

Stochastische Modelle für ATM-Konzepte

Von der
Fakultät für Elektrotechnik der
Rheinisch-Westfälischen Technischen Hochschule Aachen
zur Erlangung des akademischen Grades eines
Doktors der Ingenieurwissenschaften
genehmigte Dissertation

vorgelegt von
Diplom-Ingenieur Christoph Herrmann
aus Bensberg

Referent: Universitätsprofessor Dr.-Ing. B. H. Walke
Korreferent: Universitätsprofessor Dr. rer. nat. O. Spaniol
Tag der mündlichen Prüfung: 21. Juni 1994

AACHENER BEITRÄGE ZUR MOBIL- UND TELEKOMMUNIKATION

Herausgeber:

Universitätsprofessor Dr.-Ing. Bernhard Walke

Herrmann, Christoph:
Stochastische Modelle für ATM-Konzepte
1. Auflage Aachen:
Verlag der Augustinus Buchhandlung, 1995
(Aachener Beiträge zur Mobil- und Telekommunikation, Band 1)

ISBN 3-86073-380-X

© 1995 Christoph Herrmann

Verlag der Augustinus Buchhandlung
Pontstraße 66/68
52062 Aachen

Druck: Copy Team, Aachen
Gedruckt auf chlorfrei gebleichtem Papier

"D 82 (Diss. RWTH Aachen)"

Kurzfassung

Die Leistungsbewertung von auf ATM (Asynchronous Transfer Mode) basierenden zukünftigen Breitband-ISDN erfordert die Einbeziehung *korrelierter* stochastischer Prozesse, da einzelne Videoquellen wie auch die Überlagerung von nicht-Poissonschen (im zeitkontinuierlichen Fall) bzw. nicht-Bernoullischen (im zeitdiskreten Fall) Erneuerungsprozessen korrelierte Zellenströme erzeugen.

Zur Beurteilung von Verfahren für Nutzungsparameterüberwachung und Verbindungsannahmesteuerung werden zeitdiskrete Warteschlangenmodelle mit *korrelierten* Eingabeprozessen auf der Basis zweier völlig verschiedener Ansätze analysiert, die bekannte Modelle wesentlich erweitern: Mit der sog. "Restarbeit" können für einen beliebigen Semi-Markov-Eingabeprozess und deterministische Bedienzeit Belegung und Wartezeiten analysiert werden. Dagegen führt der sog. zeitdiskrete "Markovsche Ankunftsprozess" bei beliebigem unabhängigem Bedienprozess auf eine Systemmatrix vom M/G/1-Typ. Damit wird der Abgangsprozess vollständig als Semi-Markovsch charakterisiert und die Belegung zu Ankunftszeitpunkten bestimmt. Beide Verfahren werden zur Berechnung von Teilstrom-Dienstgüteparametern erweitert. Die Methode der Restarbeit erlaubt gegenüber dem M/G/1-Typ-Wartesystem neben der stationären Beschreibung auch die Betrachtung der *transienten* Situation, wie sie bei Zulassung einer neuen Verbindung bis zum Abklingen der "Einschwingvorgänge" zunächst auftritt. Ferner liefert sie einen exakten Algorithmus zur Bestimmung der bedingten Wahrscheinlichkeit, daß ein Nachrichtenstrom bei Verlust einer Zelle seine nächste Zelle wiederum verliert, d. h. ein Maß für die Häufigkeit des Verlustes aufeinanderfolgender Zellen einer Verbindung. Die deterministische Bedienzeit kann bei beiden Verfahren ein beliebiges Vielfaches der Zeiteinheit sein, wie zur Beschreibung der Situation an einem Konzentrador (Ausgangsbitrate kleiner als Eingangsbitrate) erforderlich. Die Wahrscheinlichkeit für gleichzeitiges Auftreten von Ankunft und Abgang wird erfaßt.

Die Methode der Restarbeit wird auch bei endlichen, zeitdiskreten Wartesystemen mit *beliebiger* Bediendauerverteilung erfolgreich eingesetzt. Der resultierende Zustandsraum ist dann allerdings nur für kleine Wartespeicher und wenige Bedienzeiten noch nicht zu groß, um eine numerische Lösung bei vertretbarem Rechenaufwand zuzulassen.

Da für einige Klassen von Eingabeprozessen beide Lösungsmethoden geeignet sind, können die aus den Modellen entwickelten Algorithmen gegenseitig "numerisch-analytisch" validiert werden: Für 16-stellige *double*-Zahlen in C-Programmen weichen die berechneten Wahrscheinlichkeiten in Exponentialdarstellung lediglich in den letzten 4 Ziffern voneinander ab. Auch in Computersimulationen wurde sehr gute Übereinstimmung gefunden.

Die entwickelten Modelle unterstützen den Vergleich von Verfahren für die Verbindungsannahmesteuerung, die effizienter als die Spitzenbitratenzuteilung sind. Es wird gezeigt, daß auch die positive Korrelation in einem Einzelstrom die Dienstgüteparameter einer Überlagerung wesentlich verschlechtern kann. Eine Pufferdimensionierung ist insbesondere für Videoquellen mit sehr niedriger Zellverlustwahrscheinlichkeit (10^{-9}) möglich.

Für Videoquellen mit variabler Bitrate (VBR) wird ein analytisches Modell zur Nachbildung der *gemessenen* periodischen Korrelationsfunktion, wie sie *ohne* das sog. *Frame-Buffering* auftritt, entwickelt. Daran zeigt sich, daß negative Werte innerhalb der ersten Periode der Korrelationsfunktion den Multiplexgewinn erheblich erhöhen können.

Da auch der sog. lokale Korrelationskoeffizient einiger wichtiger Zufallsgrößen (z. B. der Belegung) algorithmisch bestimmbar ist, dienen die Modelle als analytische Referenzen für die Computersimulation der o.g. Wartesysteme auf der Basis des LRE-Algorithmus, insbesondere zur Validierung von Verfahren zur Schätzung sehr seltener Ereignisse (z.B. des RESTART-LRE-Algorithmus).

Summary

Future broadband integrated services digital communication networks (B-ISDNs) will be based on Asynchronous Transfer Mode (ATM), a modified version of fast packet switching. In ATM networks, multiplexing is statistical so that a multiplexer can be represented by a queueing system with deterministic-service time due to the constant length of the packets (so-called *cells*) of each connection. Input streams (generated by both individual sources and superpositions thereof) mostly show cell interarrival times, which are not independent and therefore *correlated*. Network functions are necessary for guaranteeing QOS, that do not occur in STM (Synchronous Transfer Mode) applied in circuit switched networks: Preventive Congestion Control incorporated in Connection Admission Control (CAC) leading to a traffic contract, that describes the bandwidth requirements of a connection to be expected, and Policing (Usage Parameter Control) for monitoring every connection so that none of them exceeds the bandwidth agreed upon in the CAC phase. Modelling of mechanisms for those network functions is closely related to finite buffer queueing models with deterministic service times.

This thesis develops solutions for discrete-time finite buffer queueing systems with Semi-Markovian (therefore: correlated) inputs, thus extending known results considerably. These models allow the computation of QOS parameters (loss probability, mean delay and delay jitter) for traffic superpositions and for *individual* streams in a superposition, separately. Analysis is done by two completely different approaches, namely by using the "unfinished work" for a queue with deterministic service time and an arbitrary Semi-Markovian input, and by applying the M/G/1 paradigm due to Neuts, that allows dealing with the so-called Discrete-time Batch Markovian Arrival process and arbitrary renewal service process. For this class of input processes, the queue's output process is shown to be Semi-Markovian and can be characterized completely. In contrast to the M/G/1-type queue, the method of the unfinished work provides – besides stationary solutions – also a transient analysis for the situation shortly after admitting a new source until an equilibrium has been reached. Furthermore, it yields a conceptually exact algorithm for determining the probability that, given a stream loses a cell, the next cell of this stream is also lost, thus serving as a measure for the frequency of losing a string of cells. In both approaches, the deterministic service time can be an arbitrary multiple of the time-unit, as it is necessary for modelling the situation of a concentrator with an output bit rate lower than the input bit rate. Formulae for the probability of the simultaneous occurrence of an arrival and a departure are given. It is also outlined, how the method of the unfinished work can be used to deal with the more general finite queue with arbitrary Semi-Markovian input and arbitrary renewal service process. By means of input processes, that can be handled by both approaches, the algorithms derived from both models are validated numerically by each other. Numerical values of probabilities obtained by C-programs with double-numbers (16 figures) in exponential representation for both models only differ in the last 4 figures.

The models support the comparison of CAC procedures, that are more efficient than peak-rate allocation. Numerical results show that correlation in a single stream can affect QOS parameters of other streams in a superposition negatively. Buffer dimensioning is possible also for video sources with their very low loss probability of 10^{-9} . With respect to VBR video sources an analytical representation of the *measured* periodic correlation function of the counting process is given. This periodicity occurs in video sources without the so-called *Frame Buffering*. It is shown by analytical-numerical investigations that negative values of the correlation function for small lags increase the possible multiplexing gain.

The above queues serve as full reference models for simulation runs by means of the so-called LRE (Limited Relative Error) algorithm. Furthermore, implemented methods for rare event simulation (RESTART: REpetitive Simulation Trials After Reaching Threshold, combined with the LRE algorithm) have been validated by the models.

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Kommunikationsnetze der Rheinisch-Westfälischen Technischen Hochschule Aachen, gefördert durch das Graduiertenkolleg "Informatik & Technik" dieser Hochschule.

Herrn Professor Dr.-Ing. Bernhard H. Walke, dem Leiter des Instituts für Kommunikationsnetze, danke ich für die Anregung zu dieser Arbeit, ihre Betreuung und Förderung und die jederzeit gewährte Unterstützung.

Herrn Prof. Dr. rer. nat. Otto Spaniol, dem Leiter des Lehrstuhls Informatik IV und Sprecher des Graduiertenkollegs, danke ich für das Interesse an dieser Arbeit, die durch die Gegebenheiten des Graduiertenkollegs sehr intensive Zweitbetreuung und die engagierte Übernahme des Korreferats.

Analytische Modelle müssen der simulativen Überprüfung standhalten. Die intensiven Arbeiten von Herrn Prof. Dr.-Ing. (em.) Friedrich Schreiber und Frau Dr. rer. nat. Carmelita Goerg zur simulativen Erfassung von Verlustwahrscheinlichkeiten erlaubten eine sichere Abstützung der hier gewonnenen analytischen Ergebnisse und ihrer Implementierung in ausführbaren Programmen. Über den dabei entstandenen äußerst fruchtbaren Gedankenaustausch bin ich sehr froh. Herr cand.-Ing. Oliver Fuss hat mit großem Einsatz die Implementierung der Simulationsprogramme vorgenommen.

Herr Dr. rer. nat. Udo Krieger (FTZ Darmstadt) und Herr Prof. Dr. Chris Blondia (Universität Nijmegen, Niederlande) haben mir mancherlei wichtige Denkanstöße gegeben.

Allen wissenschaftlichen und nicht-wissenschaftlichen Kolleginnen und Kollegen am Institut wie auch im Graduiertenkolleg Informatik & Technik danke ich für die angenehme Arbeitsatmosphäre.

Inhaltsverzeichnis

1	Breitband-ISDN auf der Basis von ATM (Asynchronous Transfer Mode)	1
1.1	Diensteintegration	1
1.2	Übermittlungsverfahren digitaler Kommunikationsnetze	1
1.2.1	Vermittlungsverfahren	1
1.2.2	Verfahren zur Mehrfachausnutzung von physikalischen Leitungen	2
1.3	Begrenzungen des STM-Konzeptes	3
1.4	Merkmale von ATM-Netzen	4
1.5	Modifizierte und neue Netzfunktionen	5
1.5.1	Vorschläge für die Nutzungsparameterüberwachung (Usage Parameter Control)	7
1.5.2	Algorithmen für die Verbindungsannahmesteuerung (Connection Admission Control, CAC)	8
1.6	Zu erwartende Dienste in ATM-Netzen	9
2	Aspekte der stochastischen Modellierung	11
2.1	Gegenstände der Leistungsbewertung	11
2.1.1	ATM-Vermittlungsknoten	11
2.1.2	Die Verbindungsannahmesteuerung, Connection Admission Control (CAC)	12
2.1.3	Nutzungsparameterüberwachung mittels "Leaky-Bucket"	13
2.2	Zeitdiskrete und zeitkontinuierliche Modellierung	14
2.2.1	Zeitdiskrete Modellierung für ATM	16
2.3	Transiente und stationäre Modellierung	17
3	Semi-Markov-Prozesse (SMP) und Markov-Erneuerungsprozesse (MEP)	19
3.1	Stochastische Prozesse (SP)	19
3.2	Erneuerungsprozesse, Markov-Ketten und Markov-Prozesse	20
3.2.1	Die starke Markoveigenschaft, Stoppzeiten	21
3.3	Semi-Markov-Prozesse (SMP)	21
3.3.1	Zeitdiskrete SMP	23
3.3.2	Korrelationseigenschaften	25
3.3.3	Der spezielle Semi-Markov-Prozeß (SSMP)	27
3.4	Der zeitdiskrete Markovsche Ankunftsprozeß (DBMAP)	28
3.4.1	Spezialfälle	29
3.4.2	Der Zählprozeß eines DBMAP	33
3.5	Korrelationseigenschaften des DBMAP	35
3.5.1	Korrelationsfunktion des Zählprozesses	35
3.5.2	Korrelationsfunktion der Zwischenankunftszeiten	36
3.5.3	Dispersionsindizes	37
3.6	Andere Darstellungen des DMAP	37
3.6.1	Semi-Markovsche Darstellung der Überlagerung $DMAP_1 + DMAP_2$	38
3.7	Die Markov-Erneuerungsfunktion	39

4	Analyse zeitdiskreter Modelle mittels der Restarbeit	41
4.1	Das zeitdiskrete SMP ^[X] /D/1/N-Wartesystem	42
4.1.1	Die Zulassungsregeln Arrival First (AF) und Departure First (DF)	42
4.1.2	Restarbeit zu Ankunftszeitpunkten	43
4.1.3	Transiente Gleichungen	48
4.1.4	Belegung zu Ankunftszeitpunkten, Verlustwahrscheinlichkeiten und Wartezeit im stationären Fall	49
4.1.5	Ausdrücke im transienten Fall	50
4.1.6	Schätzwert für die bedingte Verlustwahrscheinlichkeit	51
4.1.7	Gleichzeitigkeit von Ankünften und Abgängen	52
4.2	Das SMP/D/1/N-System mit Null-Zwischenankunftszeiten	52
4.3	Die Restarbeit im zeitkontinuierlichen Fall	53
4.4	Ausblick: Verallgemeinerung auf die Fälle GI/GI/1/N und SMP/GI/1/N	54
4.4.1	Einige Definitionen	54
4.4.2	Restarbeit zum n -ten Ankunftszeitpunkt ohne Berücksichtigung der Ankunft	54
4.4.3	Übergangswahrscheinlichkeiten	55
4.4.4	Transiente Gleichungen	57
4.4.5	Wahrscheinlichkeitsfunktion der Belegung	59
4.4.6	Wahrscheinlichkeitsfunktion der Wartezeit	59
4.4.7	Semi-Markovscher Eingabeprozess	59
4.4.8	Diskussion der Ergebnisse	59
5	Zeitdiskrete Systeme vom M/G/1-Typ	61
5.1	Das zeitdiskrete DBMAP/G/1/N-Wartesystem	61
5.1.1	Der zu Abgangszeitpunkten eingebettete Markov-Erneuerungsprozess	62
5.1.2	Charakterisierung des Abgangsprozesses	66
5.1.3	Die Belegung zu einem endlichen Zeitpunkt t nach einem Abgang zum Zeitpunkt 0	68
5.1.4	Die Belegung bei beliebigem Zeitpunkt $t \rightarrow \infty$ im stationären Fall	69
5.1.5	Belegung zu Ankunftszeitpunkten	71
5.1.6	Gleichzeitigkeit von Ankünften und Abgängen	72
5.1.7	Die Wartezeit	75
5.2	Andere Verfahren: Bedienzeit gleich der Zeiteinheit	75
5.2.1	Blondia, Casals (1992): Departure First	76
5.2.2	Takine, Suda, Hasegawa (1993): Arrival First	76
5.2.3	Vergleich mit dem DBMAP/G/1/N-System bei $D = 1$	77
5.3	Systeme mit Bedienzeiten, die ganzzahlige Bruchteile der Zeiteinheit sind	77
5.3.1	Markov-Kette zu Ankunftszeitpunkten	78
6	Wartemodelle für die Verbindungsannahmesteuerung (CAC)	79
6.1	Die Restarbeit im $G_1^{[X]}+G_2/D/1/N$ -Wartesystem	80
6.1.1	Überlagerung zweier unabhängiger DMAPS	81
6.1.2	Transiente Gleichungen	82
6.1.3	Teilstrom-Dienstgüteparameter	84
6.1.4	Teilstrom-Verlustwahrscheinlichkeiten	85
6.1.5	Teilstrom-Wartezeiten	86
6.2	Die Überlagerung $SSMP_1^{[X]}+SSMP_2$	86
6.3	Bedingte Verlustwahrscheinlichkeit in einem ATM-Multiplexer	87

6.4	Das DBMAP ₁ +DMP2/G/1/N-Wartesystem	88
6.4.1	Teilstrom-Verlustwahrscheinlichkeiten	89
6.4.2	Teilstrom-Wartezeiten	90
6.4.3	Nachbildung von Prioritäten	90
7	Modellgestützte Untersuchung von ATM-Verkehrssituationen	91
7.1	Validierung der verschiedenen Modelle	91
7.2	Korrelationseinfluß auf QOS-Parameter eines Gesamtstromes	92
7.3	Korrelationseinfluß auf die QOS-Parameter der Teilströme	94
7.4	Korrelation des Abgangstromes	95
7.5	Effektive Bandbreite	95
7.6	Einfluß der Teilangebote auf Verluste	97
7.7	Gleichzeitigkeit von Ankünften und Abgängen	98
7.8	Schnelle Algorithmen für Markov-Ketten	100
7.8.1	Der Algorithmus von Grassmann et al. (1985)	100
7.8.2	Erweiterung des Grassmann-Algorithmus auf Blockmatrizen	101
7.8.3	Anpassung des Kramer-Algorithmus an M/G/1-Typ-Matrizen	101
8	Analytische Modelle für VBR-Videoquellen mit periodischer Korrelationsfunktion	103
8.1	Messungen an Videoquellen	103
8.2	Die Korrelationsfunktion der Anzahl von Ankünften eines DMAP ^[X]	105
8.3	Wahl der Matrizen C und D zur Erzeugung des periodischen Verlaufs	106
8.4	Korrelationsfunktionen, die den gemessenen Verläufen ähnlich sind	107
8.4.1	Korrelationsfunktion der Überlagerung zweier Quellen	107
8.4.2	Korrelationsfunktion der Zwischenankunftszeiten	108
8.5	Das Modell zum Szenenvergleich	110
8.6	Numerische Ergebnisse für (nicht-) periodische Korrelationsfunktionen	112
8.6.1	Diskussion der Ergebnisse	114
	Anhang	117
A	Beweise	118
A.1	Zum Beweis von Hilfssatz 6 (DF)	118
A.2	Formaler Beweis von Satz 2 (AF und DF)	119
A.3	Der Abgangsprozeß des DBMAP/G/1/N-Systems	120
B	Abkürzungen	122
C	Symbole	123
	Literaturverzeichnis	125

1 Breitband-ISDN auf der Basis von ATM (Asynchronous Transfer Mode)

1.1 Diensteintegration

Seit Beginn der 80er Jahre zeichnet sich im Bereich der Telekommunikation eine Entwicklung hin zu diensteintegrierenden Kommunikationsnetzen ab. Bis dahin war jedes System der Dialog- und Abrufkommunikation für eine einzige Informationsart (Sprache, Text, Daten, Bild) konzipiert: Neben dem weit verbreiteten Fernsprechnetz, das mittlerweile zum großen Teil digital arbeitet, gibt es seit ca. 50 Jahren ein weiteres öffentliches Wählnetz, das nach einem standardisierten Verfahren weltweit Textverkehr zwischen Fernschreibmaschinen erlaubt (*Telexnetz*). Dieses wurde in der Bundesrepublik zu einem eigenständigen integrierten Text- und Datennetz ausgebaut und um ein Paketvermittlungsnetz (Datex-P-Netz für Datenkommunikation) erweitert. Vor wenigen Jahren wurde der Telexdienst durch einen komfortableren Fernschreibdienst (*Teletex*) ergänzt, der mit wesentlich höherer Übertragungsgeschwindigkeit arbeitet und den Zeichenvorrat von Büromaschinen verwendet. Zwar können vorhandene Netze auch zur Übertragung von Informationsarten verwendet werden, für die sie ursprünglich nicht entwickelt wurden: Das Telefonnetz kann bei geeigneter Zusatzausstattung auch Daten, Faksimiles oder Fernwirkinformationen übertragen und einen Zugang zu Bildschirmtext-Informationsdatenbanken finden. Jedoch bleibt die Leistungsfähigkeit (Verbindungsaufbauzeit, Übertragungsdauer) der so bereitgestellten Dienste sehr beschränkt. Eine Zusammenarbeit der einzelnen Systeme zur Sprach-, Text-, Daten- und Bildkommunikation ist nur ausnahmsweise möglich, so daß z.B. die wünschenswerte Kombination von Sprach- und Bildübertragung für den Bildtelefondienst nur bei sehr geringer Bildqualität erreichbar ist.

Neben den Vorteilen für die Benutzer durch ein verbessertes Dienstangebot sprechen auch wirtschaftliche Aspekte für die Diensteintegration: Da über die Teilnehmeranschlußleitung mehrere Dienste abgewickelt werden können, ergibt sich eine bessere Ausnutzung dieser Leitungen, die bei öffentlichen Netzen einen hohen Prozentsatz der Gesamtkosten ausmachen. Verwaltung und Betrieb eines einzelnen diensteintegrierenden Netzes vereinfachen sich gegenüber mehreren Spezialnetzen.

1.2 Übermittlungsverfahren digitaler Kommunikationsnetze

1.2.1 Vermittlungsverfahren

Das Telefonnetz ist als Durchschaltenetz konzipiert: Für die Dauer einer Verbindung wird ein Kanal fest geschaltet, so daß auch in Pausen der Kanal nicht von anderen Verbindungen genutzt werden kann (geringe Kanalausnutzung). Es kommt vor, daß in einer Verbindungs-

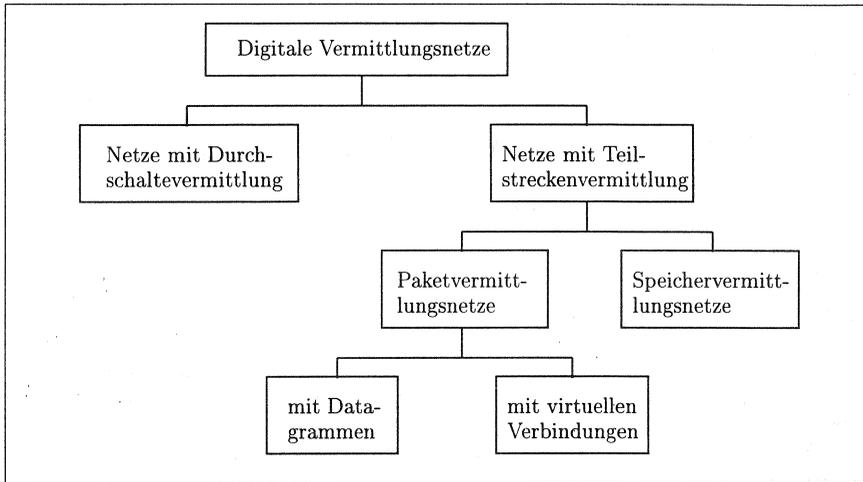


Abb. 1.1 Vermittlungsverfahren.

aufbauphase eine z.T. aufgebaute Verbindung wieder abgebaut werden muß, weil in einem nachfolgenden Übertragungsabschnitt kein Kanal frei ist.

Teilstreckenvermittlungsnetze, die vorwiegend zur Datenübertragung eingesetzt werden, umgehen diesen Nachteil dadurch, daß ganze Nachrichten (Speichervermittlung) oder Teilnachrichten (Paketvermittlung) unter Hinzufügung der Zieladresse von einem Vermittlungsknoten zum nächsten übertragen und dort zwischengespeichert werden, bis eine weitergehende Übertragung möglich ist.

Paketvermittlungsnetze können mit sog. *Datagrammen* oder mit *virtuellen Verbindungen* arbeiten: Im ersten Fall wird für jedes Paket der Weg durch das Netz neu gewählt, so daß sich die Paketreihenfolge beim Empfänger geändert haben kann. Das Netz erkennt keinen Zusammenhang zwischen den Paketen einer Verbindung, es werden keine Netzbetriebsmittel reserviert. Dagegen werden Pakete einer virtuellen Verbindung als zusammengehörig betrachtet und über einen einheitlichen Weg durch das Netz übertragen, so daß sie sich nicht überholen können. In einer Verbindungsaufbauphase werden Netzbetriebsmittel reserviert. Die Pakete einer virtuellen Verbindung in jeder der Übertragungsrichtungen erhalten eine Nummer, die den virtuellen Kanal kennzeichnet. Mittels dieser Nummern werden die Pakete in den Vermittlungsstellen weitergeleitet.

1.2.2 Verfahren zur Mehrfachausnutzung von physikalischen Leitungen

Das jeweilige Vermittlungsverfahren bestimmt weitgehend die Umsetzung der Multiplexfunktion.

Durchschaltensetze verwenden ein synchrones Zeitmultiplexverfahren (Abb. 1.2): Hierbei werden sog. Multiplexrahmen übertragen; jeder Rahmen besteht aus Zeitschlitzen, die eine feste Position innerhalb des Rahmens haben. Jedem Kanal ist ein fester Zeitschlitz zugeordnet. Dies bedeutet ein starres, "zyklisches" Zugriffsverfahren [116]. Durch Zusam-

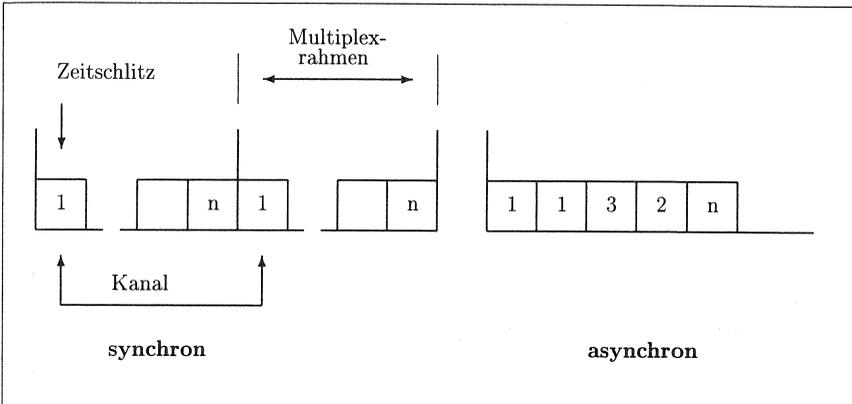


Abb. 1.2 Multiplexverfahren (Signalaufbau auf der Ausgangsleitung).

menschalten mehrerer Zeitschlitzes können Kanäle mit Bitraten bereitgestellt werden, die ganzzahlige Vielfache der Bitrate eines einzelnen Zeitschlitzes sind. Das Übermittlungsverfahren der Durchschaltenetze mit synchronem Zeitmultiplex wird mit STM (Synchronous Transfer Mode) bezeichnet.

Dagegen bedingt die Teilstreckenvermittlung ein asynchrones Zeitmultiplex [116]: Ist auf einem Teilabschnitt des Netzes eine Leitung frei, so können Nachrichten oder Pakete übertragen werden. Damit ist das Kanalzugriffsverfahren zufällig, und man spricht auch von *statistischem* Multiplex. Es können Bitraten bereitgestellt werden, die in weiten Grenzen variabel sind. Auch die Einspeisung von Quellen, die eine zeitlich veränderliche Bitrate erzeugen, ist möglich (VBR: Variable Bit Rate), während das synchrone Zeitmultiplexverfahren der Durchschaltenetze eine Umsetzung der VBR-Quelle in eine mit konstanter Bitrate erfordert (CBR-Quelle).

1.3 Begrenzungen des STM-Konzeptes

Die zur Zeit im Aufbau befindlichen schmalbandigen diensteintegrierenden Kommunikationsnetze (ISDN: Integrated Services Digital Network)(maximale Kanalbitrate 2 Mbit/s) basieren auf dem digitalisierten Fernsprechnetz mit 64 kbit/s-Fernsprechkanälen (8 bit-kodierte Tiefpaß-Signale mit einer Grenzfrequenz 4 kHz). Solche STM-Netze weisen eine Reihe von Nachteilen auf, die dem Ziel einer langfristigen Diensteintegration ("ein Netz für alle Dienste, auch für zukünftige Dienste") entgegenstehen:

- Es werden Kanäle mit verschiedenen, aber festen Bitraten bereitgestellt, die sich an den Anforderungen der vorhandenen Dienste orientieren. Die technische Entwicklung (Codierungsalgorithmen, VLSI-Technologie) reduziert fortlaufend die Bitrateanforderungen der Dienste. Das Netz kann daraus nur eingeschränkt Gewinn ziehen. Häufig werden infolge der fest vorgegebenen Kanalbitrate unnötig hohe Übertragungsraten verwendet werden müssen. Damit wird Kanalkapazität verschwendet.
- Das Durchschaltenkonzept ist sehr uneffizient, weil auch dann ein Kanal reserviert

bleibt, wenn keine Information übertragen wird (z.B. Sprechpausen beim Telefonieren in jeder Übertragungsrichtung). Dadurch werden vorhandene Leitungen unwirtschaftlich genutzt.

- Schließlich können STM-Netze nur Signale zeitlich konstanter Bitrate übertragen. In vielen Fällen liefern Kompressions- und Codierverfahren jedoch zeitlich veränderliche Bitraten, die zur Übertragung über STM-Netze durch zusätzliche Maßnahmen in eine zeitlich konstante Bitrate umgewandelt werden müssen. Gerade bei Videoquellen mit ihrem hohen Informationsgehalt lohnt sich die Ausnutzung der "natürlichen" variablen Bitrate eines Videocodecs in Verbindung mit dem statistischen Multiplex, in manchen Fällen scheint die Nutzung der variablen Bitrate sogar unerlässlich zu sein (HDTV-Verteilkommunikation mit einer erforderlichen Kanalbitrate von 130 Mbit/s). Eine LAN-Verknüpfung erfordert ebenfalls weit höhere Bitraten als die o.g. 2 Mbit/s.

Einige wesentliche nachteilige Merkmale der STM-Netze betreffen also die Teilnehmeranschlußleitung. Alle genannten Nachteile treten bei Paketvermittlungsnetzen nicht auf, und daher wurde vom CCITT¹ unter der Bezeichnung "Asynchronous Transfer Mode (ATM)" ein abgewandeltes Paketvermittlungsverfahren zur Übermittlung in zukünftigen breitbandigen ISDNs standardisiert.

1.4 Merkmale von ATM-Netzen

ATM stellt eine Variante der sog. schnellen Paketvermittlung [26] mit folgenden Hauptmerkmalen dar:

- Die Übertragung erfolgt verbindungsorientiert über virtuelle Kanäle und Gruppierungen solcher Kanäle (virtuelle Pfade). Die Zusammenfassung von Kanälen in Pfaden erlaubt die Verkleinerung der Vermittlungsinformation in jeder Zelle. Datagrammübertragung ist bisher nicht vorgesehen, allerdings gibt es Bestrebungen, ATM um einen *Connectionless Service* zu erweitern [34], um der reinen Datenübertragung näherzukommen.
- Die verwendeten Pakete (sie werden Zellen genannt) haben gleiche Längen und transportieren in einem Feld von 48 byte Nutzinformation. Der Paketkopf (5 byte) hat gegenüber herkömmlichen Paketvermittlungsnetzen stark reduzierte Aufgaben. Er trägt die Nummer des virtuellen Kanals (16 bit) (VCI: Virtual Channel Identifier), eines virtuellen Pfades (max. 12 bit) (VPI: Virtual Path Identifier), ein Prioritätsbit sowie weitere 4 bit für Konferenzschaltungen.

Die VCI und VPI werden auch zur Unterscheidung von Verkehrsklassen hinsichtlich ihrer unterschiedlichen Dienstgüteparameter eingesetzt, wodurch verschiedene Prioritätsklassen entstehen.

Verschiedene Gründe sprechen für die Verwendung relativ kurzer Zellen: Die Zellenlänge bestimmt die klein zu haltende Gesamtverzögerung mit, da der Paketierungsvorgang um so länger dauert, je größer das Informationsfeld ist. Die Zelle kann erst abgeschickt werden, wenn das Einlesen der Nutzinformation abgeschlossen ist. Bei Zellenverlust oder Fehlleitung einer Zelle bleibt der Informationsverlust bzw. die Störung durch Einfügung der Fremdzelle bei kleinen Zellen begrenzt. Schließlich macht die einheitliche Zellenlänge die Paketlängeninformation im Zellkopf überflüssig.

¹ Comité Consultatif Internationale Télégraphique et Téléphonique, seit 1992 ITU-TSS International Telecommunication Union – Telecommunication Standardization Section.

- Der Zellkopf wird fehlerkorrigierend kodiert (z.B. mit BCH-Codes). Dies ist die einzige Maßnahme zur Fehlersicherung. Auf zeitaufwendige Quittierungsmechanismen nach Übertragungen über Teilabschnitte (z.B. ARQ-Protokoll: Automatic Repeat Request) muß verzichtet werden, damit eine geringe Signalverzögerung für Sprachübertragung und Echtzeitanwendungen gewährleistet werden kann. Bei Sprachübertragung müssen Signalverzögerungen grundsätzlich unter 500 ms liegen, ohne Echoentzerrung sogar unter 25 ms. Möglich wird dies erst durch den Einsatz hochwertiger Übertragungsmedien mit sehr niedriger Bitfehlerrate (z.B. Lichtwellenleiter). Dem Nutzer bleibt überlassen, über mehrere zu sendende Zellen hinweg Fehlerkorrekturverfahren anzuwenden (FEC, *Forward Error Correction*), um eventuell verlorene Zellen zu rekonstruieren [101].
- Außer den Schutzfunktionen fehlen auch vom Netz getragene Verkehrsflußsteuerungsmechanismen etwa gegen Pufferüberlauf in Vermittlungsknoten. Schon vor der Annahme eines Verbindungswunsches wird vom Netz überprüft, ob die Zulassung einer weiteren Verbindung ohne Beeinträchtigung der übrigen Verbindungen möglich ist.
- "Header switching": Nachdem eine Zelle das Raumvielfach passiert hat, werden Teile der Vermittlungsinformation im Zellkopf durch neue Einträge ersetzt, die im nächstfolgenden Raumvielfach die Weiterleitung steuern.

1.5 Modifizierte und neue Netzfunktionen

Bei ATM-Netzen sind Schutzfunktionen in die Verbindungsaufbauphase verlagert (Präventive Schutzfunktionen gegen Stausituationen, "Preventive Congestion² Control"). Dies erfordert eine zentrale Verbindungsannahmesteuerung (CAC: Connection Admission Control), bei der das Netz aus einigen charakteristischen Kenngrößen der nachfragenden Verbindung unter Berücksichtigung der aktuellen Verkehrslage über eine Zulassung oder Ablehnung entscheidet. Hierzu müssen alle Teilabschnitte des Weges zwischen Sender und Empfänger mit einbezogen werden. Das Grundproblem der CAC-Funktion stellt die Festlegung geeigneter Kenngrößen dar, die eine neue Verbindung wie auch die bestehenden Verbindungen ausreichend beschreiben, sowie eines Kriteriums für eine Ablehnung bzw. Zulassung der neuen Verbindung.

Bei Zulassung einer Verbindung legt ein Verkehrsvertrag die Bitratenanforderungen der neuen Verbindung fest. Darin liegt eine konzeptionelle Schwierigkeit: Während der schon existierende Verkehr durch Verkehrsmessungen laufend charakterisiert werden kann, ist die neue Verbindung nur durch ihr Verhalten außerhalb des Netzes beschreibbar. Durch Verzögerungen und Zellenverlust in Vermittlungsknoten ändern sich jedoch die Quelleneigenschaften, d. h. (mehr oder weniger geringfügig) andere als die bei der Verbindungsaufnahme berücksichtigten Verkehrseigenschaften bestimmen das Verhalten der Verbindung auf dem Netz.

Das ATM-Netz überläßt einmal zugelassene Verbindungen sich selbst. Es muß jedoch sichergestellt sein, daß der vom Netz transportierte Zellenstrom jeder Verbindung die ursprünglich ausgehandelten Kenngrößen nicht überschreitet. Andernfalls kann die im Verbindungs-

²CCITT-Definition von "Congestion": "a state of network elements (e.g. switches, concentrators, transmission links) in which, due to traffic overload and/or control resource overload, the network is not able to guarantee the negotiated quality of service to the already established connections and to the new connection requests" [15].

aufbauvorgang für jede Verbindung vereinbarte Dienstgüte nicht mehr garantiert werden. Außerdem können Kunden wirtschaftliche Vorteile aus einer Kenngrößenüberschreitung ziehen, vor allem in Zeiten niedriger Netzauslastung, wenn die Überschreitung keine Beeinträchtigung der Übertragungsqualität nach sich zieht. Im Falle einer Überschreitung muß das Netz diese schnell erkennen und umgehend Maßnahmen ergreifen, die den Zellenstrom der "parasitären" Verbindung auf die vereinbarten Kenngrößen begrenzt (Nutzungsparameterüberwachung: Usage Parameter Control (UPC)³, Policing Function).

Eine naheliegende Gegenmaßnahme ist das Unterdrücken aller Zellen, die zur Kenngrößenüberschreitung beitragen, was häufig zur Minderung der Übertragungsqualität für die betreffende Verbindung führt. Es wurde auch vorgeschlagen, solche Zellen zu markieren und nur in Stausituationen zugunsten nicht markierter zu unterdrücken, dafür aber erhöhte Gebühren zu berechnen. Nach [79] ist damit eine merkliche Steigerung des Netzdurchsatzes möglich. Dem steht jedoch der erhöhte Aufwand (Gebührenberechnung, Pufferorganisation) gegenüber [99]. Diese Verkehrssteuerungsfunktion wird als "Selective Cell Discarding" bezeichnet. Die Dienstgüte einer Verbindung wird bestimmt durch:

- Zellenverlustwahrscheinlichkeit
- Zellenfehlleitungswahrscheinlichkeit
- Zellenverzögerung
- Zellenverzögerungsschwankungen.

Zellenverzögerungsschwankungen treten durch unterschiedlich lange Wartezeiten der einzelnen Zellen einer Verbindung in den Vermittlungsknoten auf. Sie können in der Depaketierungseinrichtung beim Empfänger zu einem gewissen Grad kompensiert werden. Dazu muß Synchronisationsinformation übertragen werden, etwa in Form von Zeitstempeln [26], p. 217: In festen Zeitabständen, innerhalb derer die Quellenbitrate annähernd konstant ist, schickt die Quelle einen Zeitstempel. Der Empfänger liest alle Zellen, die zwischen zwei aufeinanderfolgenden Zeitstempeln empfangen wurden, mit konstanter Zellrate aus einem Zwischenspeicher. Bei geringen Verzögerungsschwankungen auf seiten der Zeitstempel kann so beim Empfänger auch für Quellen mit variabler Bitrate die ursprüngliche "momentane" Bitrate rekonstruiert werden. Das Ausgleichen von Verzögerungsschwankungen bedeutet immer ein zusätzliches (planvolles) Verzögern des Zellenstroms, was bei Echtzeitanwendungen und Dialogkommunikation (insbesondere Sprachübertragung) die Rekonstruierbarkeit eng begrenzt. Auch die maximal tolerierbare Zellenverlustwahrscheinlichkeit ist von der Informationsart abhängig (s. Tab. 1.1).

Der Mechanismus für die Nutzungsparameterüberwachung sollte nah an der tatsächlichen Verkehrsquelle liegen, da im Laufe der Übertragung etwa durch Wartezeiten die Verkehrscharakteristika verändert werden können.

Die Funktionsfähigkeit der Verbindungsannahmesteuerung hängt wesentlich von der Leistungsfähigkeit der Nutzungsparameterüberwachung ab. Beide benötigen ausreichend genaue charakteristische Verkehrsdaten, wie [48]

³CCITT-Definition von UPC: "the set of actions taken by the network to monitor and control user traffic in terms of traffic volume and cell routing validity. Its main purpose is to protect network resources from malicious as well as *unintentional misbehaviour* which can affect the quality of service of other already established connections by detecting violations of negotiated parameters. Usage parameter control will apply only during the information transfer phase of a connection. Connection monitoring encompasses all connections crossing the user-network interface, including signalling" [15].

Dienst	Bitfehler- rate	Zellen- verlust-WKT	Zellenfehl- leitungs-WKT	Mittlere Verzögerung
Telefon	10^{-7}	10^{-3}	10^{-3}	25 ms
Datenübertragung	10^{-7}	10^{-6}	10^{-6}	50 ms
TV-Verteilkommunikation	10^{-6}	10^{-9}	10^{-9}	1000 ms
Musikübertragung	10^{-5}	10^{-7}	10^{-7}	1000 ms
Prozeßfernsteuerung	10^{-5}	10^{-3}	10^{-3}	1000 ms

Tab. 1.1 Zulässige Maximalwerte für einige Dienstgüteparameter (nach dePrycker, [26]).

Dienst	Bitrate (Mbit/s)	"Burstiness"
Datenübertragung (verbindungsorientiert)	1.5 ... 130	1 - 50
Datenübertragung (verbindungslos)	1.5 ... 130	1
Dokumentenübertragung	1.5 ... 45	1 - 20
Videotelefonie	1.5 ... 130	1 - 5
Breitband-Videotext	1.5 ... 130	1 - 20
TV-Verteilkommunikation	30 ... 130	1
HDTV-Verteilkommunikation	130	1

Tab. 1.2 Bitratenanforderungen einzelner Dienste (nach Händel, Huber, [48]).
Burstiness= Maximale Bitrate/ Mittlere Bitrate.

- mittlere Bitrate
- Spitzenbitrate
- "Burstiness"=Spitzenbitrate/mittlere Bitrate
- Dauer der Bitratenspitzen.

CCITT [16] sieht weitere Verkehrssteuerungsfunktionen vor: "Prioritätssteuerung" (Priority Control) und "Verkehrsformung" (Traffic Shaping). Durch Verkehrsformung sollen Verkehrsströme hinsichtlich ihrer kurzzeitig hohen Bitrate "geglättet" werden, um auf diese Weise Stausituationen zu vermeiden. Dazu bedarf es immer einer Zwischenspeicherung von Zellen. Die Prioritätssteuerung nutzt das Prioritätsbit zur sicheren Übertragung von "sehr wichtigen" Zellen, etwa solchen, die Zeitstempel tragen. Dazu werden in Stausituationen wartende Zellen geringer Priorität von solchen mit hoher verdrängt. Eine solche Verdrängung von Zellen kann die Zellenreihenfolge verändern [71], was durch geeignete Mechanismen unterbunden werden muß.

1.5.1 Vorschläge für die Nutzungsparameterüberwachung (Usage Parameter Control)

Verfahren zur Nutzungsparameterüberwachung müssen parasitäre Verbindungen schnell, d.h. mit wenig Meßinformation aufspüren können. Das birgt die Gefahr in sich, daß auch

korrekt arbeitende Verbindungen tangiert werden. Da jede Verbindung über ihre gesamte Dauer hinweg zu überwachen ist, muß das Verfahren einfach und kostengünstig sein. Darüber hinaus muß ein geeigneter Mechanismus die Zellen verschiedener Verbindungen unterscheiden können (Lesen von VPI und VCI).

In bisherigen Untersuchungen hat sich der sog. "Leaky-Bucket"-Mechanismus als am besten geeignet herausgestellt. Hierbei werden die von einer Verbindung eingespeisten Zellen in der Teilnehmervermittlungsstelle gezählt. Sofern der Zählerstand positiv ist, wird er in festen Zeitabständen um eine Einheit verringert (Rate r). Steigt die Zahl der eingespeisten Zellen stärker an, so wächst der Zählerstand trotz des Abwärtszählens. Ein zu definierender maximaler Zählerstand M bestimmt den Fall, daß mehr Zellen erzeugt wurden, als nach der Vereinbarung in der Verbindungsaufbauphase zulässig ist. Zellen, bei denen der Zählerstand den Wert M überschreitet, können von einer weiteren Übertragung ausgeschlossen werden. Wesentlich ist die Bestimmung einer geeigneten Kombination von r und M .

Der Leaky-Bucket-Mechanismus kann als Wartesystem mit endlichem Speicher, (geeignetem) beliebigem Ankunftsprozeß und deterministischem Bedienprozeß aufgefaßt werden. Die Bediendauer entspricht der Zeiteinheit, innerhalb der um eine Einheit abwärtsgezählt wird. Die Belegung repräsentiert den jeweiligen Zählerstand, das vollständig belegte Wartesystem bei einer Ankunft den Fall, daß Kenngrößen überschritten wurden. Der Abgangsprozeß dieses Wartesystems stimmt *nicht* überein mit dem Zellenstrom der überwachten Verbindung [99], und in einer Implementierung des Mechanismus werden zur Überwachung auch keine Zellen zwischengespeichert.

1.5.2 Algorithmen für die Verbindungsannahmesteuerung (Connection Admission Control, CAC)

Der einfachste Algorithmus benutzt als Kenngröße für eine Zulassungsentscheidung lediglich die maximale Bitrate der Verbindungen: Ist die Summe der maximalen Bitraten der auf einem Teilabschnitt bestehenden Verbindungen und der neuen Verbindung kleiner als die Bitrate des Übertragungsmediums dieses Teilabschnittes, so kann die Verbindung (auf dem Teilabschnitt) zugelassen werden. Hierzu eine beispielhafte Abschätzung: Betrachtet werde eine Übertragungsleitung, in die ein Vermittlungsknoten die Zellenströme zweier CBR-Verbindungen einspeist, von denen eine schon besteht, die andere um Zulassung nachfragt. Die Bitraten beider Verbindungen seien halb so groß wie die der Übertragungsleitung. Wenn man einen Pufferplatz vorsieht, so können beide Verbindungen ohne Zellenverlust nebeneinander bestehen. Der Speicherplatz ist notwendig, damit eine Zelle der einen Verbindung nicht verlorengeht, wenn sie während des Bedienvorganges einer Zelle der anderen Verbindung ankommt. Wegen der konstanten Bitrate tritt dies nur auf, wenn beide Quellen eine ungünstige Zeitlage zueinander haben, dann aber bei jeder Zelle. Bei b_{max} Verbindungen, die alle von verschiedenen Leitungen auf eine gemeinsame vermittelt werden und deren CBR-Summe die Bitrate der Übertragungsleitung ergibt, sind $b_{max} - 1$ Speicherplätze erforderlich, um einen Zellenverlust bei obiger CAC-Vorschrift auszuschließen. Die maximale Signalverzögerung einer Verbindung beträgt $b_{max} - 1$ Zellenlängen. Die Ausgangsleitung wird maximal ausgenutzt. Sind die oben angenommenen Bitraten maximale Bitraten von VBR-Quellen, so reichen die $b_{max} - 1$ Pufferplätze ebenfalls aus, um Verlust auszuschließen, jedoch wird die Ausgangsleitung nur teilweise genutzt. Offensichtlich kann man (unter Inkaufnahme einer geringen Zunahme der Verlustwahrscheinlichkeit) mehr als b_{max} VBR-Verbindungen auf die Ausgangsleitung vermitteln.

