
11       | // Differenz zwischen dem Eintreffen der Passagiere des
          // Zubringerzuges am Bahnsteig und der Ankuunft des
          // Abbringers:
           $t_{Anschluss} = \max\{t_{Anschluss}, t_z - E.Zeitpunkt\}$ 
12 return  $t_{Anschluss};$ 
```

Algorithmus C.1.4 : Berechnung der Wartezeit auf den letzten zu beruecksichtigenden Zubringer

Input : Zubringer z , Liste der tatsächlichen Ankünfte \mathcal{A}_z und aktuelle Wartezeit auf Zubringer $t_{Anschluss}$

```

1 forall the  $a \in \mathcal{A}_z$  do
2   if  $a.Zugnummer == E.Zugnummer + z.RelativeZugbeziehung$  then
3     ZubringerzugGefunden = true;
      // Findet die naechste Ankunft des Zubringerzuges in
      // der aktuellen Station (Standort des Abbringerzuges)
      // statt?
4     if  $a.Station == E.Station$  then
5       // Zeitpunkt, zu dem die Passagiere des Zubringers
5       // den Anschlusszug erreichen:
5        $t_z = a.Ankunftszeit + z.Umsteigezeit$ ;
6       if  $t_z > t_{AS}$  then
7         | RegistriereVerpasstenAnschluss( $E.Linienstation$ );
8       else
9         // Differenz zwischen dem Eintreffen der Passagiere
          // des
          // Zubringerzuges am Bahnsteig und der Ankunft des
          // Abbringers:
9          $t_{Anschluss} = \max\{t_{Anschluss}, t_z - E.Zeitpunkt\}$ ;
10 return  $t_{Anschluss}$ ;

```

Algorithmus C.1.5 : Suchen des aktuellen Zubringerzuges einer Linie

```

Input : Bezüglich ihrer Zeitpunkte sortierte Liste der Ereignisse  $\mathcal{E}$ 

// Initialisierung der Simulation:
// Liste der Umsteigebeziehungen  $\mathcal{C}$ 
1 forall the  $c \in \mathcal{C}$  do
2   | ErzeugeRelativeZugbeziehung( $c$ );
// Liste der Linien  $\mathcal{L}$ 
3 forall the  $l \in \mathcal{L}$  do
4   |  $e = l.$ ErzeugeStartEreignis();
5   |  $\mathcal{E}.$ SortiereEin( $e$ );
// Hauptteil der Simulation:
6  $t = 0$  ;
// Maximale Simulationszeit  $t_{max}$ 
7 while  $t < t_{max}$  do
8   | // Entnehme das erste Ereignis aus der Liste:
9   |  $aktEreignis = \mathcal{E}.$ front();
10  |  $t = aktEreignis.$ Zeitpunkt;
    | BearbeiteEreignis( $aktEreignis$ );

```

Algorithmus C.1.6 : Ablauf der ereignisgesteuerten stochastischen Simulation

Input : Ereignis E

```
// Pruefe, in welcher Phase der Simulation wir uns befinden:
1 if Messpunkt erreicht then
2   | ErzeugeStichprobenwert();
3   | if Mindestanzahl an Stichprobenwerten erreicht && Genauigkeit
   | hinreichend then
4     | return f = BerechneKostenFktswert();
5   | BeseitigeVerspaetungen();
6   | StarteVerspaetungsEntwicklung();
// Bearbeitung des aktuellen Ereignisses:
7 switch  $E.Type$  do
8   | case StartEreignis
9     | BearbeiteStartEreignis( $E$ );
10  | case AbfahrtsEreignis
11    | BearbeiteAbfahrtsEreignis( $E$ );
12  | case AnkunftsEreignis
13    | BearbeiteAnkunftsEreignis( $E$ );
```

Algorithmus C.1.7 : Bearbeitung eines Ereignisses

C.2. Komplexitätsreduzierung