Die bisherigen Forschungsarbeiten haben noch nicht zu einem standardisierungsfähigen Vorschlag für die Verbindungsannahme geführt. In der Literatur werden u.a. diskutiert die sog.

Dienst	POTS	Bild-terminal	Faksimile	Daten		
Applikation	Sprache mit Pausen- unter- drückung	<i>Colour Catalog Production</i>	<i>Remote Newspaper Printing</i>	LAN Kopplung	<i>Inter Mainframe File Transfer</i>	<i>PC File Transfer</i>
Markt	privat/ geschäftlich	geschäftlich	geschäftlich	geschäftlich	geschäftlich	privat/ geschäftlich
Interessen- tenkreis	groß	klein	klein	klein	klein	klein/groß

Dienst	Telefon	<i>Image Mail</i>	<i>Audio</i>	Standbild- über- tragung	<i>Cable Image</i>	<i>Cable Text</i>
Applikation	<i>Voice Mail</i>	Technisches Design	Unter- haltung	<i>Home Shopping</i>	<i>Colour Image</i>	Text
Markt	privat/ geschäftlich	geschäftlich	privat	privat	privat/ geschäftlich	privat/ geschäftlich
Interessen- tenkreis	groß	klein	groß	groß	klein	klein

Tab. 1.3 Übersicht einiger Dienste und ihrer Interessengruppen nach [122].

„Sigma-Regel“ unter Ausnutzung der mittleren und Spitzenbitrate, bzw. des Erwartungswertes und der Varianz der Bitrate [117] (d. h. ohne Berücksichtigung eines Maßes für die Quellenkorrelation) und die sog. „effektive Bandbreite“ [66, 102, 100, 35], bei der die Korrelation definitionsgemäß einbezogen wird.

Fairness bei CAC Nicht ATM-spezifisch, sondern ebenso von Interesse für ISDN auf der Basis von STM ist die Annahmestrategie auf Verbindungsebene („Call Level“). Man sucht nach fairen Strategien, etwa in der Situation, daß bei Zulassung einer Verbindung mit niedriger Bitratenanforderung eine andere hochbitratige Verbindung abgewiesen werden muß und diese laufend infolge „kleiner“ Verbindungen nicht zum Zuge kommt [33].

1.6 Zu erwartende Dienste in ATM-Netzen

Die folgende Klassifizierung der Dienste in einem zukünftigen Breitbandnetz basiert auf der im Auftrag der Europäischen Gemeinschaft erstellten Studie [103] sowie einigen Einzelbeiträgen anderer Autoren [122, 92, 72]. Dabei wird die Frage beleuchtet, warum ein neues Netzkonzept auch aus wirtschaftlichen Gründen sinnvoll ist. Tab. 1.3 gibt eine Übersicht über die Größe des jeweiligen Interessentenkreises.

Der einfache Telefonverkehr (*Plain old telephone service*, POTS) macht z.Z. mit ca. 85% am Gesamtverkehrsaufkommen den größten Verkehrsanteil in öffentlichen Kommunikationsnetzen aus. Für diese Dienstart allein ist wegen der niedrigen Bitrate von 64 kbit/s ein ATM-Netz nicht erforderlich. Erweiterte Sprachdienste als Teil von Multi-Media-Diensten, wie Audio-Konferenzdienste oder HiFi-Musik-Verteilkommunikation werden hauptsächlich den Signalisierungsverkehr verstärken und das Netzmanagement vor größere Anforderungen stellen. Dasselbe gilt für die Mobilisierung von Sprachdiensten mit dem Ziel der Erreichbarkeit eines Teilnehmers im Festnetz unter einer persönlichen Telefonnummer unabhängig von

seinem Aufenthaltsort: Zur Ermittlung des Aufenthaltsortes tritt zusätzlicher Netzverkehr zwischen Datenbanken auf. Die Bereitstellung von "preiswerter" Mobilkommunikation, die mit dem Festnetz kooperiert, dürfte eine weitere Zunahme des Sprachverkehrsaufkommens und anderer niederbitratiger Dienste (z.B. Faksimileübertragung) nach sich ziehen.

Dagegen erfordert die Verbindung verschiedener LAN (Local Area Networks) erheblich höhere Bitraten. Schon innerhalb eines LAN findet der Informationsaustausch mit Ethernet-Kopplung bei 10 Mbit/s statt und die nächste Generation mit FDDI-Kopplung ermöglicht 100 Mbit/s. Die bisher vorhandenen Netze zur LAN-Kopplung bilden somit einen Engpaß, etwa wenn nach dem *Client-Server-Muster*⁴ Ressourcen in einem weiter entfernt liegenden LAN genutzt werden sollen. Bei Standbildübertragung für den Wissenschaftssektor (z.B. medizinische Bildübertragung mit 50 Mbit/s für Röntgenbilder⁵) oder den technischen Bereich (CAD-Bildübertragung) werden Bitraten im Bereich von 100 Mbit/s erforderlich. Für die breite Öffentlichkeit könnte die Standbildübertragung beim *Home shopping* interessant sein.

Als Hauptdienst neuer Breitbandnetze gilt aber die Bewegtbildübertragung. Man unterscheidet *Communication Video*, d.h. zwei oder mehr Teilnehmer kommunizieren "von Angesicht zu Angesicht", und *Video Distribution*, Videobild-Verteilkommunikation. Letzteres könnte dann einen *Video-on-Demand*-Dienst ermöglichen. Für den Bildtelefondienst *Video-phone* (nur zwei Teilnehmer) reichen infolge geeigneter Kompressionsverfahren ca. 2 Mbit/s. Die HDTV-Bildübertragung benötigt dagegen 130 Mbit/s, während Videocassettenqualität wiederum mit 2 Mbit/s erreichbar ist.

⁴Das *Client-Server-Muster* geht sogar so weit, daß der Datenbestand des Client in größerer Entfernung beim Server verwaltet wird (*diskless Client*), während ein Großteil des Computing lokal beim Client erfolgt. Dies erfordert natürlich eine sehr hohe Übertragungsgeschwindigkeit zwischen Client und Server [72, 103].

⁵Ultraschallbilder können dagegen schon über den Telefonkanal übertragen werden.

2 Aspekte der stochastischen Modellierung

2.1 Gegenstände der Leistungsbewertung

Bei Paketvermittlungsnetzen sind Multiplexer (Zeitvielfache) von zentralem Interesse für die Leistungsbeurteilung. Üblicherweise werden Verkehrsströme, die einen Multiplexer durchlaufen, durch stochastische Prozesse nachgebildet. Diese erzeugen entweder die Zwischenankunftszeiten der Pakete (bei ATM: Zellen) oder die (zufällige) Anzahl von Paketen innerhalb eines Zeitintervalls. Ob ein Zählprozeß verwendet wird oder eine Folge von Zwischenankunftszeiten, hängt wesentlich von der Analysemethode des resultierenden Wartesystems ab. Obwohl jede der beteiligten Verkehrsquellen eine deterministische Folge von Paketzwischenankunftszeiten erzeugt, macht man die Annahme, diese seien zufällig. Bei einer Mischung von Quellen wird eine zufällige Komponente realistisch, weil physikalische Effekte in den zusammenlaufenden Leitungen auf zufällige Verschiebungen der Ankunftszeitpunkte führen. Auch die für jede Verbindung sehr hohe Zahl zu übertragender Zellen läßt den stochastischen Standpunkt als gerechtfertigt erscheinen.

Warteschlangenmodelle sind zur Pufferdimensionierung bei ATM von sehr großer Bedeutung, da für Videoquellen mit variabler Bitrate eine sehr niedrige Zellenverlust-WKT¹ von 10^{-9} gefordert wird, die mit üblichen Verfahren der Computersimulation nicht verifiziert werden kann. Andererseits benötigt man für neue Simulationsverfahren zur Schätzung so seltener Ereignisse (z.B. RESTART, [114]) numerisch-analytische Referenzmodelle zur Beurteilung der Schätzgüte. Schließlich sind analytische Modelle – wiederum mit Blick auf Videoquellen – hinsichtlich des Vergleichs von Verfahren für die Verbindungsannahmesteuerung von Nutzen. Die hier entwickelten Wartesysteme sind auch in konventionellen Problemstellungen (etwa Untersuchung der Wartezeit) erheblich schneller als Simulationsverfahren. Allerdings können z.B. Wartesystemkaskaden, wie sie eine einzelne virtuelle Verbindung in einem ATM-Netz sieht, (bisher) nur simulativ untersucht werden.

2.1.1 ATM-Vermittlungsknoten

In einem ATM-Vermittlungsknoten werden von verschiedenen Eingangsleitungen ankommende Zellen auf dieselbe Ausgangsleitung vermittelt. Solange die Ausgangsleitung durch die Übertragung einer Zelle blockiert ist, müssen andere für diese Leitung bestimmte Zellen in einem Speicher warten. Damit ergibt sich am Eingang einer Ausgangsleitung (dem Multiplexer) ein Wartesystem mit deterministischer Bedienzeit infolge der gleichen Länge aller Zellen. Der Bedienprozeß repräsentiert das Auslesen einer Zelle aus dem Wartespeicher. Da verschiedene Eingangsleitungen Zellen für dieselbe Ausgangsleitung erzeugen, können mehrere Zellen gleichzeitig ankommen (s. Abb. 2.1). Bei w gleichzeitig ankommenden Zellen wird hier von einem "w-Batch"² gesprochen. Der Maximalwert b_{max} der Größe eines Batch

¹WKT: Wahrscheinlichkeit

²batch, engl. Haufen, Satz, Stoß.

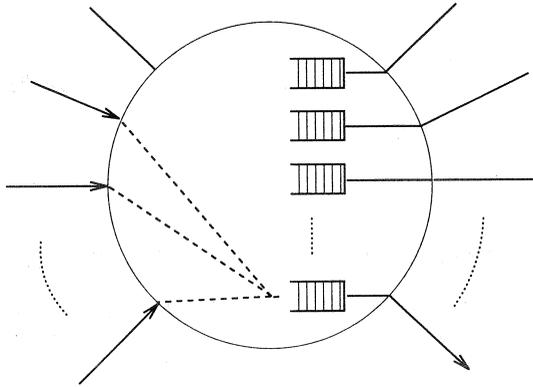


Abb. 2.1 ATM-Vermittlungsknoten mit mehreren Multiplexern.

ist vorgegeben durch die Anzahl von Eingangsleitungen, die Zellen für die betrachtete Ausgangsleitung erzeugen. Wenn Zellen gleichzeitig ankommen, muß die Reihenfolge festgelegt werden, in der sie in den Multiplexer gelangen. Aus Fairneßgründen (keine Eingangsleitung soll bevorzugt werden) sollte diese Auswahl der Reihenfolge für Zellen gleicher Priorität zufällig sein. Abb. 2.1 setzt die ausgangsseitige Pufferung innerhalb eines Vermittlungsknotens voraus. Es werden auch eingangsseitige und zentrale Speicherung diskutiert [26]. Bei eingangsseitiger Pufferung erhält man nach [65] als Durchsatz lediglich 0.586, während die ausgangsseitige Pufferung den Maximalwert von 1.0 erreicht.

Korrelierte Verkehrsströme Der Eingabeprozess des Multiplexers ist i.a. als (im stochastischen Sinne) korreliert anzusehen. Zum einen wurde für Einzelquellen (VBR Video) in Messungen positive Korrelation sowohl der Zwischenankunftszeiten als auch der Anzahl von Zellen in einem vorgegebenen Intervall gefunden [90, 46]. Andererseits führt die Überlagerung voneinander unabhängiger Erneuerungsprozesse (d.h. stochastisch unabhängige Zwischenankunftszeiten, daher unkorreliert) i.a. auf Zwischenankunftszeiten, die nicht mehr stochastisch unabhängig sind. Die einzigen Ausnahmen bilden der Poisson-Prozeß (zeitkontinuierlich, negativ-exponentiell verteilte Zwischenankunftszeiten) und der Bernoulli-Prozeß (zeitdiskret, geometrisch verteilte Zwischenankunftszeiten) [119]. Schließlich ist auch der Abgangprozeß eines Multiplexers häufig korreliert, was mit den Modellen in Kapitel 5 verifiziert wird. In Zellenströmen von Multi-Media-Diensten wie dem Videokonferenzdienst mit Sprache, Daten und Video ist infolge inhaltlicher Beziehungen ebenfalls eine Abhängigkeit im stochastischen Sinne zu erwarten.

2.1.2 Die Verbindungsannahmesteuerung, Connection Admission Control (CAC)

Ausgehend von dem Wartesystemmodell für einen ATM-Multiplexer kann das Vorgehen, das zur Annahme oder Ablehnung einer neuen Verbindung führt, folgendermaßen beschrieben werden:

Unter den topologisch möglichen Wegen zwischen zwei Teilnehmern, die über das Netz verbunden werden sollen, sind diejenigen zu bestimmen, auf denen die Dienstgüte (QOS:

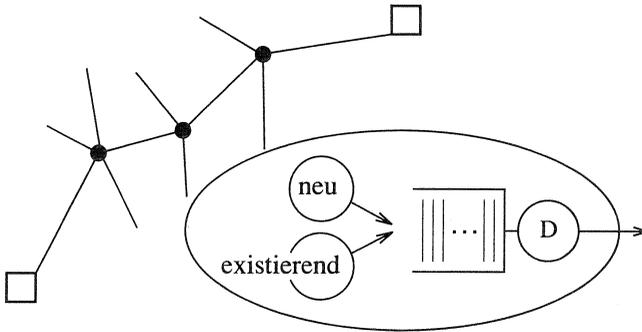


Abb. 2.2 Ein topologisch möglicher Weg zwischen zwei Teilnehmern und das resultierende Wartesystem am Eingang jedes Teilabschnitts.

Quality of Service) sowohl der schon existierenden wie auch der neuen Verbindung garantiert werden kann, wenn die neue Verbindung hinzukommt. Die Dienstgüte wird dabei charakterisiert durch Zellenverlust, mittlere Verzögerung und Verzögerungsschwankungen. Näherungsweise kann dazu jeder Teilabschnitt des Weges separat untersucht werden. Existierender Verkehr und die neue Verbindung werden durch stochastische Prozesse zur Nachbildung der Zwischenankunftszeiten repräsentiert und überlagert (s. Abb. 2.2). Zur getrennten Ermittlung der Dienstgüteparameter beider Ströme benötigt man ein Wartesystem, mit dem Verlust-WKTN und Wartezeiten für die Teilströme bestimmt werden können. Hierbei sollten auch Prioritäten berücksichtigt werden können, denn im Multiplexer mischen sich Ströme mit verschiedenen Dienstgüteanforderungen. Die Unterscheidung der Ströme erfolgt über die VCI und VPI.

Die Näherung bei dieser Vorgehensweise liegt darin, daß Einflüsse des schon existierenden Verkehrs auf die neue Verbindung nach Durchlaufen mehrerer Vermittlungsknoten ignoriert werden; jeder Zellenverlust und jede Verzögerung durch Wartezeiten in einem Vermittlungsknoten verändern den Zellenstrom der neuen Verbindung *hinter* diesem Vermittlungsknoten, also auf einem nachfolgenden Teilabschnitt. Dagegen ist der existierende Verkehr auf jedem Teilabschnitt durch Messung charakterisierbar. Allerdings dürften die Veränderungen des Zellenstroms beim Durchlaufen des Netzes klein bleiben, da sehr kleine Verlust-WKTN und Wartezeiten gefordert werden. Das so beschriebene Modell für die Vorgänge bei der Verbindungsannahme ist allerdings für einen On-line-Betrieb nicht geeignet, da Rechenzeiten für die Ermittlung der Dienstgüteparameter häufig zu groß werden. Daher müssen einfache Kenngrößen zur Schätzung der Verkehrseigenschaften verwendet werden, die mit einfachen Rechenoperationen eine Annahmeentscheidung ermöglichen. Solche einfachen Verfahren für CAC, wie die sog. "effektive Bandbreite" [100, 66, 35] oder die "Sigma-Regel" [117] beschränken sich ebenfalls auf die separate Untersuchung aller Teilabschnitte.

2.1.3 Nutzungsparameterüberwachung mittels "Leaky-Bucket"

Auch das mit "Leaky-Bucket" bezeichnete Verfahren zur Nutzungsparameterüberwachung führt auf ein Wartesystem mit deterministischer Bedienzeit. Es kann mit den in dieser Arbeit entwickelten Verfahren ebenfalls behandelt werden. Dies ist aber nicht Hauptgegenstand der Arbeit.

	zeitkontinuierlich	zeitdiskret
\mathbf{A}_k	$\mathbf{P}(k, D)$	$\mathbf{P}(k, D)$
$\mathbf{P}(k, t)$	Differentialgleichung $\frac{d}{dt}\mathbf{P}(k, t) = \sum_{j=0}^k \mathbf{P}(j, t)\mathbf{D}_{k-j}$	rekursive Summe $\mathbf{P}(k, t+1) = \begin{cases} \mathbf{P}(0, t)\mathbf{D}_0 & \text{für } k=0 \\ \sum_{v=0}^k \mathbf{P}(v, t)\mathbf{D}_{k-v} & \text{für } k \geq 1 \end{cases}$
Semi-Markov-Kern der Zwischenankunftszeiten	$\mathbf{Q}(x) = \int_0^x e^{\mathbf{D}_0 u} du (\widehat{\mathbf{D}} - \mathbf{D}_0)$	$\mathbf{Q}(\ell) = \sum_{k=1}^{\ell} \mathbf{D}_0^{k-1} (\widehat{\mathbf{D}} - \mathbf{D}_0)$
$\widehat{\mathbf{D}} = \sum_{v=0}^{b_{max}} \mathbf{D}_v$	infinitesimaler Generator eines Markov-Prozesses	Übergangsmatrix einer Markov-Kette

Tab. 2.1 Zeitkontinuierliche und zeitdiskrete Wartemodelle mit korreliertem Eingabeprozess "Markovscher Ankunftsprozess" im Vergleich. Einzelheiten über Semi-Markov-Prozesse und die hier aufgeführten Matrizen finden sich in Kapitel 3.

2.2 Zeitdiskrete und zeitkontinuierliche Modellierung

ATM-Multiplexer sind endliche Wartesysteme mit deterministischem Bedienprozeß und FIFO³-Bedienstrategie. Grundsätzlich kann ihr stochastisches Verhalten sowohl mit den "klassischen" zeitkontinuierlichen als auch mit zeitdiskreten Wartemodellen beschrieben werden. Der einfachste klassische Ansatz ist das M/G/1/N-Wartesystem, wobei der G-Bedienprozeß als deterministisch gewählt wird. Dieser Fall kann durch Bestimmung der stationären Zustands-WKTn der zu Abgangzeitpunkten eingebetteten Markov-Kette (MK) [20], p. 193 und anschließende Transformation auf Ankunftszeitpunkte [86], p. 29 numerisch behandelt werden.⁴ Damit können aber keine korrelierten Eingabeprozesse nachgebildet werden, da M für den Poisson-Prozeß steht. Ein neueres zeitkontinuierliches Modell, das korrelierte Eingabeprozesse zuläßt, stellt das BMAP/G/1/N-Wartesystem nach [75] (dort für $N = \infty$) dar.⁵

Eine Schwierigkeit des BMAP/G/1/N-Modells ist die Komplexität des Lösungsverfahrens. Diese ist beim zeitdiskreten Äquivalent, dem sog. D[B]MAP (Discrete-time [Batch] Markovian Arrival Process), der in Kapitel 5 als Eingabeprozess verwendet wird, um einiges

³First-In-First-Out.

⁴Dieses Verfahren findet auch in dieser Arbeit bei dem zeitdiskreten Modell mit *korreliertem* Eingabeprozess in Kapitel 5 Anwendung.

⁵[B]MAP: [Batch] Markovian Arrival Process. Beim BMAP wird zu einem Ankunftszeitpunkt eine Batch-ZV gezogen. Dieses Modell ist äquivalent zu dem des N/G/1/N-Wartesystems, das in der etwa 10 Jahre älteren Arbeit [96] für $N = \infty$ behandelt wurde und einen weit komplizierteren Formalismus verwendet. Dabei steht "N" für Neuts-Prozeß [87, 86]. Arbeiten für $N < \infty$ existieren bisher nur in dieser komplizierten Notation, [6].

niedriger. Die Teilmatrizen \mathbf{A}_k , aus denen die jeweilige Systemmatrix aufgebaut ist (vgl. Tab. 2.1 und die Ausführungen zum DBMAP in Kapitel 3) können im zeitkontinuierlichen System nur durch numerische Integration berechnet werden, da für ihre Bestimmung eine Differentialgleichung der Matrixfunktion $\mathbf{P}(k, t)$ gelöst werden muß. Schon hierbei treten Genauigkeitsprobleme [75] auf, die mit Blick auf Verlust-WKTswerte in der Größenordnung von 10^{-9} gravierend sein können. Im zeitdiskreten Fall wird diese Matrixfunktion durch eine einfache rekursive Summe erzeugt.

Für endliche Bedienzeiten ist die Anzahl k von Kunden (ATM-Zellen) innerhalb eines Intervalls der Länge D im zeitkontinuierlichen Fall prinzipiell unbegrenzt, während im zeitdiskreten System leicht ein Maximum vorgegeben werden kann ($b_{max} < \infty$ verlangen). Damit liegt man näher am realen System.

Infolge der Struktur des Semi-Markov-Kerns der Zwischenankunftszeiten des BMAP [84] ist es anscheinend unmöglich, einen BMAP so zu parametrisieren, daß er nur Zwischenankunftszeiten erzeugt, die ein vorgegebenes Minimum nicht unterschreiten, oder letzteres wenigstens ausreichend genau anzunähern. Der im Semi-Markov-Kern enthaltene Matrix-exponentialausdruck⁶ (vgl. Tab. 2.1)

$$e^{\mathbf{D}_0 u} := \sum_{\nu=0}^{\infty} \frac{1}{\nu!} (\mathbf{D}_0 u)^\nu \quad (2.1)$$

mit seiner spektralen Zerlegung [20], p. 368

$$e^{\mathbf{D}_0 u} = \sum_{\nu=1}^m e^{\lambda_\nu u} \mathbf{B}_\nu \quad (2.2)$$

deutet darauf hin. In (2.2) bedeuten $\lambda_\nu, \nu = 1, \dots, m$ Eigenwerte von \mathbf{D}_0 und $\mathbf{B}_\nu, \nu = 1, \dots, m$ Matrizen, die sich aus der Diagonalisierung von \mathbf{D}_0 ergeben. Die Nachbildung einer Einzelquelle mit einer *maximalen* Bitrate erfordert aber dieses *Minimum* der Zwischenankunftszeiten.

Ein anderer, konzeptioneller Nachteil besteht darin, daß die Überlagerung von zwei MAPs keinen BMAP, sondern wieder einen MAP ergibt, d.h. es können nicht gleichzeitig zwei Ankünfte auf einmal auftreten. Dagegen tritt dies bei der Überlagerung zweier realer Einzelquellen immer dann auf, wenn sie gleichzeitig eine Zelle erzeugen. Dies darf aber nicht überbewertet werden, da das Verhalten des Wartesystems sich nicht wesentlich ändert, wenn Ankünfte exakt zeitgleich sind oder einen relativ kleinen Zeitabstand zueinander haben. Es werden sogar Modelle diskutiert, die die Zwischenankunftszeiten als normalverteilt annehmen, dabei also auch negative Zwischenankunftszeiten zulassen, obwohl diese sinnlos sind. Durch geeignete Parametrisierung werden negative Zwischenankunftszeiten sehr selten [1]. Ein weiterer zeitkontinuierlicher Ansatz ist das sog. "Markov-modulated-fluid-flow"-Modell⁷ [2, 123]. Der ATM-Multiplexer wird dabei als Flüssigkeitsbehälter aufgefaßt, der mit konstanter Strömungsrate, diese entspricht der konstanten Bedienzeit, entleert wird. Der Ankunftsprozeß ist hierbei eine Strömung, deren Rate von einem Markov-Prozeß gesteuert wird. Daher gibt es in diesem Modell keine diskreten Ankunftszeitpunkte wie etwa beim Poisson-Prozeß. Dessen *Ankunftsrate* ist der Kehrwert der mittleren Zwischenankunftszeit. Man interpretiert den Flüssigkeitsstand des Behälters als Wartezeit, und deren VF ist dann Gegenstand der Analyse. Sie zeigt als asymptotisches Verhalten negativ-exponentiellen Verlauf.

⁶engl. matrix exponential

⁷bisher anscheinend nur mit unendlichem Speicher

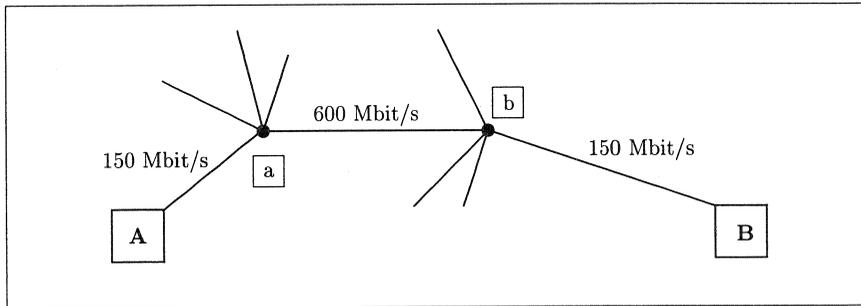


Abb. 2.3 Eine Verbindung zwischen den Teilnehmern A und B auf 3 Teilabschnitten zur Erläuterung der jeweiligen Bedienzeiten und der kleinsten auftretenden Zwischenankunftszeit.

Als technische Systeme sind ATM-Multiplexer ihrer Natur nach zeitdiskret⁸. Als Zeiteinheit kommt die Dauer eines Bit in Frage. Daher ist die zeitkontinuierliche Modellierung nur eine Näherung. Jedoch liefert auch ein praktikables zeitdiskretes System wieder nur eine Näherung, da man die Zeiteinheit im Modell erheblich größer wählen muß als die Länge eines Bit. Andernfalls wachsen die zur Lösung erforderlichen Zustandsräume auf eine Größe, die nicht mehr handhabbar ist.

Für den allgemeinen Fall des BMAP gibt es offenbar (noch) keine Verfahren zur Anpassung der Modellparameter an reale Quellen, jedoch für Sonderfälle wie den Markov-modulierten Poisson-Prozeß MMPP [81, 27]. Für das zeitdiskrete Analogon (DBMAP) wird in [10] ein Verfahren zur Parametrisierung für Videoquellen vorgeschlagen; in [58] und Kapitel 8 dieser Arbeit werden gemessene periodische Korrelationsfunktionen für VBR-Videoquellen mittels eines DMAP nachgebildet. Für einen SSMP (Special Semi Markov Process) mit zwei Zuständen steht nach [30] ein Verfahren zur Anpassung der Modellparameter zur Nachbildung einer beliebigen Quelle zur Verfügung.

2.2.1 Zeitdiskrete Modellierung für ATM

Viele Arbeiten verwenden zeitdiskrete Wartesysteme zur Untersuchung von ATM-Multiplexern [10, 13, 32, 51, 83]. Meist wird die Bedienzeit gleich der Zeiteinheit gesetzt. Dies erlaubt einfachere Lösungsverfahren.

Allerdings werden auch Bedienzeiten benötigt, die Vielfache der Zeiteinheit sind. Zur Erläuterung wird dazu eine Verbindung von A nach B in Abb. 2.3 betrachtet. Die Verbindung sieht Leitungen mit Bitraten der Größe 150 Mbit/s, 600 Mbit/s und wieder 150 Mbit/s. Für die qualitative Betrachtung wird angenommen, daß nur diese Verbindung auf den beteiligten Netzabschnitten existiert. Die Verbindung arbeite bei A mit ihrer Spitzenbitrate von 50 Mbit/s. Die Zeit für das Paketieren bestimmt den absoluten Zeitabstand zwischen zwei

⁸In [62], p. 190 wird der Standpunkt vertreten, daß der zeitdiskrete Ansatz zur Modellierung technischer Systeme meist der "natürlichere" sei. In jedem technischen System gibt es eine kleinste Zeiteinheit, die die Auflösungsgrenze darstellt, so daß zwei nicht gleichzeitige Ereignisse dennoch als gleichzeitig erscheinen. Trotzdem benötigt man prinzipiell auch kontinuierliche Zufallsvariablen, etwa zur Schätzung des kontinuierlichen Parameters, der ein (z.B. diskretes) Experiment bestimmt. Beispiel: Bestimmung des Parameters $p \in (0, 1)$ eines Bernoulli-Experiments mit unsymmetrischer Münze. Hierbei verknüpft dann die Bayessche Umkehrformel kontinuierliche und diskrete Zufallsvariablen [104].

nacheinander abgeschickten Zellen, da eine Zelle erst abgeschickt werden kann, wenn ihr Informationsfeld vollständig beschrieben worden ist. Die Zeit für das Paketieren und damit der kleinste Zeitabstand nacheinander abgeschickter Zellen beträgt $\tau = (48 \cdot 8 \text{ bit}) / (50 \text{ Mbit/s})$. Dieser Zeitabstand bleibt auf allen Teilabschnitten gleich groß, da angenommen wurde, daß sonst kein Verkehr auf den tangierten Teilabschnitten existiert. Wird als Zeiteinheit der Übertragungsstrecke die Bedienzeit $(53 \cdot 8 \text{ bit}) / (600 \text{ Mbit/s})$ im Multiplexer, der auf die Leitung mit 600 Mbit/s führt, zu $D = 1$ gewählt, so sieht dieser Multiplexer als kleinste Zwischenankunftszeit etwa 11 Zeiteinheiten $((48/50)/(53/600) \approx 11)$. Der Multiplexer am Eingang der 150-Mbit/s-Leitung, die nach B führt, hat dann eine Bedienzeit von 4 Zeiteinheiten (das Auslesen dauert hier viermal so lange wie bei der 600-Mbit/s-Leitung), sieht aber auch als kleinste Zwischenankunftszeit 11 Zeiteinheiten. Kommt im Multiplexer der nach B führenden Leitung noch Überlagerungsverkehr anderer Verbindungen hinzu, so zeigt dieser Überlagerungsprozeß aus stochastischen Gründen als kleinste Zwischenankunftszeit die *Zeiteinheit*.

Besonderheiten zeitdiskreter Wartesysteme In zeitdiskreten Wartesystemen können Abgänge und Ankünfte mit positiver WKT gleichzeitig stattfinden. Im zeitkontinuierlichen Fall ist das nur mit WKT Null möglich [44].

Im Falle der Gleichzeitigkeit muß (bei endlichem Wartesystem) zusätzlich festgelegt werden, in welcher Reihenfolge Ankunft und Abgang stattfinden sollen. Trifft etwa eine Ankunft auf ein volles System und findet gleichzeitig ein Abgang statt, so geht die Ankunft verloren, wenn sie unmittelbar vor dem Abgang stattfindet. Dagegen wird wieder ein Platz frei, wenn der Abgang als unmittelbar vor der Ankunft stattfindend definiert wird. Daher unterscheidet man die zwei Zulassungsregeln [62, 45]

- **Arrival first (AF):** Bei Gleichzeitigkeit von Ankunft und Abgang sieht der ankommende Kunde⁹ den abgehenden noch.
- **Departure first (DF):** Bei Gleichzeitigkeit von Ankunft und Abgang verläßt der abgehende Kunde das System unmittelbar vor der Ankunft.

In der älteren Arbeit [62], p. 193 wird DF mit "early arrival system" und AF mit "late arrival system" bezeichnet. Es ist erstaunlich, daß trotz der zeitdiskreten Struktur Begriffe verwendet werden, die der kontinuierlichen Welt entstammen: "unmittelbar vor, unmittelbar nach". Bei AF hinterläßt ein Abgang niemals ein volles System, während dies für DF genau dann möglich ist, wenn gleichzeitig mit einem Abgang eine (Mehrfach-) Ankunft stattfindet, die das System voll macht.

2.3 Transiente und stationäre Modellierung

Üblicherweise werden Dienstgüteparameter für einen Multiplexer im stationären Fall berechnet, d.h. die Quellen erzeugen schon seit langer Zeit Zellen. Wenn die Beschreibung des Multiplexers auf eine Markov-Kette führt, dann bestimmen ihre stationären Zustands-WKTn die Dienstgüteparameter. Insbesondere für die Modellierung der Situation bei der Verbindungsannahme von Quellen mit kurzer Dauer kann diese "stationäre" Betrachtung nicht ausreichen. Whitt in [120], p.72: "If sources come and go relatively slowly compared to the way congestion changes, this steady-state view may be appropriate. However, if sources

⁹In der Warteschlangentheorie werden die Begriffe Kunden und Ankünfte nebeneinander verwendet. Hier wird manchmal auch von Zellen gesprochen, wenn der ATM-Zusammenhang im Vordergrund steht.

come and go relatively quickly, it may be much better to base admission decisions on a transient analysis reflecting the present state of congestion in the system.”

Die im folgenden eingesetzte Methode der Restarbeit erlaubt ohne weiteres die Betrachtung auch des transienten Falls, während das Verfahren in Kapitel 5 (“M/G/1-Typ Wartesystem”) numerisch behandelbare Ausdrücke nur für den stationären Fall liefert, vgl. die transienten Ausdrücke in den Gln. (5.23), (5.24) und (5.25).

Beiträge dieser Arbeit

In den folgenden Kapiteln werden bisher bekannte Wartemodelle wesentlich erweitert: In [113] wurde mittels des Begriffes der Restarbeit ein Algorithmus für ein zeitdiskretes Wartesystem vom Typ $GI^{[X]}/D/1/N$ für DF und $D = 1$ entwickelt. Das Superskript deutet an, daß zu einem vom Erneuerungsprozeß GI (General Independent) erzeugten Ankunftszeitpunkt zusätzlich eine Zufallszahl gezogen wird, die die Anzahl w gleichzeitig ankommender Kunden bedeutet (w -Batch). Dieses Verfahren wird zum $SMP^{[X]}/D/1/N$ so erweitert, daß für AF und DF sowie $D \geq 1$ ein beliebiger zeitdiskreter Semi-Markov-Prozeß die Zwischenankunftszeiten erzeugen kann. Schließlich wird auch die Verallgemeinerung auf den Fall einer *beliebigen* Bediendauerverteilung behandelt.

In [13] wurde das $DMAP/G/1/N$ -Wartesystem für $N \leq \infty$ und AF betrachtet und mittels heuristischer Übertragung aus dem schon bekannten zeitkontinuierlichen Fall analysiert. Kapitel 5 erweitert die Ergebnisse für $N < \infty$ zum $DBMAP/G/1/N$ -System für AF¹⁰ und DF. Damit können hier auch Mehrfachankünfte nachgebildet werden. Die gegebenen Beweise basieren auf der Markov-Erneuerungstheorie. Der Lösungsansatz liefert auch die vollständige Beschreibung des Semi-Markovschen Abgangsprozesses. Die Semi-Markovsche Repräsentation der Überlagerung zweier DMAPs wird erstmals entwickelt. Die Ausführungen über den DBMAP und seinen Zählprozeß gehen über die Darstellung in [8] hinaus.

Beide Verfahren werden dann so erweitert, daß für zwei verschiedene Verkehrsströme, wie in Abb. 2.2 erläutert, Dienstgüteparameter getrennt und unter Berücksichtigung von zwei Prioritätsklassen berechnet werden können. Die implementierten Algorithmen zu beiden Verfahren können gegenseitig validiert werden, da für manche Eingabeprozesse beide Verfahren geeignet sind. Der Ansatz der Restarbeit liefert hier einen direkten Zugang zu *transienten* Dienstgüteparametern. Er erlaubt zusätzlich die Berechnung einer bedingten Verlustwahrscheinlichkeit für einen der Eingabeströme.

Abschließend werden gemessene Eigenschaften von Videoquellen mittels des DMAP-Eingabeprozesses nachgebildet und daran die Einflüsse der periodischen Korrelationsfunktion auf die Dienstgüteparameter untersucht. Ein im ATM-Zusammenhang häufig diskutiertes Wartesystem, das $DBMAP/D/1/N$ -Modell mit einer der Zeiteinheit des Systems gleichen Bedienzeit, wird so verallgemeinert, daß die Bedienzeit der f -te Bruchteil ($f \geq 1$) dieser Zeiteinheit sein kann. Damit können Schwierigkeiten umgangen werden, die bei der Untersuchung des Videoquellenmodells mit periodischer Korrelationsfunktion infolge eines zu großen Zustandsraum entstehen [57].

¹⁰In [8] werden die Formeln (5.26), (5.32) und (5.33) angegeben, aber auf anderem Weg bewiesen. Dort wird nicht explizit zwischen AF und DF unterschieden und AF betrachtet.

3 Semi-Markov-Prozesse (SMP) und Markov- Erneuerungsprozesse (MEP)

SMP und MEP repräsentieren zwei verschiedene Standpunkte für die (gleiche) einfachste Erweiterung von Erneuerungsprozessen zur Erzeugung von Zufallsvariablen (ZVN), die nicht mehr unabhängig und daher (meist) korreliert sind. Wenn die ZVN als Zwischenankunftszeiten eines Wartesystems interpretiert werden, lassen sich korrelierte Verkehrsquellen nachbilden, wie sie in ATM-Multiplexern vorkommen.

Die folgenden Kapitel behandeln zwar ausschließlich zeitdiskrete Wartesysteme; dennoch werden die (zeitkontinuierlichen) Markov-Prozesse erwähnt, um zu verdeutlichen, warum man SMP und MEP unterscheidet. Zeitdiskrete Semi-Markov-Prozesse werden auch als Semi-Markov-Ketten bezeichnet [60].

3.1 Stochastische Prozesse (SP)

Ein stochastischer Prozeß (SP) $X = \{X_t, t \in T\}$ ist eine Familie von indizierten ZVN X_t mit $t \in T$, wobei die Parametermenge T meist als Zeit interpretiert wird. Die ZVN X_t nehmen Werte im (gemeinsamen) Zustandsraum \mathcal{IE} an, der im folgenden als abzählbar (d.h. diskret) und meistens endlich angenommen wird. SP werden auch als *Zufallssequenzen* bezeichnet. Sie umfassen die zeitdynamischen Aspekte von ZVN.

Eine ZV ist eine Abbildung mit Werten in \mathcal{IE} , die jedem Elementarereignis ω in Ω einen Wert $Y(\omega)$ zuordnet. Ω heißt Stichprobenraum. Eine Teilmenge von Elementarereignissen wird als Ereignis bezeichnet. Hält man ω fest und betrachtet die Abbildung $t \mapsto X_t(\omega)$, so erhält man einen *Pfad* des SP.

SP werden vollständig beschrieben durch die Menge aller endlich-dimensionalen Verteilungsfunktionen (VFN) $P\{X_{t_1} \leq x_1, X_{t_2} \leq x_2, \dots, X_{t_n} \leq x_n\}$ [37]. Ein SP heißt im strengen Sinne stationär (SSS¹), wenn die endlich-dimensionalen VFN invariant gegen eine beliebige Zeitindexverschiebung sind [80, 91]. Zur Konstruktion von SP ist es nach [20], p. 235 üblich, den Stichprobenraum Ω selbst als Menge aller Abbildungen $\omega : [0, \infty) \rightarrow \mathcal{IE}$ aufzufassen (zeitdiskret: $\omega : \mathbb{N}_0 \rightarrow \mathcal{IE}$), und für alle $t \geq 0$ und $\omega \in \Omega$ zu definieren: $Y_t(\omega) = \omega(t)$.

Einige Bezeichnungen: $E[Y], \text{Var}[Y]$: Mittelwert bzw. Varianz der ZV Y .

$\{S_n = i\} := \{\omega \in \Omega : S_n(\omega) = i\}$.

$P\{S_n = i\} := P(\{S_n = i\})$.

Der Ausdruck $P\{S_{n+1} = j | S_n = i\}$ ist als Abbildung $\phi_j(S_n)$ aufzufassen mit

$\phi_j(i) := P\{S_{n+1} = j | S_n = i\}$, vgl. [20], p.34.

¹Strict Sense Stationary.

3.2 Erneuerungsprozesse, Markov-Ketten und Markov-Prozesse

Markov-Erneuerungsprozesse wurden erdacht als Verallgemeinerung von Erneuerungsprozessen, Markov-Ketten (MKn) und Markov-Prozessen, die nachfolgend definiert werden. Markov-Prozesse werden von manchen Autoren auch als zeitkontinuierliche Markov-Ketten bezeichnet, etwa in [17].

Definition 1 (*Erneuerungsprozeß* [20]) *Der stochastische Prozeß $\{T_n, n \in \mathbb{N}_0, T_0 = 0\}$ $T_n \in \mathbb{R}_0^+, T_n \leq T_{n+1}$ heißt Erneuerungsprozeß, wenn $T_1 - T_0, T_2 - T_1, \dots$ stochastisch unabhängige und identisch verteilte nicht-negative ZVn sind².*

Bei einem Poisson-Prozeß bildet die Folge der Zeitpunkte T_n , zu denen Ankünfte erzeugt werden, einen Erneuerungsprozeß, da die Zeitdifferenzen $T_{n+1} - T_n$ stochastisch unabhängig und negativ-exponentiell verteilt sind mit festem Ratenparameter.

Definition 2 (*Markov Kette*) *Ein stochastischer Prozeß $\{S_n, n \in \mathbb{N}_0\}$ heißt Markov-Kette mit Zustandsraum \mathcal{IE} , falls für alle $j \in \mathcal{IE}$ und alle $n \in \mathbb{N}_0$ gilt:*

$$P\{S_{n+1} = j | S_n, \dots, S_0\} = P\{S_{n+1} = j | S_n\} . \quad (3.1)$$

Die MK heißt zeithomogen, falls die Übergangs-WKtn $p_{ij} = P\{S_{n+1} = j | S_n = i\}$ nicht vom Index n abhängen.

Im Rahmen dieser Arbeit bedeutet eine MK immer die oben definierte zeitdiskrete MK. Die Übergangs-WKtn p_{ij} werden üblicherweise in einer Matrix $\mathbf{p} := (p_{ij})_{1 \leq i, j \leq m}$ zusammengefaßt, wobei $\mathcal{IE} = \{1, 2, \dots, m\}$ gewählt wurde. Die oben definierten MKn sind MKn *erster Ordnung*, da die zukünftige Entwicklung ab dem n -ten Zustand lediglich durch *eine* ZV S_n bestimmt wird. Bei MKn h -ter Ordnung bestimmen die h ZVn S_n, \dots, S_{n-h+1} das zukünftige Geschehen. Nach [20], p.142 können MKn h -ter Ordnung durch Übergang zu den Zufallsvektoren (S_n, \dots, S_{n-h+1}) immer auf eine MK erster Ordnung zurückgeführt werden.

Durch Zulassen einer zeitkontinuierlichen Parametermenge gelangt man zu den Markov-Prozessen.

Definition 3 (*Markov-Prozeß*, [20]) *Ein stochastischer Prozeß $Y = \{Y_t, t \in \mathbb{R}_0^+\}$ heißt Markov-Prozeß mit Zustandsraum \mathcal{IE} , falls für alle $s, t \geq 0$ und $j \in \mathcal{IE}$ gilt:*

$$P\{Y_{s+t} = j | Y_u; u \leq t\} = P\{Y_{s+t} = j | Y_t\} . \quad (3.2)$$

Chung [17], pp. 4, 140 definiert als Markov-Eigenschaft (sowohl im zeitdiskreten wie im zeitkontinuierlichen Fall) noch etwas allgemeiner: Für alle $n \geq 2, 0 \leq t_1 < \dots < t_n$ und alle $i_1, \dots, i_n \in \mathcal{IE}$ gilt $P\{Y_{t_n} = i_n | Y_{t_{n-1}} = i_{n-1}, \dots, Y_{t_1} = i_1\} = P\{Y_{t_n} = i_n | Y_{t_{n-1}} = i_{n-1}\}$. Daraus ergeben sich dann Gl. (3.2) bzw. Gl. (3.1).

Ein Pfad $t \rightarrow Y_t(\omega)$ eines Markov-Prozesses ergibt sich für (fast) alle $\omega \in \Omega$, wie folgt: Zum Startzeitpunkt $T_0(\omega) = 0$ befindet er sich in einem Zustand $S_0(\omega) = Y_0(\omega)$, in dem er für eine Zeit verweilt. Zu einem Zeitpunkt $T_1(\omega)$ springt er in einen *anderen* Zustand $S_1(\omega) =$

²Die angelsächsische Literatur verwendet dafür die Abkürzung i.i.d.: "independent identically distributed".

$Y_{T_1(\omega)}(\omega) \neq S_0(\omega)$, um dort zu bleiben, bis zum Zeitpunkt $T_2(\omega)$ wieder ein *anderer* Zustand $S_2(\omega) = Y_{T_2(\omega)}(\omega)$ als der gerade besuchte betreten wird, usw. Die Zeitpunkte T_0, T_1, T_2, \dots bezeichnen also echte Zustandswechsel von Y , und S_0, S_1, S_2, \dots sind die nacheinander besuchten (verschiedenen) Zustände. Gemäß *Theorem* 8.3.3 sowie *Corollaries* 8.3.10 und 8.3.11 in [20] gilt für Markov-Prozesse:

- Der SP S_0, S_1, S_2, \dots der nacheinander besuchten Zustände bildet eine MK. Diese MK heißt "zu Übergangszeitpunkten eingebettet". Ihre Übergangsmatrix hat nur Nulleinträge auf der Hauptdiagonalen, da ein Übergang immer in einen anderen Zustand als den gerade besuchten führt.
- Die Zustandsverweilzeiten $T_{n+1} - T_n$ im Zustand S_n bis zum Übergang nach S_{n+1} sind negativ-exponentiell verteilt mit dem Ratenparameter $\lambda(S_n)$, d.h. die VF hängt nur vom gerade betretenen Zustand, nicht aber vom nächsten zu besuchenden ab. Formal wird das gefaßt in der folgenden bedingten WKT

$$P\{T_{n+1} - T_n > t | S_{n+1} = j, S_n = i\} = e^{-\lambda(i)t} . \quad (3.3)$$

- Die Zeiten zwischen Übergängen sind bedingt unabhängig voneinander bei vorgegebenen aufeinanderfolgend besuchten Zuständen. Die Rate jeder der negativ-exponentiellen Verweilzeit-VFn hängt nur vom gerade besuchten Zustand ab

$$P\{T_1 - T_0 > t_0, \dots, T_{n+1} - T_n > t_n | S_{n+1} = i_{n+1}, S_n = i_n, \dots, S_0 = i_0\} = e^{-\lambda(i_0)t_0} \dots e^{-\lambda(i_n)t_n} . \quad (3.4)$$

3.2.1 Die starke Markoveigenschaft, Stoppzeiten

Die Markoveigenschaft, wie sie in den Gln. (3.1), (3.2) formuliert ist, bezieht sich auf feste (deterministische) Zeitpunkte n bzw. t . Für verschiedene Anwendungen der Theorie benötigt man die Markoveigenschaft auch für Zufallszeiten. Falls sie gilt, spricht man von der *starken Markoveigenschaft*. Insbesondere gilt die starke Markoveigenschaft bei einer MK X für Zufallszeiten T , wenn über das Eintreten des Ereignisses $\{T \leq n\}$ allein durch Betrachtung der Folge X_0, X_1, \dots, X_n entschieden werden kann. Solche Zufallszeiten heißen *Stoppzeiten*. Asmussen [3]: "The important strong Markov property states that for the sake of predicting the future development of the chain, a stopping time may be treated as a fixed deterministic point of time."

Beispiele für Stoppzeiten: Der erste Zeitpunkt, zu dem die MK X einen vorgegebenen Zustand betritt; der erste Zeitpunkt, zu dem die MK eine vorgegebene Menge von Zuständen betritt; der k -te Zeitpunkt, zu dem X eine vorgegebene Zustandsmenge betritt.