```

Input : Verteilungsfunktion  $F_X$  der Zufallsvariablen  $X$ 

1 if  $F_X$  kann verschoben werden then
2    $\lfloor F_X$ .WirdVerschoben();
3  $\mu_k := \mathbf{E}[X^k]$  für  $1 \leq i \leq 3$ ;
4 if (6.18) oder (6.19) ist verletzt then
5    $\lfloor$  Modifiziere( $\mu_2, \mu_3$ );
6  $n^* := \min \left\{ n \in \mathbb{N} \mid n > \max \left\{ \frac{\mu_1^2}{\mu_2 - \mu_1^2}, \frac{\mu_2^2}{\mu_1 \mu_3 - \mu_2^2} - 1 \right\} \right\}$ ;
   // Momentenanpassung:
   // (i) Berechnung von Hilfsgrößen:
7  $A = \mu_1/n^*$ ;
8  $B = \mu_2/(n^* \cdot (n^* + 1))$ ;
9  $C = \mu_3/(n^* \cdot (n^* + 1) \cdot (n^* + 2))$ ;
10  $D_1 = A^2 - B$ ;
11  $D_2 = A \cdot B - C$ ;
12  $D_3 = B^2 - A \cdot C$ ;
13  $\delta = (D_2^2 - 4 \cdot D_1 \cdot D_3)^{1/2}$ ;
   // (ii) Bestimmung der Parameter der Hyper-Erlangverteilung:
14  $\lambda_1 = 2 \cdot D_1/(D_2 + \delta)$ ;
15  $\lambda_2 = 2 \cdot D_1/(D_2 - \delta)$ ;
16  $\alpha_1 = \lambda_1 \cdot (1 - \lambda_2 \cdot A)/(\lambda_1 - \lambda_2)$ ;
17  $\alpha_2 = 1 - \alpha_1$ ;
   // Berechnung der approximierten Verteilungsfunktion:
18  $\tilde{F}_X = \text{BildeHyperErlangVerteilungsfunktion}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ ;
19 if  $F_X$ .WurdeVerschoben then
20    $\lfloor \tilde{F}_X$ .VerschiebungUmkehren();
21 return  $\tilde{F}_X$ ;

```

Algorithmus C.2.1 : Komplexitätsreduktion

C.3. Optimierung

```

// Fuer alle Linienstationen:
1 forall the  $L \in \mathcal{L}$  und  $(S, L) \in \mathcal{S}_L$  do
2    $(S, L).AnzahlAnkuenfte = 0$  ;
3    $(S, L).AnzahlVerpassterAnschlusse = 0$  ;
4    $(S, L).GewichteAnkunftsVerspaetung = 0$  ;
5    $(S, L).GewichteteUmsteigeVerspaetung = 0$  ;
6 while Aktueller Zyklus ist nicht abgeschlossen do
7   // Einheitliche Taktzeit  $\tau$ 
8   if aktuellesEreignis.IstAnkunftsEreignis then
9     aktuellesEreignis.LinienStation.AnzahlAnkuenfte + = 1 ;
10    aktuellesEreignis.LinienStation.GewichteAnkunftsVerspaetung + =
        aktuellesEreignis.Verspaetung ·
        aktuellesEreignis.GewichtDerAussteiger ;
11    forall the Zubringer Z von Abbringer aktuellesEreignis.Zug do
12      if Z verpasst aktuellesEreignis.Zug then
13        aktuellesEreignis.LinienStation.AnzahlVerpasster-
            Anschlusse + = 1 ;
        aktuellesEreignis.LinienStation.GewichteteUmsteige-
            Verspaetung + =  $\tau \cdot Z.GewichtDerUmsteiger$  ;

```

Algorithmus C.3.1 : Berechnung der gewichteten Verspaetungen innerhalb eines Zyklus

Input : Simulierte Verspätungen: Ankunftsverspätungen der Züge sowie durch verpasste Anschlüsse erfahrene Verspätung der Passagiere