Jede positive natürliche Konstante ist eine Stoppzeit. Die Summe zweier Stoppzeiten ist eine Stoppzeit. Die größere (kleinere) von zwei Stoppzeiten ist eine Stoppzeit. Der letzte Zeitpunkt, zu dem X einen Zustand besucht, ist *keine* Stoppzeit [20], p. 117f., [80], p. 69. Die Abgangszeitpunkte von Wartesystemen vom M/G/1-Typ in Kapitel 5 sind Stoppzeiten für den dort vorliegenden Eingabeprozess.

3.3 Semi-Markov-Prozesse (SMP)

Markov-Prozesse können durch eine MK zu Zustandswechselzeitpunkten und (bedingt unabhängige) negativ-exponentielle VFn für die Zustandsverweilzeiten vollständig beschrieben

werden. Dadurch wird eine MK und eine Folge bedingt unabhängiger Zufallszeiten verknüpft; unter der Bedingung *gleicher* besuchter Zustände sind die Verweilzeiten identisch verteilt und (bedingt) unabhängig, ähnlich den Erneuerungsprozessen. Das motiviert eine allgemeinere Verknüpfung von Erneuerungsprozessen und MKn: Man läßt die Verteilungen als beliebig (also nicht mehr nur negativ-exponentiell) zu, gibt ihnen noch eine Abhängigkeit vom nächsten zu besuchenden Zustand der MK und erlaubt auch MKn, die positive Einträge auf der Hauptdiagonalen ihrer Übergangsmatrix haben.

Dann sind einerseits Erneuerungsprozesse berücksichtigt durch den Fall einer MK mit einem einzigen Zustand und andererseits MKn bei Festlegung deterministischer Zustandsverweilzeiten der Größe 1. Dazu betrachtet man den Verbundprozeß $(S, T) := \{(S_n, T_n); n \in \mathbb{N}_0\}$, wobei S_n Werte im Zustandsraum $\mathcal{IE} = \{1, \dots, m\}$ annimmt und T_n die Zustandswechselzeitpunkte ($T_n \leq T_{n+1}$) bezeichnet.

Definition 4 (Markov-Erneuerungsprozeß) Der SP (S, T) heißt ein Markov-Erneuerungsprozeß mit Zustandsraum \mathcal{IE} , falls gilt

$$\begin{aligned} P\{S_{n+1} = j, T_{n+1} - T_n \leq t | S_n, \dots, S_0, T_n, \dots, T_0\} = \\ = P\{S_{n+1} = j, T_{n+1} - T_n \leq t | S_n\}. \end{aligned} \quad (3.5)$$

für alle $n \in \mathbb{N}$ und $t \in \mathbb{R}_0^+$ (zeitkontinuierlich) bzw. $t \in \mathbb{N}_0$ (zeitdiskret).

Zeithomogenität bedeutet wieder, daß die WKTn $Q_{ij}(t) := P\{S_{n+1} = j, T_{n+1} - T_n \leq t | S_n = i\}$ vom Index n des Zustandswechselzeitpunktes unabhängig sind. Die Menge der WKTn $\{Q_{ij}(t) : i, j \in \mathcal{IE}, t \in \mathbb{R}_0^+\}$ heißt **Semi-Markov-Kern** über \mathcal{IE} . $\{S_n; n \in \mathbb{N}_0\}$ ist eine MK mit den Übergangs-WKTn $p_{ij} = \lim_{t \rightarrow \infty} Q_{ij}(t)$, und der Quotient

$$F_{ij}(t) := \frac{Q_{ij}(t)}{p_{ij}} = P\{T_{n+1} - T_n \leq t | S_{n+1} = j, S_n = i\} \quad (3.6)$$

ist für $p_{ij} \neq 0$ eine VF der Verweilzeit im Zustand i bis zum Übergang nach j ($i = j$ zugelassen). Bekanntlich ist für $t_1 \leq t_2$ auch $Q_{ij}(t_1) \leq Q_{ij}(t_2)$ (eine größere Menge hat ein größeres WKTsmaß), d.h. $p_{ij} = \lim_{t \rightarrow \infty} Q_{ij}(t) = 0$ impliziert $Q_{ij}(t) = 0$ für alle t . Im Falle $p_{ij} = 0$ definiert man daher $F_{ij}(t) = 1$ [20].

Markov-Prozesse sind gekennzeichnet durch eine einzelne ZV mit kontinuierlicher Indizierung, Markov-Erneuerungsprozesse aber durch zwei gekoppelte, diskret indizierte ZVn, die bei geeigneten Eigenschaften einen Markov-Prozeß repräsentieren. Eine kontinuierliche Indizierung erreicht man auch für beliebige MEP, wenn man einem MEP (mit endlichem Zustandsraum) den SP $Y := \{Y_t; t \in \mathbb{R}_0^+\}$ durch die Definition

$$\text{für alle } \omega \in \Omega: Y_t(\omega) := S_n(\omega) \quad \text{für } T_n(\omega) \leq t < T_{n+1}(\omega) \quad (3.7)$$

zuordnet.³ Dieser SP heißt dann der dem MEP zugeordnete SMP. Cinlar [20], p.316: "Another convenient picture in describing a Markov Renewal process is provided by the process Y ".

Die Vorsilbe "Semi" deutet hin auf die auf Zustandswechselzeiten T_n eingeschränkte Markov-eigenschaft

$$P\{Y(T_n + t) = j | Y(T_n), Y(T_{n-1}), \dots, Y(T_0)\} = P\{Y(T_n + t) = j | Y(T_n)\}.$$

³Für unendliche Zustandsräume wird ein erweitertes Konzept [20] erforderlich; diese Arbeit benötigt nur den Fall endlicher Zustandsräume.

Cinlar, p. 337: "The expression Semi-Markov goes back to the somewhat limited Markov property for Y that the future of Y is independent of its past given the present state provided that the 'present' is a time of jump." Wenn die Zustandsverweilzeiten negativ-exponentiell verteilt sind, ergibt sich die volle Markoveigenschaft nach Gl. (3.2) infolge der sog. *Gedächtnislosigkeit* der negativ-exponentiellen Verteilung. Die Verteilung einer ZV T besitzt die Eigenschaft "Gedächtnislosigkeit", falls $P\{T > t+h | T > t\} = P\{T > h\}$ erfüllt ist [80], p. 64.

Übergänge von einem Zustand in denselben sind bei allgemeinen SMP erlaubt und beobachtbar. Sie können ausgeschlossen werden, indem die MK mit Nulleinträgen auf der Hauptdiagonalen versehen wird; sind zusätzlich die VFn der Zustandsverweilzeiten negativ-exponentiell und nicht vom nächsten zu besuchenden Zustand abhängig, so liegt ein Markov-Prozeß vor. Eine MK ergibt sich bei identischen Zustandsverweilzeiten der Größe 1, ein Erneuerungsprozeß, wenn nur ein Zustand angenommen wird (dann gibt es auch nur eine VF für die Zustandsverweilzeit und die Verweilzeiten sind "i.i.d.") oder die $F_{ij}(t)$ alle gleich sind.

Entwicklung eines Pfades Sobald ein SMP den Zustand i betreten hat, wird nach Maßgabe der Übergangs-WKT p_{ij} der nächste zu besuchende Zustand j ausgewählt und dann eine Verweilzeit gemäß der VF $F_{ij}(t)$ für das Verbleiben im Zustand i bis zum Übergang nach j gezogen. Zur Erzeugung von Zwischenankunftszeiten wird das Betreten eines Zustands als Ankunft interpretiert. Durch Zulassen von Verweilzeiten der Größe Null kann man Mehrfachankünfte ermöglichen. Zur Auswahl des Startzustandes müssen geeignete Start-WKTn vorgegeben werden. Meist wählt man die stationäre Lösung \underline{p} der eingebetteten MK.

3.3.1 Zeitdiskrete SMP

Zur weiteren Behandlung des zeitdiskreten Falls werden mittels $A_n := T_{n+1} - T_n$ die folgenden Ausdrücke definiert:

$$q_{ij}(t) := P\{S_{n+1} = j, A_n = t | S_n = i\}, \quad (3.8)$$

$$f_{ij}(t) := P\{A_n = t | S_{n+1} = j, S_n = i\} = \frac{P\{A_n = t, S_{n+1} = j, S_n = i\}}{P\{S_{n+1} = j, S_n = i\}} = \frac{q_{ij}(t)}{p_{ij}}. \quad (3.9)$$

Es werden die $(m \times m)$ -Matrizen verwendet: $\mathbf{p} = (p_{ij})_{1 \leq i, j \leq m}$, $\mathbf{q}(t) = (q_{ij}(t))_{1 \leq i, j \leq m}$ und $\mathbf{Q}(t) = (Q_{ij}(t))_{1 \leq i, j \leq m}$.

Verbund-WFn der ZVn A_n , die schließlich auf die Korrelationseigenschaften der Zustandsverweilzeiten führen, erhält man aus folgendem Zusammenhang:

Hilfssatz 1 Für alle $t_n, \dots, t_{n+\nu} \in \mathbb{N}_0$, $\nu \in \mathbb{N}$ gilt

$$\begin{aligned} P\{A_n = t_n, \dots, A_{n+\nu} = t_{n+\nu} | S_n = i_n, \dots, S_{n+\nu+1} = i_{n+\nu+1}\} &= \\ &= f_{i_n i_{n+1}}(t_n) \cdot \dots \cdot f_{i_{n+\nu} i_{n+\nu+1}}(t_{n+\nu}). \end{aligned} \quad (3.10)$$

Beweis: Der Zähler der bedingten WKT kann mittels Gl. (3.5) faktorisiert werden unter Berücksichtigung von

$$\{A_n = t_n, \dots, A_{n+\nu} = t_{n+\nu}\} = \bigcup_{\tau=0}^{\infty} \{T_n = \tau, T_{n+1} = \tau + t_n, \dots, T_{n+\nu+1} = \tau + t_n + \dots + t_{n+\nu}\} :$$

Für jedes $\nu \geq 1$ gilt:

$$\begin{aligned}
 P\{A_n = t_n, \dots, A_{n+\nu} = t_{n+\nu}, S_n = i_n, \dots, S_{n+\nu+1} = i_{n+\nu+1}\} &= \\
 &= \sum_{\tau=0}^{\infty} P\{A_{n+\nu} = t_{n+\nu}, S_n = i_n, \dots, S_{n+\nu+1} = i_{n+\nu+1}, \\
 &\quad T_n = \tau, T_{n+1} = \tau + t_n, \dots, T_{n+\nu} = \tau + t_n + \dots + t_{n+\nu-1}\} = \\
 &= P\{S_{n+\nu+1} = i_{n+\nu+1}, A_{n+\nu} = t_{n+\nu} | S_{n+\nu} = i_{n+\nu}\} \cdot \\
 &\quad \cdot \sum_{\tau=0}^{\infty} P\{S_n = i_n, \dots, S_{n+\nu} = i_{n+\nu}, T_n = \tau, T_{n+1} = \tau + t_n, \dots, \\
 &\quad T_{n+\nu} = \tau + t_n + \dots + t_{n+\nu-1}\} = \\
 &= P\{S_{n+\nu+1} = i_{n+\nu+1}, A_{n+\nu} = t_{n+\nu} | S_{n+\nu} = i_{n+\nu}\} \cdot \\
 &\quad \cdot P\{A_n = t_n, \dots, A_{n+\nu-1} = t_{n+\nu-1}, S_n = i_n, \dots, S_{n+\nu} = i_{n+\nu}\} .
 \end{aligned}$$

Insbesondere gilt für $\nu = 1$:

$$\begin{aligned}
 P\{A_{n+1} = t_{n+1}, A_n = t_n, S_{n+2} = i_{n+2}, S_{n+1} = i_{n+1}, S_n = i_n\} &= \\
 P\{S_{n+2} = i_{n+2}, A_{n+1} = t_{n+1} | S_{n+1} = i_{n+1}\} \cdot P\{S_{n+1} = i_{n+1}, A_n = t_n | S_n = i_n\} \cdot P\{S_n = i_n\} .
 \end{aligned}$$

Der Nenner ist eine Verbund-WKT einer MK, d.h.

$$P\{S_n = i_n, \dots, S_{n+\nu+1} = i_{n+\nu+1}\} = P\{S_n = i_n\} \cdot p_{i_n i_{n+1}} \cdot \dots \cdot p_{i_{n+\nu} i_{n+\nu+1}},$$

und mit Gl. (3.9) ergibt sich die Behauptung. □

Entsprechend gilt für Verbund-VFn die Faktorisierung

$$\begin{aligned}
 P\{A_n \leq t_n, \dots, A_{n+\nu} \leq t_{n+\nu} | S_n = i_n, \dots, S_{n+\nu+1} = i_{n+\nu+1}\} &= \\
 = F_{i_n i_{n+1}}(t_n) \cdot \dots \cdot F_{i_{n+\nu} i_{n+\nu+1}}(t_{n+\nu}), & \quad (3.11)
 \end{aligned}$$

und diese Beziehung beschreibt auch den zeitkontinuierlichen Fall [20]. Gl. (3.11) führt nun auf einen Ausdruck für die Verbund-VFn

$$\begin{aligned}
 P\{A_n \leq t_n, \dots, A_{n+\nu} \leq t_{n+\nu}\} &= \\
 = \sum_{i_n, \dots, i_{n+\nu+1} \in \mathcal{E}} P_{i_n}^{(n)} \cdot Q_{i_n i_{n+1}}(t_n) \cdot \dots \cdot Q_{i_{n+\nu} i_{n+\nu+1}}(t_{n+\nu}) &= \\
 = \underline{P}^{(n)} \cdot \prod_{\mu=0}^{\nu} \mathbf{Q}(t_{n+\mu}) \cdot \underline{e}, \quad \nu = 0, 1, 2, \dots & \quad (3.12)
 \end{aligned}$$

$\underline{P}^{(n)} := (P_1^{(n)}, \dots, P_m^{(n)})$ bedeutet hier den Zustands-WKTsvektor der eingebetteten MK nach n Schritten, $\underline{e} := (1, \dots, 1)^T$ mit m Komponenten.

Die Verbund-WF lautet

$$\begin{aligned}
 P\{A_n = t_n, \dots, A_{n+\nu} = t_{n+\nu}\} &= \\
 = \underline{P}^{(n)} \cdot \prod_{\mu=0}^{\nu} \mathbf{q}(t_{n+\mu}) \cdot \underline{e}, \quad \nu = 0, 1, 2, \dots & \quad (3.13)
 \end{aligned}$$

Allgemein ist das Matrixprodukt in Gl. (3.12) verschieden vom Zahlenprodukt

$$\prod_{\mu=0}^{\nu} P\{A_{n+\mu} \leq t_{n+\mu}\} = \prod_{\mu=0}^{\nu} [\underline{P}^{(n+\mu)} \cdot \mathbf{Q}(t_{n+\mu}) \cdot \underline{e}],$$

so daß die Zustandsverweilzeiten i.a. nicht mehr unabhängig sind. Wählt man für $\underline{P}^{(n)}$ die stationäre Lösung der eingebetteten MK, so erhält man im strengen Sinne stationäre Verbund-VFn $P\{A_n \leq t_n, \dots, A_{n+\nu} \leq t_{n+\nu}\}$, d.h. sie sind invariant gegen eine beliebige Verschiebung des Index n [80].

Diskretisierung eines Markov-Prozesses Es wurde schon erwähnt, daß die eingebettete MK eines Markov-Prozesses Nulleinträge auf der Hauptdiagonalen hat. Eine zeitdiskrete Approximation eines Markov-Prozesses kann durch eine MK X erzeugt werden. Dazu betrachtet man die in X zu echten⁴ Zustandswechseln eingebettete MK S . S_n beude den Zustand, der zum Zeitpunkt T_n beim n -ten echten Wechsel betreten wird. T_n ist eine Stoppzeit, weil über das Auftreten des Ereignisses $\{T_n \leq \nu\}$ allein durch Kenntnis von X_0, X_1, \dots, X_ν entschieden werden kann. $X_{T_n} = i$ impliziert $X_{T_{n-1}} \neq i$. Mit der MEP-Notation erhält man für $j \neq i$

$$P\{S_{n+1} = j, T_{n+1} - T_n = k | S_n = i\} =$$

$$P\{X_{T_n+k} = j, X_{T_n+k-1} = i, \dots, X_{T_n+1} = i | X_{T_n} = i, X_{T_{n-1}} \neq i\} =$$

$$P\{X_{T_n+k} = j, X_{T_n+k-1} = i, \dots, X_{T_n+1} = i | X_{T_n} = i\} = p_{ii}^{k-1} p_{ij}$$

und

$$P\{S_{n+1} = j | S_n = i\} = \sum_{k=1}^{\infty} P\{S_{n+1} = j, T_{n+1} - T_n = k | S_n = i\} = p_{ij} / (1 - p_{ii}),$$

so daß die WF der Zustandsverweilzeit lautet

$$P\{T_{n+1} - T_n = k | S_{n+1} = j, S_n = i\} = (1 - p_{ii}) p_{ii}^{k-1} \quad \text{für } j \neq i.$$

Sie hängt nicht vom Zielzustand j ab und ist geometrisch verteilt.

3.3.2 Korrelationseigenschaften

Globale Korrelation Üblicherweise wird als Maß für die Nicht-Unabhängigkeit eines im strengen Sinne stationären SP, hier für die Zustandsverweilzeiten, der folgende Korrelationskoeffizient h -ter Ordnung verwendet

$$\kappa_h := \text{Corr}[A_n, A_{n+h}] = \frac{\text{Cov}[A_n, A_{n+h}]}{\sqrt{\text{Var}[A_n] \text{Var}[A_{n+h}]}} \quad (3.14)$$

mit

$$\text{Cov}[A_n, A_{n+h}] = \text{E}[(A_n - \text{E}[A_n])(A_{n+h} - \text{E}[A_{n+h}])] = \text{E}[A_n \cdot A_{n+h}] - \text{E}^2[A_n]$$

und

$$\text{Var}[A_n] = \text{E}[(A_n - \text{E}[A_n])^2] = \text{E}[A_n^2] - \text{E}^2[A_n] = \text{Var}[A_{n+h}].$$

Für beliebige reelle Zahlen $a, b, c, d \in \mathbb{R}$ gilt [80]

$$\text{Corr}[aA_n + b, cA_{n+h} + d] = \text{Corr}[A_n, A_{n+h}].$$

Aus $\kappa = 0$ folgt i.a. *nicht*, daß Unabhängigkeit vorliegt. Natürlich sind aber unabhängige ZVn unkorreliert, d.h. $\kappa = 0$.

⁴d.h. der Übergang führt in einen anderen als den gerade besuchten Zustand

Wie in [30, 31, 107] wird hier κ_h zur Unterscheidung vom nachfolgend definierten lokalen Korrelationskoeffizienten als *globaler* Korrelationskoeffizient bezeichnet.

Für einen beliebigen zeitdiskreten (SSS) SMP gilt

$$E[A_n \cdot A_{n+h}] = P \cdot \sum_{t=1}^{\infty} t\mathbf{q}(t) \cdot \mathbf{p}^{h-1} \cdot \sum_{t'=1}^{\infty} t'\mathbf{q}(t') \cdot \underline{e} \quad \text{für } h \geq 1, \quad (3.15)$$

$$E[A_n^2] = P \cdot \sum_{t=1}^{\infty} t^2\mathbf{q}(t) \cdot \underline{e}, \quad (3.16)$$

$$E[A_n] = P \cdot \sum_{t=1}^{\infty} t\mathbf{q}(t) \cdot \underline{e}. \quad (3.17)$$

Lokale Korrelation Im Gegensatz zu κ_h ist der *lokale* Korrelationskoeffizient $\varrho_h(x)$ eine Funktion der möglichen Werte x , die die betrachteten ZVn annehmen können. Die Definition erfolgt (hier wieder für die Zustandsverweilzeiten) mittels der $A = \{A_n, n \in \mathbb{N}_0\}$ zugeordneten binären ZV $\zeta_n^{[A]}(x)$

$$\zeta_n^{[A]}(x) = \begin{cases} 1 & \text{if } A_n > x, \\ 0 & \text{if } A_n \leq x \end{cases} \quad (3.18)$$

unter Verwendung des globalen Korrelationskoeffizienten durch

$$\varrho_h(x) := \text{Corr}[\zeta_n^{[A]}(x), \zeta_{n+h}^{[A]}(x)]. \quad (3.19)$$

Nach [30], p.22 läßt sich Gl. (3.19) überführen in

$$\varrho_h(x) = 1 - \frac{P\{A_n > x, A_{n+h} \leq x\}}{P\{A_n \leq x\} \cdot (1 - P\{A_n \leq x\})}, \quad (3.20)$$

und dieser Ausdruck ist äquivalent zu dem in [108] gegebenen. Der lokale Korrelationskoeffizient erlaubt die Bestimmung aller einen SSMP (s. u.) mit *zwei* Zuständen beschreibenden Parameter aus Meßwerten [30], so daß im Hinblick auf die ATM-Problematik die Parameter geeigneter SSMPs auf reale Quellen angepaßt werden können. Der globale Korrelationskoeffizient läßt das nicht zu.

Dispersionsindex (Index of dispersion) In [111] wird ein weiteres Korrelationsmaß verwendet, der sog. Dispersionsindex für Zeitintervalle nach [25]. Er setzt $\text{Var}[T_k]$ in Beziehung zum entsprechenden Wert, der sich für den Poisson-Prozeß ergibt; nach [25], p. 71 sind $\text{Var}[T_k]$ und $\text{Corr}[A_n, A_{n+k}]$ äquivalent. Weitere Eigenschaften finden sich in [111].

$$c_k^2 := \frac{k\text{Var}[T_k]}{E^2[T_k]} = \frac{k\text{Var}[A_0] + 2\sum_{j=1}^{k-1}(k-j)\text{Cov}[A_0, A_j]}{kE^2[A_0]} \quad (3.21)$$

3.3.3 Der spezielle Semi-Markov-Prozeß (SSMP)

Sind die VFn $F_{ij}(t)$ unabhängig vom nächsten zu besuchenden Zustand j , d.h. kann man auf diesen Index verzichten, so wird der SMP bei [30, 32] als spezieller Semi-Markov-Prozeß bezeichnet. In einer älteren Arbeit [60] findet sich die Bezeichnung "unabhängiger Semi-Markov-Prozeß". Infolge dieser Eigenschaft vereinfachen sich einige Rechenausdrücke. Der Semi-Markov-Kern kann als Matrixprodukt geschrieben werden

$$\mathbf{Q}(t) = \mathbf{Diag}(F_1(t), \dots, F_m(t)) \cdot \mathbf{p} \quad , \quad (3.22)$$

und die (stationäre) VF der Zustandsverweilzeit ergibt sich als Linearkombination der VFn der Verweilzeit in den einzelnen Zuständen

$$F(t) := P\{A_n \leq t\} = \underline{P} \cdot \mathbf{Diag}(F_1(t), \dots, F_m(t)) \cdot \underline{e} = \sum_{i=1}^m P_i F_i(t) \quad . \quad (3.23)$$

Globaler Korrelationskoeffizient ([29], p.93)

$$\text{Corr}[A_n, A_{n+h}] = \frac{\sum_{i=1}^m \sum_{j=1}^m P_i E[A^{(i)}] (p_{ij}^{[h]} - P_j) E[A^{(j)}]}{\text{Var}[A_n]} \quad , \quad (3.24)$$

wobei $A^{(i)}$ die ZV der Verweilzeit im Zustand i bezeichnet. Diese ZV wird nicht mit n indiziert, da es keinen direkten Zusammenhang zur n -ten Ankunft gibt.

Weitere Vereinfachungen ergeben sich für den Fall von $m = 2$ Zuständen (SSMP(2)). Dann sind die globalen Korrelationskoeffizienten von Zustandsverweilzeiten A und eingebetteter MK S proportional, und beide fallen mit wachsender Ordnung h geometrisch

$$\text{Corr}[A_{n+h}, A_n] = (\kappa_1^{[S]})^h \cdot \left(1 + \frac{P_1 \text{Var}[A^{(1)}] + P_2 \text{Var}[A^{(2)}]}{P_1 P_2 (E[A^{(1)}] - E[A^{(2)}])^2} \right)^{-1} \quad , \quad (3.25)$$

$$\kappa_1^{[S]} = 1 - p_{12} - p_{21}, \quad \kappa_h^{[S]} = (\kappa_1^{[S]})^h \quad .$$

Ein Beweis für den Zusammenhang $\text{Corr}[S_n, S_{n+h}] = \text{Corr}[S_n, S_{n+1}]^h = (\kappa_1^{[S]})^h$ bei $m = 2$ ist in [106] enthalten. Infolge des geometrischen Abfalls mit der Ordnung h kann ein SSMP(2) die in [90] für Zwischenankunftszeiten gemessenen (schwach) periodischen Korrelationsfunktionen nur bedingt repräsentieren. Für eine Nachbildung der periodischen Korrelationsfunktion der Anzahl von Ankünften in einem vorgegebenen Intervall scheint er ungeeignet, da es keine kompakte Form des Zählprozesses eines beliebigen SSMP gibt. Der im nächsten Abschnitt behandelte DBMAP ist in dieser Hinsicht sehr viel leistungsfähiger. Dagegen umfaßt der SSMP(2) offensichtlich die Videodaten-Charakterisierung aus [78] mit geometrisch abfallender Korrelationsfunktion.

Lokaler Korrelationskoeffizient Nach [30] gilt

$$\varrho_h^{[A]}(x) = \frac{\sum_{i=1}^m \sum_{j=1}^m P_i F_i(x) (P_i - p_{ij}^{[h]}) (1 - F_j(x))}{F(x)(1 - F(x))} \quad , \quad (3.26)$$

und aus " $P_i = p_{ij}^{[h]}$ für alle $i \in \mathbb{IE}$ " oder " $F_j = F$ für alle $j \in \mathbb{IE}$ " folgt $\varrho_h^{[A]} = 0$. Bei nur zwei Zuständen ergibt sich der vereinfachte Ausdruck

$$\varrho_h^{[A]}(x) = (\kappa_1^{[S]})^h \cdot P_1 P_2 \frac{(F_1(x) - F_2(x))^2}{F(x)(1 - F(x))} \quad . \quad (3.27)$$

Es folgt für $P_1 \neq 0$ und $P_2 \neq 0$: $\varrho_n^{[A]} = 0 \iff [\kappa_1^{[S]} = 0 \text{ oder } F_1 = F_2]$.

Alle Markov-Prozesse sind SSMPs mit negativ-exponentiell verteilten Zustandsverweilzeiten. Bemerkenswert ist, daß es für den sog. "Markov-modulierten Poisson-Prozeß" mit zwei Zuständen (MMPP(2)) eine äquivalente Darstellung durch einen Markov-Prozeß mit zwei Zuständen gibt. Obwohl der MMPP(2) für die Beschreibung von Datenverkehr und paketierter Sprache schon mehrere Jahre in der Literatur diskutiert wurde [50], fiel dieser Zusammenhang erst in [30] auf und wurde dann in [74] (durch Betrachtung der endlich-dimensionalen Verteilungsfunktionen) streng bewiesen.

3.4 Der zeitdiskrete Markovsche Ankunftsprozeß (DBMAP)

Als zeitdiskretes Analogon des Markovschen Ankunftsprozesses nach Lucantoni [75] wird in [10] der DBMAP ("Discrete Batch Markovian Arrival Process") mittels einer MK mit m Zuständen und *verschiedenen Typen von Übergängen*⁵ definiert. Ein Übergang vom Zustand $S(t) = i$ nach $S(t+1) = j$, $t = 0, 1, \dots$, wird bestimmt durch die vom Typ $w = 0, 1, 2, \dots$ abhängigen WKTn $[\mathbf{D}_w]_{ij}$. Dabei bedeutet \mathbf{D}_w die Übergangsmatrix für Typ w und $[\cdot]_{ij}$ den Matrixeintrag in der i -ten Zeile und der j -ten Spalte. Ein Übergang vom Typ $w > 0$ wird mit Bezug auf ein Wartesystem als Ankunft von $w > 0$ Kunden aufgefaßt. Ein Übergang vom Typ $w = 0$ bedeutet keine Ankunft, d.h. lediglich ein Fortzählen der Zeiteinheit. Solch ein Modell beinhaltet also die zur Modellierung von ATM-Multiplexern wichtigen Mehrfachankünfte sowie Zwischenankunftszeiten, die – wie nachfolgend erläutert – nicht unabhängig sind. Zur Unterscheidung von anderen Zustands-ZVn heißen die Zustände der den DBMAP erzeugenden MK auch *Phasen*. Durch die Vorgabe von verschiedenen Übergangstypen wird der übliche Begriff der MK erweitert. Die Matrix

$$\widehat{\mathbf{D}} := \sum_{w=0}^{\infty} \mathbf{D}_w \quad (3.28)$$

ist stochastisch und repräsentiert die (konventionelle) MK ohne Unterscheidung der Übergangstypen.

Als (erweiterte) Markoveigenschaft erhält man nun mit der Typ-ZV $X(t) \in \mathbb{N}_0$ als Übergangstyp vom $(t-1)$ -ten zum t -ten Schritt:

$$\begin{aligned} P\{S(t+1) = j, X(t+1) = v | S(t), X(t), \dots, S(1), X(1), S(0)\} = \\ = P\{S(t+1) = j, X(t+1) = v | S(t)\} \end{aligned} \quad (3.29)$$

Dieser Ausdruck hat Ähnlichkeit mit Gl. (3.5). Obwohl im folgenden eine Maximalzahl b_{max} der gleichzeitig ankommenden Kunden angenommen wird, wird $X(t)$ auf ganz \mathbb{N}_0 erklärt und $\mathbf{D}_w = \mathbf{0}$ gesetzt für $w > b_{max}$.

Hilfssatz 2 *Der Punktprozeß der Ankunftszeitpunkte eines DBMAP ist Semi-Markovsch mit dem folgenden Semi-Markov-Kern*

$$P\{S(T_{n+1}) = j, T_{n+1} - T_n = k | S(T_n) = i\} = [\mathbf{D}_0^{k-1} \cdot (\widehat{\mathbf{D}} - \mathbf{D}_0)]_{ij}. \quad (3.30)$$

Dabei bedeutet T_n den Ankunftszeitpunkt der n -ten Ankunft.

⁵Das in [10] behandelte Wartesystem DBMAP/D/1/N mit $D = 1$ ist nicht identisch mit dem in Kapitel 5; es wird in Abschnitt 5.2 erwähnt.

Beweis: T_n ist eine Stoppzeit, da über das Auftreten von $\{T_n \leq k\}$ durch Inspektion von $S(0), X(1), S(1), \dots, X(k), S(k)$ entschieden werden kann. Die Markov-Erneuerungseigenschaft (3.5) folgt aus (3.29). Den Semi-Markov-Kern erhält man wie folgt:

$$P\{S(T_{n+1}) = j, T_{n+1} - T_n = k | S(T_n) = i\} =$$

$$\sum_{i_1, \dots, i_{k-1} \in \mathbb{E}} P\{S(T_n + k) = j, X(T_n + k) \neq 0, S(T_n + k - 1) = i_{k-1}, \\ X(T_n + k - 1) = 0, \dots, S(T_n + 1) = i_1, X(T_n + 1) = 0 | S(T_n) = i\} =$$

$$\sum_{i_1, \dots, i_{k-1} \in \mathbb{E}} P\{S(T_n + k) = j, X(T_n + k) \neq 0 | S(T_n + k - 1) = i_{k-1}\} \cdot \\ P\{S(T_n + k - 1) = i_{k-1}, X(T_n + k - 1) = 0 | S(T_n + k - 2) = i_{k-2}\} \cdot \dots \\ \cdot P\{S(T_n + 1) = i_1, X(T_n + 1) = 0 | S(T_n) = i\} = \\ = [\mathbf{D}_0^{k-1} \cdot (\widehat{\mathbf{D}} - \mathbf{D}_0)]_{ij} . \quad \square$$

Weiter kann noch zwischen der Art des Übergangs unterschieden werden, und man erhält

$$P\{S(T_{n+1}) = j, T_{n+1} - T_n = k, X(T_{n+1}) = w | S(T_n) = i\} = \\ = \begin{cases} [\mathbf{D}_0^{k-1} \cdot \mathbf{D}_w]_{ij} & \text{für } w > 0, \\ 0 & \text{für } w = 0. \end{cases} \quad (3.31)$$

Sei $\underline{\pi}$ der Vektor der stationären Phasen-WKTN: $\underline{\pi} \widehat{\mathbf{D}} = \underline{\pi}$. Die zu Ankunftszeitpunkten eingebettete MK hat die Übergangsmatrix

$$\sum_{k=1}^{\infty} \mathbf{D}_0^{k-1} (\widehat{\mathbf{D}} - \mathbf{D}_0) = (\mathbf{I} - \mathbf{D}_0)^{-1} (\widehat{\mathbf{D}} - \mathbf{D}_0) . \quad (3.32)$$

Für diese MK erhält man als stationäre Lösung den Phasenvektor bei Ankunftszeitpunkten zu

$$\underline{P} = \frac{1}{\underline{\pi} (\mathbf{I} - \mathbf{D}_0) \underline{e}} \underline{\pi} (\widehat{\mathbf{D}} - \mathbf{D}_0) = \underline{\pi} (\mathbf{I} - \mathbf{D}_0) / [1 - \underline{\pi} \mathbf{D}_0 \underline{e}] , \quad (3.33)$$

und mit dieser Lösung ergibt sich nach (3.13) mit $\nu = 0$ die mittlere Zwischenankunftszeit zu

$$E[T_A] = \underline{P} \sum_{k=1}^{\infty} k \mathbf{D}_0^{k-1} (\widehat{\mathbf{D}} - \mathbf{D}_0) \underline{e} = \underline{P} (\mathbf{I} - \mathbf{D}_0)^{-1} \underline{e} = 1 / [1 - \underline{\pi} \mathbf{D}_0 \underline{e}] . \quad (3.34)$$

Allgemein ist

$$E[T_A]^{-1} \neq \lambda := \underline{\pi} \sum_{w=1}^{\infty} w \mathbf{D}_w \underline{e} , \quad (3.35)$$

wobei λ die Ankunftsrate bedeutet.

3.4.1 Spezialfälle

Discrete Markovian Arrival Process (DMAP) Für $b_{max} = 1$ wird der SP in der Literatur als DMAP bezeichnet, und es ist üblich, $\mathbf{C} := \mathbf{D}_0$ und $\mathbf{D} := \mathbf{D}_1$ zu schreiben.

Beim DMAP ist natürlich $\lambda = E[T_A]^{-1} = \underline{\pi} \mathbf{D} \underline{e}$.

Gibt es nur eine Phase, d.h. \mathbf{C} und \mathbf{D} sind (reelle) Zahlen, so repräsentiert der DMAP den Bernoulli-Prozeß GEO (geometrisch verteilte Zwischenankunftszeiten⁶). Der Bernoulli-Prozeß ist das zeitdiskrete Analogon des Poisson-Prozesses; wie die negativ-exponentielle Verteilung ist auch die geometrische Verteilung *gedächtnislos*.

⁶Die Anzahl der Ankünfte gehorcht der Bernoulli-Verteilung, die auch Binomial-Verteilung heißt.

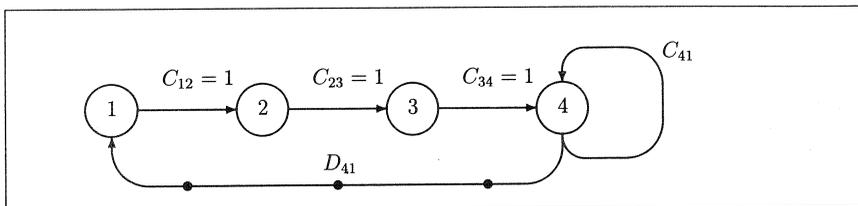


Abb. 3.1 3-fach verschobener GEO: $f(k) = \mathbf{1}_{k \geq 4} D_{41} C_{41}^{k-4}$.

DMAP^[X] Ein DMAP erzeugt Ankunftszeitpunkte; der davon unabhängige SP X unabhängiger, identisch verteilter ZVN X_n bestimmt beim n -ten Ankunftszeitpunkt die Größe des ankommenden Batch nach Maßgabe der WF $x(v) = P\{X_n = v\}$. Der DMAP^[X] ist vollständig durch die Matrizen $D_0 = C$ und $D_w = x(w) \cdot D$, $1 \leq w \leq b_{max}$ gegeben [10].

Die nachfolgenden Fälle des zeitverschobenen DMAP und des DMAP mit Vielfachen der Zeiteinheit sind mit Blick auf die Nachbildung realer Einzelquellen von großer Bedeutung, da die kleinste mögliche Zwischenankunftszeit größer als die Zeiteinheit ist. Eine Videoquelle mit maximaler Bitrate 100 Mbit/s, die auf eine 150-Mbit/s-Leitung eingespeist wird, erzeugt als kleinste Zwischenankunftszeit (etwa) 6 Zeiteinheiten, wenn als Zeiteinheit die Länge einer Zelle auf einer 600-Mbit/s-Leitung festgelegt wird, wie in Abb. 2.3 erörtert. Hat die Videoquelle nur Bitraten, die ganzzahlige Bruchteile der maximalen Bitrate (100, 50, 33, 25, 20, ... Mbit/s) sind, so ist der DMAP mit Vielfachen der Zeiteinheit ein gültiges Modell. Ist eine feinere Bitrateneinteilung vorhanden, so beschreibt der zeitverschobene DMAP die Quelle (unter diesem Gesichtspunkt) in geeigneter Weise. In beiden Fällen bestimmt die kleinste mögliche Zwischenankunftszeit die Anzahl der Phasen. Gegenüber den allgemeinen SMP ist das ein Nachteil. Bei diesen hat die Wahl der Zustandsverweilzeit-WFn keinen Einfluß auf die Anzahl der Zustände des Prozesses und damit auf die Größe der Matrix, die letztlich das gespeiste Wartesystem beschreibt.

Zeitverschobener DMAP Sei ein DMAP durch seine Matrizen C und D gegeben. Dann bestimmen die Matrizen \tilde{C} und \tilde{D} mit τ Hauptzeilen und -spalten

$$\tilde{C} = \begin{pmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{C} \end{pmatrix}, \quad \tilde{D} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix} \quad (3.36)$$

einen DMAP mit Zwischenankunftszeiten, die um $\tau - 1$ größer sind als diejenigen des von C und D erzeugten DMAP. Der Korrelationskoeffizient der Zwischenankunftszeiten wird durch diese Verschiebung nicht beeinflusst, wohl aber derjenige des Zählprozesses. Für GEO zeigt Abb. 3.1, wie es zu dieser Zeitverschiebung kommt.

DMAP mit Vielfachen der Zeiteinheit Ganz ähnlich läßt sich ein DMAP, bestimmt durch \tilde{C} und \tilde{D} (mit f Hauptzeilen und -spalten), angeben, der gegenüber dem von C und

\mathbf{D} bestimmten DMAP jeweils nur f -fache Vielfache der Zeiteinheit als mögliche Zwischenankunftszeiten erzeugt:

$$\tilde{\mathbf{C}} = \begin{pmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \\ \mathbf{C} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \tilde{\mathbf{D}} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}. \quad (3.37)$$

Für GEO erläutert Abb. 3.2 die Entstehung der Vielfachen. In diesem Fall ändert sich der Korrelationskoeffizient der Zwischenankunftszeiten, während derjenige des Zählprozesses mit einem Bezugsintervall der f -fachen Länge erhalten bleibt.

Beweis: a) zeitverschobener DMAP Sei $\tilde{\pi} = (\pi_1, \dots, \pi_{\tau-1}, \pi_\tau)$. $\tilde{\pi}(\tilde{\mathbf{C}} + \tilde{\mathbf{D}}) = \tilde{\pi}$ liefert $\pi_\tau \mathbf{D} = \pi_1, \pi_1 = \pi_2, \dots, \pi_{\tau-2} = \pi_{\tau-1}, \pi_{\tau-1} + \pi_\tau \mathbf{C} = \pi_\tau$, und dies führt zu $\pi_\tau(\mathbf{C} + \mathbf{D}) = \pi_\tau$. Da $\tilde{\pi}$ in $\tilde{\pi}(\mathbf{C} + \mathbf{D}) = \tilde{\pi}$ mit $\tilde{\pi} \mathbf{e} = 1$ die eindeutig bestimmte stochastische Lösung ist, gibt es eine positive Konstante α , so daß $\pi_\tau = \alpha \pi$. Mit $1 = \tilde{\pi} \tilde{\mathbf{e}} = (\pi_1 + \dots + \pi_{\tau-1} + \pi_\tau) \mathbf{e} = [(\tau - 1)\alpha \pi \mathbf{D} + \alpha \pi] \mathbf{e}$, folgt schließlich

$$\alpha = 1/(1 + (\tau - 1)\lambda), \quad \lambda = \pi \mathbf{D} \mathbf{e}. \quad (3.38)$$

Die Ankunftsrate $\lambda^{(s)} = \tilde{\pi} \tilde{\mathbf{D}} \tilde{\mathbf{e}}$ (s für *shifted*) des zeitverschobenen DMAP ergibt sich zu $\lambda^{(s)} = \pi_\tau \mathbf{D} \mathbf{e} = \alpha \cdot \lambda$, und es ist

$$\frac{\pi_\tau}{\lambda^{(s)}} = \frac{\pi}{\lambda}. \quad (3.39)$$

Für den stationären Phasenvektor $\tilde{\mathbf{P}}$ zu Ankunftszeitpunkten erhält man $\tilde{\mathbf{P}} = \tilde{\pi} \tilde{\mathbf{D}} / \lambda^{(s)} = (\pi \mathbf{D} / \lambda, \underline{0}, \dots, \underline{0}) = (\underline{P}, \underline{0}, \dots, \underline{0})$. Der SM-Kern lautet (Matrizen mit τ Hauptzeilen und -spalten)

$$\tilde{\mathbf{C}}^{k-1} \tilde{\mathbf{D}} = \begin{cases} \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & & & \\ \mathbf{D} & & & \\ \mathbf{C} \mathbf{D} & & & \\ \vdots & \vdots & & \vdots \\ \mathbf{C}^{k-1} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix} & \text{für } 1 \leq k \leq \tau - 1, \\ \begin{pmatrix} \mathbf{C}^{k-\tau} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}^{k-\tau+1} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}^{k-1} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix} & \text{für } k \geq \tau. \end{cases}$$

Diese Matrizen führen auf die WF der Zwischenankunftszeiten

$$P\{A_n = k\} = \tilde{\mathbf{P}} \tilde{\mathbf{C}}^{k-1} \tilde{\mathbf{D}} \tilde{\mathbf{e}} = \begin{cases} 0 & \text{für } 1 \leq k \leq \tau - 1, \\ \mathbf{P} \mathbf{C}^{k-\tau} \mathbf{D} \mathbf{e} & \text{für } k \geq \tau. \end{cases}$$

Entsprechende Ausdrücke ergeben sich für die Verbund-WF'n nach Gl. (3.13). Da der Korrelationskoeffizient verschiebungsinvariant ist, bleibt derjenige der Zwischenankunftszeiten

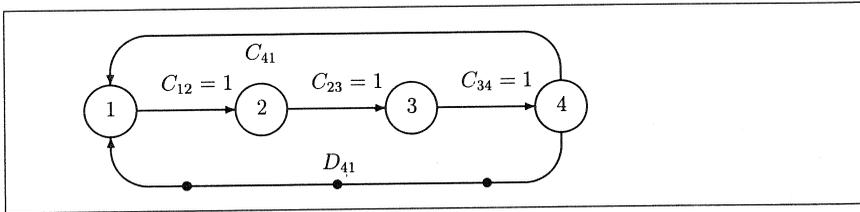


Abb. 3.2 GEO mit vervierfachen Zwischenankunftszeiten.

nach Zeitverschiebung unverändert. Die zu Ankunftszeiten eingebettete MK hat die Übergangsmatrix:

$$(\tilde{\mathbf{I}} - \tilde{\mathbf{C}})^{-1} \tilde{\mathbf{D}} = \begin{pmatrix} (\mathbf{I} - \mathbf{C})^{-1} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} \\ (\mathbf{I} - \mathbf{C})^{-1} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ (\mathbf{I} - \mathbf{C})^{-1} \mathbf{D} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}$$

b) DMAP mit Vielfachen der Zeiteinheit Es ist für $\alpha \in \mathbb{N}$

$$\check{\mathbf{C}}^{\alpha f-1} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{C}^{\alpha-1} \\ \mathbf{C}^\alpha & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^\alpha & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{C}^\alpha & \mathbf{0} \end{pmatrix}, \quad \check{\mathbf{C}}^{\alpha f} = \begin{pmatrix} \mathbf{C}^\alpha & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^\alpha & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \mathbf{0} \\ \vdots & & & & \mathbf{C}^\alpha & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{C}^\alpha \end{pmatrix}$$

$\tilde{\pi}(\tilde{\mathbf{C}} + \tilde{\mathbf{D}}) = \tilde{\pi}$ mit $\tilde{\pi} = (\pi_1, \dots, \pi_{f-1}, \pi_f)$ ergibt $\pi_1 = \pi/f$ mit $\pi(\mathbf{C} + \mathbf{D}) = \pi, \pi \underline{e} = 1$, so daß $\tilde{P} = \tilde{\pi} \tilde{\mathbf{D}} / \lambda^{(m)} = (\underline{P}, \underline{0}, \dots, \underline{0})$ mit $\lambda^{(m)} = \lambda/f$ (m für *multiplied*), $\lambda = \pi \mathbf{D} \underline{e}$ und $\underline{P} = \pi \mathbf{D} / \lambda$. Daher

$$\tilde{P} \check{\mathbf{C}}^{k-1} \tilde{\mathbf{D}} \underline{e} = \begin{cases} \underline{P} \mathbf{C}^{k/f-1} \mathbf{D} \underline{e} & \text{falls } k \text{ Vielfaches von } f \text{ ist,} \\ 0 & \text{sonst.} \end{cases} \quad \square$$

Überlagerung zweier unabhängiger DMAP^[X] Die Überlagerung von DMAP^[X₁] und DMAP^[X₂], vorgegeben durch $(\mathbf{C}^{(1)}, \mathbf{D}^{(1)}, x_1(v))$ und $(\mathbf{C}^{(2)}, \mathbf{D}^{(2)}, x_2(v))$ ergibt einen DBMAP mit folgenden Matrizen (\otimes : Kronecker-Produkt)

$$\mathbf{D}_v = \begin{cases} \mathbf{C}^{(1)} \otimes \mathbf{C}^{(2)} & \text{für } v = 0 \\ x_1(v) [\mathbf{D}^{(1)} \otimes \mathbf{C}^{(2)}] + \sum_{\ell=1}^{v-1} x_1(\ell) x_2(v-\ell) [\mathbf{D}^{(1)} \otimes \mathbf{D}^{(2)}] \\ + x_2(v) [\mathbf{C}^{(1)} \otimes \mathbf{D}^{(2)}] & \text{für } 1 \leq v \leq b_{max1} + b_{max2} \end{cases} \quad (3.40)$$

Markov-modulierter Bernoulli-Prozeß (MMBP) Das zeitdiskrete Analogon des MMPP [50] ergibt sich, wenn anstelle des Markov-Prozesses, der die Rate des Poisson-Prozesses steuert, eine MK die Rate eines Bernoulli-Prozesses bestimmt. Ist \mathbf{p} die Übergangsmatrix der steuernden MK und $1 - q_i$ die Rate (Kehrwert der mittleren Zwischenankunftszeit) des Bernoulli-Prozesses, solange die MK ununterbrochen den Zustand i besucht,

so erzeugt der DMAP mit $\mathbf{C} := \mathbf{Diag}(q_1, \dots, q_m) \cdot \mathbf{p}$ und $\mathbf{D} := \mathbf{Diag}(1 - q_1, \dots, 1 - q_m) \cdot \mathbf{p}$ den MMBP. In [47] wird (bei etwas anderer Notation) $\mathbf{C} := \mathbf{Diag}(q_1, \dots, q_m)$ gesetzt und der so erzeugte SP als MMBP bezeichnet. Bei dieser Wahl sind aber nicht alle Bernoulli-Muster (Folgen von 0 und 1) während des ununterbrochenen Verweilens der MK in einem Zustand möglich. Entweder wird ein solches Muster immer mit einer 0 begonnen oder immer mit einer 0 beendet (je nachdem ob man einen Übergang vom Typ 0 oder 1 der MK in einen anderen Zustand $j \neq i$ noch zum Bernoulli-Muster des Zustandes i zählt oder schon zu dem des neuen Zustandes j).

3.4.2 Der Zählprozeß eines DBMAP

Alternativ zu den Zwischenankunftszeiten kann ein stochastischer Prozeß auch charakterisiert werden, indem man für alle t nach der Anzahl der Ankünfte im Intervall $(0, t]$ fragt. Der so bestimmte SP heißt dann Zählprozeß. Ein Erneuerungsprozeß mit negativ-exponentiellen Zwischenankunftszeiten hat als Zählprozeß einen Poisson-Prozeß. Bei geometrischer Verteilung dieser Zeiten ist der Zählprozeß ein Bernoulli-Prozeß.

Nach [10] ist der Zählprozeß eines DBMAP eine zweidimensionale MK $\{(N(0, t], S(t)), t \in \mathbb{N}_0\}$, wobei $N(0, t]$ die Anzahl der Ankünfte im Intervall $(0, t]$ und $S(t) \in \{1, 2, \dots, m\}$ die Phase des DBMAP repräsentiert. Die Übergangsmatrix \mathbf{T} ist unendlich-dimensional und hat Blockstruktur

$$\mathbf{T} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{D}_1 & \mathbf{D}_2 & \mathbf{D}_3 & \cdots \\ \mathbf{0} & \mathbf{D}_0 & \mathbf{D}_1 & \mathbf{D}_2 & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_0 & \mathbf{D}_1 & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Ein Übergang vom Verbundzustand (v, i) nach (w, j) ($w \geq v$) entspricht einer Ankunft von $w - v \geq 0$ Kunden und einem Phasenübergang von i nach j . Die Übergangs-WKTN lauten also

$$\begin{aligned} P\{N(0; t+1] = w, S(t+1) = j \mid N(0; t] = v, S(t) = i\} = \\ = \left[[\mathbf{T}]_{vw} \right]_{ij} = \begin{cases} [\mathbf{D}_{w-v}]_{ij} & \text{für } w \geq v \\ 0 & \text{sonst} \end{cases}, \end{aligned}$$

wobei für $t = 0$ sinnvoll definiert wird durch (für die Indikatorfunktion $\mathbf{1}$ vgl. Anhang C)

$$P\{N(0; 0] = v, S(0) = i\} := \mathbf{1}_{v=0} \cdot P\{S(0) = i\}.$$

Zur Berechnung der Übergangs-WKTN nach t_0 Schritten kann ausgenutzt werden, daß die w -te Hauptzeile ($w = 0, 1, 2, \dots$) von \mathbf{T} aus der nullten durch Verschiebung um w Hauptspalten nach rechts und Auffüllen mit $\mathbf{0}$ -Matrizen hervorgeht. Diese Eigenschaft setzt sich auf jede Potenz von \mathbf{T} fort, d.h.

$$\left[[\mathbf{T}^{t_0}]_{vw} \right]_{ij} = \left[[\mathbf{T}^{t_0}]_{0w-v} \right]_{ij} \quad \text{für } w \geq v. \quad (3.41)$$

Mit der Markoveigenschaft und der Zeithomogenität erhält man daher:

$$\begin{aligned} P\{N(0; t+t_0] = w, S(t+t_0) = j \mid N(0; t] = v, S(t) = i\} = \\ = \left[[\mathbf{T}^{t_0}]_{vw} \right]_{ij} = \left[[\mathbf{T}^{t_0}]_{0w-v} \right]_{ij} = \end{aligned}$$

$$= P\{N(0; t_0] = w - v, S(t_0) = j \mid N(0; 0] = 0, S(0) = i\} . \quad (3.42)$$

Dieser Zusammenhang ist für die Lösung des Wartesystems in Kapitel 5 sehr wichtig. Er besagt, daß die WKT für $w - v \geq 0$ Zuwächse im Intervall $(t, t + t_0]$ nur abhängt von der Länge des Intervalls und den Phasen $S(t)$ zu Beginn und $S(t + t_0)$ am Ende des Intervalls, kurz:

Die Zuwächse sind unabhängig unter der Bedingung vorgegebener Phasen.

Nach Gl. (3.41) und der oberen Dreiecksstruktur von \mathbf{T} und jeder Potenz von \mathbf{T} reicht es aus, nur die erste Hauptzeile von \mathbf{T}^t zu berechnen, um \mathbf{T}^t ganz zu kennen. Zur (rekursiven) Berechnung der Einträge werden folgende Abkürzungen definiert:

$$[\mathbf{P}(w, t_0)]_{ij} := P\{N(0; t_0] = w, S(t_0) = j \mid N(0; 0] = 0, S(0) = i\} ,$$

$$\mathbf{P}(w, 0) = \mathbf{1}_{w=0} \cdot \mathbf{I} ,$$

\mathbf{I} : $(m \times m)$ -Einheitsmatrix .

Die Betrachtung des Produktes $\mathbf{T}^{t+1} = \mathbf{T}^t \cdot \mathbf{T}$ liefert die Rekursion

$$\mathbf{P}(w, t+1) = [\mathbf{T}^{t+1}]_{0w} = \begin{cases} [\mathbf{T}^t]_{00} \cdot \mathbf{D}_0 = \mathbf{P}(0, t) \cdot \mathbf{D}_0 & \text{für } w = 0, \\ \sum_{v=0}^w [\mathbf{T}^t]_{0v} \mathbf{D}_{w-v} = \sum_{v=0}^w \mathbf{P}(v, t) \mathbf{D}_{w-v} & \text{für } w \geq 1. \end{cases} \quad (3.43)$$

Mittels vollständiger Induktion ergibt sich aus Gl. (3.43) die z -Transformierte gemäß

Hilfssatz 3 Für alle $t \geq 0$ gilt:

$$\sum_{w=0}^{\infty} \mathbf{P}(w, t) z^w = \left[\sum_{w=0}^{b_{\max}} \mathbf{D}_w z^w \right]^t . \quad (3.44)$$

Im Spezialfall des DMAP^[X] hat man

$$\sum_{w=0}^{\infty} \mathbf{P}(w, t) z^w = (\mathbf{C} + \sum_{w=1}^{b_{\max}} \mathbf{D}_w z^w)^t = (\mathbf{C} + \widehat{X}(z) \mathbf{D})^t . \quad (3.45)$$

Dabei ist $\widehat{X}(z) = \sum_{v=1}^{\infty} x(v) z^v$ die erzeugende Funktion der Batchgröße.

Momentenmatrizen für den Zählprozeß ergeben sich durch Anwendung des Grenzwertsatzes der z -Transformation. Für die Matrixdifferenzierung gilt die Produktregel:

$\frac{d}{dz} [\mathbf{M}_1(z) \cdot \mathbf{M}_2(z)] = \frac{d}{dz} \mathbf{M}_1(z) \cdot \mathbf{M}_2(z) + \mathbf{M}_1(z) \cdot \frac{d}{dz} \mathbf{M}_2(z)$. Sie liefert wieder mittels vollständiger Induktion und der Identität $\sum_{w=0}^{\infty} \mathbf{P}(w, t+1) z^w = [\sum_{w=0}^{\infty} \mathbf{P}(w, t) z^w] \cdot [\sum_{w=0}^{\infty} \mathbf{D}_w z^w]$

$$\begin{aligned} \sum_{w=1}^{\infty} w \mathbf{P}(w, t) z^{w-1} &= \frac{d}{dz} \left[\sum_{w=0}^{\infty} \mathbf{D}_w z^w \right]^t = \\ &= \sum_{i=1}^t \left(\sum_{w=0}^{\infty} \mathbf{P}(w, i-1) z^w \right) \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, 1) z^{v-1} \right] \left(\sum_{w=0}^{\infty} \mathbf{P}(w, t-i) z^w \right) , \end{aligned} \quad (3.46)$$

$$\begin{aligned} \sum_{w=2}^{\infty} w(w-1) \mathbf{P}(w, t) z^{w-2} &= \frac{d^2}{dz^2} \left[\sum_{w=0}^{\infty} \mathbf{D}_w z^w \right]^t = \\ &= \sum_{i=1}^t \left\{ \left(\sum_{w=1}^{\infty} w \mathbf{P}(w, i-1) z^{w-1} \right) \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, 1) z^{v-1} \right] \left(\sum_{w=0}^{\infty} \mathbf{P}(w, t-i) z^w \right) + \right. \end{aligned}$$

$$\begin{aligned}
& + \left(\sum_{w=0}^{\infty} \mathbf{P}(w, i-1) z^w \right) \left[\sum_{v=2}^{\infty} v(v-1) \mathbf{P}(v, 1) z^{v-2} \right] \left(\sum_{w=0}^{\infty} \mathbf{P}(w, t-i) z^w \right) + \\
& + \left(\sum_{w=0}^{\infty} \mathbf{P}(w, i-1) z^w \right) \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, 1) z^{v-1} \right] \left(\sum_{w=1}^{\infty} w \mathbf{P}(w, t-i) z^{w-1} \right) \Big\}. \quad (3.47)
\end{aligned}$$

Die Ausdrücke (3.46) und (3.47) werden benötigt, um die Korrelationsfunktion der Anzahl von Ankünften im nachfolgenden Abschnitt zu berechnen.

3.5 Korrelationseigenschaften des DBMAP

3.5.1 Korrelationsfunktion des Zählprozesses

Hilfssatz 4 Für die Anzahl der Ankünfte $N(t, t + \tau]$ eines DBMAP im Intervall $(t, t + \tau]$ gelten die folgenden (stationären) WFn:

$$P\{N(t, t + \tau] = k\} = \underline{\pi} \mathbf{P}(k, \tau) \underline{e}, \quad (3.48)$$

$$P\{N(t, t + \tau] = k, N(t + h\tau, t + (h+1)\tau] = \ell\} = \underline{\pi} \mathbf{P}(k, \tau) \widehat{\mathbf{D}}^{(h-1)\tau} \mathbf{P}(\ell, \tau) \underline{e}. \quad (3.49)$$

$\underline{\pi}$ bezeichnet den stationären Phasenvektor des DBMAP.

Beweis: a) $P\{N(t, t + \tau] = k\} =$

$$\begin{aligned}
& = \sum_{i, j \in \mathbb{E}} \sum_{\ell=0}^{\infty} P\{S(t + \tau) = j, N(0, t + \tau] = k + \ell | S(t) = i, N(0, t] = \ell\} \cdot \\
& \quad \cdot P\{S(t) = i, N(0, t] = \ell\} =
\end{aligned}$$

$$= \sum_{i, j \in \mathbb{E}} [\mathbf{P}(k, \tau)]_{ij} \cdot \sum_{\ell=0}^{\infty} P\{S(t) = i, N(0, t] = \ell\} = \underline{\pi} \mathbf{P}(k, \tau) \underline{e}$$

b) $P\{N(t, t + \tau] = k, N(t + h\tau, t + (h+1)\tau] = \ell\} =$

$$= \sum_{\ell', \ell''} P\{N(0, t] = \ell', N(0, t + \tau] = k + \ell', N(0, t + h\tau] = k + \ell' + \ell'', N(0, t + (h+1)\tau] = k + \ell + \ell' + \ell''\} =$$

$$= \sum_{\ell', \ell''} \sum_{i_1, i_2, i_3, i_4 \in \mathbb{E}} P\{N(0, t + (h+1)\tau] = k + \ell + \ell' + \ell'', S(t + (h+1)\tau) = i_4 | N(0, t + h\tau] = k + \ell' + \ell'', S(t + h\tau) = i_3\} \cdot$$

$$\cdot P\{N(0, t + h\tau] = k + \ell' + \ell'', S(t + h\tau) = i_3 | N(0, t + \tau] = k + \ell', S(t + h\tau) = i_2\} \cdot$$

$$\cdot P\{N(0, t + \tau] = k + \ell', S(t + h\tau) = i_2 | N(0, t] = \ell', S(t) = i_1\} \cdot$$

$$\cdot P\{N(0, t] = \ell', S(t) = i_1\} =$$

$$= \sum_{\ell'} \sum_{i_1, i_2, i_3, i_4 \in \mathbb{E}} [\mathbf{P}(\ell, \tau)]_{i_3, i_4} [\mathbf{P}(\ell', (h-1)\tau)]_{i_2, i_3} [\mathbf{P}(k, \tau)]_{i_1, i_2} [\underline{\pi}]_{i_1} =$$

$$= \underline{\pi} \mathbf{P}(k, \tau) \widehat{\mathbf{D}}^{(h-1)\tau} \mathbf{P}(\ell, \tau) \underline{e} \quad \square$$

Mit den aus Gl.(3.46) folgenden Identitäten

$$\underline{\pi} \sum_{v=1}^{\infty} v \mathbf{P}(v, \tau) = \underline{\pi} \sum_{v=1}^{\infty} v \mathbf{P}(v, 1) \sum_{i=1}^{\tau} \widehat{\mathbf{D}}^{\tau-i} = \underline{\pi} \sum_{v=1}^{\infty} v \mathbf{D}_v \sum_{i=1}^{\tau} \widehat{\mathbf{D}}^{\tau-i}, \quad (3.50)$$

$$\sum_{v=1}^{\infty} v \mathbf{P}(v, \tau) \underline{e} = \sum_{i=1}^{\tau} \widehat{\mathbf{D}}^{i-1} \sum_{v=1}^{\infty} v \mathbf{P}(v, 1) \underline{e} = \sum_{i=1}^{\tau} \widehat{\mathbf{D}}^{i-1} \sum_{v=1}^{\infty} v \mathbf{D}_v \underline{e} \quad (3.51)$$

ergeben sich die die Korrelationsfunktion bestimmenden Erwartungswerte mit der Abkürzung $\mathcal{N}_t := N(t, t + \tau)$ unter Verwendung der Gln. (3.46), (3.47), (3.48) und (3.49) zu:

$$\begin{aligned} E[\mathcal{N}_t \cdot \mathcal{N}_{t+h\tau}] &= \underline{\pi} \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, \tau) \right] \widehat{\mathbf{D}}^{(h-1)\tau} \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, \tau) \right] \underline{e} = \\ &= \underline{\pi} \left[\sum_{v=1}^{\infty} v \mathbf{D}_v \sum_{i=1}^{\tau} \widehat{\mathbf{D}}^{\tau-i} \right] \widehat{\mathbf{D}}^{(h-1)\tau} \left[\sum_{i=1}^{\tau} \widehat{\mathbf{D}}^{i-1} \sum_{v=1}^{\infty} v \mathbf{D}_v \right] \underline{e} \quad , \end{aligned} \quad (3.52)$$

$$E[\mathcal{N}_t^2] = 2\underline{\pi} \left[\sum_{v=1}^{\infty} v \mathbf{D}_v \right] \sum_{i=0}^{\tau-2} \sum_{\xi=0}^i \widehat{\mathbf{D}}^{\xi} \left[\sum_{v=1}^{\infty} v \mathbf{D}_v \right] \underline{e} + \tau \underline{\pi} \sum_{v=1}^{\infty} v^2 \mathbf{D}_v \underline{e} \quad , \quad (3.53)$$

$$E[\mathcal{N}_t] = \tau \underline{\pi} \sum_{v=1}^{\infty} v \mathbf{D}_v \underline{e} \quad . \quad (3.54)$$

Der Korrelationskoeffizient h -ter Ordnung lautet

$$\text{Corr}[\mathcal{N}_t, \mathcal{N}_{t+h\tau}] = \frac{E[\mathcal{N}_t \cdot \mathcal{N}_{t+h\tau}] - E^2[\mathcal{N}_t]}{E[\mathcal{N}_t^2] - E^2[\mathcal{N}_t]} \quad (3.55)$$

und ergibt als Funktion von h die Korrelationsfunktion.

3.5.1.1 Der Fall des DMAP^[X]

$$\begin{aligned} E[\mathcal{N}_t \cdot \mathcal{N}_{t+h\tau}] &= \underline{\pi} \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, \tau) \right] (\mathbf{C} + \mathbf{D})^{(h-1)\tau} \left[\sum_{v=1}^{\infty} v \mathbf{P}(v, \tau) \right] \underline{e} = \\ &= [\bar{x}^{(1)}]^2 \underline{\pi} \left[\mathbf{D} \sum_{i=1}^{\tau} (\mathbf{C} + \mathbf{D})^{\tau-i} \right] (\mathbf{C} + \mathbf{D})^{(h-1)\tau} \left[\sum_{i=1}^{\tau} (\mathbf{C} + \mathbf{D})^{i-1} \mathbf{D} \right] \underline{e} \quad \text{für } h \geq 1, \end{aligned} \quad (3.56)$$

$$E[\mathcal{N}_t^2] = 2[\bar{x}^{(1)}]^2 \underline{\pi} \mathbf{D} \sum_{i=0}^{\tau-2} \sum_{\xi=0}^i (\mathbf{C} + \mathbf{D})^{\xi} \mathbf{D} \underline{e} + \bar{x}^{(2)} \tau \underline{\pi} \mathbf{D} \underline{e} \quad , \quad (3.57)$$

$$E[\mathcal{N}_t] = \bar{x}^{(1)} \tau \underline{\pi} \mathbf{D} \underline{e} \quad . \quad (3.58)$$

Unter Verwendung der Identitäten $\underline{\pi}(\mathbf{C} + \mathbf{D}) = \underline{\pi}$ und $(\mathbf{C} + \mathbf{D})\underline{e} = \underline{e}$ ergibt sich leicht, daß (3.55) invariant gegen eine Vertauschung von \mathbf{C} und \mathbf{D} ist, so daß es immer zwei DMAPs mit demselben Verlauf der Korrelationsfunktion gibt, aber unterschiedlicher Ankunftsrate $\underline{\pi} \mathbf{D} \underline{e}$ bzw. $1 - \underline{\pi} \mathbf{D} \underline{e}$, falls $\underline{\pi} \mathbf{D} \underline{e} \neq 0.5$.

3.5.2 Korrelationsfunktion der Zwischenankunftszeiten

Die Zwischenankunftszeiten eines DBMAP stimmen mit denen eines DMAP mit $\mathbf{C} = \mathbf{D}_0$ und $\mathbf{D} = \widehat{\mathbf{D}} - \mathbf{D}_0$ überein. Daher kann man sich auf den Fall eines DMAP beschränken. Mit der Übergangsmatrix $(\mathbf{I} - \mathbf{C})^{-1} \mathbf{D}$ der zu Ankunftszeiten eingebetteten MK liefert Gl. (3.13)

$$P\{A_n = t, A_{n+h} = t'\} = \underline{P} \mathbf{q}(t) [(\mathbf{I} - \mathbf{C})^{-1} \mathbf{D}]^{h-1} \mathbf{q}(t') \underline{e} \quad , \quad h \geq 1 \quad ,$$

und das ergibt

$$\begin{aligned} E[A_n \cdot A_{n+h}] &= \frac{1}{\underline{\pi} \mathbf{D} \underline{e}} \cdot \underline{\pi} [(\mathbf{I} - \mathbf{C})^{-1} \mathbf{D}]^h (\mathbf{I} - \mathbf{C})^{-1} \underline{e} \quad , \\ E[A_n^2] &= \frac{1}{\underline{\pi} \mathbf{D} \underline{e}} \left[2\underline{\pi} (\mathbf{I} - \mathbf{C})^{-1} \underline{e} - 1 \right] \quad , \quad E[A_n] = \frac{1}{\underline{\pi} \mathbf{D} \underline{e}} \quad . \end{aligned}$$

3.5.3 Dispersionsindizes

Für den Zählprozeß ergibt sich als Dispersionsindex $I(t) := \frac{\text{Var}[N(0, t)]}{\mathbb{E}[N(0, t)]}$ nach [103, 25]

$$I(t) = \frac{2\pi \left[\sum_{v=1}^{\infty} v \mathbf{D}_v \right] \sum_{i=0}^{t-2} \sum_{\xi=0}^i \widehat{\mathbf{D}}^\xi \left[\sum_{v=1}^{\infty} v \mathbf{D}_v \right] \mathbf{e} + t\pi \sum_{v=1}^{\infty} v^2 \mathbf{D}_v \mathbf{e}}{t\pi \sum_{v=1}^{\infty} v \mathbf{D}_v \mathbf{e}} - t\pi \sum_{v=1}^{\infty} v \mathbf{D}_v \mathbf{e} \quad (3.59)$$

Die Folge der Zwischenankunftszeiten besitzt den Dispersionsindex

$$c_k^2 = \frac{k \text{Var}[A_0] + 2 \sum_{h=1}^{k-1} (k-h) \text{Cov}[A_0, A_h]}{k \mathbb{E}^2[A_0]} \quad (3.60)$$

mit den Teilausdrücken $\text{Var}[A_0] = \frac{1}{\pi \mathbf{D} \mathbf{e}} \left[2\pi (\mathbf{I} - \mathbf{C})^{-1} \mathbf{e} - 1 - \frac{1}{\pi \mathbf{D} \mathbf{e}} \right]$ und $\text{Cov}[A_0, A_h] = \frac{1}{\pi \mathbf{D} \mathbf{e}} \left[\pi [(\mathbf{I} - \mathbf{C})^{-1} \mathbf{D}]^h (\mathbf{I} - \mathbf{C})^{-1} \mathbf{e} - \frac{1}{\pi \mathbf{D} \mathbf{e}} \right]$, $h \geq 1$.

3.6 Andere Darstellungen des DMAP

In [52] wurde der DMAP mittels einer absorbierenden MK definiert. Diese Definition erlaubt eine Semi-Markovsche Darstellung der Überlagerung zweier unabhängiger DMAPs im folgenden Abschnitt. Die Definition ist angelehnt an diejenige des zeitkontinuierlichen Falles in [76]⁷ und geht zurück auf die sog. "phase-type distributions" in [85].

Gegeben sei eine MK $\widehat{\mathbf{p}}$ mit Zustandsraum $\mathbb{I}E_t \cup \mathbb{I}E_a$, mit $\mathbb{I}E_t = \{t(1), t(2), \dots, t(m)\}$ und $\mathbb{I}E_a = \{a(1), a(2), \dots, a(m)\}$. Die Zustände von $\mathbb{I}E_t$ seien alle transient, diejenigen von $\mathbb{I}E_a$ alle absorbierend. Mit positiver WKT tritt nach endlicher Zeit eine Absorption auf. Die Übergangsmatrix hat bei einer Zeilennumerierung $t(1), t(2), \dots, t(m), a(1), a(2), \dots, a(m)$ die Struktur

$$\widehat{\mathbf{p}} = \begin{pmatrix} \mathbf{C} & \mathbf{D} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \quad (3.61)$$

\mathbf{C} enthält die Übergangs-WKTN zwischen transienten Zuständen, \mathbf{D} diejenigen zwischen transienten und absorbierenden. \mathbf{I} ist eine $(m \times m)$ -Einheitsmatrix, in der sich die Eigenschaft "absorbierend" widerspiegelt.

Unmittelbar nach einer Absorption in einem Zustand $a(i)$ wird die MK sofort neu gestartet im transienten Zustand $t(i)$ ($i = 1, 2, \dots, m$). Jede Absorption repräsentiert eine Ankunft, so daß sich die resultierenden Zwischenankunftszeiten als "Ersteintrittszeit" der MK in einen absorbierenden Zustand ergeben. Nach [61], p. 205 ergibt sich die WKT $\varphi_{t(i)a(j)}^{(k)}$, den absorbierenden Zustand $a(j)$ nach $k \geq 1$ Schritten und Start in $t(i)$ zu erreichen, zu

$$\varphi_{t(i)a(j)}^{(k)} = \left[\widehat{\mathbf{p}}^k - \widehat{\mathbf{p}}^{k-1} \right]_{t(i)a(j)}. \quad (3.62)$$

Dabei bedeutet $[\]_{t(i)a(j)}$ das Matrixelement in der $t(i)$ -ten Zeile und in der $a(j)$ -ten Spalte. In der sofort nach Absorption neugestarteten MK ist $\varphi_{t(i)a(j)}^{(k)}$ gleichbedeutend mit der bedingten WKT, daß unter der Bedingung einer Absorption (d. h. Ankunft) in $a(i)$ die nächste

⁷Allerdings wird hier gegenüber [76] nicht nur ein absorbierender Zustand angenommen, sondern genauso viele absorbierende wie transiente Zustände.

Absorption (Ankunft) nach k Schritten in $a(j)$ erfolgt. Damit ergibt sich als andere Sicht des Punktprozesses der Ankünfte ein (zeithomogener) SMP mit dem SM-Kern

$$P\{S_n = a(j), A_{n-1} = k | S_{n-1} = a(i)\} = \varphi_{t(i)a(j)}^{(k)}. \quad (3.63)$$

In Gl. (3.62) ist

$$\hat{\mathbf{p}}^k - \hat{\mathbf{p}}^{k-1} = \hat{\mathbf{p}}^{k-1} \cdot (\hat{\mathbf{p}} - \mathbf{I}) = \begin{pmatrix} \mathbf{C}^{k-1}(\mathbf{C} - \mathbf{I}) & \mathbf{C}^{k-1}\mathbf{D} \\ \mathbf{0} & \mathbf{0} \end{pmatrix},$$

und als SM-Kern ergibt sich der rechte obere Haupteintrag (Übergänge $t(i) \rightarrow a(j)$)

$$P\{S_n = a(j), A_{n-1} = k | S_{n-1} = a(i)\} = [\mathbf{C}^{k-1}\mathbf{D}]_{ij}, \quad 1 \leq i, j \leq m.$$

Die Markov-Erneuerungseigenschaft (3.5) folgt direkt aus der Markoveigenschaft der MK.

3.6.1 Semi-Markovsche Darstellung der Überlagerung DMAP₁+DMAP₂

Durch Überlagerung der absorbierenden MKn kann DMAP₁+DMAP₂ so beschrieben werden, daß der resultierende Zustandsraum nach Art der Ankünfte zerlegt ist.

Die Überlagerung zweier unabhängiger MKn $\{Y_n; n \in \mathbb{N}_0\}$, $\hat{\mathbf{p}}_1$, Zustandsraum \mathbb{E}_1 und $\{Z_n; n \in \mathbb{N}_0\}$, $\hat{\mathbf{p}}_2$, Zustandsraum \mathbb{E}_2 ergibt wiederum eine MK $\{(Y_n, Z_n); n \in \mathbb{N}_0\}$ mit dem Zustandsraum $\mathbb{E}_1 \times \mathbb{E}_2$. Die Übergangs-WKtn ergeben sich infolge der Unabhängigkeit zu

$$\begin{aligned} P\{Y_{n+1} = j_1, Z_{n+1} = j_2 | Y_n = i_1, Z_n = i_2\} = \\ P\{Y_{n+1} = j_1 | Z_{n+1} = j_2, Y_n = i_1, Z_n = i_2\} \cdot P\{Z_{n+1} = j_2 | Y_n = i_1, Z_n = i_2\} = \\ = P\{Y_{n+1} = j_1 | Y_n = i_1\} \cdot P\{Z_{n+1} = j_2 | Z_n = i_2\}, \end{aligned}$$

und daher kann die Übergangsmatrix durch das Kroneckerprodukt [59] $\hat{\mathbf{p}}_1 \otimes \hat{\mathbf{p}}_2$ (s. auch Anhang C) beschrieben werden. In der Überlagerung zweier absorbierender MKn sind nur diejenigen Zustände von $\mathbb{E}_1 \times \mathbb{E}_2$ absorbierend, die aus zwei absorbierenden Zuständen gebildet wurden.

In der Überlagerung zweier DMAPs, gegeben durch die absorbierenden MKn nach Gl. (3.61) mit $\mathbf{C}^{(1)}, \mathbf{D}^{(1)}, \mathbf{I}^{(1)}$ und $\mathbf{C}^{(2)}, \mathbf{D}^{(2)}, \mathbf{I}^{(2)}$, ist das wegen des sofortigen Neustarts einer der Ketten, wenn sie einen absorbierenden Zustand betreten hat, anders: Erreicht die Überlagerung etwa den Verbundzustand $(a(i_1), t(i_2))$, so geht es unmittelbar weiter mit $(t(i_1), t(i_2))$. Da Ersteintrittszeiten Stoppzeiten sind [20], kann man den Neustart eines der DMAPs als Neustart der gesamten Überlagerung in $(t(i_1), t(i_2))$ auffassen. Damit ist jeder Zustand der die DMAP-Überlagerung beschreibenden Kette $\mathbb{E}_1 \times \mathbb{E}_2$ absorbierend, der aus wenigstens einem absorbierenden (Teil-) Zustand gebildet wird. Als Übergangsmatrix einer absorbierenden MK, die die Überlagerung der beiden DMAPs bestimmt, ergibt sich somit:

$$\hat{\mathbf{P}} = \begin{pmatrix} \mathbf{C}^{(1)} \otimes \mathbf{C}^{(2)} & \mathbf{D}^{(1)} \otimes \mathbf{C}^{(2)} & \mathbf{C}^{(1)} \otimes \mathbf{D}^{(2)} & \mathbf{D}^{(1)} \otimes \mathbf{D}^{(2)} \\ \mathbf{0} & \mathbf{I}^{(1)} \otimes \mathbf{I}^{(2)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}^{(1)} \otimes \mathbf{I}^{(2)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}^{(1)} \otimes \mathbf{I}^{(2)} \end{pmatrix}.$$

Da Doppelankünfte auftreten, ist die Überlagerung kein DMAP mehr; ein SM-Kern für die Überlagerung ergibt sich aber wieder nach Gl. (3.62) zu⁸

$$\mathbf{q}(k) = \begin{pmatrix} \mathbf{C}^{k-1} \mathbf{D}_1^{spl} & \mathbf{C}^{k-1} \mathbf{D}_2^{spl} & \mathbf{C}^{k-1} \mathbf{D}^{ba} \\ \mathbf{C}^{k-1} \mathbf{D}_1^{spl} & \mathbf{C}^{k-1} \mathbf{D}_2^{spl} & \mathbf{C}^{k-1} \mathbf{D}^{ba} \\ \mathbf{C}^{k-1} \mathbf{D}_1^{spl} & \mathbf{C}^{k-1} \mathbf{D}_2^{spl} & \mathbf{C}^{k-1} \mathbf{D}^{ba} \end{pmatrix}, \quad (3.64)$$

mit $\mathbf{C} := \mathbf{C}^{(1)} \otimes \mathbf{C}^{(2)}$, $\mathbf{D}_1^{spl} := \mathbf{D}^{(1)} \otimes \mathbf{C}^{(2)}$, $\mathbf{D}_2^{spl} := \mathbf{C}^{(1)} \otimes \mathbf{D}^{(2)}$, $\mathbf{D}^{ba} := \mathbf{D}^{(1)} \otimes \mathbf{D}^{(2)}$. Die Matrix (3.64) hat identische Hauptzeilen, weil nach einer Absorption in $(a(j_1), t(j_2))$, $(t(j_1), a(j_2))$ oder $(a(j_1), a(j_2))$ die erzeugende MK in demselben transienten Zustand $(t(j_1), t(j_2))$ neugestartet wird. Die Zustände der ersten [zweiten] Hauptspalte in (3.64) repräsentieren Einfachankünfte von DMAP₁ [DMAP₂] und bilden den Teilraum $\mathbb{I}E_1^{spl}$ [$\mathbb{I}E_2^{spl}$]. Diejenigen der dritten Hauptspalte beziehen sich auf Doppelankünfte von beiden DMAPs und bilden den Teilraum $\mathbb{I}E^{ba}$, $\mathbb{I}E := \mathbb{I}E_1^{spl} \cup \mathbb{I}E_2^{spl} \cup \mathbb{I}E^{ba}$. (*spl* steht für *simple*, *ba* für *batch*). Die so gewonnene SM-Darstellung der Überlagerung zweier DMAPs wird in Kapitel 6 für die Gewinnung von Dienstgüteparametern des Wartesystems DMAP₁^[X]+DMAP₂/D/1/N mittels der Restarbeit für beide Teilströme getrennt benötigt. Durch die getrennte Behandlung der beiden Ströme steht damit ein stochastisches Modell für die Verbindungsannahmesteuerung zur Verfügung.

3.7 Die Markov-Erneuerungsfunktion

In Kapitel 5 werden für den in der Analyse des DBMAP/G/1/N-Wartesystems resultierenden Markov-Erneuerungsprozeß $\{(S_n, T_n); n \in \mathbb{N}_0\}$ mit *positiven* Zustandsverweilzeiten die folgenden WKTn benötigt

$$\hat{r}_{ij}(t) := P\{\text{nach } t \text{ Zeiteinheiten wird } j \text{ besucht} | S_0 = i\}. \quad (3.65)$$

Sie sind wegen der positiven Zustandsverweilzeiten gegeben durch

$$\hat{r}_{ij}(t) := \sum_{n=0}^t P\{S_n = j, T_n - T_0 = t | S_0 = i\}, \quad \hat{r}_{ij}(0) := \mathbf{1}_{i=j}. \quad (3.66)$$

Werden auch Zustandsverweilzeiten der Größe 0 zugelassen, so muß sich die Summation von 0 bis ∞ erstrecken.

Sei

$$q_{ij}^{(n)}(t) := P\{S_n = j, T_n - T_0 = t | S_0 = i\}. \quad (3.67)$$

Mit der Identität

$$\begin{aligned} P\{T_{n+1} - T_n = t_n, T_n - T_0 = t | S_{n+1} = i_{n+1}, S_n = i_n, S_0 = i_0\} = \\ P\{T_{n+1} - T_n = t_n | S_{n+1} = i_{n+1}, S_n = i_n\} \cdot P\{T_n - T_0 = t | S_n = i_n, S_0 = i_0\}, \end{aligned} \quad (3.68)$$

die leicht aus Gl. (3.10) und der Markov-Erneuerungseigenschaft Gl. (3.5) folgt, erhält man die Rekursion

$$q_{ij}^{(n)}(t) := \sum_{k \in \mathbb{E}} \sum_{s=0}^t q_{ik}^{(n-1)}(s) q_{kj}(t-s), \quad (3.69)$$

⁸Dazu berechnet man $\hat{\mathbf{p}}^{k-1} \cdot (\hat{\mathbf{p}} - \mathbf{I})$ durch $\hat{\mathbf{p}}^{k-2} \cdot [\hat{\mathbf{p}}(\hat{\mathbf{p}} - \mathbf{I})]$.

die üblicherweise als n -fache Matrixfaltung $[q^{*n}(t)]_{ij} := q_{ij}^{(n)}(t)$ [37] bezeichnet wird. Der Ausdruck

$$\hat{R}_{ij}(t) := \sum_{s=0}^t \hat{r}_{ij}(s) \quad \text{mit} \quad \hat{r}_{ij}(t) = \sum_{n=1}^t q_{ij}^{(n)}(t) \quad (3.70)$$

heißt "Markov-Erneuerungsfunktion", weil man zeigen kann, daß $\hat{R}_{ij}(t)$ gleich der mittleren Anzahl der Besuche zum Zustand j nach Start in i und Ablauf von t Zeiteinheiten ist, wobei für $i = j$ der Startzustand konventionsgemäß nicht mitgezählt wird [37], pp. 165, 170.

In Kapitel 5 wird außerdem der folgende Grenzwertsatz benötigt, der aus dem sog. "key renewal theorem" [20] hervorgeht. Er wird hier nur für den sog. d -periodischen MEP zitiert, da in dieser Arbeit ausschließlich zeitdiskrete MEP betrachtet werden. d -periodisch bedeutet bei einem (irreduziblen) MEP, daß $Q_{ij}(t)$ für alle $i, j \in IE$ eine Treppenfunktion mit Sprüngen bei $d_{ij}, d_{ij} + d, \dots$ ist. Darin ist $d_{ij} \geq 0$ die erste Sprungstelle [37], p. 168. Der Ausdruck "periodisch" mit Bezug auf einen MEP bezeichnet also nicht eine Eigenschaft der eingebetteten MK, sondern die Art des SM-Kerns.

Satz 1 (Pyke 1961)

Ist der MEP (S, T) d -periodisch und wird $\sum_{k \in \mathcal{K}(x, d, d_{ij})} g_j(x - d_{ij} + kd) < \infty$ für jede Funktion $g_j : \mathbb{R} \rightarrow \mathbb{R}$ erfüllt, so gilt für alle $i, j \in IE$, $x \in [0, d)$

$$\lim_{t \rightarrow \infty} \hat{R}_{ij} * g_j(x + td) = \frac{x_j}{E^*} \sum_{k \in \mathcal{K}(x, d, d_{ij})} g_j(x - d_{ij} + kd) . \quad (3.71)$$

Darin ist $\mathcal{K}(x, d, d_{ij}) := \{k \in \mathbb{N}_0 : x + kd \geq d_{ij}\}$, x_j bedeutet die stationäre Zustands-WKT der eingebetteten MK im Zustand j . E^* ist der stationäre Mittelwert der Zustandsverweilzeiten des MEP. $\hat{R} * g$ bezeichnet eine Faltung, wie sie im Anhang C definiert ist.

Ein Beweis des d -periodischen Falls findet sich nur in [95], p. 1251f. In [20]⁹ werden d -periodischer und aperiodischer Fall angegeben, aber nur der aperiodische Fall bewiesen, d.h. die Funktionen $Q_{ij}(t)$ sind stetig oder gemischt stetig-diskret [37], p. 139.

Der für die folgenden Kapitel ausschließlich erforderliche zeitdiskrete Fall ist enthalten für $d = 1$. Mit $x = 0$ und $d_{ij} \geq 1$ (positive Zustandsverweilzeiten) erhält man für Funktionen $g_j : \mathbb{N}_0 \rightarrow \mathbb{R}$, $j \in IE$ aus Gl. (3.71)

$$\lim_{t \rightarrow \infty} \sum_{k=0}^t \hat{r}_{ij}(k) g_j(t - k) = \frac{x_j}{E^*} \sum_{k=0}^{\infty} g_j(k) . \quad (3.72)$$

⁹Dort ist der d -periodische Fall mit einem Druckfehler aufgeführt: Mit den dort verwendeten Symbolen muß anstelle von $\frac{\delta}{\nu m} \sum_{k=0}^{\infty} \nu g(x - \delta_{ij} + k\delta)$ für die rechte Seite $\frac{\delta}{\nu m} \sum_{k=0}^{\infty} \sum_{j \in E} \nu_j g(j, x - \delta_{ij} + k\delta)$ erscheinen.

In [21], p.736 ist dies berücksichtigt.

4 Analyse zeitdiskreter Modelle mittels der Restarbeit

Bei unendlichem Wartespeicher ist die Wartezeit der Kunden von zentralem Interesse. In [22], p. 298 wird gezeigt, daß die Belegung zu Ankunftszeitpunkten auf die Wartezeit zurückgeführt werden kann. Die Entwicklung der Warteschlange über der Zeit kann formal durch die *virtuelle Wartezeit* zum Zeitpunkt t beschrieben werden. Diese ist die Wartezeit bei FIFO¹-Strategie, die sich für einen Kunden ergäbe, wenn er zum Zeitpunkt t ankommen würde. Die virtuelle Wartezeit als Funktion der Zeit hat einen charakteristischen sägezahnartigen Verlauf (Abb. 4.1). Ankunftszeitpunkte sind Sprungstellen. Die Sprunghöhe repräsentiert (bei Einfachankünften) die Bedienzeit der jeweiligen Ankunft. Für die virtuelle Wartezeit W_n beim n -ten Ankunftszeitpunkt T_n (das ist dann auch die eigentliche Wartezeit) gilt die bekannte Rekursionsvorschrift

$$W_n = \max[W_{n-1} + B_{n-1} - A_{n-1}, 0] , \quad (4.1)$$

wobei $A_{n-1} := T_n - T_{n-1}$ die Zwischenankunftszeit bedeutet und B_n die Bedienzeit der n -ten Ankunft. Aus dieser Beziehung läßt sich für das zeitkontinuierliche GI/GI/1/∞-Wartesystem die sog. Lindley-Gleichung [73] ableiten, deren Lösung die Wartezeit-VF ergibt. In [32] wird die Rekursion (4.1) benutzt, um die Wartezeit-WF des zeitdiskreten SSMP/GI/1-Systems iterativ zu bestimmen. Obwohl die Autoren sich auf den Fall des SSMP-Eingabeprozesses beschränken, ist die vorgestellte Lösung ohne prinzipielle Veränderungen auch für einen beliebigen SMP gültig.

Bei *endlichem* Wartespeicher muß der Begriff der virtuellen Wartezeit leicht verändert werden, da das Wartesystem voll besetzt sein kann. In Zeitintervallen mit vollem System ist es nicht mehr sinnvoll, von einer virtuellen Wartezeit zu sprechen, da eine Ankunft innerhalb dieses Zeitintervalls verlorengehe und man ihr dann üblicherweise keine Wartezeit zuordnet. An die Stelle der virtuellen Wartezeit tritt die sog. *Restarbeit* $U(t)$ zum Zeitpunkt t . Bei zeitkontinuierlichen Wartesystemen gibt sie die Zeit an, die von t ab vergehen muß, bis das Wartesystem zum ersten Mal wieder leer ist, wenn alle Ankünfte nach t ignoriert werden. Diese Definition bleibt im zeitdiskreten Fall (mit den beiden Möglichkeiten AF und DF, s. Abschnitt 2.2.1) weiterhin gültig. Solange das Wartesystem nicht voll besetzt ist, stimmt die Restarbeit mit der virtuellen Wartezeit überein.

Die nachfolgend entwickelte Analyse von endlichen Wartesystemen wurde durch [113] angeregt. Dort wird der Begriff der Restarbeit zur Analyse eines zeitdiskreten endlichen Wartesystems mit Erneuerungseingabeprozess und deterministischer Bedienzeit, die mit der Zeiteinheit des Systems übereinstimmt, für den Fall DF eingesetzt. [113] wird hier in verschiedenen Aspekten erweitert (AF und DF, beliebiger Semi-Markov-Eingabeprozess, schließlich auch Erweiterung auf einen beliebigen unabhängigen Bedienprozess). Neben den Methoden, die für Wartesysteme vom M/G/1-Typ geeignet sind, scheint der Begriff der Restarbeit der *einzige* Zugang zur Lösung *endlicher* Wartesysteme bei nicht zu starken Einschränkungen für den Eingabeprozess zu sein.

¹First-In-First-Out¹: Abfertigung der Kunden gemäß ihrer Ankunftsreihenfolge.

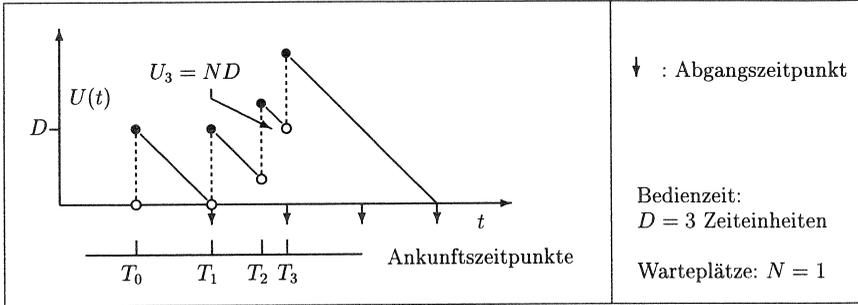


Abb. 4.1 Pfad der Restarbeit eines zeitkontinuierlichen Wartesystems.

4.1 Das zeitdiskrete SMP^[X]/D/1/N-Wartesystem

Es wird das SMP^[X]/D/1/N-Wartesystem mit FIFO-Bedienstrategie betrachtet [53] : Die (positiven) Zwischenankunftszeiten werden repräsentiert durch die Zustandsverweilzeiten $A_n := T_{n+1} - T_n$ eines SMP, es gibt N Warteplätze, und der Bedienprozeß $\{B_n; n \in \mathbb{N}_0\}$ ist deterministisch: $B_n = D$. Der SP $X := \{X_n; n \in \mathbb{N}_0\}$ aus unabhängigen, identisch verteilten ZVN X_n bestimmt die Anzahl der zum n -ten Ankunftszeitpunkt T_n gleichzeitig ankommenden Kunden²; $x(v) := P\{X_n = v\}$, $x(v) = 0$ für $v > b_{max}$, $X(v) := P\{X_n \leq v\}$, $X^c(v) := P\{X_n \geq v + 1\} = 1 - X(v)$. Zwischenankunftszeiten der Größe Null würden ebenfalls gleichzeitige Ankünfte erlauben. Dann kann aber nur DF behandelt werden, s. Abschnitt 4.2. Immer wenn mehrere Zellen gleichzeitig ankommen, werden soviele wie möglich ins System zugelassen.

Es werden die ZVN $U_n :=$ Restarbeit zu T_n ohne Berücksichtigung der Ankunft und $U_n^+ :=$ Restarbeit zu T_n mit Berücksichtigung der Ankunft betrachtet. Es wird auch die Sprechweise "die Ankunft sieht den Wert U_n der Restarbeit" benutzt. Im zeitkontinuierlichen Fall kann alternativ U_n als Restarbeit unmittelbar vor und U_n^+ als diejenige unmittelbar nach dem Ankunftszeitpunkt definiert werden, s. Abb. 4.1, also als links- bzw. rechtsseitiger Grenzwert der Zeitfunktion $U(t)$.

4.1.1 Die Zulassungsregeln Arrival First (AF) und Departure First (DF)

Die Restarbeit in einem zeitdiskreten System wird durch Treppenfunktionen repräsentiert, die für eine Zeiteinheit konstant bleiben und zu diskreten Zeitpunkten springen, Abb. 4.2. Es zeigt sich, daß die beiden Möglichkeiten der Festlegung der Werte an den Sprungstellen durch rechts- bzw. linksseitig stetige Treppenfunktionen zur Repräsentation von DF bzw. AF verwendet werden können: DF wird durch rechtsseitig stetige Treppenfunktionen, AF durch linksseitig stetige bestimmt. Dazu müssen definiert werden der Anfangswert $U_0 := 0$ und unter Verwendung der Grenzwerte $U(T+0) := \lim_{t \uparrow T} U(t)$, $U(T-0) := \lim_{t \downarrow T} U(t)$:

$$\text{für AF: } \begin{matrix} U_n := U(T_n) \\ U_n^+ := U(T_n + 0) \end{matrix} \quad \text{und für DF: } \begin{matrix} U_n := \max(U(T_n - 0) - 1, 0) \\ U_n^+ := U(T_n) \end{matrix}$$

²Wegen der ATM-Thematik und der deterministischen Bedienzeit wird im folgenden von Zellen und nicht von Kunden gesprochen.

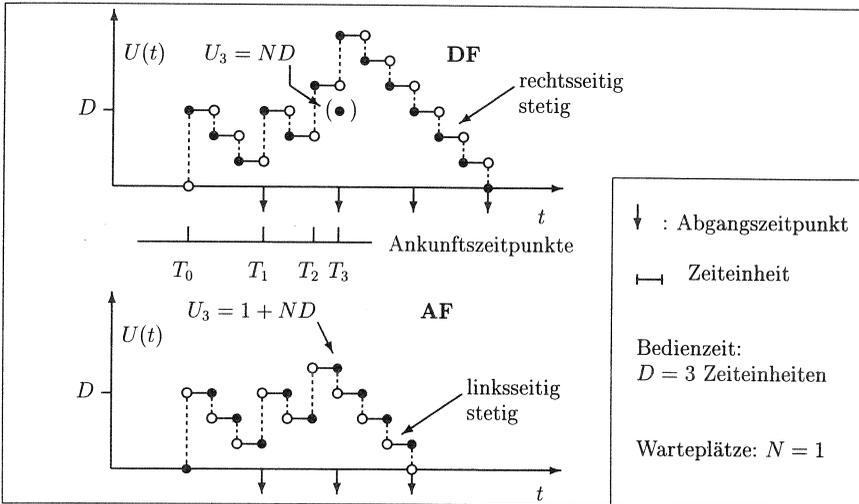


Abb. 4.2 Pfade der Restarbeit $U(t)$ für DF und AF: Die zu T_3 ankommende Zelle geht bei AF verloren, weil sie (anders als bei DF) noch die zu T_1 angekommene Zelle sieht.