```

1  $Y_{\Sigma} = 0$ ;
  // Fuer alle Linienstationen:
2 forall the  $L \in \mathcal{L}$  und  $(S, L) \in \mathcal{S}_L$  do
3   if  $(S, L).AnzahlAnkuenfte > 0$  then
4      $Y_{\Sigma} += (S, L).GewichteteAnkunftsVerspaetung /$ 
       $(S, L).AnzahlAnkuenfte$  ;
5   if  $(S, L).AnzahlVerpassterAnschluesse > 0$  then
6      $Y_{\Sigma} += (S, L).GewichteteUmsteigeVerspaetung /$ 
       $(S, L).AnzahlVerpassterAnschluesse$  ;
7 return  $Y_{\Sigma}$ ;

```

Algorithmus C.3.2 : Auswertung eines Zyklus

Input : Berechnete Verteilungsfunktionen und zugehoerige Erwartungswerte

```

1  $Y_{\Sigma} = 0$ ;
2 forall the  $a = change(S, L, L') \in \mathcal{A}^{change}$  do
  // Einheitliche Taktzeit  $\tau$ 
3    $Y_{\Sigma} += \tau \cdot \mathbf{P}(\text{Anschluss wird verpasst}) \cdot a.Gewicht$  ;
4 forall the  $a = arr(S, L) \in \mathcal{A}^{arr}$  do
5    $Y_{\Sigma} += \mathbf{E}[Y(a)] \cdot (S, L).GewichtDerAussteiger$  ;
6 return  $Y_{\Sigma}$ ;

```

Algorithmus C.3.3 : Auswertung der analytischen Verspätungsberechnung

```

Input : Graph  $G(\mathcal{V}, \mathcal{E})$  und Startknoten  $v^* \in \mathcal{V}$ 

// Initialisierung des Graphen:
1 forall the  $v \in \mathcal{V}$  do
   // Länge des kürzesten Weges  $d_{\mathcal{V}}(v)$ ,  $v \in \mathcal{V}$ 
2    $d_{\mathcal{V}}(v) = \infty$ ;
   // Liste der Vorgänger auf kuerzesten Wegen  $L(v)$ ,  $v \in \mathcal{V}$ 
3    $L(v).LoescheElemente()$ ;
4    $AnzahlKuerzesteWege(v) = 0$ ;
5  $d_{\mathcal{V}}(v^*) = 0$ ;
   // Liste  $K$  aller Ecken des Graphen
6  $K.SortiereListeAufsteigend(d_{\mathcal{V}})$ ;
   // Hauptteil des Algorithmus:
7 while  $K.IstNichtLeer()$  do
8    $u = K.GibErstesElement()$ ;
9    $K.EntferneErstesElement()$ ;
   // Menge der Nachbarn  $\mathcal{N}(v)$ ,  $v \in \mathcal{V}$ 
10  forall the  $v \in \mathcal{N}(u)$  do
11  |   // Aktualisiere Nachfolger:
   |   Algorithmus C.3.5( $u, v$ );

```

Algorithmus C.3.4 : Modifizierter DIJKSTRA-Algorithmus zum Auffinden aller kürzester Wege

Input : Aktueller Knoten u und Nachfolgeknoten v sowie
Kantenbewertung $d_{\mathcal{E}} : \mathcal{E} \rightarrow \mathbb{R}_{\geq 0}$

// Wenn der aktuelle Weg kürzer ist:

```
1 if  $d_{\mathcal{V}}(u) + d_{\mathcal{E}}(u, v) < d_{\mathcal{V}}(v)$  then
2   |  $d_{\mathcal{V}}(v) = d_{\mathcal{V}}(u) + d_{\mathcal{E}}(u, v);$ 
3   |  $L(v).LoescheElemente();$ 
4   |  $L(v).FuegeHinzu(u);$ 
5   |  $H(v).LoescheElemente();$ 
6   |  $H(v).FuegeHinzu((u, v));$ 
7   |  $AnzKW(v) = AnzKW(u);$ 
8   |  $K.SortiereListeAufsteigend(d_{\mathcal{V}});$ 
9 else
10  | // Wenn der aktuelle Weg die gleiche Länge besitzt:
11  | if  $d_{\mathcal{V}}(u) + d_{\mathcal{E}}(u, v) == d_{\mathcal{V}}(v)$  then
12  |   |  $L(v).FuegeHinzu(u);$ 
13  |   |  $H(v).FuegeHinzu((u, v));$ 
13  |   |  $AnzKW(v) = AnzKW(v) + AnzKW(u);$ 
```

Algorithmus C.3.5 : Aktualisierung des Nachfolgeknotens $v \in \mathcal{N}(u)$