Im zeitkontinuierlichen Fall werden diese beiden Fälle nicht unterschieden, weil Ankünfte und Abgänge nur mit WKT Null auftreten können [44].

Auch bei AF bleibt die o.g. Definition der Restarbeit gültig, s. Abb. 4.2: $U(T_0 + 1) = D$, der nächste Abgangszeitpunkt ist $T_0 + 3$, und das System ist zum ersten Mal leer bei $T_0 + 4 = U(T_0 + 1)$, wenn man die Ankünfte bei T_1, T_2, \dots ignoriert. Dieses etwas befremdliche Verhalten von $U(t)$ (nur bei AF) macht jedoch keine Schwierigkeiten, weil die zeitliche Entwicklung des Wartesystems nur durch U_n und U_n^+ bestimmt wird.

Es bleibt noch zu bemerken, daß infolge der Mehrfachankünfte hier $U(t)$ das Systemverhalten nur wegen der deterministischen Bedienzeit vollständig beschreibt. Man kann nämlich aus dem Wert von $U(t)$ auf die Belegung $\lceil (U(t)/D) \rceil^3$ schließen und damit entscheiden, ob das System voll ist oder nicht. Bei beliebigem Bedienprozeß und Mehrfachankünften ist das nicht mehr möglich.

4.1.2 Restarbeit zu Ankunftszeitpunkten

Zwischen U_{n-1} und U_{n-1}^+ besteht für DF die Beziehung

$$U_{n-1}^+ = \begin{cases} \left. \begin{aligned} &U_{n-1} + X_{n-1}D & 0 \leq U_{n-1} \leq (N - X_{n-1} + 1)D, \\ &U_{n-1} + \xi D & 1 + (N - \xi)D \leq U_{n-1} \leq (N - \xi + 1)D \\ && \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq X_{n-1} - 1 \end{aligned} \right\} & \text{für } 1 \leq X_{n-1} \leq N, \\ \left. \begin{aligned} &U_{n-1} + ND + D & U_{n-1} = 0, \\ &U_{n-1} + \xi D & 1 + (N - \xi)D \leq U_{n-1} \leq (N - \xi + 1)D \\ && \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq N \end{aligned} \right\} & \text{für } X_{n-1} \geq 1 + N. \end{cases} \quad (4.2)$$

Die ZVn U_n und U_{n-1}^+ sind verknüpft durch

$$U_n = \begin{cases} U_{n-1}^+ - A_{n-1} & \text{für } U_{n-1}^+ > A_{n-1}, \\ 0 & \text{sonst.} \end{cases} \quad (4.3)$$

³ $\lceil (U(t)/D) \rceil$ bezeichnet die nächst größere ganze Zahl, s. Symbolverzeichnis.

AF ist gegeben durch

$$U_{n-1}^+ = \left\{ \begin{array}{ll} U_{n-1} + X_{n-1}D & U_{n-1} = 0, \\ U_{n-1} + X_{n-1}D - 1 & 1 \leq U_{n-1} \leq (N - X_{n-1} + 1)D, \\ U_{n-1} + \xi D - 1 & 1 + (N - \xi)D \leq U_{n-1} \leq (N - \xi + 1)D \\ & \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq X_{n-1} - 1 \end{array} \right\} \text{für } 1 \leq X_{n-1} \leq N, \quad (4.4)$$

$$\left\{ \begin{array}{ll} U_{n-1} + ND + D & U_{n-1} = 0, \\ U_{n-1} + \xi D - 1 & 1 + (N - \xi)D \leq U_{n-1} \leq (N - \xi + 1)D \\ & \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq N \end{array} \right\} \text{für } X_{n-1} \geq 1 + N.$$

und

$$U_n = \begin{cases} U_{n-1}^+ - A_{n-1} + 1 & \text{für } U_{n-1}^+ \geq A_{n-1}, \\ 0 & \text{sonst.} \end{cases} \quad (4.5)$$

Wegen (4.2), (4.3) für DF und (4.4), (4.5) für AF ist $\{(U_n, S_n, X_n); n \in \mathbb{N}_0\}$ eine MK mit den Übergangs-WKTn $P\{U_n = \ell, S_n = j, X_n = w | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} = P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} \cdot x(w)$ gemäß dem folgenden Hilfssatz 6. Zunächst wird benötigt

Hilfssatz 5 Für alle $n \in \mathbb{N}$ gilt:

$$P\{S_n = j, A_{n-1} = t | S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} = P\{S_n = j, A_{n-1} = t | S_{n-1} = i\}.$$

Beweis: Von zentraler Bedeutung ist die Markov-Erneuerungseigenschaft Gl. (3.5). Jede Realisation $((t_0, i_0, v_0), \dots, (t_{n-2}, i_{n-2}, v_{n-2}), (t_{n-1}, i, \cdot))$ des Zufallsvektors $((T_0, S_0, X_0), \dots, (T_{n-2}, S_{n-2}, X_{n-2}), (T_{n-1}, S_{n-1}, \cdot))$ bestimmt genau einen Pfad des SP der Restarbeit zu einem beliebigen Zeitpunkt bis zur $(n-1)$ -ten Ankunft und daher auch eine Realisation k von U_{n-1} . Die Menge der Realisationen des Zufallsvektors zerfällt in Klassen C_k mit denselben Realisationen k von U_{n-1} . Zerlegt man noch weiter nach Zuständen i des SMP zu t_{n-1} mittels $C_k^i := \{(t_0, i_0, v_0, \dots, t_{n-2}, i_{n-2}, v_{n-2}, t_{n-1}) | (t_0, i_0, v_0, \dots, t_{n-2}, i_{n-2}, v_{n-2}, t_{n-1}, i) \in C_k\}$, so folgt

$$\begin{aligned} & \{S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} = \\ & = \bigcup_{(t_0, i_0, v_0, \dots, t_{n-2}, i_{n-2}, v_{n-2}, t_{n-1}) \in C_k^i} \{(T_0, S_0, X_0) = (t_0, i_0, v_0), \dots, \\ & \quad (T_{n-2}, S_{n-2}, X_{n-2}) = (t_{n-2}, i_{n-2}, v_{n-2}), \\ & \quad (T_{n-1}, S_{n-1}, X_{n-1}) = (t_{n-1}, i, v)\}. \end{aligned} \quad (4.6)$$

Wegen Gl. (3.5) und der Unabhängigkeit von X erhält man

$$\begin{aligned} P\{S_n = j, A_{n-1} = t, S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} = \\ = P\{S_n = j, A_{n-1} = t | S_{n-1} = i\} \cdot P\{S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} \quad \square \end{aligned}$$

Zu beachten ist, daß Hilfssatz 5 auf Gl. (3.5) zurückgeht. Eine Begründung wie: " U_n beschreibt den Systemzustand und kann deshalb nur unabhängig von den Zwischenankunftszeiten sein", ist nicht zulässig. U_n und A_n sind nicht unabhängig, weil die Zwischenankunftszeiten nicht unabhängig sind. Denn einerseits ist z.B. für DF und $x(v) = 1_{v=1}$

$$P\{A_1 = t, U_1 = k\} = \mathbf{1}_{k=0} \cdot P\{A_1 = t, A_0 \geq D\} + \mathbf{1}_{k=1}^D \cdot P\{A_1 = t, A_0 = D - k\}.$$

Andererseits gilt $P\{U_1 = k\} = \mathbf{1}_{k=0} \cdot P\{A_0 \geq D\} + \mathbf{1}_{k=1}^D \cdot P\{A_0 = D - k\}$ und A_0, A_1 sind wegen Gl. (3.13) nicht unabhängig. $\mathbf{1}$ bezeichnet die Indikatorfunktion gemäß Anhang C.

$$\begin{aligned}
P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} = & \\
= \left\{ \begin{array}{l} \left. \begin{array}{l} q_{ij}(k - \ell + vD) \quad 0 \leq k \leq (N+1-v)D, \\ q_{ij}(k - \ell + \xi D) \quad 1 + (N-\xi)D \leq k \leq (N-\xi+1)D \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq v-1, \end{array} \right\} & \text{für } 1 \leq v \leq N \\ \left. \begin{array}{l} q_{ij}(-\ell + ND + D) \quad k = 0, \\ q_{ij}(k - \ell + \xi D) \quad 1 + (N-\xi)D \leq k \leq (N-\xi+1)D \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq N, \end{array} \right\} & \text{für } v \geq 1 + N \end{array} \right\} \quad \ell > 0, \\
= \left\{ \begin{array}{l} \left. \begin{array}{l} p_{ij} - Q_{ij}(k - 1 + vD) \quad 0 \leq k \leq (N+1-v)D, \\ p_{ij} - Q_{ij}(k - 1 + \xi D) \quad 1 + (N-\xi)D \leq k \leq (N-\xi+1)D, \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq v-1, \end{array} \right\} & \text{für } 1 \leq v \leq N \\ \left. \begin{array}{l} p_{ij} - Q_{ij}(-1 + ND + D) \quad k = 0, \\ p_{ij} - Q_{ij}(k - 1 + \xi D) \quad 1 + (N-\xi)D \leq k \leq (N-\xi+1)D, \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq N, \end{array} \right\} & \text{für } v \geq 1 + N \end{array} \right\} \quad \ell = 0.
\end{aligned}$$

Abb. 4.3 Übergangswahrscheinlichkeiten für DF.

Hilfssatz 5 liefert zwei Beziehungen, die bei der Ableitung des Hilfssatzes 6 benötigt werden:

$$\begin{aligned}
P\{S_n = j | S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} = & \\
= \sum_{t=0}^{\infty} P\{S_n = j, A_{n-1} = t | S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} = p_{ij} \quad , & \quad (4.7)
\end{aligned}$$

$$\begin{aligned}
P\{A_{n-1} = t | S_n = j, S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\} = & \\
= \frac{P\{S_n = j, A_{n-1} = t | S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\}}{P\{S_n = j | S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\}} = & \\
= P\{A_{n-1} = t | S_n = j, S_{n-1} = i\} = f_{ij}(t) \quad . & \quad (4.8)
\end{aligned}$$

Sie liefern

Hilfssatz 6 Für alle $n \in \mathbb{N}$, $k, \ell \in \{0, \dots, ND + D\}$, $i, j \in \mathbb{IE}$, $v \in \mathbb{IN}$ sind die bedingten WKTn $P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\}$ allein durch den Semi-Markov-Kern $\mathbf{q}(t)$ bestimmt. Für DF lauten sie, wie in Abb. 4.3 angegeben, für AF wie in Abb. 4.4.

Bis auf $k = 0$, stimmen die Übergangs-WKTn für AF und DF überein. Der Beweis für AF erfordert aber andere Fallunterscheidungen.

Beweis für DF: X_n, X_{n-1} und der SMP sind nach Voraussetzung unabhängig, d.h. (S_n, A_n) und X_n, X_{n-1} sind unabhängig. U_n und X_{n-1} sind nicht unabhängig, weil die Anzahl gleichzeitig ankommender Zellen zum $(n-1)$ -ten Ankunftszeitpunkt die Restarbeit U_n beim folgenden Ankunftszeitpunkt mitbestimmt; dagegen sind (S_n, U_n) und X_n unabhängig. Mit Gl. (4.7) hat man

$$\begin{aligned}
P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} = & \\
= P\{U_n = \ell | S_n = j, U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} \cdot p_{ij} \quad . & \quad (4.9)
\end{aligned}$$

Für die bedingten WKTn $P\{\cdot | S_n = j, S_{n-1} = i, U_{n-1} = k, X_{n-1} = v\}$ wird im folgenden die Abkürzung $P_{ijkv}\{\cdot\}$ verwendet.

Zunächst ergibt sich mittels Gl. (4.3):

$$P_{ijkv}\{U_n = \ell\} = \mathbf{1}_{\ell=0} \cdot P_{ijkv}\{U_{n-1}^+ \leq A_{n-1}\} + \mathbf{1}_{\ell>0} \cdot P_{ijkv}\{U_{n-1}^+ - A_{n-1} = \ell\} \quad (4.10)$$

Nun wird die Menge $\{U_{n-1}^+ = \ell', S_{n-1} = i, X_{n-1} = v\}$ betrachtet. Als Beispiel sei der Fall $1 \leq v \leq N$ erläutert: Für $\ell' \leq ND$ werden alle Zellen eines bei T_n ankommenden v -Batch zugelassen, denn das System ist – mit Berücksichtigung der Ankunft – noch nicht voll. Außerdem ist $\ell' \geq vD$, da ohne Zellverlust die Restarbeit mit Berücksichtigung der Ankunft wenigstens derjenigen Arbeit entspricht, die der zugelassene v -Batch mitbringt. Wenn $1 + ND \leq \ell' \leq D + ND$ gilt, so gehen $v, v-1, \dots$, oder 0 Zellen verloren, und U_{n-1} ist so groß, daß U_{n-1}^+ , die zugelassenen Zellen beinhaltend, einen Wert zwischen $1 + ND$ und $D + ND$ annimmt. Die anderen Fälle ergeben sich mit ähnlicher Begründung, s. auch Gl. (4.2). Mit der weiteren Abkürzung $\{ \cdot \}_{iv} := \{ \cdot, S_{n-1} = i, X_{n-1} = v \}$ erhält man für beliebige $i \in \mathbb{E}$:

$$\{U_{n-1}^+ = \ell'\}_{iv} = \begin{cases} \{U_{n-1} = \ell' - vD\}_{iv} & \text{für } vD \leq \ell' \leq ND \\ \bigcup_{\xi=0}^v \{U_{n-1} = \ell' - \xi D\}_{iv} & \text{für } 1 + ND \leq \ell' \leq ND + D \\ & \text{für } 1 \leq v \leq N, \\ \bigcup_{\xi=0}^N \{U_{n-1} = \ell' - \xi D\}_{iv} & \text{für } 1 + ND \leq \ell' \leq ND + D - 1 \\ \bigcup_{\xi=0}^{N+1} \{U_{n-1} = (N+1 - \xi)D\}_{iv} & \text{für } \ell' = ND + D \\ & \text{für } 1 + N \leq v. \end{cases} \quad (4.11)$$

Weiter ergibt sich⁴ mit den disjunkten Mengen aus (4.11) in (4.10) im Falle $\ell > 0$:

a) Für $1 \leq v \leq N$:

$$\begin{aligned} P_{ijkv}\{U_{n-1}^+ - A_{n-1} = \ell\} &= \sum_{\ell'=vD}^{ND+D} P_{ijkv}\{U_{n-1}^+ = \ell', A_{n-1} = \ell' - \ell\} = \\ &= \mathbf{1}_{k=0}^{(N-v+1)D} P_{ijkv}\{A_{n-1} = k - \ell + vD\} + \sum_{\xi=0}^{v-1} \mathbf{1}_{k=1+(N-\xi)D}^{(N-\xi+1)D} P_{ijkv}\{A_{n-1} = k - \ell + \xi D\} = \\ &\stackrel{\text{mit (4.8)}}{=} \mathbf{1}_{k=0}^{ND-vD+D} f_{ij}(k - \ell + vD) + \sum_{\xi=0}^{v-1} \mathbf{1}_{k=1+(N-\xi)D}^{(N-\xi+1)D} f_{ij}(k - \ell + \xi D). \end{aligned}$$

b) Für $1 + N \leq v < \infty$:

$$\begin{aligned} P_{ijkv}\{U_{n-1}^+ - A_{n-1} = \ell\} &= \sum_{\ell'=1+ND}^{D+ND} P_{ijkv}\{U_{n-1}^+ = \ell', A_{n-1} = \ell' - \ell\} = \\ &= \sum_{\xi=0}^N \mathbf{1}_{k=1+(N-\xi)D}^{(N+1-\xi)D} P_{ijkv}\{A_{n-1} = k - \ell + \xi D\} + \mathbf{1}_{k=0} P_{ijkv}\{A_{n-1} = ND + D - \ell\} = \\ &\stackrel{\text{mit (4.8)}}{=} \sum_{\xi=0}^N \mathbf{1}_{k=1+(N-\xi)D}^{(N+1-\xi)D} f_{ij}(k - \ell + \xi D) + \mathbf{1}_{k=0} f_{ij}(ND + D - \ell). \end{aligned}$$

Die Ausdrücke für $\ell = 0$ folgen aus der Beziehung

$$P_{ijkv}\{U_{n-1}^+ \leq A_{n-1}\} = 1 - \sum_{\ell=1}^{ND+D} P_{ijkv}\{U_{n-1}^+ - A_{n-1} = \ell\}.$$

Mit Gl. (4.9) erhält man dann die gesuchten Übergangs-WKTN. □

Beweis für AF: Mit Gl. (4.5) folgt im Unterschied zu Gl. (4.10)

⁴Zwischenschritte sind im Anhang A.1 aufgeführt.

$$\begin{aligned}
& P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} = \\
& = \left\{ \begin{array}{l} \left. \begin{array}{l} q_{ij}(vD - \ell + 1) \quad k = 0, \\ q_{ij}(k - \ell + vD) \quad 1 \leq k \leq (N + 1 - v)D, \\ q_{ij}(k - \ell + \xi D) \quad 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq v - 1, \end{array} \right\} \text{ für } 1 \leq v \leq N \\ \left. \begin{array}{l} q_{ij}(-\ell + ND + D + 1) \quad k = 0, \\ q_{ij}(k - \ell + \xi D) \quad 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq N, \end{array} \right\} \text{ für } v \geq N + 1 \end{array} \right\} \ell > 0, \\
& \left\{ \begin{array}{l} \left. \begin{array}{l} p_{ij} - Q_{ij}(vD) \quad k = 0, \\ p_{ij} - Q_{ij}(k - 1 + vD) \quad 1 \leq k \leq (N + 1 - v)D, \\ p_{ij} - Q_{ij}(k - 1 + \xi D) \quad 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D, \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq v - 1, \end{array} \right\} \text{ für } 1 \leq v \leq N \\ \left. \begin{array}{l} p_{ij} - Q_{ij}(ND + D) \quad k = 0, \\ p_{ij} - Q_{ij}(k - 1 + \xi D) \quad 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D, \\ \text{für jedes } \xi \text{ mit } 0 \leq \xi \leq N, \end{array} \right\} \text{ für } v \geq 1 + N \end{array} \right\} \ell = 0.
\end{aligned}$$

Abb. 4.4 Übergangswahrscheinlichkeiten für AF.

$$P_{ijkv}\{U_n = \ell\} = \mathbf{1}_{\ell=0} \cdot P_{ijkv}\{U_{n-1}^+ < A_{n-1}\} + \mathbf{1}_{\ell>0} \cdot P_{ijkv}\{U_{n-1}^+ - A_{n-1} + 1 = \ell\}. \quad (4.12)$$

Für die Menge $\{U_{n-1}^+ = \ell', S_{n-1} = i, X_{n-1} = v\}$ ergibt sich wieder mit der Abkürzung $\{\cdot\}_{iv} := \{\cdot, S_{n-1} = i, X_{n-1} = v\}$:

$$\text{a) } 1 \leq v \leq N - 1: \quad \{U_n^+ = \ell'\}_{iv} = \begin{cases} \{U_n = 1\}_{iv} \cup \{U_n = 0\}_{iv} & \text{für } \ell' = vD \\ \{U_n = \ell' - vD + 1\}_{iv} & \text{für } 1 + vD \leq \ell' \leq ND - 1 \\ \bigcup_{\xi=0}^v \{U_n = (N - \xi)D + 1\}_{iv} & \text{für } \ell' = ND \\ \bigcup_{\xi=0}^v \{U_n = \ell' - \xi D + 1\}_{iv} & \text{für } 1 + ND \leq \ell' \leq ND + D - 1, \end{cases} \quad (4.13)$$

$$\text{b) } v = N: \quad \{U_n^+ = \ell'\}_{iv} = \begin{cases} \bigcup_{\xi=0}^N \{U_n = (N - \xi)D + 1\}_{iv} \cup \{U_n = 0\} & \text{für } \ell' = ND \\ \bigcup_{\xi=0}^N \{U_n = \ell' - \xi D + 1\}_{iv} & \text{für } 1 + ND \leq \ell' \leq ND + D - 1, \end{cases} \quad (4.14)$$

$$\text{c) } v \geq N + 1: \quad \{U_n^+ = \ell'\}_{iv} = \begin{cases} \bigcup_{\xi=0}^N \{U_n = \ell' - \xi D + 1\}_{iv} & \text{für } ND \leq \ell' \leq ND + D - 1 \\ \{U_n = 0\}_{iv} & \text{für } \ell' = ND + D. \end{cases} \quad (4.15)$$

Gl. (4.12) kann mit den Mengen in (4.13), (4.14), (4.15) für jeden der Fälle berechnet werden unter Beachtung von $P_{ijkv}\{U_{n-1} = \ell, \cdot\} = 0$ für $k \neq \ell$ und Gl. (4.8). Im Fall $\ell > 0$ ist $P_{ijkv}\{U_{n-1}^+ - A_{n-1} + 1 = \ell\}$ zu berechnen, b) geht in a) auf, und die Ausdrücke von Hilfssatz 6 folgen; diejenigen für $\ell = 0$ ergeben sich mittels $P_{ijkv}\{U_{n-1}^+ < A_{n-1}\} = 1 - \sum_{\ell=1}^{ND+D} P_{ijkv}\{U_{n-1}^+ - A_{n-1} + 1 = \ell\}$. \square

Der nächste Abschnitt zeigt, wie man die Zustands-WKTn $P\{U_n = \ell, S_n = j, X_n = w\} =$

$P\{U_n = \ell, S_n = j\} \cdot x(v)$ der MK zum n -ten Ankunftszeitpunkt mittels einer (stochastischen) Matrix berechnen kann, die nur $P\{U_n = \ell, S_n = j\}$ mit $P\{U_{n-1} = \ell, S_{n-1} = j\}$ verknüpft, also nur $m(ND + D + 1)$ Zeilen und Spalten hat gegenüber der um den Faktor b_{max} größeren Anzahl für die ursprüngliche MK.

4.1.3 Transiente Gleichungen

Wie üblich bei MKn gelten die transienten Gleichungen

$$u_n(\ell, j, w) = x(w) \cdot \sum_{k=0}^{ND+D} \sum_{i=1}^m \sum_{v=1}^{b_{max}} u_{n-1}(k, i, v) \cdot P_{ikv}\{U_n = \ell, S_n = j\}.$$

Dabei bedeuten $u_n(\ell, j, w) := P\{U_n = \ell, S_n = j, X_n = w\}$ und $P_{ikv}\{U_n = \ell, S_n = j\} := P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\}$.

Wegen $P\{U_n = \ell, S_n = j, X_n = w\} = P\{U_n = \ell, S_n = j\} \cdot x(w)$ können die Gleichungen so umgeformt werden, daß nur noch $u_n(\ell, j) := \sum_{w=1}^{\infty} u_n(\ell, j, w)$ und $u_{n-1}(k, i) := \sum_{v=1}^{\infty} u_{n-1}(k, i, v)$ verknüpft werden. Einige etwas umfangreichere Umformungen führen auf die Gleichung $\underline{u}_n = \underline{u}_{n-1} \mathbf{F}$. Darin ist $\underline{u}_n := (\underline{u}_n(0), \dots, \underline{u}_n(ND+D))$ und $[\underline{u}_n(k)]_i := u_n(k, i)$ sowie

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_0^0 & \dots & \mathbf{F}_0^{ND+D} \\ \vdots & & \vdots \\ \mathbf{F}_{ND+D}^0 & \dots & \mathbf{F}_{ND+D}^{ND+D} \end{pmatrix}$$

mit den folgenden Teilmatrizen \mathbf{F}_k^ℓ , die sich für AF und DF (wieder) nur für $k = 0$ unterscheiden. Es ist für $\ell = 0$

$$\mathbf{F}_k^0 = \begin{cases} \begin{cases} \mathbf{p} - \sum_{v=1}^N x(v) \mathbf{Q}(vD) \\ -X^c(N) \mathbf{Q}(ND+D) \end{cases} & \text{für } k=0 \text{ und AF} \\ \begin{cases} \mathbf{p} - \sum_{v=1}^N x(v) \mathbf{Q}(vD-1) \\ -X^c(N) \mathbf{Q}(ND+D-1) \end{cases} & \text{für } k=0 \text{ und DF} \\ \begin{cases} \mathbf{p} - \sum_{v=1}^{\xi-1} x(v) \mathbf{Q}(k-1+vD) \\ -X^c(\xi-1) \mathbf{Q}(k-1+\xi D) \end{cases} & \text{für } 1+(N-\xi)D \leq k \leq (N+1-\xi)D, \\ & \text{für jedes } \xi \text{ mit } 2 \leq \xi \leq N \\ \mathbf{p} - \mathbf{Q}(k-1+D) & \text{für } 1+ND-D \leq k \leq ND \\ \mathbf{p} - \mathbf{Q}(k-1) & \text{für } 1+ND \leq k \leq ND+D \end{cases} \quad (4.16)$$

und für $\ell > 0$

$$\mathbf{F}_k^\ell = \begin{cases} \begin{cases} \sum_{v=1}^N x(v) \mathbf{q}(-\ell+vD+1) \\ +X^c(N) \mathbf{q}(-\ell+ND+D+1) \end{cases} & \text{für } k=0 \text{ und AF} \\ \begin{cases} \sum_{v=1}^N x(v) \mathbf{q}(-\ell+vD) \\ +X^c(N) \mathbf{q}(-\ell+ND+D) \end{cases} & \text{für } k=0 \text{ und DF} \\ \begin{cases} \sum_{v=1}^{\xi-1} x(v) \mathbf{q}(k-\ell+vD) \\ +X^c(\xi-1) \mathbf{q}(k-\ell+\xi D) \end{cases} & \text{für } 1+(N-\xi)D \leq k \leq (N+1-\xi)D, \\ & \text{für jedes } \xi \text{ mit } 2 \leq \xi \leq N \\ \mathbf{q}(k-\ell+D) & \text{für } 1+ND-D \leq k \leq ND \\ \mathbf{q}(k-\ell) & \text{für } 1+ND \leq k \leq ND+D \end{cases} \quad (4.17)$$

Jeder der Fälle in (4.16), (4.17) repräsentiert eine Anzahl (z.B. ξ) noch freier Plätze zum $(n-1)$ -ten Ankunftszeitpunkt. Da \mathbf{F} stochastisch ist, können die Algorithmen von Grassmann et al. [43] bzw. Kramer [70] unter Ausnutzung der Blockstruktur zur Bestimmung der stationären Lösung $\underline{u} := \lim_{n \rightarrow \infty} \underline{u}_n$ verwendet werden, falls sie existiert. Diese sind direkte Verfahren, die sich dadurch auszeichnen, daß in jedem Iterationsschritt nicht-negative Zahlen auftreten, so daß sie numerisch sehr stabil sind. Letzteres gilt nur bedingt für [70], weil dort als Umformungsschritt die Matrixinversion vorkommt. Die üblichen Algorithmen zur Matrixinversion verwenden Subtraktionen und führen infolge von Rundungsfehlern häufig doch wieder auf negative Werte, obwohl die auftretenden invertierten Matrizen theoretisch [70] nicht-negative Einträge haben sollten, s. auch Abschnitt 7.8.

Zu bemerken bleibt, daß $\{(U_n, S_n); n \in \mathbb{N}_0\}$ für $b_{max} > 1$ keine MK darstellt, obwohl die stochastische Matrix \mathbf{F} nur $P\{U_n = \ell, S_n = j\}$ und $P\{U_{n-1} = \ell, S_{n-1} = j\}$ verknüpft. Für $b_{max} = 1$ können Mehrfachankünfte durch Zwischenankunftszeiten der Größe Null berücksichtigt werden. Die Systemmatrix hat dann viele Nulleinträge, und die Anzahl der Operationen zur Bestimmung der stationären Zustands-WKTN mittels des Grassmann-Algorithmus reduziert sich auf $m^3(1+D)D^2N^2$, s. [51] und Abschnitt 4.2.

4.1.4 Belegung zu Ankunftszeitpunkten, Verlustwahrscheinlichkeiten und Wartezeit im stationären Fall

Aus dem Wert der Restarbeit zu einem Ankunftszeitpunkt kann auf die gesehene Belegung geschlossen werden. Außerdem bestimmt die Restarbeit auch den Wert der Wartezeit, sofern kein Verlust auftritt; hierbei unterscheiden sich jedoch AF und DF, wie unten erläutert wird. Mit der stationären Lösung \underline{u} von \mathbf{F} lassen sich die WFN von Belegung und Wartezeit angeben. Mit $\underline{e} := (1, \dots, 1)^T$, m Komponenten, und dem Skalarprodukt $\underline{u}(k)\underline{e} := \sum_{i=1}^m u(k, i)$ erhält man für die Belegung L_n zum n -ten Ankunftszeitpunkt

$$P\{L_n = \mu\} = \begin{cases} \underline{u}(0)\underline{e} & \text{für } \mu = 0, \\ \sum_{\ell=1+(\mu-1)D}^{\mu D} \underline{u}(\ell)\underline{e} & \text{für } 1 \leq \mu \leq N+1. \end{cases} \quad (4.18)$$

Das Ereignis "Verlust" ist direkt verknüpft mit der Belegungs-WKTN. Infolge der Batch-ZV gibt es jedoch verschiedene Möglichkeiten, Verlust zu definieren bzw. zu quantifizieren: Verlust wenigstens einer Zelle, Verlust einer "Testzelle".

a) Verlust wenigstens einer Zelle

$$P_{\text{Verlust}} = \sum_{\mu=0}^{N+1} P\{X_n \geq N+2-\mu\}P\{L_n = \mu\}. \quad (4.19)$$

b) Verlust einer Testzelle

Die Zellen eines ankommenden w -Batch werden in völlig zufälliger Reihenfolge ins Wartesystem eingebracht. Unter den Zellen des w -Batch ist eine sog. Testzelle markiert. Für diese Testzelle ist die WKTN als ν -te Zelle des Batch ins System zu gelangen (oder "in ν -ter Position für die Zulassung ausgewählt zu werden") $\frac{1}{w}$, und die Testzelle geht verloren, wenn ν die Zahl $N+1-L_n$ der noch freien Plätze im Wartesystem übersteigt. Die Anzahl der Auswahlpositionen, in denen die Testzelle verlorengeht, beträgt $\max[w - (N+1-L_n), 0]$. Es ergibt sich daher

$$P_{\text{Verlust}}^{\text{Testzelle}} = \sum_{\mu=0}^{N+1} \sum_{w=N+2-\mu}^{\infty} \frac{w - (N+1-\mu)}{w} P\{X_n = w, L_n = \mu\}, \quad (4.20)$$

wobei

$$P\{X_n = w, L_n = \mu\} = x(w)P\{L_n = \mu\}$$

zu berücksichtigen ist.

Offenbar kann die Wartezeit nur als diejenige einer Testzelle definiert werden. Für DF ergibt sich die Wartezeit einer Testzelle, die als ν -te Zelle des Batch ins System gelangt, aus der Restarbeit zum Ankunftszeitpunkt *plus* der Summe $(\nu - 1)D$ der Bedienzeiten der $\nu - 1$ Zellen desselben Batch, die vor der Testzelle zugelassen werden.

Bei AF wird die (positive) Restarbeit unmittelbar nach der Ankunft um eine Zeiteinheit abwärts gezählt. Daher ist die Wartezeit der nicht-verlorenen Testzelle um eine Zeiteinheit kleiner als die Summe aus gesehener (positiver) Restarbeit und Bedienzeiten der Zellen desselben Batch, die vor der Testzelle zugelassen werden. Bei leerem Wartesystem sieht die als ν -te zugelassene Testzelle dagegen wie bei DF eine Wartezeit der Größe $(\nu - 1)D$.

Es macht keinen Sinn (wie etwa bei der Restarbeit), von der Wartezeit zum n -ten Ankunftszeitpunkt zu sprechen, da nicht klar ist, ob zu diesem Zeitpunkt Verlust auftritt. Einer verlorenen Zelle wird üblicherweise keine Wartezeit zugeordnet. Daher wird als ZV für die Wartezeit W ohne Index verwendet. $\mathbf{1}$ bezeichnet die Indikatorfunktion gemäß Anhang C. $C := \sum_{\ell=0}^{ND} P\{W = \ell\}$ ist eine Normierungskonstante, die die WKTn auf Ankunftsereignisse ohne Verlust (der Testzelle) einschränkt.

Für DF ergibt sich mit den obigen Ausführungen

$$P\{W = \ell \mid \text{kein Verlust}\} = \frac{\mathbf{1}_{\ell=0}^{ND}}{C} \cdot \left[\sum_{w=1}^{N+1} \sum_{\nu=1}^w \frac{x(w)}{w} \cdot \underline{u}(\ell - (\nu - 1)D) \underline{e} + \sum_{w=N+2}^{\infty} \sum_{\nu=1}^{N+1} \frac{x(w)}{w} \cdot \underline{u}(\ell - (\nu - 1)D) \underline{e} \right] \quad (4.21)$$

und für AF

$$P\{W = \ell \mid \text{kein Verlust}\} = \frac{1}{C} \cdot \left[\sum_{w=1}^{N+1} \sum_{\nu=1}^w \frac{x(w)}{w} \cdot \left(\mathbf{1}_{\ell=0}^{ND-1} \cdot \mathbf{1}_{\ell+1 \neq (\nu-1)D} \cdot \underline{u}(\ell + 1 - (\nu - 1)D) \underline{e} + \mathbf{1}_{\ell=0}^{ND} \cdot \mathbf{1}_{\ell=(\nu-1)D} \cdot \underline{u}(0) \underline{e} \right) + \sum_{w=N+2}^{\infty} \sum_{\nu=1}^{N+1} \frac{x(w)}{w} \cdot \left(\mathbf{1}_{\ell=0}^{ND-1} \cdot \mathbf{1}_{\ell+1 \neq (\nu-1)D} \cdot \underline{u}(\ell + 1 - (\nu - 1)D) \underline{e} + \mathbf{1}_{\ell=0}^{ND} \cdot \mathbf{1}_{\ell=(\nu-1)D} \cdot \underline{u}(0) \underline{e} \right) \right]. \quad (4.22)$$

4.1.5 Ausdrücke im transienten Fall

Im ATM-Umfeld sind auch Verlust-WKTn und Wartezeiten im transienten Fall, d.h. zum n -ten Ankunftszeitpunkt mit $n < \infty$, von Interesse, wenn eine Quelle gerade zu senden beginnt. Ausdrücke dafür erhält man, wenn man oben $\underline{u}(k)$ durch $\underline{u}_n(k)$ ersetzt und einen geeigneten Anfangsvektor \underline{u}_0 wählt, z. B. $u_0(k, i) := P_i \cdot \mathbf{1}_{k=0}$. Für das Wartesystem vom M/G/1-Typ (s. Kapitel 5) sind Ausdrücke für den transienten Fall erheblich komplexer.

4.1.6 Schätzwert für die bedingte Verlustwahrscheinlichkeit

In [10] wird für ein Wartesystem (mit Bedienzeit gleich Zeiteinheit) als wichtige Leistungs-kenngröße eines Multiplexers die WKT des Ereignisses betrachtet: "Unter der Bedingung, daß der Multiplexer zu $t = 0$ Zellen verliert, verliert er nach r Zeiteinheiten immer noch Zellen". Die Autoren setzen r als mittleren Zeitabstand für die Zellen einer Einzelverbindung an und benutzen die WKT des obigen Ereignisses als obere Grenze für die WKT, daß eine einzelne Quelle aufeinanderfolgende Zellen verliert. Angewendet auf den Fall des hier betrachteten Wartesystems wird das Verlieren aufeinanderfolgender Zellen (näherungsweise) durch das Ereignis erfaßt: "Unter der Bedingung, daß zum n -ten Ankunftszeitpunkt eine Zelle verlorengelht, gehen zum $(n + r)$ -ten Ankunftszeitpunkt wieder Zellen verloren". Da $\{(U_n, S_n, X_n); n \in \mathbb{N}_0\}$ eine MK ist, kann diese bedingte Verlust-WKT hier (auch für $D > 1$) leicht berechnet werden durch

$$P_{\text{Verlust}}^{\text{bedingt}} \approx P\{\Gamma_{n+r} = 1 | \Gamma_n = 1\}$$

$$\text{mit } \Gamma_n = \begin{cases} 1 & \text{Verlust wenigstens einer Zelle bei } n\text{-ter Ankunft} \\ 0 & \text{kein Verlust bei } n\text{-ter Ankunft} \end{cases}$$

Dabei ist

$$P\{\Gamma_{n+r} = 1, \Gamma_n = 1\} = \sum_{w=N+2}^{\infty} \sum_{j=1}^m x(w) P\{U_{n+r} = 0, S_{n+r} = j, \Gamma_n = 1\} + \\ + \sum_{\mu=1}^{N+1} \sum_{\ell=1+(\mu-1)D}^{\mu D} \sum_{w=N+2-\mu}^{\infty} \sum_{j=1}^m x(w) P\{U_{n+r} = \ell, S_{n+r} = j, \Gamma_n = 1\}$$

$$\text{mit } P\{U_{n+r} = \ell, S_{n+r} = j, \Gamma_n = 1\} =$$

$$= \sum_{k'=0}^{ND+D} \sum_{i'=1}^m \left[[\mathbf{F}^{r-1}]_{k'\ell} \right]_{i'j} P\{U_{n+1} = k', S_{n+1} = i', \Gamma_n = 1\}.$$

Darin wiederum bedeuten $[[\mathbf{F}]_{k'\ell}]_{i'j} := [\mathbf{F}_{k'j}^{\ell}]_{i'j}$ und $[\mathbf{q}(\cdot)]_{i'j} := q_{i'j}(\cdot)$, und es ist

$$P\{U_{n+1} = k', S_{n+1} = i', \Gamma_n = 1\} =$$

$$= \sum_{v=N+2}^{\infty} \sum_{i=1}^m u(0, i) x(v) P_{i0v}\{U_{n+1} = k', S_{n+1} = i'\} + \\ + \sum_{\mu=1}^{N+1} \sum_{k=1+(\mu-1)D}^{\mu D} \sum_{v=N+2-\mu}^{\infty} \sum_{i=1}^m u(k, i) x(v) P_{ikv}\{U_{n+1} = k', S_{n+1} = i'\}.$$

$P\{\Gamma_n = 1\}$ ergibt sich nach (4.19); $P_{ikv}\{\cdot\}$ ist in Abschnitt 4.1.3 definiert.

Eine hohe bedingte Verlust-WKT bedeutet, daß die Einzelquelle häufig mehrere aufeinanderfolgende Zellen verliert. Dies kann die Dienstgüte verschlechtern, auch wenn die Verlust-WKT bezogen auf alle verlorenen Zellen der Verbindung sehr klein ist. Die hier betrachtete bedingte Verlust-WKT ist vom Konzept her nur eine Näherung. In Kapitel 6 wird für eine in einer Überlagerung noch zu identifizierende Einzelquelle ein Algorithmus angegeben, mit dem eine konzeptionell exakte bedingte Verlust-WKT für diese Einzelquelle berechnet werden kann.

4.1.7 Gleichzeitigkeit von Ankünften und Abgängen

Die Gleichzeitigkeit von Ankünften und Abgängen kann ebenfalls mittels der Restarbeit quantifiziert werden.

Departure First Für DF sieht man leicht, daß eine Ankunft, die positive ganzzahlige Vielfache von D sieht, gleichzeitig mit einem Abgang stattfindet (wobei nach Definition von DF der Abgang unmittelbar vor der Ankunft erfolgt). Dagegen wird das Ereignis "die n -te Ankunft trifft bei gleichzeitigem Abgang ein leeres System" durch den komplizierteren Ausdruck $\{U_{n-1}^+ - A_{n-1} = 0\}$ beschrieben. Wegen DF kann bei Gleichzeitigkeit von Ankunft und Abgang das System *nicht* voll besetzt sein, wenn nur positive Zwischenankunftszeiten zugelassen sind.

$P\{L_n = \ell, \text{ Abgang gleichzeitig}\} =$

$$= \begin{cases} P\{U_{n-1}^+ = A_{n-1}\} & \text{für } \ell = 0 \\ \underline{u}(\ell D)\underline{e} & \text{für } 1 \leq \ell \leq N \\ 0 & \text{für } \ell = N + 1 \end{cases} \quad (4.23)$$

mit

$$P\{U_{n-1}^+ = A_{n-1}\} = \sum_{v=1}^N \left[\sum_{k=0}^{(N-v)D} x(v)\underline{u}(k)\mathbf{q}(k+vD)\underline{e} + \sum_{\xi=0}^v \sum_{k=1+(N-\xi)D}^{(N+1-\xi)D} x(v)\underline{u}(k)\mathbf{q}(k+\xi D)\underline{e} \right] \\ + \sum_{v=N+1}^{b_{max}} x(v) \left[\underline{u}(0)\mathbf{q}(ND+D)\underline{e} + \sum_{\xi=0}^N \sum_{k=1+(N-\xi)D}^{(N+1-\xi)D} \underline{u}(k)\mathbf{q}(k+\xi D)\underline{e} \right]. \quad (4.24)$$

Arrival First Bei AF kann das System nicht leer sein, wenn Ankunft und Abgang gleichzeitig auftreten. Wenn eine Ankunft die Restarbeit $1 + (\ell - 1)D$ sieht, so sind ℓ Zellen im System, und unmittelbar danach verläßt eine Zelle das System.

$P\{L_n = \ell, \text{ Abgang gleichzeitig}\} =$

$$= \begin{cases} 0 & \text{für } \ell = 0 \\ \underline{u}(1 + (\ell - 1)D)\underline{e} & \text{für } 1 \leq \ell \leq N + 1 \end{cases} \quad (4.25)$$

4.2 Das SMP/D/1/N-System mit Null-Zwischenankunftszeiten

Schon mehrfach wurde erwähnt, daß auch Zwischenankunftszeiten der Größe Null für den Eingabeprozess zugelassen werden können, so daß auch ohne Batch-ZV Mehrfachankünfte ermöglicht werden. Dann kann allerdings nur DF behandelt werden. Folgende Gleichungen bestimmen die zeitliche Entwicklung des Systems:

$$U_n^+ = \begin{cases} U_n + D & \text{für } 0 \leq U_n \leq ND, \\ U_n & \text{für } 1 + ND \leq U_n \leq D + ND, \end{cases} \quad (4.26)$$

$$U_{n+1} = \begin{cases} U_n^+ - A_n & \text{für } U_n^+ > A_n, \\ 0 & \text{sonst.} \end{cases} \quad (4.27)$$

$$\begin{pmatrix} \mathbf{F}_0^0 & \mathbf{F}_0^1 & \cdot & \cdot & \mathbf{F}_0^D & \mathbf{0} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \mathbf{0} \\ \mathbf{F}_1^0 & \mathbf{F}_1^1 & \cdot & \cdot & \cdot & \mathbf{F}_1^{1+D} & \mathbf{0} & \cdot & \cdot & \cdot & \cdot & \cdot & \mathbf{0} \\ \cdot & \cdot & \mathbf{F}_2^2 & \cdot & \cdot & \cdot & \mathbf{F}_2^{2+D} & \mathbf{0} & \cdot & \cdot & \cdot & \cdot & \mathbf{0} \\ \cdot & \cdot \\ \cdot & \mathbf{0} \\ \mathbf{F}_0^{ND} & \mathbf{F}_1^{ND} & \cdot & \cdot & \cdot & \mathbf{F}_{1+D}^{ND} & \mathbf{F}_{ND}^{ND} & \cdot & \cdot & \cdot & \cdot & \cdot & \mathbf{F}_{ND+D}^{ND} \\ \mathbf{F}_0^{1+ND} & \mathbf{F}_1^{1+ND} & \cdot & \cdot & \cdot & \cdot & \cdot & \mathbf{F}_{1+ND}^{1+ND} & \mathbf{0} & \cdot & \cdot & \cdot & \mathbf{0} \\ \cdot & \cdot \\ \mathbf{F}_0^{ND+D} & \mathbf{F}_1^{ND+D} & \cdot & \mathbf{F}_{ND+D}^{ND+D} \end{pmatrix}$$

Abb. 4.5 Struktur der Übergangsmatrix für DF.

Man erhält in der Übergangsmatrix (Abb. 4.5) der MK $\{(U_n, S_n), n \in \mathbb{N}_0\}$ viele Nulleinträge. Die Matrix hat Blockstruktur mit den folgenden Untermatrizen

$$\mathbf{F}_k^\ell := \begin{cases} \mathbf{p} - \mathbf{Q}(k - 1 + D) & \text{für } 0 \leq k \leq ND \text{ und } \ell = 0 \\ \mathbf{p} - \mathbf{Q}(k - 1) & \text{für } 1 + ND \leq k \leq ND + D \text{ und } \ell = 0 \\ \mathbf{q}(k - \ell + D) & \text{für } 0 \leq k \leq ND \text{ und } 1 \leq \ell \leq k + D \\ \mathbf{q}(k - \ell) & \text{für } 1 + ND \leq k \leq ND + D \text{ und } 1 \leq \ell \leq k \end{cases}$$

Die Nulleinträge können im Algorithmus von Grassmann ausgenutzt werden für eine Reduktion der Anzahl von Operationen auf die Ordnung von $m^3(1+D)D^2N^2$. Mit ausschließlich positiven Zwischenankunftszeiten wird AF durch die Gleichungen

$$U_n^+ = \begin{cases} D & \text{für } U_n = 0, \\ U_n + D - 1 & \text{für } 1 \leq U_n \leq ND, \\ U_n - 1 & \text{für } 1 + ND \leq U_n \leq D + ND \end{cases} \quad (4.28)$$

und

$$U_{n+1} = \begin{cases} U_n^+ - A_n + 1 & \text{für } U_n^+ \geq A_n, \\ 0 & \text{sonst} \end{cases} \quad (4.29)$$

vollständig beschrieben. Diese Gleichungen verdeutlichen, daß Null-Zwischenankunftszeiten auf eine Inkonsistenz führen, wenn die n -te Ankunft ein leeres System sieht und die $(n+1)$ -te Ankunft nach Zwischenankunftszeit Null auftritt: $U_n = 0$, $U_n^+ = D$, $U_{n+1} = D + 1$ gemäß Gln. (4.28), (4.29), während die $(n+1)$ -te Ankunft die Restarbeit $U_{n+1} = D$ sehen sollte.

4.3 Die Restarbeit im zeitkontinuierlichen Fall

Im zeitkontinuierlichen Falle erhält man eine MK mit kontinuierlichem Zustandsraum, da die Restarbeit jeden reellen Wert zwischen 0 und $ND + D$ annehmen kann. Schon im Falle des unendlichen Puffers ergibt sich als Bestimmungsgleichung für die (stationäre) VF der

Wartezeit, die dann mit der Restarbeit zum n -ten Ankunftszeitpunkt übereinstimmt, eine Integralgleichung, die sog. "Lindley-Gleichung" [73]. Diese ist meistens wieder nur durch Diskretisierung zu lösen. Es ist zu erwarten, daß der Fall des endlichen Wartespeichers, wenn er denn formal zu beschreiben ist, auch nur durch Diskretisierung lösbar ist, so daß man stattdessen wohl besser direkt vom zeitdiskreten System ausgeht. Möglicherweise liefert der zeitkontinuierliche Ansatz aber Grenzfallaussagen wie in [73].