```

Input : Startknoten  $v_s$  (Metaknoten), Endknoten  $v_e$  (Metaknoten),
        Aktueller Knoten  $v_a$ , Verzweigungsgewicht  $g$ , relatives
        OD-Gewicht  $G_{v_s, v_e}$ 

// Update des aktuellen Verzweigungsgewichtes:
1  $g = g / v_a.$ AnzahlVorgaengerImKuerzestenWeg;
2 forall the Vorgaenger  $v_v$  des Knoten  $v_a$  do
    // Falls der Beginn der (kürzesten) Wege noch nicht erreicht
    // wurde:
3   if  $v_a \neq v_s$  then
        // Vom Vorgaenger zum aktuellen Knoten hinfuehrende Kante:
4      $e_a = (v_v, v_a)$ ;
        // Berechnung des Gewichts der Aussteiger:
5     if  $e_a.$ IstAussteigeKante then
6        $e_a.$ Gewicht =
        |  $e_a.$ Gewicht +  $G_{v_s, v_e} / v_a.$ AnzahlVorgaengerImKuerzestenWeg;
7     else
8       if  $e_a.$ IstChangeKante ||  $e_a.$ IstStoppKante ||  $e_a.$ IstFahrKante
        then
        | // Berechnung des Gewichts der Umsteiger bzw.
        | // Sitzenbleiber:
9       |  $e_a.$ Gewicht =  $e_a.$ Gewicht +  $G_{v_s, v_e} \cdot g$ ;
        // Rekursiver Aufruf:
10    | Algorithmus C.3.6( $v_s, v_e, v_v, g, G_{v_s, v_e}$ );

```

Algorithmus C.3.6 : Umlegen der Gewichte aus einer OD-Matrix

C.4. Verspätungsfortpflanzung