4.4 Ausblick: Verallgemeinerung auf die Fälle GI/GI/1/N und SMP/GI/1/N

Bei Einfachankünften liefert die Methode der Restarbeit einen Zugang auch zur Behandlung der allgemeineren endlichen Wartesysteme GI/GI/1/N und SMP/GI/1/N. Für DF wird dies im folgenden skizzenhaft erläutert.

Bei deterministischer Bedienzeit ist die Größe der Bedienzeit jedes Kunden im vorhinein bekannt. Der Übergang zu beliebigen Bedienzeitverteilungen gelingt, indem man den Verbundprozeß aus der Restarbeit bei der n -ten Ankunft und aus den Bedienzeiten der letzten $N + 1$ Kunden, die bis T_n das System betreten haben, betrachtet.

4.4.1 Einige Definitionen

1. $\Gamma_1^{(n)}$ bezeichne die Nummer des – von T_n aus gesehen – letzten zugelassenen Kunden, $\Gamma_2^{(n)}$ diejenige des vorletzten usw. In jedem Pfad der Restarbeit bis zum n -ten Ankunftszeitpunkt sind die Nummern der letzten $N + 1$ Kunden eindeutig bestimmt. $\underline{\Gamma}^{(n)} := (\Gamma_1^{(n)}, \dots, \Gamma_{N+1}^{(n)})$. Da der $(n + 1)$ -te Kunde erst nach dem n -ten eintreffen kann, ist

$$n > \Gamma_1^{(n)} > \Gamma_2^{(n)} > \dots > \Gamma_{N+1}^{(n)}.$$

2. $B_{\Gamma_1^{(n)}}$ bezeichne die Bedienzeit desjenigen Kunden, der – von T_n aus betrachtet – als letzter ins System gelangt ist. $\underline{B}_{\underline{\Gamma}^{(n)}} := (B_{\Gamma_1^{(n)}}, \dots, B_{\Gamma_{N+1}^{(n)}})$ bezeichne den Vektor der Bedienzeiten der – von T_n aus betrachtet – letzten $N + 1$ zugelassenen Kunden. Der n -te Kunde wird dabei noch nicht berücksichtigt. Die Indizes dieser Bedienzeiten sind selbst zufällig, so daß man z.B.

$$\{\omega \in \Omega : B_{\Gamma_1^{(n)}} = \beta_1\} = \bigcup_{1 \leq i \leq n-1} \{\omega \in \Omega : B_i = \beta_1, \Gamma_1^{(n)} = i\} \quad (4.30)$$

und allgemein

$$\{\omega \in \Omega : \underline{B}_{\underline{\Gamma}^{(n)}} = \underline{\beta}\} = \bigcup_{\underline{i}} \{\omega \in \Omega : \underline{B}_{\underline{i}} = \underline{\beta}, \underline{\Gamma}^{(n)} = \underline{i}\} \quad (4.31)$$

erhält, wobei $\underline{i} = (i_1, \dots, i_{N+1})$ alle hier möglichen Werte annimmt. Z.B. ist $n > i_1 > i_2 > \dots > i_{N+1}$. Bei $n \leq N$ sind einige der Indizes $\Gamma_i^{(n)}$ nicht definiert. Für den stationären Fall $n \rightarrow \infty$ ist das aber ohne Bedeutung. $b(\beta)$ bezeichne die Bedienzeit-WF.

4.4.2 Restarbeit zum n -ten Ankunftszeitpunkt ohne Berücksichtigung der Ankunft

Kennt man die Bedienzeiten – der von T_n aus gesehen – letzten $N + 1$ zugelassenen Kunden, so kann man durch Vergleich des Wertes von U_n mit den Bedienzeiten darüber entscheiden,

ob das System zu T_n voll ist oder nicht. Es gilt für $n > N + 2$

$$U_n = \begin{cases} \max[U_{n-1} + B_{n-1} - A_{n-1}, 0] \\ \quad \text{für } U_{n-1} \leq B_{\Gamma_1^{(n-1)}} + \dots + B_{\Gamma_N^{(n-1)}} , \\ \max[U_{n-1} - A_{n-1}, 0] \\ \quad \text{für } U_{n-1} \geq 1 + B_{\Gamma_1^{(n-1)}} + \dots + B_{\Gamma_N^{(n-1)}} . \end{cases} \quad (4.32)$$

Im ersten Fall von (4.32) ist $\Gamma_1^{(n)} = n - 1$, weil die $(n - 1)$ -te Ankunft zugelassen werden konnte ($B_{\Gamma_1^{(n)}} = B_{n-1}$ wird addiert). Dann gilt: $\Gamma_2^{(n)} = \Gamma_1^{(n-1)}, \dots, \Gamma_{N+1}^{(n)} = \Gamma_N^{(n-1)}$.

Im zweiten Fall ist $\Gamma_1^{(n)} < n - 1$, die $(n - 1)$ -te Ankunft geht verloren, weil die Bedienzeit $B_{\Gamma_{N+1}^{(n-1)}}$ noch nicht ganz abgelaufen ist, d.h. es sind noch $N + 1$ Kunden im System und $\underline{\Gamma}^{(n)} = \underline{\Gamma}^{(n-1)}$. Daher ist

$$(B_{\Gamma_1^{(n)}}, \dots, B_{\Gamma_{N+1}^{(n)}}) = \begin{cases} (B_{n-1}, B_{\Gamma_1^{(n-1)}}, \dots, B_{\Gamma_N^{(n-1)}}) \\ \quad \text{für } U_{n-1} \leq B_{\Gamma_1^{(n-1)}} + \dots + B_{\Gamma_N^{(n-1)}} , \\ (B_{\Gamma_1^{(n-1)}}, \dots, B_{\Gamma_{N+1}^{(n-1)}}) \\ \quad \text{für } U_{n-1} \geq 1 + B_{\Gamma_1^{(n-1)}} + \dots + B_{\Gamma_N^{(n-1)}} . \end{cases} \quad (4.33)$$

U_n und auch $\underline{B}_{\underline{\Gamma}^{(n)}}$ hängen also nur von U_{n-1} und von $\underline{B}_{\underline{\Gamma}^{(n-1)}}$ ab, so daß $\{(U_n, \underline{B}_{\underline{\Gamma}^{(n)}}); n \in \mathbb{N}_0, n > N + 2\}$ eine MK ist.

4.4.3 Übergangswahrscheinlichkeiten

Gesucht werden die Übergangs-WKTn $P\{U_n = \ell, \underline{B}_{\underline{\Gamma}^{(n)}} = \underline{\beta} | U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = \underline{\gamma}\}$. Für ihre Bestimmung werden zunächst die Ausdrücke

$$P\{U_n = \ell, \underline{B}_j = \underline{\beta}, \underline{\Gamma}^{(n)} = j | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \underline{\Gamma}^{(n-1)} = i\} \quad \text{mit } \ell > 0$$

betrachtet. Im ersten Fall $\Gamma_1^{(n)} = n - 1$, d.h. Kunde $n - 1$ geht nicht verloren, ist

$$[\Gamma_1^{(n)} = n - 1 \wedge U_n = \ell] \iff [\Gamma_1^{(n)} = n - 1 \wedge \underline{*}\Gamma^{(n)} = \underline{\Gamma}^{*(n-1)} \wedge U_{n-1} + B_{n-1} + A_{n-1} = \ell],$$

und es gilt⁵

$$\begin{aligned} P\{U_n = \ell, \underline{B}_j = \underline{\beta}, \underline{\Gamma}^{(n)} = j | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \underline{\Gamma}^{(n-1)} = i\} &= \\ = P\{U_{n-1} + B_{n-1} - A_{n-1} = \ell, \underline{B}_{(n-1, i^*)} = (\beta_1, \underline{\gamma}^*), \underline{\Gamma}^{(n)} = (n - 1, i^*) | \\ &\quad U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \underline{\Gamma}^{(n-1)} = i\} = \\ = P\{A_{n-1} = k - \ell + \beta_1, \underline{B}_{(n-1, i^*)} = (\beta_1, \underline{\gamma}^*), \underline{\Gamma}^{(n)} = (n - 1, i^*) | \\ &\quad U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \underline{\Gamma}^{(n-1)} = i\}. \end{aligned} \quad (4.34)$$

⁵ **Abkürzungen:**

$\underline{\Gamma}^{*(n-1)} = (\Gamma_1^{(n-1)}, \dots, \Gamma_N^{(n-1)})$	$j := (j_1, \dots, j_{N+1})$
$\underline{*}\Gamma^{(n)} = (\Gamma_2^{(n)}, \dots, \Gamma_{N+1}^{(n)})$	$\underline{*}j := (j_2, \dots, j_{N+1})$
$\underline{\gamma}^* = (\gamma_1, \dots, \gamma_N)$	$\underline{j}^* := (j_1, \dots, j_N)$
$(\underline{\beta}_1, \underline{\gamma}^*) = (\beta_1, \gamma_1, \dots, \gamma_N)$	$(\underline{j}^*, i_{N+1}) := (j_1, \dots, j_N, i_{N+1})$

$$\begin{aligned} \text{Mit } \{\omega : \underline{B}_{(n-1, i_*)} = (\beta_1, \underline{\gamma}^*), \Gamma^{(n)} = (n-1, i_*)\} \cap \{\omega : \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} = \\ = \{\omega : B_{n-1} = \beta_1, \underline{B}_i = \underline{\gamma}, \Gamma_1^{(n)} = n-1, \Gamma^{(n-1)} = i\} \end{aligned}$$

und der Unabhängigkeit von Zwischenankunftszeiten und Bedienzeiten kann Gl. (4.34) weiter umgeformt werden zu ($a(k)$ bezeichnet die WF der Zwischenankunftszeiten)

$$\begin{aligned} a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot P\{\Gamma_1^{(n)} = n-1 | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} = \\ = a(k - \ell + \beta_1) \cdot b(\beta_1), \end{aligned} \quad (4.35)$$

$$\text{da } P\{\Gamma_1^{(n)} = n-1 | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} =$$

$$= \begin{cases} 1 & \text{für } 0 \leq k \leq \sum_{i=1}^N \gamma_i \\ 0 & \text{für } 1 + \sum_{i=1}^N \gamma_i \leq k \leq \sum_{i=1}^{N+1} \gamma_i, \\ & \text{was Verlust der } (n-1)\text{-ten Ankunft impliziert.} \end{cases}$$

Im zweiten Fall geht der Kunde $n-1$ verloren, d.h. $\Gamma_1^{(n)} < n-1$, und mit

$$[\Gamma_1^{(n)} < n-1 \wedge U_n = \ell] \iff [\Gamma_1^{(n)} < n-1 \wedge \Gamma^{(n)} = \Gamma^{(n-1)} \wedge U_{n-1} - A_{n-1} = \ell]$$

$$\begin{aligned} \text{folgt } P\{U_n = \ell, \underline{B}_j = \underline{\beta}, \Gamma^{(n)} = j | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} = \\ = P\{U_{n-1} - A_{n-1} = \ell, \underline{B}_i = \underline{\beta}, \Gamma^{(n)} = i | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} = \\ = P\{A_{n-1} = k - \ell | U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} = a(k - \ell). \end{aligned} \quad (4.36)$$

Insgesamt ergibt sich also

$$\begin{aligned} P\{U_n = \ell, \underline{B}_{(j_1, \dots, j_{N+1})} = (\beta_1, \dots, \beta_{N+1}), \Gamma^{(n)} = (j_1, \dots, j_{N+1}) | \\ U_{n-1} = k, \underline{B}_{(i_1, \dots, i_{N+1})} = (\gamma_1, \dots, \gamma_{N+1}), \Gamma^{(n-1)} = (i_1, \dots, i_{N+1})\} = \\ = \begin{cases} \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{j_1 = i_1} \cdot \mathbf{1}_{j_2 = \dots = j_N = \dots} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) & \text{für } j_1 = n-1 \\ \mathbf{1}_{k \geq 1 + \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{j_1 = i_1} \cdot \mathbf{1}_{j_2 = \dots = j_N = \dots} \cdot a(k - \ell) & \text{für } j_1 < n-1 \\ 0 & \text{sonst.} \end{cases} \end{aligned} \quad (4.37)$$

Mit den Ausdrücken (4.37) können die Übergangs-WKTN formuliert werden, indem über i und j summiert wird.

$$\begin{aligned} P\{U_n = \ell, \underline{B}_{\Gamma^{(n)}} = \underline{\beta}, U_{n-1} = k, \underline{B}_{\Gamma^{(n-1)}} = \underline{\gamma}\} = \\ \sum_{i, j} P\{U_n = \ell, \underline{B}_j = \underline{\beta}, \Gamma^{(n)} = j, U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} = \\ \sum_{i, j} \sum_{j_1} \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{j_1 = n-1} \cdot \mathbf{1}_{j_2 = \dots = j_N = \dots} \cdot \mathbf{1}_{i_1 = \dots = i_N = \dots} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot \\ \cdot P\{U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \Gamma^{(n-1)} = i\} + \end{aligned} \quad (4.38)$$

$$+ \sum_{\substack{i \geq j \\ \mathbf{1}_{k \geq 1 + \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{j_1 < n-1} \cdot \mathbf{1}_{\underline{j} = i} \cdot \mathbf{1}_{\underline{\beta} = \underline{\gamma}}}} a(k - \ell) \cdot P\{U_{n-1} = k, \underline{B}_i = \underline{\gamma}, \underline{\Gamma}^{(n-1)} = \underline{i}\} . \quad (4.39)$$

Die erste Summe (4.38) liefert:

$$\begin{aligned} & \sum_{j_1, i_{N+1}} \sum_{\substack{i \geq * \\ j}} \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{j_1 = n-1} \cdot \mathbf{1}_{\underline{j} = i \underline{*}} \cdot \mathbf{1}_{*\underline{\beta} = \underline{\gamma} *} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot \\ & \quad \cdot P\{U_{n-1} = k, \underline{B}_{i \underline{*}} = \underline{\gamma} *, \underline{\Gamma}^{(n-1)} = i \underline{*}, B_{i_{N+1}} = \gamma_{N+1}, \Gamma_{N+1}^{(n-1)} = i_{N+1}\} = \\ & = \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{*\underline{\beta} = \underline{\gamma} *} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot \\ & \quad \cdot \sum_{i_{N+1}} \sum_{i \underline{*}} P\{U_{n-1} = k, \underline{B}_{i \underline{*}} = \underline{\gamma} *, \underline{\Gamma}^{(n-1)} = i \underline{*}, B_{i_{N+1}} = \gamma_{N+1}, \Gamma_{N+1}^{(n-1)} = i_{N+1}\} = \\ & = \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{*\underline{\beta} = \underline{\gamma} *} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot \\ & \quad \cdot \sum_{i_{N+1}} P\{U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = \underline{\gamma} *, B_{i_{N+1}} = \gamma_{N+1}, \Gamma_{N+1}^{(n-1)} = i_{N+1}\} = \\ & = \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{*\underline{\beta} = \underline{\gamma} *} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot P\{U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = \underline{\gamma}\} , \end{aligned}$$

die zweite Summe (4.39) ergibt:

$$\begin{aligned} & \sum_{\underline{i}} \mathbf{1}_{k \geq 1 + \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{i_1 < n-1} \cdot \mathbf{1}_{\underline{\beta} = \underline{\gamma}} \cdot a(k - \ell) \cdot P\{U_{n-1} = k, \underline{B}_{\underline{i}} = \underline{\gamma}, \underline{\Gamma}^{(n-1)} = \underline{i}\} = \\ & = \mathbf{1}_{k \geq 1 + \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{\underline{\beta} = \underline{\gamma}} \cdot a(k - \ell) \cdot P\{U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = \underline{\gamma}\} . \end{aligned}$$

(Nach Definition ist $\Gamma_1^{(n-1)} < n - 1$, daher kann $\mathbf{1}_{i_1 < n-1}$ entfallen.)

Durch Division ergeben sich schließlich die gesuchten Übergangs-WKTn für $\ell > 0$

$$\begin{aligned} & P\{U_n = \ell, \underline{B}_{\underline{\Gamma}^{(n)}} = \underline{\beta} | U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = \underline{\gamma}\} = \\ & = \mathbf{1}_{k \leq \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{*\underline{\beta} = \underline{\gamma} *} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) + \mathbf{1}_{k \geq 1 + \gamma_1 + \dots + \gamma_N} \cdot \mathbf{1}_{\underline{\beta} = \underline{\gamma}} \cdot a(k - \ell) \end{aligned}$$

4.4.4 Transiente Gleichungen

Wie üblich erhält man die Zustands-WKTn der MK wie folgt:

Für $\ell > 0$:

$$\begin{aligned} & P\{U_n = \ell, \underline{B}_{\underline{\Gamma}^{(n)}} = (\beta_1, \dots, \beta_{N+1})\} = \\ & \sum_{(\gamma_1, \dots, \gamma_{N+1})} \sum_{k=0}^{\gamma_1 + \dots + \gamma_{N+1}} P\{U_n = \ell, \underline{B}_{\underline{\Gamma}^{(n)}} = (\beta_1, \dots, \beta_{N+1}), U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = (\gamma_1, \dots, \gamma_{N+1})\} = \\ & \sum_{(\gamma_1, \dots, \gamma_{N+1})} \sum_{k=0}^{\gamma_1 + \dots + \gamma_N} \mathbf{1}_{*\underline{\beta} = \underline{\gamma} *} \cdot a(k - \ell + \beta_1) \cdot b(\beta_1) \cdot P\{U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = (\gamma_1, \dots, \gamma_{N+1})\} + \\ & \quad + \sum_{k=1 + \gamma_1 + \dots + \gamma_N}^{\gamma_1 + \dots + \gamma_{N+1}} \mathbf{1}_{\underline{\beta} = \underline{\gamma}} \cdot a(k - \ell) \cdot P\{U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = (\gamma_1, \dots, \gamma_{N+1})\} = \\ & = b(\beta_1) \cdot \sum_{k=0}^{\beta_2 + \dots + \beta_{N+1}} a(k - \ell + \beta_1) \sum_{\gamma_{N+1}} P\{U_{n-1} = k, \underline{B}_{\underline{\Gamma}^{(n-1)}} = (\beta_2, \dots, \beta_{N+1}, \gamma_{N+1})\} + \end{aligned}$$

$$+ \sum_{k=1+\beta_1+\dots+\beta_N}^{\beta_1+\dots+\beta_{N+1}} a(k-\ell) P\{U_{n-1} = k, \underline{B}_{\Gamma^{(n-1)}} = (\beta_1, \dots, \beta_{N+1})\}.$$

Mit den Abkürzungen $u_{n-1}(k, \beta_1, \dots, \beta_{N+1}) := P\{U_{n-1} = k, \underline{B}_{\Gamma^{(n-1)}} = (\beta_1, \dots, \beta_{N+1})\}$ erhält man für alle gemäß $b(\beta)$ möglichen $\beta_1, \dots, \beta_{N+1}$ und $1 \leq \ell \leq \beta_1 + \dots + \beta_{N+1}$

$$\begin{aligned} u_n(\ell, \beta_1, \dots, \beta_{N+1}) &= \\ & b(\beta_1) \cdot \sum_{k=0}^{\beta_2+\dots+\beta_{N+1}} a(k-\ell+\beta_1) \cdot \sum_{\gamma_{N+1}} u_{n-1}(k, \beta_2, \dots, \beta_{N+1}, \gamma_{N+1}) + \\ & + \sum_{k=1+\beta_1+\dots+\beta_N}^{\beta_1+\dots+\beta_{N+1}} a(k-\ell) \cdot u_{n-1}(k, \beta_1, \dots, \beta_{N+1}). \end{aligned} \quad (4.40)$$

Die fehlenden Gleichungen für $\ell = 0$ ergeben sich aus der Beziehung

$$\sum_{k=0}^{\beta_1+\dots+\beta_{N+1}} u_n(k, \beta_1, \dots, \beta_{N+1}) = b(\beta_1) \cdot \dots \cdot b(\beta_{N+1}). \quad (4.41)$$

Sie läßt sich folgendermaßen begründen: Bei N Warteplätzen ist $0 \leq U_n \leq \beta_1 + \dots + \beta_{N+1}$. Es bleibt zu zeigen

$$P\{\underline{B}_{\Gamma^{(n)}} = \underline{\beta}\} = b(\beta_1) \cdot \dots \cdot b(\beta_{N+1}).$$

Ausgeschrieben heißt das $P\{\underline{B}_{\Gamma^{(n)}} = \underline{\beta}\} =$

$$\sum_{i_1, \dots, i_{N+1}} P\{B_{i_1} = \beta_1, \dots, B_{i_{N+1}} = \beta_{N+1}, \Gamma_1^{(n)} = i_1, \dots, \Gamma_{N+1}^{(n)} = i_{N+1}\} \quad (4.42)$$

Der letzte Ausdruck kann mit folgenden Argumenten weiter umgeformt werden:

- Die Bedienzeiten sind voneinander und von allen anderen ZVn unabhängig.
- Die Zulassung des Kunden $\Gamma_{N+1}^{(n)}$ kann nur durch die Nummern und Bedienzeiten von Kunden, die vor dem Kunden $\Gamma_{N+1}^{(n)}$ zugelassen wurden, beeinflusst werden, nicht aber von den Nummern und Bedienzeiten später zugelassener Kunden (diese haben die Nummern $\Gamma_N^{(n)}, \dots, \Gamma_1^{(n)}$).
- Die Bedienzeiten sind im strengen Sinne stationär, d.h. ihre WFn sind unabhängig von ihren Indizes.

Die Umformung ergibt aus Gl. (4.42):

$$\begin{aligned} & b(\beta_{N+1}) \cdot \sum_{i_1, \dots, i_{N+1}} P\{B_{i_1} = \beta_1, \dots, B_{i_N} = \beta_N, \Gamma_1^{(n)} = i_1, \dots, \Gamma_{N+1}^{(n)} = i_{N+1}\} = \\ & = b(\beta_{N+1}) \cdot \sum_{i_1, \dots, i_N} P\{B_{i_1} = \beta_1, \dots, B_{i_N} = \beta_N, \Gamma_1^{(n)} = i_1, \dots, \Gamma_N^{(n)} = i_N\}. \end{aligned}$$

Dieselben Argumente gelten auch für

$$\sum_{i_1, \dots, i_N} P\{B_{i_1} = \beta_1, \dots, B_{i_N} = \beta_N, \Gamma_1^{(n)} = i_1, \dots, \Gamma_N^{(n)} = i_N\}$$

und die sukzessiv verkleinerten Ausdrücke, so daß die Behauptung folgt.

Im stationären Fall $n \rightarrow \infty$ werden die Unbekannten $u_{n-1}(k, \beta_1, \dots, \beta_{N+1}) = u_n(k, \beta_1, \dots, \beta_{N+1})$ unabhängig von n , und es ergibt sich ein System linearer Gleichungen in diesen Unbekannten.

4.4.5 Wahrscheinlichkeitsfunktion der Belegung

$$P\{L_n = 0\} = \sum_{\underline{\beta}} u_n(0, \beta_1, \dots, \beta_{N+1})$$

$$P\{L_n = 1\} = \sum_{\underline{\beta}} \sum_{\ell=1}^{\beta_1} u_n(\ell, \beta_1, \dots, \beta_{N+1})$$

$$P\{L_n = i\} = \sum_{\underline{\beta}} \sum_{\ell=1+\beta_1+\dots+\beta_{i-1}}^{\beta_1+\dots+\beta_i} u_n(\ell, \beta_1, \dots, \beta_{N+1}), \quad 2 \leq i \leq N+1$$

$\sum_{\underline{\beta}}$ bedeutet dabei die Summation über alle möglichen Vektoren $\underline{\beta} = (\beta_1, \dots, \beta_{N+1})$. Diese haben als Komponenten Werte β_i , $1 \leq i \leq N+1$, die nach der vorgegebenen WF $b(\beta)$ der Bedienzeit mit positiver WKT auftreten.

4.4.6 Wahrscheinlichkeitsfunktion der Wartezeit

$$P\{W = \ell \mid \text{kein Verlust}\} = \frac{\sum_{\underline{\beta}} \mathbf{1}_{\ell \leq \beta_1 + \dots + \beta_N} \cdot u_n(\ell, \beta_1, \dots, \beta_{N+1})}{\sum_{\underline{\beta}} \sum_{\ell=0}^{\beta_1 + \dots + \beta_N} u_n(\ell, \beta_1, \dots, \beta_{N+1})}$$

4.4.7 Semi-Markovscher Eingabeprozess

Durch ähnliches Vorgehen, das formal allerdings noch unübersichtlicher ist, erhält man für einen Semi-Markovschen Eingabeprozess mit SM-Kern $\mathbf{q}(k)$ und Zustandsraum \mathbb{IE} mit der Abkürzung $u_{n-1}(k, i, \beta_1, \dots, \beta_{N+1}) := P\{U_{n-1} = k, S_{n-1} = i, \underline{B}_{\Gamma(n-1)} = (\beta_1, \dots, \beta_{N+1})\}$ für $j \in \mathbb{IE}$, $\beta_1, \dots, \beta_{N+1}$ und $\ell > 0$

$$u_n(\ell, j, \beta_1, \dots, \beta_{N+1}) =$$

$$\begin{aligned} & b(\beta_1) \cdot \sum_{k=0}^{\beta_2 + \dots + \beta_{N+1}} \sum_{i \in \mathbb{IE}} q_{ij}(k - \ell + \beta_1) \cdot \sum_{\gamma_{N+1}} u_{n-1}(k, i, \beta_2, \dots, \beta_{N+1}, \gamma_{N+1}) \\ & + \sum_{k=1+\beta_1+\dots+\beta_N}^{\beta_1+\dots+\beta_{N+1}} \sum_{i \in \mathbb{IE}} q_{ij}(k - \ell) \cdot u_{n-1}(k, i, \beta_1, \dots, \beta_{N+1}) \end{aligned} \quad (4.43)$$

und ($\ell = 0$ repräsentierend)

$$\sum_{k=0}^{\beta_1+\dots+\beta_{N+1}} u_n(k, j, \beta_1, \dots, \beta_{N+1}) = P_j \cdot b(\beta_1) \cdot \dots \cdot b(\beta_{N+1}) . \quad (4.44)$$

Zwar kann man die Zustands-WKTn auch allein aus den Übergangs-WKTn bestimmen (mittels Grassmann-Algorithmus bei vorgegebener Übergangsmatrix). Doch ist wegen der Vielfachindizierung der Unbekannten eine Blockmatrizendarstellung der Übergangsmatrix äußerst unübersichtlich.

4.4.8 Diskussion der Ergebnisse

Mittels der obigen Gleichungssysteme können für die endlichen zeitdiskreten Wartesysteme vom Typ GI/GI/1/N und SMP/GI/1/N WFn der Belegung und der Wartezeit in ähnlicher

Weise berechnet werden, wie es für das $\text{SMP}^{[X]}/D/1/N$ -System schon dargestellt wurde. Dabei wächst allerdings die Anzahl der Unbekannten gewaltig. Der Anstieg der Anzahl wird hauptsächlich durch die Anzahl C_β derjenigen Bedienzeiten bestimmt, die mit positiver WKT auftreten können. Schon im $GI/GI/1/N$ -Fall hat man $\sum_\beta (1 + \beta_1 + \dots + \beta_{N+1}) > (C_\beta)^{N+1}$ Unbekannte⁶; bei Semi-Markovschem Eingabeprozess ist diese Zahl nochmals um die Anzahl der Zustände des Eingabeprozesses multiplikativ höher.

Von praktischem Nutzen sind die obigen Verfahren somit nur für kleine N und C_β . Für $C_\beta = 1$ erhält man dieselben Gleichungen wie für das $\text{SMP}/D/1/N$ -Wartesystem [51].

Validierung Eine Validierung der Gleichungen (sowohl für Erneuerungs- als auch für Semi-Markov-Eingabeprozess) ist möglich mittels des DMAP/G/1/N-Wartesystems, s. Kapitel 5. Durch geeignete Wahl der Matrizen \mathbf{C} und \mathbf{D} kann nach [13] ein beliebiger Erneuerungsprozess mit vorgegebener maximaler Zwischenankunftszeit durch einen DMAP repräsentiert werden. Für kleine N und C_β wurde für die WF der Belegung nach Abschnitt 4.4.5 genauso gute Übereinstimmung gefunden wie für $\text{SMP}^{[X]}/D/1/N$ und $\text{DBMAP}/G/1/N$, s. Kapitel 7. Dazu wurde das Gleichungslösungstool des MAPLE-Programmpaketes⁷ eingesetzt.

Verwendet man mit geeignetem DMAP das DMAP/G/1/N-Modell zur Behandlung des $GI/GI/1/N$ -Systems, so bestimmt der maximal mögliche Wert der Zwischenankunftszeiten im wesentlichen die Anzahl der Gleichungen (und zwar linear), während es bei der Methode der Restarbeit die Anzahl der Bedienzeiten ist, die mit positiver WKT auftreten (dann aber exponentiell).

Da die Bedienzeiten aller wartenden Kunden bei der Methode der Restarbeit bekannt sind, dürfte das Verfahren auch für die SPT-Strategie (shortest processing time first) geeignet sein. Dies bleibt zukünftigen Untersuchungen vorbehalten.

⁶Es ist $0 \leq k \leq \beta_1 + \dots + \beta_{N+1}$ für jeden Vektor $\beta = (\beta_1, \dots, \beta_{N+1})$.

⁷MAPLE ist ein eingetragenes Warenzeichen der WATERLOO MAPLE SOFTWARE.

5 Zeitdiskrete Systeme vom M/G/1-Typ

Eine bekannte Analysemethode für das M/G/1-Wartesystem zu Abgangszeiten nutzt die Eigenschaft des Eingabeprozesses aus, daß der zugehörige Zählprozeß (der Poisson-Prozeß) als Markov-Prozeß aufgefaßt werden kann [20], p. 235. Die Lösungsmethode läßt sich auf andere Eingabeprozesse mit Markovschem Zählprozeß übertragen. Für diese erweiterte Gruppe von Wartesystemen, deren Eingabeprozeß "Markovian Arrival Process" (kurz: MAP) genannt wird, hat sich die Bezeichnung "Wartesystem vom M/G/1-Typ" etabliert. Die (meist mehr interessierenden) Wartesystemgrößen zu *Ankunftszeitpunkten* werden durch eine Transformation aus denjenigen zu Abgangszeitpunkten gewonnen. Die Untersuchung der allgemeinen Gruppe der M/G/1-Typ-Wartesysteme wurde von Neuts [85, 86] initiiert und von Ramaswami [96, 97] und Lucantoni [75, 76] für zeitkontinuierliche unendliche Systeme verfeinert. Blondia [6, 7] behandelt den Fall eines endlichen zeitkontinuierlichen Systems.

Nachfolgend werden die zugehörigen zeitdiskreten Systeme analysiert. Infolge der beiden Zulassungsregeln AF und DF geht der zeitdiskrete Fall nicht einfach durch "Diskretisierung" aus dem zeitkontinuierlichen hervor.

5.1 Das zeitdiskrete DBMAP/G/1/N-Wartesystem

Wie im letzten Kapitel wird die FIFO-Strategie angenommen. Der Wartespeicher hat N Plätze, d.h. die Gesamtkapazität einschließlich Bediener beträgt $N + 1$. Die Zwischenankunftszeiten werden von einem DBMAP¹ erzeugt. Die (positiven) Bedienzeiten sind stochastisch unabhängig und identisch verteilt, ihre WF wird mit $h(t)$ bezeichnet. Da eine beliebige Bedienzeitverteilung vorausgesetzt werden kann, wird auch der Fall einer deterministischen Bedienzeit behandelbar, die größer ist als die Zeiteinheit. Letzteres wird zur Modellierung eines ATM-Konzentrators erforderlich, wenn die Ausgangsbitrate kleiner ist als diejenige am Eingang. Diese Situation tritt auf in der Teilnehmervermittlungsstelle des Empfängers, wo die Bitrate von 600 Mbit/s auf 150 Mbit/s abfällt.

Im Falle des M/G/1/N-Systems bildet der SP der Belegung zu Abgangszeitpunkten bekanntlich einen Markov-Erneuerungsprozeß (MEP) ([20], p. 317), weil infolge des Poisson-Eingabeprozesses die WKT für k Ankünfte in einem Intervall der Länge τ nur von τ abhängt. Der Zustandsraum des MEP entspricht der Menge der Belegungswerte, die Zeit-ZVn repräsentieren die Abgangszeitpunkte.

Beim Zählprozeß des DBMAP hängt die WKT für k Ankünfte in einem Zeitintervall der Länge τ von τ und den Phasen des DBMAP zu Beginn und am Ende des Intervalls ab, s. Gl. (3.42). Da die Phasen eine MK bilden, ergibt sich wiederum ein MEP für den Verbundprozeß von Belegung (des Wartesystems) und Phase (des Eingabeprozesses) zu Abgangszeitpunkten. Dieser MEP ist zu Abgangszeitpunkten *eingebettet*, da der Belegungszustand des Wartesystems sich zwischen zwei Abgängen durch weitere Ankünfte ändern

¹"Discrete Batch Markovian Arrival Process", wie in Abschnitt 3.4 erläutert.

kann. Der resultierende SM-Kern $\tilde{q}(t)$ ist vom M/G/1-Typ [86], d.h. er bildet eine *obere Block-Hessenberg-Matrix*². Die Matrixeinträge³ $[\tilde{q}(t)]_{k\ell}$ bezeichnen die WKT, daß unter der Bedingung, daß ein Abgang bei Phase i des Eingabeprozesses k Kunden im System zurückläßt, der nächste Abgang t Zeiteinheiten später geschieht, der Eingabeprozess dann in Phase j weilt und der Abgang ℓ Kunden im System zurückläßt. Die Übergangsmatrix der eingebetteten MK des MEP wird mit

$$\tilde{q} := \sum_{t=1}^{\infty} \tilde{q}(t) \quad (5.1)$$

bezeichnet. Mit ihrer stationären Lösung können die WFn der Belegung zu beliebigen Zeitpunkten und zu Ankunftszeitpunkten bestimmt werden und damit die Verlust-WKT. Außerdem charakterisiert der MEP den Abgangsprozeß, d.h. den SP der Zwischenabgangszeiten vollständig. Die Berechnung von Verbund-WFn der Zwischenabgangszeiten erfolgt mittels bekannter Methoden der Semi-Markov-Prozesse an Hand von Gl. (3.13).

Zunächst wird der Abgangsprozeß für AF und DF analysiert, um dann, ausgehend von der Belegung zu Abgangszeiten, zu Ausdrücken für die Belegung zu beliebigen Zeitpunkten und schließlich zu Ankunftszeitpunkten zu gelangen. Hierbei wird nur der stationäre Fall berücksichtigt. Geschlossene Ausdrücke für die transiente Situation scheinen mit dieser Methode unmöglich zu sein, s. Abschnitt 5.1.4.

5.1.1 Der zu Abgangszeitpunkten eingebettete Markov-Erneuerungsprozeß

Sowohl für AF als auch für DF bedeutet die Sprechweise "der Abgang hinterläßt k Kunden im System", daß unmittelbar nach dem Abgang die Belegung des Systems den Wert k hat, wobei eine Ankunft, die gleichzeitig mit dem Abgang stattfindet, bei AF und DF darin mitberücksichtigt ist. Die MEP für AF bzw. DF sind verschieden. Bei AF kann ein Abgang höchstens N Kunden im System zurücklassen, wie es auch für den zeitkontinuierlichen Fall zutrifft, denn dort treten Ankünfte und Abgänge nur mit WKT Null gleichzeitig auf. Bei DF hinterläßt ein Abgang ein volles System (d.h. Belegung $N + 1$) genau dann, wenn eine Ankunft gleichzeitig mit diesem Abgang stattfindet *und* diese Ankunft das System voll macht.

Abb. 5.1 verdeutlicht die Entwicklung des Wartesystems mittels der Restarbeit. Wie in Abschnitt 4.1.1 erläutert, kann bei DF die Restarbeit $U(t)$ durch rechtsseitig stetige Treppenfunktionen beschrieben werden, während sie für AF linksseitig stetig sind. Ein Pfad von $U(t)$ wird wie folgt konstruiert: Beim n -ten Ankunftszeitpunkt T_n entscheidet die Belegung *ohne Berücksichtigung der Ankunft* über die Zulassung des ankommenden Kunden. Wird der Kunde zugelassen, so springt die Restarbeit bei T_n um die Bedienzeit des Kunden. Geht der Kunde verloren, so bleibt die Restarbeit *mit Berücksichtigung der Ankunft* unverändert gegenüber derjenigen *ohne Berücksichtigung der Ankunft*. Die hierfür nötige Formalisierung findet sich in Abschnitt 4.1.1. Sie ist hier nicht erforderlich, weil die Beschreibung des Systems sich (an dieser Stelle) auf Abgangszeitpunkte beschränkt, wofür die Belegung ausreicht.

²Bei einer oberen Hessenberg-Matrix (h_{ij}) ist $h_{ij} = 0$ für $i > j + 1$ [39], p. 153, bei der Block-Hessenberg-Matrix gilt die entsprechende Beziehung hinsichtlich der Untermatrizen.

³mit Bezug auf Abb. 5.2 bzw. 5.3; die innere Klammer bezieht sich auf die k -te Hauptzeile und ℓ -te Hauptspalte, die äußere auf die i -te Zeile und j -te Spalte der jeweiligen Teilmatrix

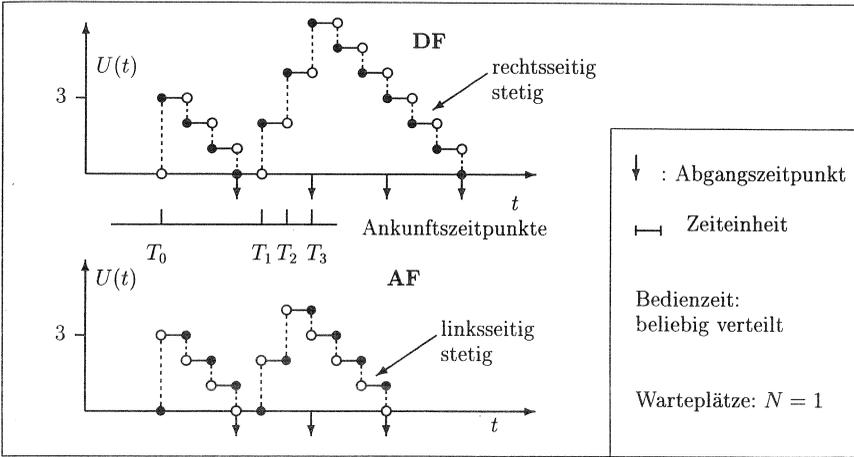


Abb. 5.1 Pfade der Restarbeit $U(t)$ für AF und DF: Der bei T_3 ankommende Kunde geht bei AF verloren, weil er noch den Kunden antrifft, der bei T_1 das System betreten hat und es bei T_3 wieder verläßt.

5.1.1.1 Arrival First

Für AF ergibt sich ein SM-Kern $\tilde{q}(t)$ ähnlich demjenigen des zeitkontinuierlichen Systems, s. Abb. 5.2. Als Untermatrizen treten auf

- $\mathbf{A}_k(t)$, wobei $[\mathbf{A}_k(t)]_{ij}$ die (bedingte) WKT bedeutet, daß unter der Bedingung eines Abgangs, der das System nicht leer hinterläßt und bei dem der Eingabeprozess in Phase i weilt, in der folgenden t Zeiteinheiten dauernden Bedienzeit k Kunden ankommen und der Eingabeprozess beim nächsten Abgang sich in Phase j befindet. Der Bedienprozeß ist von allem anderen unabhängig, daher gilt $\mathbf{A}_k(t) = \mathbf{P}(k, t)h(t)$. Für $\mathbf{P}(k, t)$ vgl. Abschnitt 3.4.2.
- $\mathbf{B}_k(t)$, wobei $[\mathbf{B}_k(t)]_{ij}$ die (bedingte) WKT bedeutet, daß unter der Bedingung eines Abgangs, der das System leer verläßt und bei dem der Eingabeprozess in Phase i weilt, der nächste Abgang nach t Zeiteinheiten stattfindet, in der Bedienzeit, die dann abgelaufen ist, k Kunden angekommen sind und der Eingabeprozess sich zu diesem Abgangszeitpunkt in Phase j befindet. Die Zeit t beinhaltet also immer genau eine Freiperiode⁴ und die erste Bedienzeit einer Betriebsperiode⁵. Man erhält

$$\begin{aligned} \mathbf{B}_k(t) &= \sum_{w=1}^{k+1} \sum_{v=1}^t \mathbf{D}_0^{v-1} \mathbf{D}_w \mathbf{P}(k-w+1, t-v) h(t-v) = \\ &= \sum_{w=1}^{k+1} \sum_{v=1}^t \mathbf{D}_0^{v-1} \mathbf{D}_w \mathbf{A}_{k-w+1}(t-v) . \end{aligned}$$

Begründung: Jeder Abgangszeitpunkt ist für den Eingabeprozess eine Stoppzeit. Nach einem Abgang, der eine Betriebsperiode beendet, muß wenigstens eine Zeiteinheit

⁴engl.: idle period.
⁵engl.: busy period.

$$\tilde{\mathbf{q}}(t) = \begin{pmatrix} \mathbf{B}_0(t) & \mathbf{B}_1(t) & \mathbf{B}_2(t) & \cdots & \mathbf{B}_{N-1}(t) & \sum_{k=N}^{\infty} \mathbf{B}_k(t) \\ \mathbf{A}_0(t) & \mathbf{A}_1(t) & \mathbf{A}_2(t) & \cdots & \mathbf{A}_{N-1}(t) & \sum_{k=N}^{\infty} \mathbf{A}_k(t) \\ \mathbf{0} & \mathbf{A}_0(t) & \mathbf{A}_1(t) & \cdots & \mathbf{A}_{N-2}(t) & \sum_{k=N-1}^{\infty} \mathbf{A}_k(t) \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_0(t) & \cdots & \mathbf{A}_{N-3}(t) & \sum_{k=N-2}^{\infty} \mathbf{A}_k(t) \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_0(t) & \sum_{k=1}^{\infty} \mathbf{A}_k(t) \end{pmatrix}$$

Abb. 5.2 Der Semi-Markov-Kern $\tilde{\mathbf{q}}(t)$ für AF.

verstreichen, bis der nächste w -Batch ankommt ($\mathbf{D}_0^{v-1} \mathbf{D}_w, v \geq 1$), andernfalls wäre die Betriebsperiode noch nicht abgeschlossen. Innerhalb der Bedienzeit des ersten zugelassenen Kunden des w -Batch, die $t - v$ Zeiteinheiten dauert, müssen weitere $k - w + 1$ Kunden ankommen, damit innerhalb dieser Bedienzeit (nach Abzug des bei Ablauf dieser Bedienzeit abgehenden Kunden) insgesamt k Ankünfte gezählt werden.

Die Übergangsmatrix $\tilde{\mathbf{q}}$ der dem MEP eingebetteten MK setzt sich zusammen aus den Blockmatrizen

$$\mathbf{B}_k := \sum_{t=1}^{\infty} \mathbf{B}_k(t) \quad \text{und} \quad \mathbf{A}_k := \sum_{t=1}^{\infty} \mathbf{A}_k(t), \quad k \geq 0,$$

wobei folgende Beziehung besteht

$$\mathbf{B}_k = \sum_{w=1}^{k+1} \sum_{v=1}^{\infty} \mathbf{D}_0^{v-1} \mathbf{D}_w \sum_{t=1}^{\infty} \mathbf{A}_{k-w+1}(t-v) = (\mathbf{I} - \mathbf{D}_0)^{-1} \sum_{w=0}^k \mathbf{D}_{w+1} \mathbf{A}_{k-w}. \quad (5.2)$$

Gl. (5.2) ergibt sich, wie folgt: $\sum_{v=1}^t \mathbf{A}_{k-w+1}(t-v) = \sum_{v=1}^{\infty} \mathbf{A}_{k-w+1}(t-v)$. Wenn \mathbf{B}_k für eine Summationsreihenfolge konvergiert, so nach dem *großen Umordnungssatz* [14], p.129, auch für jede Umordnung der Reihenfolge, da nur nicht-negative Summanden auftreten, also absolute Konvergenz vorliegt. Insbesondere kann die Summation über t und v, w vertauscht werden. Mit $\sum_{t=1}^{\infty} \mathbf{A}_{k-w+1}(t-v) = \sum_{t=1}^{\infty} \mathbf{A}_{k-w+1}(t) = \mathbf{A}_{k-w+1}$ für alle $v > 0$ folgt Gl. (5.2).

5.1.1.2 Departure First

Der Besonderheit bei DF, daß das System nach einem Abgang voll sein kann, wird durch Zerlegungen

$$\mathbf{A}_k(t) = \mathbf{A}'_k(t) + \mathbf{A}''_k(t) \quad \text{und} \quad \mathbf{B}_k(t) = \mathbf{B}'_k(t) + \mathbf{B}''_k(t) \quad (5.3)$$

Rechnung getragen. Dabei beziehen sich $\mathbf{A}'_k(t)$ und $\mathbf{B}'_k(t)$ auf den Fall, daß keine der Ankünfte innerhalb der Bedienzeit gleichzeitig mit dem Abgang bei Ablauf der Bedienzeit stattfindet. $\mathbf{A}''_k(t)$ und $\mathbf{B}''_k(t)$ repräsentieren die Situation, daß wenigstens eine der Ankünfte gleichzeitig mit dem Abgang auftritt. Formal heißt das ($k \geq 0$):

$$\mathbf{A}'_k(t) := \mathbf{P}(k, t-1) \mathbf{D}_0 h(t), \quad (5.4)$$

$$\tilde{\mathbf{q}}(t) = \begin{pmatrix} \mathbf{B}_0(t) & \mathbf{B}_1(t) & \mathbf{B}_2(t) & \cdots & \mathbf{B}_{N-1}(t) & \mathbf{B}_N(t) + \sum_{k=N+1}^{\infty} \mathbf{B}'_k(t) & \sum_{k=N+1}^{\infty} \mathbf{B}''_k(t) \\ \mathbf{A}_0(t) & \mathbf{A}_1(t) & \mathbf{A}_2(t) & \cdots & \mathbf{A}_{N-1}(t) & \mathbf{A}_N(t) + \sum_{k=N+1}^{\infty} \mathbf{A}'_k(t) & \sum_{k=N+1}^{\infty} \mathbf{A}''_k(t) \\ \mathbf{0} & \mathbf{A}_0(t) & \mathbf{A}_1(t) & \cdots & \mathbf{A}_{N-2}(t) & \mathbf{A}_{N-1}(t) + \sum_{k=N}^{\infty} \mathbf{A}'_k(t) & \sum_{k=N}^{\infty} \mathbf{A}''_k(t) \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_0(t) & \cdots & \mathbf{A}_{N-3}(t) & \mathbf{A}_{N-2}(t) + \sum_{k=N-1}^{\infty} \mathbf{A}'_k(t) & \sum_{k=N-1}^{\infty} \mathbf{A}''_k(t) \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_0(t) & \mathbf{A}_1(t) + \sum_{k=2}^{\infty} \mathbf{A}'_k(t) & \sum_{k=2}^{\infty} \mathbf{A}''_k(t) \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{A}_0(t) + \sum_{k=1}^{\infty} \mathbf{A}'_k(t) & \sum_{k=1}^{\infty} \mathbf{A}''_k(t) \end{pmatrix}$$

Abb. 5.3 Der Semi-Markov-Kern $\tilde{\mathbf{q}}(t)$ für DF.

$$\mathbf{A}''_k(t) := \sum_{\mu=1}^k \mathbf{P}(k - \mu, t - 1) \mathbf{D}_{\mu} h(t), \quad \mathbf{A}''_0(t) := \mathbf{0}, \quad (5.5)$$

$$\mathbf{B}'_k(t) := \sum_{w=1}^{k+1} \sum_{v=1}^t \mathbf{D}_0^{v-1} \mathbf{D}_w \mathbf{A}'_{k-w+1}(t-v), \quad (5.6)$$

$$\mathbf{B}''_k(t) := \sum_{w=1}^{k+1} \sum_{v=1}^t \mathbf{D}_0^{v-1} \mathbf{D}_w \mathbf{A}''_{k-w+1}(t-v). \quad (5.7)$$

Der SM-Kern ist in Abb. 5.3 dargestellt. Die Übergangsmatrix der eingebetteten MK enthält neben den oben für AF definierten Untermatrizen \mathbf{B}_k und \mathbf{A}_k zusätzlich noch

$$\mathbf{A}'_k := \sum_{t=1}^{\infty} \mathbf{A}'_k(t), \quad \mathbf{A}''_k := \sum_{t=1}^{\infty} \mathbf{A}''_k(t), \quad \mathbf{B}'_k := \sum_{t=1}^{\infty} \mathbf{B}'_k(t), \quad \mathbf{B}''_k := \sum_{t=1}^{\infty} \mathbf{B}''_k(t) \quad (5.8)$$

mit den der Gl. (5.2) ähnlichen Verknüpfungen:

$$\mathbf{B}'_k = (\mathbf{I} - \mathbf{D}_0)^{-1} \sum_{w=0}^k \mathbf{D}_{w+1} \mathbf{A}'_{k-w} \quad \text{und} \quad (5.9)$$

$$\mathbf{B}''_k = (\mathbf{I} - \mathbf{D}_0)^{-1} \sum_{w=0}^k \mathbf{D}_{w+1} \mathbf{A}''_{k-w}. \quad (5.10)$$

Die Berechnung der Matrizen $\mathbf{A}_k, \mathbf{A}'_k$ erfolgt über Gl. (3.43), \mathbf{A}''_k ist durch $\mathbf{A}_k = \mathbf{A}'_k + \mathbf{A}''_k$ bestimmt und die anderen Matrizen ergeben sich dann aus (5.2), (5.9), (5.10).