```
Input : Knotenvektor fuer sortierte Knoten  $\sigma$ 

// Initialisierung:
1 forall the  $V \in \mathcal{V}_M$  do
2    $\lfloor V.Berechne\_a_V()$ ;
3  $\mathcal{U} = \mathcal{V}_M$ ;
   // Es wird solange fortgefahren, bis alle Knoten einsortiert
   wurden:
4 while  $\mathcal{U} \neq \emptyset$  do
   | // Zunaechst waehlen wir einen noch nicht einsortierten
   | Knoten:
5    $V^* = \arg \min_{V \in \mathcal{U}} a_V$ ;
6    $\mathcal{U} = \mathcal{U} \setminus \{V^*\}$ ;
7    $\sigma.pushback(V^*)$ ;
   | // Update des Anteils noch unsortierter Vorgaenger:
8   forall the  $V \in \mathcal{S}uc(V^*)$  do
9      $\lfloor V.Berechne\_a_V()$ ;
10 return  $\sigma$ ;
```

Algorithmus C.4.1 : Berechnungsreihenfolge fuer Kreisvereinigungen

Literaturverzeichnis

- [Alt85] ALTIOK, T.: On the Phase-Type Approximations of General Distributions. In: *IIE Transactions* 17 (1985), Nr. 2, S. 110 – 116
- [Bau13] BAUM, D.: *Grundlagen der Warteschlangentheorie*. Springer Verlag, 2013
- [BB05] BREUER, L. ; BAUM, D.: *An Introduction to Queueing Theory and Matrix-Analytic Methods*. Springer, 2005
- [BCOQ92] BACCELLI, F.L. ; COHEN, G. ; OLDSER, G.J. ; QUADRAT, J.-P.: *Synchronization and Linearity: An Algebra for Discrete Event Systems*. Chichester : Wiley, 1992
- [BDS08] BOOM, T. Van d. ; DE SCHUTTER, B.: Model predictive control for switching max-plus-linear systems with random and deterministic switching. In: *Proceedings of the 17th IFAC World Congress*, 2008, S. 7660–7665
- [BGMHO11] BERGER, A. ; GEBHARDT, A. ; MÜLLER-HANNEMANN, M. ; OSTROWSKI, M.: Stochastic Delay Prediction in Large Train Networks / University Halle-Wittenberg, Institute of Computer Science. 2011. – Forschungsbericht
- [BHT05] BOBBIO, A. ; HORVATH, A. ; TELEK, M.: Matching three moments with minimal acyclic phase type distributions. In: *Stochastic Models* (2005), Nr. 21, S. 303 – 326
- [Bil86] BILLINGSLEY, P.: *Probability and Measure*. Wiley, 1986
- [BW] BÜKER ; WENDLER: Efficient modelling of delay distribution functions.
- [Bü10] BÜKER, T.: *Ausgewählte Aspekte der Verspätungsfortpflanzung in Netzen*, Rheinisch-Westfälischen Technischen Hochschule Aachen, Diss., 2010
- [Car99] CAREY, M.: Ex ante heuristic measures of schedule reliability. In: *Transportation Research* 33 (1999), S. 473 – 494

- [CC00] CAREY, M. ; CARVILLE, S.: Testing schedule Performance and Reliability for train stations. In: *Journal of the Operational Research Society* 51 (2000), S. 666 – 682
- [Cox55] COX, D.R.: A Use of Complex Probabilities In The Theory of Stochastic Processes. In: *Mathematical Proceedings of the Cambridge Philosophical Society* Bd. 51, 1955, S. 313 – 319
- [Dij59] DIJKSTRA, E. W.: A Note on Two Problems in Connexion with Graphs. In: *Numerische Mathematik* 1 (1959), S. 269–271
- [DLR77] DEMPSTER, A.P. ; LAIRD, N.M. ; RUBIN, D.B.: Maximum Likelihood from Incomplete Data via the EM Algorithm. In: *J. Royal Statistical Soc.* 39 (1977), S. 1 – 38
- [EF02] ENGELHARDT-FUNKE, O.: *Stochastische Modellierung und Simulation von Verspätungen in Verkehrsnetzen für die Anwendung bei der Fahrplanoptimierung*, Technische Universität Clausthal, Diss., 2002
- [Fel71] FELLER, W.: *An Introduction to Probability Theory and Its Applications*. Bd. 2. Wiley, 1971
- [Fle90] FLEISCHMANN, G.: *Leistungsbewertung paralleler Programme für MIMD Architekturen: Modellbildung und mathematische Analyse*, Friedrich Alexander-Universität Erlangen-Nürnberg, Diss., 1990
- [Fuh07] FUHR, Brigitte: *Robuste Flugplanung mittels stochastischer Modellierung*, Diss., 2007
- [FW98] FELDMANN, A. ; WHITT, W.: Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. In: *Performance Evaluation* 31 (1998), S. 245 – 279
- [GHM08] GOVERDE, R. M. ; HEIDERGOTT, B. ; MERLET, G.: A fast approximation algorithm for the Lyapunov exponent of stochastic max-plus systems. In: *Proceedings of the 9th International Workshop on Discrete Event Systems (WODES'08)*. Piscataway, 2008, S. 49 – 54
- [GHM09] GOVERDE, R. M. ; HEIDERGOTT, B. ; MERLET, G.: Railway Timetable Stability Analysis Using Stochastic Max-Plus Linear Systems. In: *Operations Research* (2009), S. 1 – 18
- [Gov98] GOVERDE, Rob M.: Synchronization Control of Scheduled Train Services to Minimize Passenger Waiting Times. In: *Proceedings of the 4th TRAIL Annual Congress*. Delft, 1998 (2)

- [Gov07] GOVERDE, Rob M.: Railway Timetable Stability Analysis Using Max-Plus System Theory. In: *Transportation Research* 41 (2007), Nr. 2, S. 179 – 201
- [GS11] GOERIGK, M. ; SCHÖBEL, A.: A Scenario-Based Approach for Robust Linear Optimization. In: *Proceedings of the 1st International ICST Conference on Practice and Theory of Algorithms in (Computer) Systems*, 2011
- [HKKS14] HARBERING, J. ; KIRCHHOFF, F. ; KOLONKO, M. ; SCHÖBEL, A.: Delay Propagation in Public Transport - Stochastic Modeling Meets Scenario Approach. 2014. – Forschungsbericht
- [HMG⁺06] HOFMAN, M. ; MADSEN, L. ; GROTH, J.J. ; CLAUSEN, J. ; LARSEN, J.: *Robustness and Recovery in Train Scheduling - a Simulation study from DSB S-tog a/s*. 2006
- [HS92] HERTEL, G. ; STECKEL, J.: Fahrzeitberechnung unter stochastischem Aspekt. In: *Eisenbahningenieur* 43 (1992), Nr. 5
- [Joh93] JOHNSON, M.A.: Selecting Parameters of Phase Distributions: Combining Nonlinear Programming, Heuristics, and Erlang Distributions. In: *ORSA Journal on Computing* 5 (1993), S. 69 – 83
- [JT88] JOHNSON, M. ; TAAFFE, M.: *The Denseness of Phase Distributions*. Research Memorandum No. 88-20, 1988
- [JT89] JOHNSON, M.A. ; TAAFFE, M.R.: Matching Moments to Phase Distributions: Mixtures of Erlang Distributions of Common Order. In: *Communications in Statistics - Stochastic Models* 5 (1989), S. 711 – 743
- [JT90a] JOHNSON, M.A. ; TAAFFE, M.R.: Matching Moments to Phase Distributions: Density Function Shapes. In: *Communications in Statistics - Stochastic Models* 6 (1990), S. 283 – 306
- [JT90b] JOHNSON, M.A. ; TAAFFE, M.R.: Matching Moments to Phase Distributions: Nonlinear Programming Approaches. In: *Communications in Statistics - Stochastic Models* 6 (1990), S. 259 – 281
- [JT91a] JOHNSON, M. A. ; TAAFFE, M. F.: A Graphical Investigation Of Error Bounds For moment-based Queueing Approximations. In: *Queueing Systems* 8 (1991), S. 295 – 312

- [JT91b] JOHNSON, M.A. ; TAAFFE, M.R.: An Investigation of Phase-Distribution Moment-Matching Algorithms for Use In Queueing Models. In: *Queueing Systems* 8 (1991), S. 129 – 148
- [Kle08] KLENKE, A.: *Wahrscheinlichkeitstheorie*. Springer Verlag, 2008
- [KM97] KRISHNAN, T. ; MCLACHLAN, G.J.: *The EM Algorithm and Extensions*. John Wiley and Sons, 1997
- [Loy62] LOYNES, R.M.: The stability of a queue with non-independent inter-arrival and service times. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 58 (1962), S. 497–520
- [LR99] LATOUCHE, G. ; RAMASWAMI, V.: *Introduction fo Matrix Analytic Methods in Stochastic Modeling*. ASA, 1999
- [LSS⁺07] LIEBCHEN, C. ; SCHACHTEBECK, M. ; SCHÖBEL, A. ; STILLER, S. ; PRIGGE, A.: *Computing Delay Resistant Railway Timetables*. 2007