Stationäre Zustandswahrscheinlichkeiten Zur Bestimmung der stationären WKTn $[\underline{x}_r]_i$, $r = 0, 1, \dots, N, N+1$ für DF], daß ein Abgang r Kunden im System hinterläßt und sich der Eingabeprozess in Phase i befindet, wurden die Algorithmen aus [43, 70] verwendet, s. Abschnitt 7.8. Diese direkten Verfahren arbeiten nur mit der Übergangsmatrix.

Mit $\underline{\mathbf{x}} := (\underline{x}_0, \dots, \underline{x}_N)$ für AF und $\underline{\mathbf{x}} := (\underline{x}_0, \dots, \underline{x}_N, \underline{x}_{N+1})$ für DF erhält man aus der Matrixgleichung $\underline{\mathbf{x}}\tilde{\mathbf{q}} = \underline{\mathbf{x}}$:

AF:

$$\underline{x}_r = \underline{x}_0 \mathbf{B}_r + \sum_{\ell=1}^{r+1} \underline{x}_\ell \mathbf{A}_{r+1-\ell} \quad \text{für } 0 \leq r \leq N-1, \quad (5.11)$$

$$\underline{x}_N = \underline{x}_0 \sum_{k=N}^{\infty} \mathbf{B}_k + \sum_{\ell=1}^N \underline{x}_\ell \sum_{k=N+1-\ell}^{\infty} \mathbf{A}_k. \quad (5.12)$$

DF:

$$\underline{x}_r = \underline{x}_0 \mathbf{B}_r + \sum_{\ell=1}^{r+1} \underline{x}_\ell \mathbf{A}_{r+1-\ell} \quad \text{für } 0 \leq r \leq N-1, \quad (5.13)$$

$$\underline{x}_N + \underline{x}_{N+1} = \underline{x}_0 \sum_{k=N}^{\infty} \mathbf{B}_k + \sum_{\ell=1}^{N+1} \underline{x}_\ell \sum_{k=N+1-\ell}^{\infty} \mathbf{A}_k \quad (5.14)$$

und

$$\underline{x}_{N+1} = \underline{x}_0 \sum_{k=N+1}^{\infty} \mathbf{B}_k'' + \sum_{\ell=1}^{N+1} \underline{x}_\ell \sum_{k=N+2-\ell}^{\infty} \mathbf{A}_k''. \quad (5.15)$$

5.1.2 Charakterisierung des Abgangsprozesses

Mit dem SM-Kern $\tilde{\mathbf{q}}(t)$ können (Verbund-)WFn der Zwischenabgangszeiten E_n nach Gl. (3.13) berechnet werden. Da $\underline{\mathbf{x}}$ die stationäre Lösung der eingebetteten MK ist, sind die Verbund-WFn ebenfalls stationär, also invariant gegen eine beliebige Indexverschiebung.

$$P\{E_n = t_n, \dots, E_{n+\nu} = t_{n+\nu}\} = \underline{\mathbf{x}} \cdot \prod_{\mu=0}^{\nu} \tilde{\mathbf{q}}(t_{n+\mu}) \cdot \mathbf{e}, \quad \nu = 0, 1, 2, \dots \quad (5.16)$$

mit $\mathbf{e} = (1, 1, \dots, 1)^T$, $m \cdot (N+1)$ bzw. $m \cdot (N+2)$ Komponenten für AF bzw. DF und $\underline{\mathbf{x}} = (\underline{x}_0, \dots, \underline{x}_N, [\underline{x}_{N+1}])$.

5.1.2.1 WFn der Zwischenabgangszeiten

Für den Fall der deterministischen Bedienzeit werden explizit die WF sowie die Verbund-WFn erster und zweiter Ordnung angegeben, da letztere zur Berechnung des lokalen Korrelationskoeffizienten (erster und zweiter Ordnung) der Zwischenabgangszeiten benötigt werden, um damit den Abgangsprozeß durch einen SSMP(2) zu approximieren [30].

$$f(t) := P\{E_n = t\} = \underline{\mathbf{x}} \cdot \mathbf{q}(t) \cdot \mathbf{e}$$

Verbund-WF "erster Ordnung"

$$f_1(t', t'') := P\{E_n = t', E_{n+1} = t''\} = \underline{\mathbf{x}} \cdot \mathbf{q}(t') \mathbf{q}(t'') \cdot \mathbf{e}$$

Verbund-WF "zweiter Ordnung"

$$f_2(t', t'') := P\{E_n = t', E_{n+2} = t''\} = \underline{x} \cdot \mathbf{q}(t') \mathbf{q}\mathbf{q}(t'') \cdot \underline{e}$$

Im Anhang A.3 sind die Ausdrücke $\mathbf{q}(t) \cdot \underline{e}$, $\mathbf{q}(t_1)\mathbf{q}(t_2) \cdot \underline{e}$ und $\mathbf{q}(t_1)\mathbf{q}\mathbf{q}(t_2) \cdot \underline{e}$ für den allgemeinen Fall einschließlich der WF_n f , f_1 und f_2 berechnet. Im deterministischen Fall erhält man:

$$f(t) = \mathbf{1}_{t \geq 1+D} \underline{x}_0 \mathbf{D}_0^{t-(1+D)} (\mathbf{I} - \mathbf{D}_0) \underline{e} + \mathbf{1}_{t=D} (1 - \underline{x}_0 \underline{e}) \quad , \quad (5.17)$$

$$\begin{aligned} f_1(t_1, t_2) = & \mathbf{1}_{t_1 \geq 1+D} \left[\mathbf{1}_{t_2 \geq 1+D} \underline{x}_0 \mathbf{D}_0^{t_1-(1+D)} \mathbf{D}_1 \mathbf{D}_0^{t_2-1} (\mathbf{I} - \mathbf{D}_0) \underline{e} \right. \\ & \left. + \mathbf{1}_{t_2=D} \underline{x}_0 \mathbf{D}_0^{t_1-(1+D)} (\mathbf{I} - \mathbf{D}_0 - \mathbf{D}_1 \mathbf{D}_0^D) \underline{e} \right] \\ & + \mathbf{1}_{t_1=D} \left[\mathbf{1}_{t_2 \geq 1+D} \underline{x}_1 \mathbf{D}_0^{t_2-1} (\mathbf{I} - \mathbf{D}_0) \underline{e} + \mathbf{1}_{t_2=D} (1 - \underline{x}_0 \underline{e} - \underline{x}_1 \mathbf{D}_0^D \underline{e}) \right] \end{aligned} \quad (5.18)$$

und

$$\begin{aligned} f_2(t_1, t_2) = & \mathbf{1}_{t_1 \geq 1+D} \left[\mathbf{1}_{t_2 \geq 1+D} \underline{x}_0 \mathbf{D}_0^{t_1-(1+D)} \mathbf{H}_1 \mathbf{D}_0^{t_2-1} (\mathbf{I} - \mathbf{D}_0) \underline{e} \right. \\ & \left. + \mathbf{1}_{t_2=D} \underline{x}_0 \mathbf{D}_0^{t_1-(1+D)} (\mathbf{I} - \mathbf{D}_0 - \mathbf{H}_1 \mathbf{D}_0^D) \underline{e} \right] \\ & + \mathbf{1}_{t_1=D} \left[\mathbf{1}_{t_2 \geq 1+D} (\underline{x}_1 \mathbf{H}_2 + \underline{x}_2 \mathbf{D}_0^D) \mathbf{D}_0^{t_2-1} (\mathbf{I} - \mathbf{D}_0) \underline{e} \right. \\ & \left. + \mathbf{1}_{t_2=D} (1 - \underline{x}_0 \underline{e} - \underline{x}_1 \mathbf{H}_2 \mathbf{D}_0^D \underline{e} - \underline{x}_2 \mathbf{D}_0^{2D} \underline{e}) \right] \end{aligned} \quad (5.19)$$

mit den Hilfsausdrücken

$$\mathbf{H}_1 = \mathbf{D}_1 \mathbf{D}_0^D (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_1 + \mathbf{D}_1 \mathbf{A}_1(D) + \mathbf{D}_2 \mathbf{D}_0^D \quad ,$$

$$\mathbf{H}_2 = \mathbf{D}_0^D (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_1 + \mathbf{A}_1(D) \quad .$$

5.1.2.2 Geschlossene Ausdrücke für Mittelwert, Varianz und Kovarianz der Zwischenabgangszeiten

Zur Berechnung des Mittelwertes und der Varianz wird $\underline{x} \cdot \tilde{\mathbf{q}}(t) \cdot \underline{e}$ benötigt, für die Kovarianz $\text{Cov}[E_n, E_{n+1}]$ zusätzlich $\underline{x} \cdot \tilde{\mathbf{q}}(t_1) \tilde{\mathbf{q}}(t_2) \cdot \underline{e}$. Die Ausdrücke für AF und DF unterscheiden sich nur darin, daß bei DF die Summationen bis $N+1$ laufen, bei AF dagegen nur bis N . Der Mittelwert $E^* = \mathbb{E}[E_n]$ heißt in [86] "fundamental mean of the MEP", weil er im Grenzwertsatz Gl. (3.72) und damit in den im folgenden Hilfssatz 7 berechneten Grenzwerten auftritt. In den nachfolgenden Formeln bedeuten

$$\mathcal{M} = \begin{cases} N & \text{für AF} \\ N+1 & \text{für DF} \end{cases} \quad , \quad \bar{h}_1 := \sum_{t=1}^{\infty} t h(t) \quad \text{und} \quad \bar{h}_2 := \sum_{t=1}^{\infty} t^2 h(t).$$

$$E^* = \mathbb{E}[E_n] = \sum_{t=1}^{\infty} t \left[\underline{x}_0 \cdot \sum_{k=0}^{\infty} \mathbf{B}_k(t) \cdot \underline{e} + \left(\sum_{r=1}^{\mathcal{M}} \underline{x}_r \right) \cdot \sum_{k=0}^{\infty} \mathbf{A}_k(t) \cdot \underline{e} \right] = \underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-1} \underline{e} + \bar{h}_1 \quad (5.20)$$

$$\text{Var}[E_n] = \bar{h}_2 - (E^*)^2 + \underline{x}_0 \left[2\bar{h}_1 (\mathbf{I} - \mathbf{D}_0)^{-1} + 2(\mathbf{I} - \mathbf{D}_0)^{-2} - (\mathbf{I} - \mathbf{D}_0)^{-1} \right] \underline{e} \quad (5.21)$$

$$\begin{aligned} \text{Cov}[E_n, E_{n+1}] = & \left[\underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-2} \mathbf{D}_1 \mathbf{A}_0 + \right. \\ & \left. + \left(\underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_1 + \underline{x}_1 \right) \sum_{t=1}^{\infty} t \mathbf{A}_0(t) - E^* \underline{x}_0 \right] \cdot (\mathbf{I} - \mathbf{D}_0)^{-1} \underline{e} \end{aligned} \quad (5.22)$$

5.1.2.3 Höhere Momente der Zwischenabgangszeiten

$$M[E_n^\mu] = \sum_{t=1}^{\infty} t^\mu P\{E_n = t\} = \bar{h}_\mu + \underline{x}_0 \sum_{\xi=0}^{\mu-1} \bar{h}_\xi \binom{\mu}{\xi} \sum_{v=1}^{\infty} v^{\mu-\xi} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0) \underline{e}$$

mit $\bar{h}_\mu := \sum_{t=1}^{\infty} t^\mu h(t)$, $\mu \in \mathbb{N}$. $\sum_{v=1}^{\infty} v^\xi \mathbf{D}_0$ erhält man nach [13] aus

$$\sum_{k=1}^{\infty} k(k-1)(k-2) \cdots (k-n+1) \mathbf{D}_0^{k-1} = n! (\mathbf{I} - \mathbf{D}_0)^{-(n+1)} \mathbf{D}_0^{n-1}.$$

5.1.3 Die Belegung zu einem endlichen Zeitpunkt t nach einem Abgang zum Zeitpunkt 0

Um von (stationären) Abgangszeitpunkten auf (stationäre) Ankunftszeitpunkte transformieren zu können, benötigt man zunächst die Belegung zu einem beliebigen Zeitpunkt im stationären Fall, d.h. $t \rightarrow \infty$. Diese erhält man durch Grenzübergang aus der nachfolgend betrachteten Belegung zu $t < \infty$.

Wie am Beispiel des M/G/1-Wartesystems in [86] erläutert, wird der Ausdruck

$$y_{kr}^{ij}(t) := P\{\text{Belegung bei } t \text{ ist } r, S(t) = j \text{ unter der Bedingung eines Abgangs bei } 0, \text{ der } k \text{ Kunden im System hinterläßt, und } S(0) = i\}$$

ausgewertet. Die Belegung bei t beinhaltet schon Ankünfte oder Abgänge, die bei t stattfinden; t kann Ankunftszeitpunkt, Abgangszeitpunkt, Ankunfts- und Abgangszeitpunkt gleichzeitig oder keines davon sein.

Die nachfolgenden Ausdrücke für $y_{kr}^{ij}(t)$ erhält man, indem der **letzte** Abgangszeitpunkt u **vor** dem Beobachtungszeitpunkt t betrachtet **und** $u = t$ zugelassen wird. Der SM-Kern berücksichtigt Ankünfte, die gleichzeitig mit dem Abgang stattfinden, so daß für $t = u$ in $y_{kr}^{ij}(t)$ eine Ankunft gleichzeitig mit einem Abgang, wie auch ein alleiniger Abgang schon berücksichtigt sind. Daher braucht für $t > u$ ein Abgang zu t nicht betrachtet zu werden. In den Formeln (5.23), (5.24) und (5.25) ist das berücksichtigt durch die komplementären Verteilungsfunktionen $H^c(t)$ der Bedienzeit und Summation von u bis t . Ein nicht-leeres System bei t kann in zwei disjunkten Fällen zustandekommen, s. Abb. 5.4 und [86], pp. 24ff (dort für M/G/1):

- t fällt in die erste Bedienzeit einer Betriebsperiode, d.h. bei u ist das System leer verlassen worden (das impliziert $t > u$),
- t fällt in die zweite oder eine spätere Bedienzeit einer Betriebsperiode, d.h. bei u wird das System nicht leer verlassen.

Weiter wird die WKT benötigt, daß unter der Bedingung eines Abgangs bei 0, der k Kunden zurückläßt und bei dem der Eingabeprozess in Phase i weilt, ein Abgang bei u dann ℓ Kunden hinterläßt und der Eingabeprozess sich bei u in Phase j befindet. Diese WKT ist durch die der Markov-Erneuerungsfunktion (3.70) zugeordnete WF

$$\hat{r}_{k\ell}^{ij}(u) := \left[[\hat{\mathbf{r}}(u)]_{k\ell} \right]_{ij}, \quad \hat{r}_{k\ell}^{ij}(0) := \mathbf{1}_{k=\ell} \mathbf{1}_{i=j}$$

gegeben, die über

$$\hat{\mathbf{r}}(t) = \sum_{n=1}^t \tilde{\mathbf{q}}^{n*}(t)$$

mit dem SM-Kern $\tilde{\mathbf{q}}(t)$ verknüpft ist. Dabei bedeutet n^* die n -fache Matrixfaltung, s. Gl. (3.69). Unter Berücksichtigung der (erweiterten) Markoveigenschaft (3.29) für den DBMAP

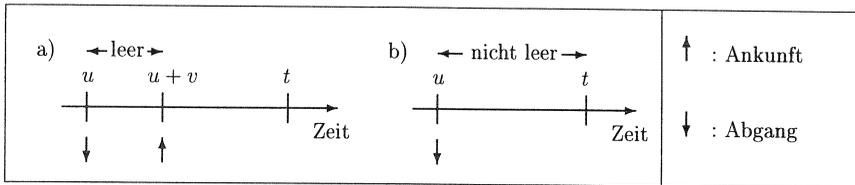


Abb. 5.4 Zwei disjunkte Fälle für eine Belegung $r > 0$ bei t .

erhält man nun die gesuchten Ausdrücke für $y_{kr}^{ij}(t)$.

Für $r = 0$:

$$y_{k0}^{ij}(t) = \sum_{j' \in \mathcal{E}} \sum_{u=0}^t \hat{r}_{k0}^{ij'}(u) [\mathbf{P}(0, t-u)]_{j'j}, \quad (5.23)$$

für $1 \leq r \leq N$:

$$\begin{aligned} y_{kr}^{ij}(t) &= \sum_{j' \in \mathcal{E}} \sum_{u=0}^{t-1} \hat{r}_{k0}^{ij'}(u) \sum_{v=1}^{t-u} \sum_{w=1}^r [\mathbf{D}_0^{v-1} \mathbf{D}_w \mathbf{P}(r-w, t-u-v) H^c(t-u-v)]_{j'j} + \\ &+ \sum_{j' \in \mathcal{E}} \sum_{u=0}^t \sum_{\ell=1}^r \hat{r}_{k\ell}^{ij'}(u) [\mathbf{P}(r-\ell, t-u) H^c(t-u)]_{j'j} \end{aligned} \quad (5.24)$$

und für $r = N+1$:

$$\begin{aligned} y_{k(N+1)}^{ij}(t) &= \sum_{j' \in \mathcal{E}} \sum_{u=0}^{t-1} \hat{r}_{k0}^{ij'}(u) \sum_{v=1}^{t-u} \sum_{w=1}^{b_{max}} \sum_{r'=N+1-w}^{\infty} [\mathbf{D}_0^{v-1} \mathbf{D}_w \mathbf{P}(r', t-u-v) H^c(t-u-v)]_{j'j} + \\ &+ \sum_{j' \in \mathcal{E}} \sum_{u=0}^t \sum_{\ell=1}^M \hat{r}_{k\ell}^{ij'}(u) \sum_{r'=N+1-\ell}^{\infty} [\mathbf{P}(r', t-u) H^c(t-u)]_{j'j}. \end{aligned} \quad (5.25)$$

5.1.4 Die Belegung bei beliebigem Zeitpunkt $t \rightarrow \infty$ im stationären Fall

Für die Bildung des Grenzwertes $t \rightarrow \infty$ wird der Grenzwertsatz für zeitdiskrete Markov-Erneuerungsprozesse Gl. (3.72) benötigt. Die Grenzwerte hängen nicht mehr von k und i ab, daher wird $[\underline{y}_r]_j := \lim_{t \rightarrow \infty} y_{kr}^{ij}(t)$ definiert.

Für die Konvergenz der Reihe $\sum_{k=0}^{\infty} \mathbf{D}_0^k = (\mathbf{I} - \mathbf{D}_0)^{-1}$ vgl. [41].

Hilfssatz 7 Sei $H^c(t) := 1 - H(t)$ die komplementäre Verteilungsfunktion der Bedienzeit, $\mathcal{M} = N$ für AF und $\mathcal{M} = N+1$ für DF. Die stationäre WKT $[\underline{y}_r]_i$, daß zu einem beliebigen⁶ Zeitpunkt r Kunden im System sind und der Eingabeprozess sich in Phase i befindet, ist folgendermaßen mit dem Vektor $\underline{x} = (\underline{x}_0, \underline{x}_1, \dots, \underline{x}_{\mathcal{M}})$ der stationären Belegungs-WKTn zu Abgangszeitpunkten verknüpft:

$$\underline{y}_0 = \frac{1}{E^*} \underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-1}, \quad (5.26)$$

⁶Beliebig heißt: Ankunftszeitpunkt, Abgangszeitpunkt, Ankunfts- und Abgangszeitpunkt oder keines davon.

für $1 \leq r \leq N$

$$\underline{y}_r = \frac{1}{E^*} \sum_{\ell=1}^r \left[(\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell) \sum_{\sigma=0}^{\infty} \mathbf{P}(r - \ell, \sigma) H^c(\sigma) \right], \quad (5.27)$$

$$\begin{aligned} \underline{y}_{N+1} &= \frac{1}{E^*} \sum_{\ell=1}^M (\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell) \sum_{r'=N+1-\ell}^{\infty} \sum_{\sigma=0}^{\infty} \mathbf{P}(r', \sigma) H^c(\sigma) + \\ &+ \frac{1}{E^*} \underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \sum_{w=\mathcal{M}+1}^{\infty} \mathbf{D}_w \sum_{r'=N+1-w}^{\infty} \sum_{\sigma=0}^{\infty} \mathbf{P}(r', \sigma) H^c(\sigma). \end{aligned} \quad (5.28)$$

Beweis: Durch Anwenden von Gl. (3.72) auf die Gln. (5.23), (5.24), (5.25). Z.B. ist für $r = 0$ in der Notation des Grenzwertsatzes Gl. (3.72) $g_j(k) = [\mathbf{P}(0, k)]_{j'j}$ für jedes j . \square

Bemerkung: Für *transiente* Ausdrücke könnte man von einem leeren System bei 0 ausgehen und von da die Belegungs-WKTn zum Zeitpunkt $t < \infty$ formulieren. Da dann kein Grenzübergang $t \rightarrow \infty$ erfolgt, verschwindet der für jedes u sehr komplexe Ausdruck $\hat{r}_{0\ell}(u)$ nicht mehr. Kompakte geschlossene Ausdrücke für den transienten Fall scheinen daher nicht möglich.

Mit den folgenden Identitäten

$$\sum_{k=1}^r \mathbf{P}(k - \ell, \sigma) \mathbf{D}_{r+1-k} = \sum_{k=1}^{r+1-\ell} \mathbf{P}(k - 1, \sigma) \mathbf{D}_{r+2-\ell-k} \quad \text{für } r \geq \ell, \quad (5.29)$$

$$\mathbf{P}(r, \sigma)(\mathbf{I} - \mathbf{D}_0) - \sum_{r'=1}^r \mathbf{P}(r' - 1, \sigma) \mathbf{D}_{r+1-r'} = \mathbf{P}(r, \sigma) - \mathbf{P}(r, \sigma + 1) \quad \text{für } r \geq 1, \quad (5.30)$$

$$\sum_{\sigma=0}^{\infty} [\mathbf{P}(r, \sigma) - \mathbf{P}(r, \sigma + 1)] H^c(\sigma) = \mathbf{1}_{r=0} \cdot \mathbf{I} - \mathbf{A}_r \quad \text{für } r \geq 0 \quad (5.31)$$

erhält man einen einfachen Zusammenhang zwischen \underline{y}_{r+1} und $\underline{y}_0, \dots, \underline{y}_r$ für $0 \leq r \leq N-1$, der demjenigen im kontinuierlichen Fall ähnelt ([75], dort allerdings für $N = \infty$).

Hilfssatz 8 Für $0 \leq r \leq N-1$ gilt

$$\underline{y}_{r+1} = \left[\frac{1}{E^*} (\underline{x}_{r+1} - \underline{x}_r) + \sum_{k=0}^r \underline{y}_k \mathbf{D}_{r+1-k} \right] (\mathbf{I} - \mathbf{D}_0)^{-1}. \quad (5.32)$$

Zusammen mit

$$\underline{y}_{N+1} = \underline{\pi} - \sum_{r=0}^N \underline{y}_r \quad (5.33)$$

können die Vektoren $\underline{y}_r, 0 \leq r \leq N+1$ rekursiv mittels $\underline{x}_r, \underline{y}_r, 0 \leq r' \leq r-1$ und dem stationären Phasenvektor $\underline{\pi}$ des DBMAP berechnet werden.

$$\begin{aligned}
\text{Beweis: } E^* \cdot \left[\underline{y}_{r+1}(\mathbf{I} - \mathbf{D}_0) - \sum_{k=0}^r \underline{y}_k \mathbf{D}_{r+1-k} \right] &= \\
= \sum_{\ell=1}^{r+1} \left[\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell \right] \sum_{\sigma=0}^{\infty} \mathbf{P}(r+1-\ell, \sigma) H^c(\sigma) (\mathbf{I} - \mathbf{D}_0) \\
- \sum_{k=1}^r \sum_{\ell=1}^k \left[\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell \right] \sum_{\sigma=0}^{\infty} \mathbf{P}(k-\ell, \sigma) H^c(\sigma) \mathbf{D}_{r+1-k} - \underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_{r+1} &=
\end{aligned}$$

(unter Berücksichtigung der Tatsache, daß die zweite Summation über ℓ bis $r+1$ ausgedehnt werden kann, und der Gln. (5.29), (5.30) und (5.31))

$$= \sum_{\ell=1}^{r+1} \left[\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell \right] \cdot [\mathbf{1}_{\ell=r+1} \mathbf{I} - \mathbf{A}_{r+1-\ell}] - \underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_{r+1} =$$

(und mit (5.11) für AF und (5.13) für DF)

$$= - \sum_{\ell=0}^r \underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_{\ell+1} \mathbf{A}_{r-\ell} - \sum_{\ell=1}^{r+1} \underline{x}_\ell \mathbf{A}_{r-\ell+1} + \underline{x}_{r+1} = \underline{x}_{r+1} - \underline{x}_r .$$

Summation über alle möglichen Belegungswerte zu beliebigen Zeitpunkten liefert den stationären Phasenvektor, d.h. Gl. (5.33). \square

5.1.5 Belegung zu Ankunftszeitpunkten

Unter der Belegung zu einem Ankunftszeitpunkt wird hier die Belegung verstanden, die eine Ankunft "sieht", d.h. ohne daß sie selbst mitberücksichtigt wird. Dieser Belegungswert entscheidet über die Zulassung oder Ablehnung der Ankunft. Als eine Ankunft wird hier ein w -Batch betrachtet. Bei Überlagerung von zwei Strömen wie in Kapitel 6 werden andere Ankunftsereignisse notwendig, etwa die Ereignisse, daß ein Strom allein oder beide zusammen eine Ankunft erzeugen.

Die Ausdrücke für die Belegung zu Ankunftszeitpunkten unterscheiden sich für AF und DF strukturell. Bei AF sind sie in einfacher Weise mit den Belegungsvektoren \underline{y}_r zu einem beliebigen Zeitpunkt verknüpft, während bei DF der gleichzeitige Abgang getrennt berücksichtigt werden muß.

Satz 2 Seien \underline{y}_r , $r = 0, 1, \dots, N+1$ die Vektoren der stationären Belegungs-WKTN zu beliebigem Zeitpunkt, die sich für AF und DF unterscheiden. Sei $[z_r(w)]_i := P\{\text{ein } w\text{-Batch kommt an und sieht } r \text{ Kunden im System, und der Eingabeprozess ist in Phase } i\}$. Für AF gilt:

$$z_r(w) = \underline{y}_r \mathbf{D}_w, \quad r = 0, 1, \dots, N+1 .$$

Für DF gilt:

$$z_r(w) = \begin{cases} \left[\underline{y}_0 + \underline{\Delta}_0 \right] \mathbf{D}_w & r = 0 \\ \left[\underline{y}_r - \underline{\Delta}_{r-1} + \underline{\Delta}_r \right] \mathbf{D}_w & 1 \leq r \leq N \\ \left[\underline{y}_{N+1} - \underline{\Delta}_N \right] \mathbf{D}_w & r = N+1 \end{cases} .$$

Dabei ist $[\underline{\Delta}_r \mathbf{D}_w]_i := P\{\text{ein } w\text{-Batch kommt an und sieht } r \text{ Kunden im System, gleichzeitig findet ein Abgang statt, und der Eingabeprozess ist in Phase } i\}$, wie im nachfolgenden Satz 3 in Abschnitt 5.1.6 für AF und DF gegeben.

Beweis: $[y_r \mathbf{D}_w]_i$ gibt die WKT an, daß zu einem beliebigen stationären Zeitpunkt r Kunden im System sind und *eine Zeiteinheit später* ein w -Batch ankommt und dann die Phase des Eingabeprozesses i ist. Zu diesem Ankunftszeitpunkt kann ein Abgang gleichzeitig stattfinden, muß aber nicht.

Bei AF macht es für die Belegung, die der w -Batch sieht, keinen Unterschied, ob ein Abgang stattfindet oder nicht, da er immer erst unmittelbar nach der Ankunft geschieht. Der Batch sieht also die Belegung r .

Bei DF sieht der w -Batch die Belegung r , wenn kein Abgang stattfindet, sonst $r - 1$, sofern $r > 0$. Daher ergibt $[y_r \mathbf{D}_w]_i - [\Delta_{r-1} \mathbf{D}_w]_i$ für $1 \leq r \leq N$ die WKT, daß der w -Batch r Kunden sieht und kein Abgang stattfindet. Addition von $[\Delta_r \mathbf{D}_w]_i$, d.h. gesehene Belegung r und gleichzeitig ein Abgang, liefert dann die Behauptung. $[y_0 \mathbf{D}_w]_i$ bezeichnet die WKT, daß der w -Batch ein leeres System antrifft, ohne daß ein Abgang gleichzeitig stattfinden kann, da das System spätestens eine Zeiteinheit vorher leer geworden ist. Da bei DF der w -Batch auch dann ein leeres System sieht, wenn gleichzeitig mit seiner Ankunft ein Abgang das System leert, benötigt man zusätzlich den Summanden $[\Delta_0 \mathbf{D}_w]_i$. Weiterhin kann bei DF eine Ankunft kein volles System sehen und gleichzeitig ein Abgang stattfinden. Daher fehlt für $r = N + 1$ ein Summand $[\Delta_{N+1} \mathbf{D}_w]_i$; diese WKT ist Null. \square

Ein formaler Beweis, der die Ausdrücke (5.23), (5.24), (5.25) für endliche *Ankunftszeitpunkte* t_A formuliert und dann den Grenzwertsatz (3.72) anwendet, findet sich im Anhang A.2.

5.1.5.1 WF der Belegung zu einem Ankunftszeitpunkt

Für Ausdrücke wie $P\{\text{eine Ankunft sieht } r \text{ Kunden im System}\}$ müssen die $z_r(w)$, die ihrer Natur nach Verbund-WKTN sind, auf Ankunftsereignisse eingeschränkt werden. Mit $P\{\text{eine Ankunft findet statt}\} = 1 - \pi \mathbf{D}_0 \underline{e}$ folgt dann:

$$P\{\text{eine Ankunft sieht } r \text{ Kunden im System}\} = \frac{1}{1 - \pi \mathbf{D}_0 \underline{e}} \sum_{w=1}^{b_{\max}} z_r(w) \underline{e}$$

und

$$P\{\text{ein } w\text{-Batch sieht } r \text{ Kunden im System}\} = \frac{1}{1 - \pi \mathbf{D}_0 \underline{e}} z_r(w) \underline{e} .$$

5.1.6 Gleichzeitigkeit von Ankünften und Abgängen

Gleichzeitigkeit von Ankünften und Abgängen kann durch Modifizierung der Gln. (5.23), (5.24) und (5.25) erfaßt werden. Dies wird hier wieder für einen ankommenden w -Batch formuliert. Die wesentlichen Änderungen in den Gln. (5.23), (5.24) und (5.25) sind das Ausschließen von $t_A = u$ für den hier betrachteten *Ankunftszeitpunkt* t_A , indem die Summation über u nur bis $t_A - 1$ läuft, und das Ersetzen von $H^c(t)$ durch die WF $h(t)$, wodurch sichergestellt ist, daß zu t_A auch ein Abgang stattfindet. Der ankommende w -Batch wird repräsentiert durch Ersetzen von $\mathbf{P}(r, t)$ durch $\mathbf{P}(r, t - 1) \mathbf{D}_w$.

Es bezeichne

${}^w \Delta_{k,r}^{i,j}(t_A) := P\{\text{ein } w\text{-Batch kommt zu } t_A \text{ an, sieht } r \text{ Kunden im System, gleichzeitig findet ein Abgang statt, und der Ankunftsprozeß ist in Phase } j \text{ unter der Bedingung, daß ein Abgang bei } 0 \text{ das System mit } k \text{ Kunden hinterläßt und der Eingabeprozess sich in Phase } i \text{ befindet}\}.$

5.1.6.1 Ausdrücke für AF

$r = 0$:

$$w\Delta_{k,r}^{i,j}(t_A) = 0 .$$

$1 \leq r \leq N$:

$$w\Delta_{k,r}^{i,j}(t_A) = \sum_{j' \in \mathcal{E}} \left[\sum_{u=0}^{t_A-2} \hat{r}_{k,0}^{i,j'}(u) \sum_{v=1}^{t_A-u-1} \sum_{w'=1}^r \left[\mathbf{D}_0^{v-1} \mathbf{D}_{w'} \mathbf{P}(r-w', t_A-u-v-1) \cdot \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u-v) + \sum_{u=0}^{t_A-1} \sum_{\ell=1}^r \hat{r}_{k,\ell}^{i,j'}(u) \left[\mathbf{P}(r-\ell, t_A-u-1) \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u) \right] ,$$

$r = N + 1$

$$w\Delta_{k,r}^{i,j}(t_A) = \sum_{j' \in \mathcal{E}} \sum_{r'=N+1}^{\infty} \left[\sum_{u=0}^{t_A-2} \hat{r}_{k,0}^{i,j'}(u) \sum_{v=1}^{t_A-u-1} \sum_{w'=1}^{r'} \left[\mathbf{D}_0^{v-1} \mathbf{D}_{w'} \mathbf{P}(r'-w', t_A-u-v-1) \cdot \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u-v) + \sum_{u=0}^{t_A-1} \sum_{\ell=1}^{r'} \hat{r}_{k,\ell}^{i,j'}(u) \left[\mathbf{P}(r'-\ell, t_A-u-1) \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u) \right] .$$

5.1.6.2 Ausdrücke für DF

$0 \leq r \leq N - 1$:

$$w\Delta_{k,r}^{i,j}(t_A) = \sum_{j' \in \mathcal{E}} \left[\sum_{u=0}^{t_A-2} \hat{r}_{k,0}^{i,j'}(u) \sum_{v=1}^{t_A-u-1} \sum_{w'=1}^{r+1} \left[\mathbf{D}_0^{v-1} \mathbf{D}_{w'} \mathbf{P}(r+1-w', t_A-u-v-1) \cdot \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u-v) + \sum_{u=0}^{t_A-1} \sum_{\ell=1}^{r+1} \hat{r}_{k,\ell}^{i,j'}(u) \left[\mathbf{P}(r+1-\ell, t_A-u-1) \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u) \right] ,$$

$r = N$:

$$w\Delta_{k,r}^{i,j}(t_A) = \sum_{j' \in \mathcal{E}} \sum_{r'=N}^{\infty} \left[\sum_{u=0}^{t_A-2} \hat{r}_{k,0}^{i,j'}(u) \sum_{v=1}^{t_A-u-1} \sum_{w'=1}^{r'+1} \left[\mathbf{D}_0^{v-1} \mathbf{D}_{w'} \mathbf{P}(r'+1-w', t_A-u-v-1) \cdot \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u-v) + \sum_{u=0}^{t_A-1} \sum_{\ell=1}^{r'+1} \hat{r}_{k,\ell}^{i,j'}(u) \left[\mathbf{P}(r'+1-\ell, t_A-u-1) \mathbf{D}_w \right]_{j',j} \cdot h(t_A-u) \right] .$$

$r = N + 1 :$

$$w\Delta_{k,r}^{i,j}(t_A) = 0 .$$

Es ist zu beachten, daß $\hat{r}_{k,\ell}^{i,j'}(u) = 0$ für $\ell > \mathcal{M}$ und $\mathbf{D}_{w'} = \mathbf{0}$ für $w' > b_{max}$.

Mit der Definition $[\underline{\Delta}_r, \mathbf{D}_w]_j := \lim_{t_A \rightarrow \infty} w\Delta_{k,r}^{i,j}(t_A)$ liefert der Grenzübergang nach Gl. (3.72) für $t_A \rightarrow \infty$ den

Satz 3 Sei $[\underline{\Delta}_r, \mathbf{D}_w]_i := P\{\text{ein } w\text{-Batch kommt an und sieht } r \text{ Kunden im System, gleichzeitig findet ein Abgang statt, und der Eingabeprozess befindet sich in Phase } i\}$. Die Vektoren $\underline{\Delta}_r$ lauten wie folgt:

Für AF:

$$\underline{\Delta}_r = \begin{cases} \underline{0} & \text{für } r = 0 \\ \frac{1}{E^*} \sum_{\ell=1}^r [\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell] \sum_{\sigma=0}^{\infty} \mathbf{P}(r - \ell, \sigma) h(\sigma + 1) & \text{für } 1 \leq r \leq N \\ \frac{1}{E^*} \sum_{r'=N+1}^{\infty} \left[\sum_{\ell=1}^{r'} [\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell] \sum_{\sigma=0}^{\infty} \mathbf{P}(r' - \ell, \sigma) h(\sigma + 1) \right] & \text{für } r = N + 1 . \end{cases} \quad (5.34)$$

Für DF:

$$\underline{\Delta}_r = \begin{cases} \frac{1}{E^*} \sum_{\ell=1}^{r+1} [\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell] \sum_{\sigma=0}^{\infty} \mathbf{P}(r + 1 - \ell, \sigma) h(\sigma + 1) & \text{für } 0 \leq r \leq N - 1 \\ \frac{1}{E^*} \sum_{r'=N}^{\infty} \left[\sum_{\ell=1}^{r'+1} [\underline{x}_0(\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell] \sum_{\sigma=0}^{\infty} \mathbf{P}(r' + 1 - \ell, \sigma) h(\sigma + 1) \right] & \text{für } r = N \\ \underline{0} & \text{für } r = N + 1 \end{cases} \quad (5.35)$$

Ein rekursiver Zusammenhang wie für die \underline{y}_r gemäß Hilfssatz 8 scheint nicht zu existieren, da man für seine Berechnung die Eigenschaft $H^c(t + 1) = H^c(t) - h(t + 1)$ benutzt. Eine kompaktere Formulierung wird durch die Definition der Matrizen \mathbf{A}_r''' und \mathbf{B}_r''' mit

$$\mathbf{A}_r''' := \sum_{\sigma=0}^{\infty} \mathbf{P}(r, \sigma) h(\sigma + 1) , \quad (5.36)$$

$$\mathbf{B}_r''' := (\mathbf{I} - \mathbf{D}_0)^{-1} \sum_{\ell=1}^{r+1} \mathbf{D}_\ell \mathbf{A}_{r+1-\ell}''' \quad (5.37)$$

ermöglicht. Diese hängen mit \mathbf{A}'_r bzw. \mathbf{B}'_r (vgl. Gln. (5.4), (5.6) und (5.8)) vermittels $\mathbf{A}'_r = \mathbf{A}_r''' \mathbf{D}_0$ und $\mathbf{B}'_r = \mathbf{B}_r''' \mathbf{D}_0$ zusammen. Da zur Erzeugung der eingebetteten MK des MEP für DF \mathbf{A}'_r bzw. \mathbf{B}'_r ohnehin berechnet werden müssen, ist die Umformulierung durch die Gln. (5.36) und (5.37) schon von erheblichem Nutzen für eine Programm-Implementierung. Man hat somit

für AF:

$$\Delta_r = \begin{cases} 0 & \text{für } r = 0 \\ \frac{1}{E^*} \left(x_0 \mathbf{B}_{r-1}''' + \sum_{\ell=1}^r x_\ell \mathbf{A}_{r-\ell}''' \right) & \text{für } 1 \leq r \leq N \\ \frac{1}{E^*} \left(x_0 \sum_{r'=N}^{\infty} \mathbf{B}_{r'}''' + \sum_{\ell=1}^N x_\ell \sum_{r'=N+1-\ell}^{\infty} \mathbf{A}_{r'}''' \right) & \text{für } r = N + 1, \end{cases} \quad (5.38)$$

für DF:

$$\Delta_r = \begin{cases} \frac{1}{E^*} \left(x_0 \mathbf{B}_r''' + \sum_{\ell=1}^{r+1} x_\ell \mathbf{A}_{r+1-\ell}''' \right) & \text{für } 0 \leq r \leq N - 1 \\ \frac{1}{E^*} \left(x_0 \sum_{r'=N}^{\infty} \mathbf{B}_{r'}''' + \sum_{\ell=1}^{N+1} x_\ell \sum_{r'=N+1-\ell}^{\infty} \mathbf{A}_{r'}''' \right) & \text{für } r = N \\ 0 & \text{für } r = N + 1. \end{cases} \quad (5.39)$$

Schließlich bleibt noch die WKT anzugeben, daß eine Ankunft bei gleichzeitigem Abgang r Kunden im System sieht. Diese erfordert wieder die Einschränkung der obigen Ausdrücke auf Ankunftszeitpunkte.

$$P\{\text{w-Batch sieht bei gleichzeitigem Abgang } r \text{ Kunden im System}\} = \frac{1}{1 - \pi \mathbf{D}_0 \underline{e}} \Delta_r \mathbf{D}_w \underline{e}$$

und

$$P\{\text{Ankunft und Abgang finden gleichzeitig statt}\} = \frac{1}{1 - \pi \mathbf{D}_0 \underline{e}} \sum_{r=0}^{N+1} \Delta_r (\widehat{\mathbf{D}} - \mathbf{D}_0) \underline{e}.$$

5.1.7 Die Wartezeit

Die WF der Wartezeit ist bisher nicht zugänglich. Zwar kann der Ansatz in [6] für das (zeitkontinuierliche) N/G/1/N-System übertragen werden und eine z -Transformierte der WF (gegenüber der Laplace-Transformierten der Dichtefunktion in [6]) unter der Annahme formuliert werden, daß eine Reihenentwicklung der inversen Matrix $(\mathbf{I}z_1 - \sum_{w=0}^{b_{\max}} \mathbf{D}_w z_2^w)^{-1}$ nach Potenzen in z_2 gegeben ist. Aber es fehlen eben die Koeffizienten dieser Reihenentwicklung für eine vollständige Beschreibung. Für den Fall des DMAP/G/1/N-Systems dürften sie mit dem in [7] verwendeten "Trick" jedoch zu gewinnen sein.

5.2 Andere Verfahren: Das DBMAP/D/1/N-Wartesystem für eine der Zeiteinheit gleiche Bedienzeit

Unter der Voraussetzung, daß die Bedienzeit gleich der Zeiteinheit ist, findet man für das Wartesystem mit DBMAP-Eingabeprozess eine MK von Belegung und Phase am Ende einer Zeiteinheit ("slot"). Für DF wurde dies in [10, 9] behandelt, für AF in [112]. Wegen $D = 1$ ergibt sich die Wartezeit einfach aus der Belegung zu Ankunftszeitpunkten.

5.2.1 Blondia, Casals (1992): Departure First

Ankünfte und Abgänge finden am Ende eines Slot statt, und eine abgehende Zelle verläßt das System, unmittelbar bevor eine ankommende Zelle das System betritt, s. Abb. 5.5. Bezeichne $L^{sl}(t)$ die Belegung *unmittelbar nach* dem Ende des t -ten Slot und $S(t)$ die Phase des DBMAP, so ist $(L^{sl}(t), S(t))$ eine MK mit der folgenden Übergangsmatrix (wiederum vom M/G/1-Typ)

$$\begin{pmatrix} \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \mathbf{D}_{N-1} & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \mathbf{D}_{N-1} & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \mathbf{0} & \mathbf{D}_0 & \cdots & \mathbf{D}_{N-2} & \mathbf{D}_{N-1} & \sum_{w=N}^{\infty} \mathbf{D}_w \\ \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{D}_0 & \mathbf{D}_1 & \sum_{w=2}^{\infty} \mathbf{D}_w \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{D}_0 & \sum_{w=1}^{\infty} \mathbf{D}_w \end{pmatrix}. \tag{5.40}$$

Ihre stationäre Lösung $[x_k]_i, k = 0, \dots, N + 1, i = 1, \dots, m$ bezeichnet die WKT, daß *unmittelbar nach* dem Ende eines Slot (zum Beobachtungszeitpunkt in Abb. 5.5, d.h. es sind mögliche Ankünfte und Abgänge am Ende des Slot berücksichtigt) k Zellen im System sind und der Eingabeprozess in Phase i ist. Eine Ankunft am Ende des nächsten Slot sieht daher $\max(k - 1, 0)$ Zellen im System. Wegen DF sieht eine Ankunft niemals ein volles System. Damit ist Verlust aber nicht ausgeschlossen, weil Mehrfachankünfte stattfinden können.

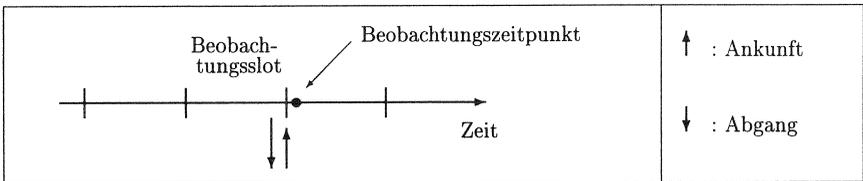


Abb. 5.5 Slotstruktur für "Departure First".

Belegung zu Ankunftszeitpunkten

$P\{\text{ein } v\text{-Batch kommt an und sieht eine Belegung der Größe } \mu\} =$

$$= \begin{cases} x_{\mu+1} \mathbf{D}_v \mathbf{e} & \text{für } 1 \leq \mu \leq N \\ (x_1 + x_0) \mathbf{D}_v \mathbf{e} & \text{für } \mu = 0 \end{cases}. \tag{5.41}$$

Die Einschränkung auf Ankunftszeitpunkte erfolgt wieder durch Division durch $1 - \pi \mathbf{D}_0 \mathbf{e}$.