- [LSS⁺10] LIEBCHEN, C. ; SCHACHTEBECK, M. ; SCHÖBEL, A. ; STILLER, S. ; PRIGGE, A.: Computing Delay Resistant Railway Timetables. In: *Computers & Operations Research* 37 (2010), S. 857 – 868
- [Mar80] MARIE, R.: Calculating Equilibrium Probabilities for $\lambda(n)/c_k/1/n$ Queues. In: *Proceedings of Performance*, 1980, S. 117 – 125
- [MB01] MIDDELKOOP, D. ; BOUWMAN, M.: Large Scale Train Network Simulations. In: *Proceedings of the 2001 Winter Simulation Conference*, 2001
- [MM07] MEESTER ; MUNS: Stochastic delay propagation in railway networks and phase-type distributions. In: *Transportation Research* 41 (2007), Nr. 2, S. 218 – 230
- [Mü90] MÜHLHANS, E.: Berechnung der Verspätungsentwicklung bei Zugfahrten. In: *Eisenbahntechnische Rundschau* (1990), Nr. 39, S. 465 – 468
- [Neu77] NEUMANN, Klaus: Dynamische Optimierung, Lagerhaltung, Simulation, Warteschlangen. In: *Operations Research Verfahren*. Carl Hanser Verlag, 1977
- [Neu81] NEUTS, Marcel F.: *Matrix-Geometric Solutions in Stochastic Models*. The Johns Hopkins University Press, 1981

- [Noo96] NOORDEEN, M.: *Stability analysis of cyclic timetables for a highly interconnected rail network*. Lausanne, EPFL, Diss., 1996
- [OHB03] OSOGAMI, T. ; HARCHOL-BALTER, M.: A closed-form solution for mapping general distributions to minimal PH distributions / School of Computer Science, Carnegie Mellon University. 2003. – Forschungsbericht
- [OHB04] OSOGAMI, T. ; HARCHOL-BALTER, M.: Closed Form Solutions for Mapping General Distributions to Minimal PH Distributions. In: *Performance Evaluation* (2004)
- [Rac91] RACHEV, S.T.: *Probability metrics and stability of stochastic models*. Jo, 1991
- [SBI12] SALIDO, M.A. ; BARBER, F. ; INGOLOTTI, L.: Robustness for a single railway line: Analytical and simulation methods. In: *Expert Systems with Applications* 39 (2012), S. 13305 – 13327
- [SC75] SAUER, C. ; CHANDY, K.: Approximate Analysis of Central Server Models. In: *IBM Journal of Research and Development* 19 (1975), S. 301 – 313
- [Sch74] SCHWANHÄUSSER, W.: *Die Bemessung der Pufferzeiten im Fahrplangefüge der Eisenbahn*, RWTH Aachen, Diss., 1974
- [Sch07] SCHÖBEL, A.: Integer programming approaches for solving the delay management problem. In: *Algorithmic Methods for Railway Optimization* 4359 (2007), S. 145–170
- [SS07] SHAKED, M. ; SHANTHIKUMAR, J.G.: *Stochastic Orders*. Springer Science & Business Media, 2007
- [TBT06] THÜMMLER, Axel ; BUCHHOLZ, Peter ; TELEK, Miklos: A Novel Approach for Phase-Type Fitting with the EM Algorithm. In: *IEEE Transactions on Dependable And Secure Computing* 3 (2006), Nr. 3, S. 245 – 258
- [TG97] TROGEMANN, G. ; GENT, M.: Performance analysis of parallel programs based on directed acyclic graphs. In: *Acta Informatica* (1997), Nr. 34, S. 411 – 428
- [TTH07] TAKEUCHI, Y. ; TOMII, N. ; HIRAI, C.: Evaluation method of robustness of train schedules. In: *QR of RTRI* 48 (2007), Nr. 4

- [Whi82] WHITT, W.: Approximating a Point Process by a Renewal Process, I: Two Basic Methods. In: *Operations Research* 30 (1982), Nr. 1, S. 125 – 147
- [Whi84] WHITT, W.: On Approximations for Queues, III: Mixtures of Exponential Distributions. In: *AT&T Bell Laboratories Technical Journal* 63 (1984), S. 163 – 175
- [WN04] WENDLER, E. ; NAEHRIG, M.: Statistische Auswertung von Verspätungsdaten. In: *Eisenbahn-Ingenieur-Kalender* (2004), S. 321–331
- [Yua06] YUAN, J.: *Stochastic modelling of train delays and delay propagation in stations*, Technische Universität Delft, Diss., 2006