5.2.2 Takine, Suda, Hasegawa (1993): Arrival First

Nun finden Ankünfte und Abgänge am Ende eines Slot statt, und der Abgang geschieht unmittelbar nach der Ankunft, s. Abb. 5.6. Daher kann eine Ankunft ein volles System sehen, und zwar dann, wenn das System einen Slot vorher voll geworden ist. Letzteres ist aber nur möglich, wenn ein v -Batch mit $v \geq N + 1$ auf ein *leeres* System trifft. In allen

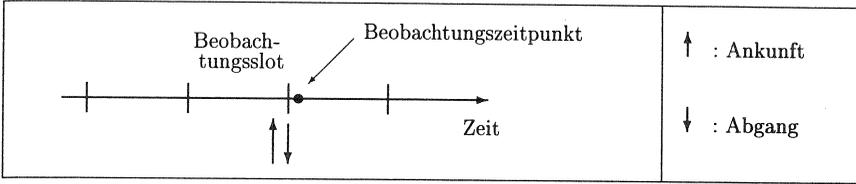


Abb. 5.6 Slotstruktur für "Arrival First".

anderen Fällen sind unmittelbar nach dem Ende eines Slot höchstens N Zellen im System. Man erhält für die MK folgende Übergangsmatrix

$$\begin{pmatrix}
 D_0 & D_1 & \cdots & D_{N-1} & D_N & \sum_{w=N+1}^{\infty} D_w \\
 D_0 & D_1 & \cdots & D_{N-1} & \sum_{w=N}^{\infty} D_w & 0 \\
 0 & D_0 & \cdots & D_{N-2} & \sum_{w=N-1}^{\infty} D_w & 0 \\
 \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\
 \vdots & \ddots & \ddots & D_0 & \sum_{w=1}^{\infty} D_w & 0 \\
 0 & \cdots & \cdots & 0 & \sum_{w=0}^{\infty} D_w & 0
 \end{pmatrix} \quad (5.42)$$

Belegung bei Ankunftszeitpunkten

$$\begin{aligned}
 &P\{\text{ein } v\text{-Batch kommt an und sieht eine Belegung der Größe } \mu\} = \\
 &= \underline{x}_\mu D_v \underline{e} \quad \text{für } 0 \leq \mu \leq N + 1 \quad (5.43)
 \end{aligned}$$

Die Einschränkung auf Ankunftszeitpunkte erfolgt durch Division durch $1 - \underline{x} D_0 \underline{e}$.

5.2.3 Vergleich mit dem DBMAP/G/1/N-System bei $D = 1$

Die Übergangsmatrizen der zu Abgangszeiten eingebetteten MKn, die sich für $D = 1$ aus Abb. 5.2 bzw. 5.3 ergeben, unterscheiden sich von den obigen, weil die Ausdrücke (5.40) und (5.42) sich auf das Ende eines Slot beziehen, bei dem nicht unbedingt ein Abgang stattfinden muß.

5.3 Systeme mit Bedienzeiten, die ganzzahlige Bruchteile der Zeiteinheit sind

Eine einfache, in der Literatur bisher offenbar noch nicht diskutierte Erweiterung der Systeme des letzten Abschnitts stellt die Vergrößerung der Abgangsrate auf den f -fachen Wert dar. Die Zwischenankunftszeiten sind weiterhin Vielfache der Zeiteinheit. f -fache Rate bedeutet, daß eine Belegung von μ unmittelbar nach dem Ende eines Slot auf den Wert $\max(\mu - f, 0)$ am Ende des nächsten Slot führt, wenn keine Ankünfte stattfinden. Ferner wird definiert, daß die innerhalb eines Slot stattfindenden Mehrfachabgänge nacheinander über den Slot gleichmäßig verteilt erfolgen. Diese Definition ist allerdings nur für AF zur Vermeidung von Mehrdeutigkeiten erforderlich. Das hat zur Folge, daß das System, wenn es zu einem Beobachtungszeitpunkt eine Zelle enthielt, schon nach $1/f$ -tel der folgenden Slotlänge wieder leer ist. Für DF erhält man wie bei $f = 1$ eine $(N + 2) \times (N + 2)$ -Matrix,

bei der die ersten $f + 1$ Hauptzeilen übereinstimmen (daher liegt keine M/G/1-Typ-Matrix mehr vor)

$$\begin{pmatrix} \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \cdots & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \vdots & \vdots & & & \vdots & \vdots \\ \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \cdots & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \cdots & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \mathbf{0} & \mathbf{D}_0 & \cdots & \cdots & \mathbf{D}_{N-1} & \sum_{w=N}^{\infty} \mathbf{D}_w \\ \vdots & \ddots & \ddots & & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \mathbf{D}_{f-1} & \sum_{w=f}^{\infty} \mathbf{D}_w \end{pmatrix}. \quad (5.44)$$

Belegung bei Ankunftszeitpunkten

$$\begin{aligned} P\{\text{ein } v\text{-Batch kommt an und sieht eine Belegung der Größe } \mu\} = \\ = \begin{cases} x_{\mu+f} \mathbf{D}_v \underline{\ell} & \text{für } 1 \leq \mu \leq N + 1 - f \\ (\sum_{k=0}^f x_k) \mathbf{D}_v \underline{\ell} & \text{für } \mu = 0 \end{cases}. \end{aligned} \quad (5.45)$$

Wegen der Festlegung, daß die innerhalb eines Slot stattfindenden "Mehrfachabgänge" nacheinander über den Slot gleichmäßig verteilt erfolgen, wirkt sich die AF-Zulassungsregel nur aus, wenn zu einem Beobachtungszeitpunkt mindestens f Zellen im System sind. Dann findet am Ende des folgenden Slot ein Abgang statt, und zwar AF-definitionsgemäß unmittelbar nach einer möglichen Ankunft zu diesem Zeitpunkt. In der resultierenden $(N + 2) \times (N + 2)$ -Übergangsmatrix stimmen die ersten f Hauptzeilen überein:

$$\begin{pmatrix} \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \cdots & \mathbf{D}_{N-1} & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \vdots & \vdots & & & \vdots & \vdots & \vdots \\ \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \cdots & \mathbf{D}_{N-1} & \mathbf{D}_N & \sum_{w=N+1}^{\infty} \mathbf{D}_w \\ \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \cdots & \mathbf{D}_{N-1} & \sum_{w=N}^{\infty} \mathbf{D}_w & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_0 & \cdots & \cdots & \mathbf{D}_{N-2} & \sum_{w=N-1}^{\infty} \mathbf{D}_w & \mathbf{0} \\ \vdots & \ddots & \ddots & & \vdots & \vdots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{D}_0 & \cdots & \mathbf{D}_{f-2} & \sum_{w=f-1}^{\infty} \mathbf{D}_w & \mathbf{0} \end{pmatrix}. \quad (5.46)$$

Belegung zu Ankunftszeitpunkten

$$\begin{aligned} P\{\text{ein } v\text{-Batch kommt an und sieht eine Belegung der Größe } \mu\} = \\ = \begin{cases} x_{\mu+f-1} \mathbf{D}_v \underline{\ell} & \text{für } 1 \leq \mu \leq N + 2 - f \\ (\sum_{k=0}^{f-1} x_k) \mathbf{D}_v \underline{\ell} & \text{für } \mu = 0 \end{cases}. \end{aligned} \quad (5.47)$$

5.3.1 Markov-Kette zu Ankunftszeitpunkten

Die o.g. MKn beziehen sich auf das Slotende mit Berücksichtigung aller Ankünfte und Abgänge. Dagegen wird in [57] für dasselbe System eine MK zu *Ankunftszeitpunkten* (ohne Berücksichtigung der Ankünfte) angegeben. Diese hat gegenüber derjenigen zu Abgangszeitpunkten eine um f Zeilen und Spalten kleinere Übergangsmatrix, ist aber ebenso nicht mehr vom M/G/1-Typ.

6 Wartemodelle für die Verbindungsannahmesteuerung (CAC)

Die Wartemodelle dieses Kapitels unterscheiden sich von den bisher behandelten darin, daß der Eingabeprozess eine Überlagerung $G_1^{[X]}+G_2$ zweier (korrelierter) Ströme darstellt. Beide Ströme sind innerhalb der Überlagerung identifizierbar. Dadurch wird es möglich, die Dienstgüteparameter *Verlustwahrscheinlichkeit* und *Momente der Wartezeit* für jeden Strom getrennt zu berechnen. Wie in Kapitel 2 erläutert, repräsentiert ein solches Wartesystem die Situation bei der Verbindungsannahme auf einem einzelnen Teilabschnitt eines Weges durch das Netz, auf dem eine neue Verbindung zugelassen werden soll. Wiederum mit Blick auf Videoquellen mit der geforderten sehr niedrigen Verlust-WKT von 10^{-9} ist ein solches Modell von Bedeutung: In Computersimulationen sind so geringe Verlust-WKTn sehr schwierig und nur unter großem Zeitaufwand zu verifizieren.

Strom $G_1^{[X]}$ repräsentiert den auf dem betrachteten Teilabschnitt schon vorhandenen Verkehr, wobei der SP X der Batches der Möglichkeit von Mehrfachankünften Rechnung trägt; denn der schon existierende Verkehr besteht aus einer Überlagerung von Zellenströmen verschiedener Eingangsleitungen, die auf dieselbe Ausgangsleitung weitervermittelt werden. $G_1^{[X]}$ ist durch Messungen auf jedem Teilabschnitt zugänglich. G_2 bildet den Zellenstrom der neuen Verbindung nach, wie sie sich am Eingang des Netzes darstellt. Da die Zellenströme der durch $G_1^{[X]}$ repräsentierten Überlagerung aus verschiedenen Richtungen kommen, ist die Annahme gerechtfertigt, daß sie stochastisch unabhängig sind. Dann sollte die WF der Batchgröße monoton fallen mit einem absoluten Maximum beim Wert 1.

Die Berücksichtigung von Einflüssen des existierenden Verkehrs auf den Zellenstrom der neuen Verbindung *während der Verbindungsannahmephase* ist prinzipiell ausgeschlossen, da die Annahmemeascheidung erfolgen muß, *bevor* die neue Verbindung ins Netz gelangt. Die Betrachtung eines isolierten Teilabschnitts, wie an Hand von Abb. 2.2 erläutert, bedeutet also immer eine Näherung, da G_2 die neue Verbindung *ohne* diese Einflüsse nachbildet. Jeder Vermittlungsknoten, den die neue Verbindung passiert, ändert die Eigenschaften der Quelle durch zusätzliche Verzögerung der Zellen oder Zellenverlust.

Auch Verfahren für die Verbindungsannahmesteuerung, die mit einfachen Schätzwerten eine Zulassungsentscheidung treffen, ignorieren den Einfluß schon durchlaufener Vermittlungsknoten [100, 66, 35, 117]. Allerdings dürften die sehr strengen Forderungen an die Dienstgüteparameter die Veränderungen klein halten.

Ein Wartesystem vom Typ $G_1^{[X]}+G_2/D/1/N$ erlaubt noch weitergehende Untersuchungen: So kann der Einfluß der Korrelation in einem der Ströme auf die Dienstgüteparameter beider Teilströme und der Überlagerung als Ganzes studiert werden, wie auch der Einfluß der jeweiligen Teilangebote. ATM sieht Verkehrsklassen mit verschiedenen Dienstgüteanforderungen vor: etwa Verlust-WKTn 10^{-3} , 10^{-6} bzw. 10^{-9} für Sprache, Daten bzw. Video. Treten diese Verkehrsklassen am Eingang desselben Multiplexers auf, so kann ohne Priorisierung – wie die numerischen Ergebnisse in Kapitel 7 andeuten – nicht gleichzeitig etwa für Videozellen ein Wert von 10^{-9} , aber für Sprache 10^{-3} als Verlust-WKT erreicht werden.

Ohne Priorisierung liegen die Verlust-WKTn offenbar nie um Größenordnungen auseinander, auch dann nicht, wenn ein Strom ein sehr hohes, der andere ein sehr niedriges Angebot hat; allerdings sind die Teilstrom-Verlust-WKTn i.a. auch nicht exakt gleich (es sei denn, die Teilströme sind exakt gleich). Schließlich können die Dienstgüteparameter der Teilströme mit denjenigen der Überlagerung als Ganzes verglichen werden.

Relativ leicht lassen sich im Modell zwei Prioritätsklassen unterscheiden, wobei eine Zelle des priorisierten Stromes bei gleichzeitiger Ankunft von Zellen des anderen Stromes immer zuerst ins System gelangt, sofern es nicht voll ist.

Eine (näherungsweise) Erweiterung dahingehend, daß der priorisierte Strom wartende Zellen des anderen Stromes verdrängt, ist ebenfalls möglich.

Es werden nun die Verfahren aus Kapitel 4 und Kapitel 5 für die Überlagerung zweier Ströme erweitert. Im Falle der Restarbeit ändert sich die MK (U_n, S_n, X_n) , während beim M/G/1-Typ-Wartesystem die MK zu Abgangszeitpunkten direkt übernommen werden kann.

Die Ergebnisse dieses Kapitels umfassen den in [83] behandelten Fall $\text{GEO}^{[X]} + \text{GI}/D/1/N$ mit $D = 1$.

Die Methode der Restarbeit erlaubt für beliebige deterministische Bedienzeiten $D \geq 1$ eine exakte Berechnung der Teilstrom-Wartezeit-WFn sowie einen vom Konzept her exakten Algorithmus zur Bestimmung der bedingten Verlust-WKT des G_2 -Stromes. Mit beiden Verfahren können für $D \geq 1$ die WFn der Belegung zu Ankunftszeitpunkten nach Strömen getrennt berechnet werden, damit also auch die Teilstrom-Verlust-WKTn. Die Methode der Restarbeit liefert wiederum transiente WKTn, liegt damit also noch näher an dem "Einschaltvorgang", der durch das Zulassen einer neuen Verbindung entsteht.

6.1 Die Restarbeit im $G_1^{[X]} + G_2/D/1/N$ -Wartesystem

Die Verknüpfung der Ergebnisse aus Kapitel 4 mit denen aus [52] liefert den Zugang zur numerischen Behandlung einer Überlagerung von zwei Strömen $G_1^{[X]} + G_2$ im Eingabeprozess. Die einzige Voraussetzung ist eine äquivalente Semi-Markovsche Darstellung der Überlagerung der Punktprozesse G_1 und G_2 , die die Ankunftszeitpunkte erzeugen. In dieser Darstellung muß der Zustandsraum als disjunkte Zerlegung $\mathcal{IE} = \mathcal{IE}_1 \cup \mathcal{IE}_2 \cup \mathcal{IE}_3$ vorgegeben sein, wobei die Zustände von \mathcal{IE}_1 [\mathcal{IE}_2] Ankünfte von G_1 [G_2] und diejenigen von \mathcal{IE}_3 (gleichzeitige) Ankünfte von beiden repräsentieren. Erst mit dieser Zerlegung sind Dienstgüteparameter für Einzelströme berechenbar. Die Überlagerung $\text{DMP}_1 + \text{DMP}_2$ hat eine solche Semi-Markovsche Darstellung (Abschnitt 3.6.1).

Beim n -ten Ankunftszeitpunkt wird eine Batchgröße X_n gezogen. Erzeugt nur G_2 eine Ankunft, d.h. der Eingabeprozess befindet sich in einem Zustand $S_n \in \mathcal{IE}_2$, so wird die Batchgröße ignoriert. Erzeugt G_1 eine Ankunft, d.h. $S_n \in \mathcal{IE}_1$, so wird die Batchgröße nach Maßgabe einer WF $x(v)$ mit $x(v) \neq 0$ für $1 \leq v \leq b_{max}$ gezogen. Erzeugen dagegen sowohl G_1 als auch G_2 eine Ankunft, d.h. $S_n \in \mathcal{IE}_3$, so wird eine modifizierte WF $\hat{x}(v)$ mit $\hat{x}(v) \neq 0$ für $1 \leq v \leq b_{max} - 1$ verwendet. Damit ist sichergestellt, daß die Maximalzahl der (im Modell) gleichzeitig erzeugten Zellen b_{max} nicht überschreitet, so daß b_{max} die Gesamtzahl derjenigen Eingangsleitungen repräsentiert, die Zellen für dieselbe Ausgangsleitung transportieren. Formal heißt das:

$$P\{X_n = v | S_n = i\} = \begin{cases} x(v), & \text{falls } i \in \mathcal{IE}_1 \\ \text{bel. WF, da } X_n \text{ ignoriert wird,} & \text{falls } i \in \mathcal{IE}_2 \\ \hat{x}(v), & \text{falls } i \in \mathcal{IE}_3 \end{cases} \quad (6.1)$$

mit

$$\hat{x}(v) = \begin{cases} x(v)/(\sum_{w=1}^{b_{max}-1} x(w)) & 1 \leq v \leq b_{max} - 1, \\ 0 & \text{sonst.} \end{cases} \quad (6.2)$$

$X = \{X_n; n \in \mathbb{N}_0\}$ repräsentiert also die Batchgröße von Strom 1; dieser bildet den – abgesehen von Strom 2 – auf die MUX-Ausgangsleitung weitervermittelten Verkehr nach. Nachfolgend sind die Hilfssatz 6 entsprechenden Übergangs-WKTN für DF und $\ell > 0$ aufgeführt. Sie können in ähnlicher Weise wie im Beweis von Hilfssatz 6 berechnet werden, und die Ausdrücke für $\ell = 0$ ergeben sich analog aus denen für $\ell > 0$:

$$P\{U_n = \ell, S_n = j | U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\} =$$

$$= \left\{ \begin{array}{l} \left. \begin{array}{l} q_{ij}(k - \ell + vD) \quad \text{für } 0 \leq k \leq ND - vD + D \\ q_{ij}(k - \ell + \xi D) \quad \text{für } 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D \\ \quad 0 \leq \xi \leq v - 1 \end{array} \right\} \quad \left. \begin{array}{l} 1 \leq v \leq N \\ \\ \end{array} \right\} \\ \left. \begin{array}{l} q_{ij}(-\ell + ND + D) \quad \text{für } k = 0 \\ q_{ij}(k - \ell + \xi D) \quad \text{für } 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D \\ \quad 0 \leq \xi \leq N \end{array} \right\} \quad \left. \begin{array}{l} 1 + N \leq v \\ \\ \end{array} \right\} \\ \text{für } i \in \mathbb{IE}_1, G_1^{[X]} \text{ allein,} \\ \\ \left. \begin{array}{l} q_{ij}(k - \ell + D) \quad \text{für } 0 \leq k \leq ND \\ q_{ij}(k - \ell) \quad \text{für } 1 + ND \leq k \leq ND + D \end{array} \right\} \quad \left. \begin{array}{l} 1 \leq v \\ \\ \end{array} \right\} \\ \text{für } i \in \mathbb{IE}_2, G_2 \text{ allein,} \\ \\ \left. \begin{array}{l} q_{ij}(k - \ell + vD + D) \quad \text{für } 0 \leq k \leq ND - vD \\ q_{ij}(k - \ell + \xi D) \quad \text{für } 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D \\ \quad 0 \leq \xi \leq v \end{array} \right\} \quad \left. \begin{array}{l} 1 \leq v \\ v \leq N - 1 \end{array} \right\} \\ \left. \begin{array}{l} q_{ij}(-\ell + ND + D) \quad \text{für } k = 0 \\ q_{ij}(k - \ell + \xi D) \quad \text{für } 1 + (N - \xi)D \leq k \leq (N - \xi + 1)D \\ \quad 0 \leq \xi \leq N \end{array} \right\} \quad \left. \begin{array}{l} N \leq v \\ \\ \end{array} \right\} \\ \text{für } i \in \mathbb{IE}_3, G_1^{[X]} \text{ und } G_2 . \end{array}$$

6.1.1 Überlagerung zweier unabhängiger DMAPs

Wie in Abschnitt 3.6.1 erläutert, besitzt die Überlagerung von zwei unabhängigen DMAPs eine äquivalente Semi-Markovsche Darstellung mit dem folgenden SM-Kern

$$\mathbf{q}(k) = \begin{pmatrix} \mathbf{q}_1(k) & \mathbf{q}_2(k) & \mathbf{q}_3(k) \\ \mathbf{q}_1(k) & \mathbf{q}_2(k) & \mathbf{q}_3(k) \\ \mathbf{q}_1(k) & \mathbf{q}_2(k) & \mathbf{q}_3(k) \end{pmatrix}, \quad (6.3)$$

$$\mathbf{q}_\nu(k) = \mathbf{C}^{k-1} \mathbf{D}^{(\nu)} \quad \text{für } \nu = 1, 2, 3,$$

$$\text{mit } \mathbf{C} := \mathbf{C}_1 \otimes \mathbf{C}_2, \quad \mathbf{D}^{(1)} := \mathbf{D}_1 \otimes \mathbf{C}_2, \quad \mathbf{D}^{(2)} := \mathbf{C}_1 \otimes \mathbf{D}_2, \quad \mathbf{D}^{(3)} := \mathbf{D}_1 \otimes \mathbf{D}_2.$$

Es werden im folgenden noch die Matrizen

$$\mathbf{Q}_\nu(k) = \sum_{\ell=1}^k \mathbf{q}_\nu(\ell), \quad \mathbf{p}_\nu = \mathbf{Q}_\nu(\infty)$$

verwendet. Die Teilmengen $\mathbb{IE}_\nu, \nu = 1, 2, 3$ des Zustandsraumes \mathbb{IE} sind gleichmächtig. Sind m_1, m_2 die Anzahlen der Phasen der beiden DMAPs, so enthält jede Teilmenge $m_1 \cdot m_2$ Zustände.

6.1.2 Transiente Gleichungen

In ähnlicher Weise wie in Abschnitt 4.1.3 können die transienten Gleichungen der MK $\{(U_n, S_n, X_n), n \in \mathbb{N}_0\}$ auf solche zurückgeführt werden, die nur die WKTn $u_n(k, i) = P\{U_n = k, S_n = i\}$ und $u_{n+1}(k, i)$ über eine stochastische Matrix verknüpfen, die implizit die WF der Batchgröße berücksichtigt. Mit den Abkürzungen

$$\begin{aligned} \underline{u}_n &:= (\underline{u}_n^{(1)}, \underline{u}_n^{(2)}, \underline{u}_n^{(3)}), \\ \underline{u}_n^{(\nu)} &= (\underline{u}_n^{(\nu)}(0), \underline{u}_n^{(\nu)}(1), \dots, \underline{u}_n^{(\nu)}(ND + D)) \text{ für } \nu = 1, 2, 3, \\ [\underline{u}_n^{(\nu)}(k)]_i &= u_n(k, i) = P\{U_n = k, S_n = i\} \text{ für } i \in \mathbb{I}E_\nu \end{aligned}$$

sowie

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}[1] & \mathbf{F}[2] & \mathbf{F}[3] \\ \mathbf{G}[1] & \mathbf{G}[2] & \mathbf{G}[3] \\ \mathbf{H}[1] & \mathbf{H}[2] & \mathbf{H}[3] \end{pmatrix} \quad (6.4)$$

ergibt die Matrixgleichung $\underline{u}_n \mathbf{F} = \underline{u}_{n+1}$ für $\nu = 1, 2, 3$ und $0 \leq \ell \leq ND + D$

$$\underline{u}_{n+1}^{(\ell)} = \sum_{k=0}^{ND+D} \underline{u}_n^{(1)}(k) \mathbf{F}_k^\ell[\nu] + \underline{u}_n^{(2)}(k) \mathbf{G}_k^\ell[\nu] + \underline{u}_n^{(3)}(k) \mathbf{H}_k^\ell[\nu] \quad (6.5)$$

Die Blockmatrizen $\mathbf{F}[\nu]$, $\mathbf{G}[\nu]$, $\mathbf{H}[\nu]$ ähneln denjenigen in den Gln. (4.16), (4.17) :

$$\mathbf{F}_k^0[\nu] = \begin{cases} \mathbf{p}_\nu - \sum_{v=1}^N x(v) \mathbf{Q}_\nu(vD) \\ \quad - X^c(N) \mathbf{Q}_\nu(ND + D) & \text{für } k = 0 \text{ und AF} \\ \\ \mathbf{p}_\nu - \sum_{v=1}^N x(v) \mathbf{Q}_\nu(vD - 1) \\ \quad - X^c(N) \mathbf{Q}_\nu(ND + D - 1) & \text{für } k = 0 \text{ und DF} \\ \\ \mathbf{p}_\nu - \sum_{v=1}^{\xi-1} x(v) \mathbf{Q}_\nu(k - 1 + vD) \\ \quad - X^c(\xi - 1) \mathbf{Q}_\nu(k - 1 + \xi D) \\ \quad \text{für } 1 + (N - \xi)D \leq k \leq (N + 1 - \xi)D, \\ \quad \text{für jedes } \xi \text{ mit } 2 \leq \xi \leq N \\ \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k - 1 + D) & \text{für } 1 + ND - D \leq k \leq ND \\ \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k - 1) & \text{für } 1 + ND \leq k \leq ND + D \end{cases} \quad (6.6)$$

und für $\ell > 0$

$$\mathbf{F}_k^\ell[\nu] = \begin{cases} \sum_{v=1}^N x(v) \mathbf{q}_\nu(-\ell + vD + 1) \\ \quad + X^c(N) \mathbf{q}_\nu(-\ell + ND + D + 1) \\ \quad \text{für } k = 0 \text{ und AF} \\ \\ \sum_{v=1}^N x(v) \mathbf{q}_\nu(-\ell + vD) \\ \quad + X^c(N) \mathbf{q}_\nu(-\ell + ND + D) \\ \quad \text{für } k = 0 \text{ und DF} \\ \\ \sum_{v=1}^{\xi-1} x(v) \mathbf{q}_\nu(k - \ell + vD) \\ \quad + X^c(\xi - 1) \mathbf{q}_\nu(k - \ell + \xi D) \\ \quad \text{für } 1 + (N - \xi)D \leq k \leq (N + 1 - \xi)D \\ \quad \text{für jedes } \xi \text{ mit } 2 \leq \xi \leq N \\ \\ \mathbf{q}_\nu(k - \ell + D) & \text{für } 1 + ND - D \leq k \leq ND \\ \\ \mathbf{q}_\nu(k - \ell) & \text{für } 1 + ND \leq k \leq ND + D \end{cases} \quad (6.7)$$

$$\mathbf{G}_k^0[\nu] = \begin{cases} \mathbf{p}_\nu - \mathbf{Q}_\nu(D) & \text{für } k = 0 \text{ und AF} \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(D-1) & \text{für } k = 0 \text{ und DF} \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k-1+D) & \text{für } 1 \leq k \leq ND \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k-1) & \text{für } 1+ND \leq k \leq ND+D \end{cases} \quad (6.8)$$

und für $\ell > 0$

$$\mathbf{G}_k^\ell[\nu] = \begin{cases} \mathbf{q}_\nu(-\ell+D+1) & \text{für } k = 0 \text{ und AF} \\ \mathbf{q}_\nu(-\ell+D) & \text{für } k = 0 \text{ und DF} \\ \mathbf{q}_\nu(k-\ell+D) & \text{für } 1 \leq k \leq ND \\ \mathbf{q}_\nu(k-\ell) & \text{für } 1+ND \leq k \leq ND+D \end{cases} \quad (6.9)$$

$$\mathbf{H}_k^0[\nu] = \begin{cases} \mathbf{p}_\nu - \sum_{v=1}^{N-1} \hat{x}(v) \mathbf{Q}_\nu((v+1)D) \\ \quad - \hat{X}^c(N-1) \mathbf{Q}_\nu(ND+D) & \text{für } k = 0 \text{ und AF} \\ \mathbf{p}_\nu - \sum_{v=1}^{N-1} \hat{x}(v) \mathbf{Q}_\nu((v+1)D-1) \\ \quad - \hat{X}^c(N-1) \mathbf{Q}_\nu(ND+D-1) & \text{für } k = 0 \text{ und DF} \\ \mathbf{p}_\nu - \sum_{v=1}^{\xi-2} \hat{x}(v) \mathbf{Q}_\nu(k-1+(v+1)D) \\ \quad - \hat{X}^c(\xi-2) \mathbf{Q}_\nu(k-1+\xi D) \\ \quad \text{für } 1+(N-\xi)D \leq k \leq (N+1-\xi)D, \\ \quad \text{für jedes } \xi \text{ mit } 3 \leq \xi \leq N \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k-1+2D) \\ \quad \text{für } 1+ND-2D \leq k \leq ND-D \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k-1+D) \\ \quad \text{für } 1+ND-D \leq k \leq ND \\ \mathbf{p}_\nu - \mathbf{Q}_\nu(k-1) \\ \quad \text{für } 1+ND \leq k \leq ND+D \end{cases} \quad (6.10)$$

und für $\ell > 0$

$$\mathbf{H}_k^\ell[\nu] = \begin{cases} \sum_{v=1}^{N-1} \hat{x}(v) \mathbf{q}_\nu(-\ell+(v+1)D+1) \\ \quad + \hat{X}^c(N-1) \mathbf{q}_\nu(-\ell+ND+D+1) \\ \quad \text{für } k = 0 \text{ und AF} \\ \sum_{v=1}^{N-1} \hat{x}(v) \mathbf{q}_\nu(-\ell+(v+1)D) \\ \quad + \hat{X}^c(N-1) \mathbf{q}_\nu(-\ell+ND+D) \\ \quad \text{für } k = 0 \text{ und DF} \\ \sum_{v=1}^{\xi-2} \hat{x}(v) \mathbf{q}_\nu(k-\ell+(v+1)D) \\ \quad + \hat{X}^c(\xi-2) \mathbf{q}_\nu(k-\ell+\xi D) \\ \quad \text{für } 1+(N-\xi)D \leq k \leq (N+1-\xi)D, \\ \quad \text{für jedes } \xi \text{ mit } 3 \leq \xi \leq N \\ \mathbf{q}_\nu(k-\ell+2D) \\ \quad \text{für } 1+ND-2D \leq k \leq ND-D \\ \mathbf{q}_\nu(k-\ell+D) \\ \quad \text{für } 1+ND-D \leq k \leq ND \\ \mathbf{q}_\nu(k-\ell) \\ \quad \text{für } 1+ND \leq k \leq ND+D \end{cases} \quad (6.11)$$

Die Fallunterscheidungen bei $\mathbf{F}_k^\ell[\nu]$ und $\mathbf{H}_k^\ell[\nu]$ spiegeln (auch durch ξ) jeweils die Anzahl freier Plätze im System wider.

6.1.3 Teilstrom-Dienstgüteparameter

Die Restarbeit U_n , die die n -te Ankunft sieht, entscheidet über Zulassung oder Verlust und bestimmt die WFn der Belegung und der Wartezeit. Diese Größen können aus den WKTn $u_n(k, i)$ nach Strömen getrennt berechnet werden. Dies ist nachfolgend dargestellt.

Zunächst sind die Begriffe Verlust und Wartezeit für die beiden Ströme zu definieren. Erzeugt G_2 allein eine Zelle, so geht sie verloren, wenn das System voll ist, d.h. $U_n \geq 1 + ND$. Kann die Zelle zugelassen werden, so ist die Wartezeit durch die gesehene Restarbeit gegeben. Erzeugen dagegen G_2 und G_1 gleichzeitig Ankünfte, so kann über den Verlust etwa der G_2 -Zelle nicht ohne weiteres nur aus der gesehenen Restarbeit entschieden werden. Dies gelingt erst durch die Festlegung einer Zulassungsreihenfolge der gleichzeitig ankommenden Zellen.

Da Prioritäten zunächst ausgeschlossen werden, ist die zufällige Auswahl der Zellen für die Zulassung zum System am sinnvollsten. Kein Strom wird bevorzugt. Für die Zellen von Strom G_1 werden entsprechend Abschnitt 4.1.4 Verlust und Wartezeit durch eine Testzelle (s. [83]) gemessen. Dazu wird angenommen, daß unter den von $G_1^{[X]}$ gleichzeitig erzeugten Zellen eine markiert ist.

Wenn $G_1^{[X]}$ und G_2 gleichzeitig Zellen erzeugen, enthält der ankommende Batch eine markierte Testzelle und die ebenfalls unterscheidbare Zelle des G_2 -Stromes.

Zur weiteren Formalisierung wird neben der Sprechweise "als ψ -te zugelassen" auch "in Position ψ zugelassen" verwendet. Bezeichne $\Gamma_n^{(1)}$ [$\Gamma_n^{(2)}$] die Nummer der Zulassungsposition der Testzelle von Strom 1 [Zelle von Strom 2] beim n -ten Ankunftszeitpunkt. Dann gilt wegen der zufälligen Auswahl

$$P\{\Gamma_n^{(1)} = \psi | X_n = w, S_n = i\} = \begin{cases} \frac{1}{w} & \text{für } i \in \mathbb{E}_1, 1 \leq \psi \leq w, \\ \frac{w}{w(w+1)} = \frac{1}{w+1} & \text{für } i \in \mathbb{E}_3, 1 \leq \psi \leq w+1, \\ 0 & \text{sonst.} \end{cases} \quad (6.12)$$

$$P\{\Gamma_n^{(2)} = \psi | X_n = w, S_n = i\} = \begin{cases} 1 & \text{für } i \in \mathbb{E}_2, \psi = 1, \\ \frac{w}{w(w+1)} = \frac{1}{w+1} & \text{für } i \in \mathbb{E}_3, 1 \leq \psi \leq w+1, \\ 0 & \text{sonst.} \end{cases} \quad (6.13)$$

Begründung: $S_n \in \mathbb{E}_1$: Für einen w -Batch gibt es w Zulassungspositionen der Testzelle, die wegen der zufälligen Auswahl alle gleichwahrscheinlich sind. $S_n \in \mathbb{E}_3$: Enthält ein $(w+1)$ -Batch neben der Testzelle auch eine Zelle von Strom 2, so gibt es w Möglichkeiten der Wahl einer Position für eine der beiden, wenn die Position der anderen schon festliegt und insgesamt $(w+1)w$ Möglichkeiten, Positionen für die Testzelle und die Zelle von Strom 2 auszuwählen.

In den folgenden Ausdrücken wird zur Verkürzung das Skalarprodukt

$$\underline{u}^{(\nu)}(k)\underline{e} := \sum_{i \in \mathbb{E}_\nu} u(k, i), \quad \nu = 1, 2, 3$$

verwendet.

6.1.4 Teilstrom-Verlustwahrscheinlichkeiten

Infolge der Mehrfachankünfte kann Verlust bei jedem möglichen Belegungswert L_n , selbst bei leerem System auftreten. Verlust tritt auf, wenn

1. $i \in \mathbb{IE}_1$: $X_n = w$, $L_n + w \geq N + 2$; es gibt $w - (N + 1 - L_n)$ Verlustpositionen der Testzelle. Sie haben die Nummern $N + 2 - L_n, \dots, w$.
2. $i \in \mathbb{IE}_2$: $L_n = N + 1$.
3. $i \in \mathbb{IE}_3$: $X_n = w$, $L_n + w + 1 \geq N + 2$, denn auch Strom 2 erzeugt eine Zelle. Es gibt $w + 1 - (N + 1 - L_n)$ Verlustpositionen der Testzelle, bzw. der Zelle von Strom 2. Sie haben die Nummern $N + 2 - L_n, \dots, w + 1$.

Die WFn $\mathcal{L}(\mu, \nu) := P\{L_n = \mu, S_n \in \mathbb{IE}_\nu\}$ der Belegung zum Ankunftszeitpunkt nach Zustandsteilmengen getrennt ($\nu = 1, 2, 3$) ergeben sich zu

$$\mathcal{L}(\mu, \nu) = \begin{cases} \underline{u}^{(\nu)}(0)\underline{e} & \text{für } \mu = 0 \\ \sum_{k=1+(\mu-1)D}^{\mu D} \underline{u}^{(\nu)}(k)\underline{e} & \text{für } 1 \leq \mu \leq N + 1 \end{cases} \quad \nu = 1, 2, 3. \quad (6.14)$$

Die Definition der Verlust-WKT für Strom $i = 1, 2$ als relative Häufigkeit meint den Quotienten aus der Anzahl verlorener Zellen (Testzellen) von Strom i und der Anzahl der insgesamt erzeugten Zellen (Testzellen) dieses Stromes. Um dieser Definition gerecht zu werden, müssen die Verbundereignisse in den Belegungs-WKTn (6.14) bei den Teilstrom-Verlust-WKTn auf Ankunftsereignisse des jeweiligen Stromes eingeschränkt werden, d.h. durch

$$P\{\text{Strom } i \text{ erzeugt eine Ankunft}\} = \sum_{k=0}^{ND+D} [\underline{u}^{(i)}(k) + \underline{u}^{(3)}(k)]\underline{e}, \quad i = 1, 2$$

dividiert werden. Daher ist i.a. die Verlust-WKT $P_{\text{Verlust}[1+2]}$ des Gesamtstromes nicht gleich der Summe $P_{\text{Verlust}1} + P_{\text{Verlust}2}$.¹ Die beiden Ausdrücke für Strom $i = 1, 2$ enthalten jeweils eine gleich aufgebaute Summe, die den Fall repräsentiert, daß beide Ströme Ankünfte erzeugen. Diese Gemeinsamkeit tritt auch bei den WFn der Wartezeit auf.

Verlustwahrscheinlichkeit des $G_1^{[X]}$ -Stromes

$$\begin{aligned} P_{\text{Verlust}1} \cdot \left[\sum_{k=0}^{ND+D} [\underline{u}^{(1)}(k) + \underline{u}^{(3)}(k)]\underline{e} \right] &= \\ = \sum_{\mu=0}^{N+1} \sum_{w=N+2-\mu}^{\infty} \frac{w - (N + 1 - \mu)}{w} [x(w)\mathcal{L}(\mu, 1) + \hat{x}(w - 1)\mathcal{L}(\mu, 3)] & \quad (6.15) \end{aligned}$$

Verlustwahrscheinlichkeit des G_2 -Stromes

$$\begin{aligned} P_{\text{Verlust}2} \cdot \left[\sum_{k=0}^{ND+D} [\underline{u}^{(2)}(k) + \underline{u}^{(3)}(k)]\underline{e} \right] &= \\ = \mathcal{L}(N + 1, 2) + \sum_{\mu=0}^{N+1} \sum_{w=N+2-\mu}^{\infty} \frac{w - (N + 1 - \mu)}{w} \hat{x}(w - 1)\mathcal{L}(\mu, 3) & \quad (6.16) \end{aligned}$$

¹Wegen der Ungleichung $\frac{a}{c} \leq \frac{a+b}{c+d} \leq \frac{b}{d}$, $c, d \neq 0$, gilt bei $P_{\text{Verlust}1} \leq P_{\text{Verlust}2}$ auch $P_{\text{Verlust}1} \leq P_{\text{Verlust}[1+2]} \leq P_{\text{Verlust}2}$, vgl. Abb. 7.3.

Verlustwahrscheinlichkeit des Gesamtstromes

$$P_{\text{Verlust}[1+2]} = \mathcal{L}(N + 1, 2) + P_{\text{Verlust}1} \cdot \left[\sum_{k=0}^{ND+D} [\underline{u}^{(1)}(k) + \underline{u}^{(3)}(k)] \underline{e} \right] \quad (6.17)$$

6.1.5 Teilstrom-Wartezeiten

Eine Zelle, die nicht verlorenght, sieht als Wartezeit die Restarbeit zum jeweiligen Ankunftszeitpunkt plus der Summe der Bedienzeiten aller Zellen, die gleichzeitig angekommen und vor der betrachteten Zelle ins System gelangt sind. Da einer verlorenen Zelle keine Wartezeit zugeordnet wird, nimmt die Wartezeit nur Werte aus $\{0, 1, \dots, ND\}$ an. Da sie auf die Restarbeit zurückgeführt wird, benötigt man Normierungskonstanten $N_1, N_2, N_{[1+2]}$. Für $k < 0$ wird $u(k, i) := 0$ definiert. Die nachfolgenden Ausdrücke gelten für DF. Für AF muß neben kleinen Modifikationen der Fall des leeren Systems getrennt betrachtet werden, wie in Abschnitt 4.1.4 dargestellt.

Für eine Testzelle des $G_1^{[X]}$ -Stromes (DF)

$$P\{W_1 = \ell\} = \frac{\mathbf{1}_{\ell=0}^{ND}}{N_1} \left[\sum_{w=1}^{b_{\max}} \frac{x(w)}{w} \sum_{\psi=1}^w \underline{u}^{(1)}(\ell - (\psi - 1)D) \underline{e} + \sum_{w=2}^{b_{\max}} \frac{\hat{x}(w-1)}{w} \sum_{\psi=1}^w \underline{u}^{(3)}(\ell - (\psi - 1)D) \underline{e} \right], \quad (6.18)$$

N_1 ist gegeben durch $\sum_{\ell=0}^{ND} P\{W_1 = \ell\} = 1$.

Für den G_2 -Strom (DF)

$$P\{W_2 = \ell\} = \frac{\mathbf{1}_{\ell=0}^{ND}}{N_2} \left[\underline{u}^{(2)}(\ell) \underline{e} + \sum_{w=2}^{b_{\max}} \frac{\hat{x}(w-1)}{w} \sum_{\psi=1}^w \underline{u}^{(3)}(\ell - (\psi - 1)D) \underline{e} \right], \quad (6.19)$$

N_2 ist gegeben durch $\sum_{\ell=0}^{ND} P\{W_2 = \ell\} = 1$.

Für den Gesamtstrom (DF)

$$P\{W_{[1+2]} = \ell\} = \frac{1}{N_{[1+2]}} \left[N_1 \cdot P\{W_1 = \ell\} + \mathbf{1}_{\ell=0}^{ND} \cdot \underline{u}^{(2)}(\ell) \underline{e} \right], \quad (6.20)$$

$N_{[1+2]}$ ist gegeben durch $\sum_{\ell=0}^{ND} P\{W_{[1+2]} = \ell\} = 1$.

6.2 Die Überlagerung $\text{SSMP}_1^{[X]} + \text{SSMP}_2$

Das oben dargestellte Verfahren benötigt die Semi-Markovsche Repräsentation der Überlagerung $G_1 + G_2$. Für zwei DMAPs wurde diese angegeben. Andere SP als DMAPs zur Nachbildung der Teilströme sind natürlich von Interesse, insbesondere der $\text{SSMP}(2)$, da für diesen eine Methode zur Parametrisierung aus Meßdaten bereitsteht [30]. Es gibt einige Hinweise darauf, daß auch für $\text{SSMP}_1(2) + \text{SSMP}_2(2)$ eine solche SMP-Darstellung existiert. Die Arbeiten hierzu sind noch nicht abgeschlossen.

6.3 Bedingte Verlustwahrscheinlichkeit in einem ATM-Multiplexer

Infolge der MK (U_n, S_n, X_n) liefert die Methode der Restarbeit einen exakten Algorithmus zur Berechnung der WKT, daß "unter der Bedingung, daß Strom 2 eine Zelle verliert, er seine nächste Zelle ebenfalls verliert". Diese *bedingte Verlustwahrscheinlichkeit* dient für diesen Strom als Schätzwert für die Häufigkeit, mehrere Zellen nacheinander zu verlieren. Ein solches Verlustereignis kann die Dienstgüte erheblich verschlechtern, auch wenn die (auf alle Zellen bezogene) Verlust-WKT sehr klein ist. Die in [10] betrachtete bedingte Verlust-WKT bezieht sich auf einen Gesamtstrom, in dem eine Einzelquelle nicht erkennbar ist, und ist somit vom Konzept her nur eine Näherung, s. Abschnitt 4.1.6.

Mit dem Ereignis $\mathcal{E}(\nu) :=$ "Eine Zelle von Strom 2 geht bei der ν -ten Ankunft verloren" ergibt sich die bedingte Verlust-WKT zu folgendem Ausdruck (stationärer Fall):

$$P_{\text{Verlust}}^{\text{bedingt}} := P\{\mathcal{E}(n+1)|\mathcal{E}(n)\} + \sum_{\xi=2}^{\infty} P\{\mathcal{E}(n+\xi), S_{n+\xi-1} \in \mathcal{IE}_1, \dots, S_{n+1} \in \mathcal{IE}_1 | \mathcal{E}(n)\} \quad (6.21)$$

Diese unendliche Summe kann durch eine endliche Summe beliebig genau angenähert werden. Die WKTn $P\{\mathcal{E}(n+\xi), S_{n+\xi-1} \in \mathcal{IE}_1, \dots, S_{n+1} \in \mathcal{IE}_1 | \mathcal{E}(n)\}$ in Gl. (6.21) ergeben sich mittels der rekursiv berechenbaren Ausdrücke $\Theta_{\xi}(\ell, j) := P\{U_{n+\xi} = \ell, S_{n+\xi} = j, S_{n+\xi-1} \in \mathcal{IE}_1, \dots, S_{n+1} \in \mathcal{IE}_1, \mathcal{E}(n)\}$ mit $\xi > 1$ aus den Startwerten $\Theta_1(\ell, j) := P\{U_{n+1} = \ell, S_{n+1} = j, \mathcal{E}(n)\}$. Mit der Abkürzung

$$P_{ikv}\{U_{n+1} = \ell, S_{n+1} = j\} := P\{U_{n+1} = \ell, S_{n+1} = j | U_n = k, S_n = i, X_n = v\}$$

erhält man:

$$\begin{aligned} \Theta_1(\ell, j) &= P\{U_{n+1} = \ell, S_{n+1} = j, \mathcal{E}(n)\} = \\ &= \sum_{i \in \mathcal{IE}_2} \sum_{v=1}^{\infty} \sum_{k=1+ND}^{ND+D} x(v)u(k, i)P_{ikv}\{U_{n+1} = \ell, S_{n+1} = j\} + \\ &+ \sum_{i \in \mathcal{IE}_3} \sum_{\mu=1}^{N+1} \sum_{v=N+2-\mu}^{\infty} \hat{x}(v-1) \frac{v-(N+1-\mu)}{v} \cdot \sum_{k=1+(\mu-1)D}^{\mu D} P_{ik(v-1)}\{U_{n+1} = \ell, S_{n+1} = j\}u(k, i) \\ &+ \sum_{i \in \mathcal{IE}_3} \sum_{v=N+2}^{\infty} \hat{x}(v-1) \cdot \frac{v-(N+1)}{v} P_{i0(v-1)}\{U_{n+1} = \ell, S_{n+1} = j\}u(0, i) \quad . \end{aligned}$$

Die $\Theta_{\xi}(\ell, j)$ ergeben sich aus $\Theta_{\xi-1}(\ell, j)$ für $\xi > 1$ gemäß

$$\begin{aligned} \Theta_{\xi}(\ell, j) &= P\{U_{n+\xi} = \ell, S_{n+\xi} = j, S_{n+\xi-1} \in \mathcal{IE}_1, \dots, S_{n+1} \in \mathcal{IE}_1, \mathcal{E}(n)\} = \\ &= \sum_{i \in \mathcal{IE}_1} \sum_{k=0}^{ND+D} \sum_{v=1}^{\infty} P_{ikv}\{U_{n+\xi} = \ell, S_{n+\xi} = j\}x(v)\Theta_{\xi-1}(k, i) \quad . \end{aligned}$$

Schließlich ist

$$\begin{aligned}
& P\{\mathcal{E}(n+\xi), S_{n+\xi-1} \in \mathbb{IE}_1, \dots, S_{n+1} \in \mathbb{IE}_1, \mathcal{E}(n)\} = \\
& = \sum_{i \in \mathbb{IE}_2} \sum_{k=1+ND}^{ND+D} \Theta_{n+\xi}(k, i) + \sum_{i \in \mathbb{IE}_3} \sum_{\mu=1}^{N+1} \sum_{w=N+2-\mu}^{\infty} \hat{x}(w-1) \cdot \frac{w-(N+1-\mu)}{w} \sum_{k=1+(\mu-1)D}^{\mu D} \Theta_{n+\xi}(k, i) + \\
& \quad + \sum_{i \in \mathbb{IE}_3} \sum_{w=N+2}^{\infty} \hat{x}(w-1) \cdot \frac{w-(N+1)}{w} \Theta_{n+\xi}(0, i) .
\end{aligned}$$

Der Nenner $P\{\mathcal{E}(n)\}$ in den bedingten WKTn der Gl. (6.21) berechnet sich zu

$$P\{\mathcal{E}(n)\} = \sum_{i \in \mathbb{IE}} \sum_{k=0}^{ND+D} \Theta_1(k, i) .$$

6.4 Das DBMAP₁+DMAP₂/G/1/N-Wartesystem

Die Überlagerung DBMAP₁+DMAP₂ ergibt wiederum einen DBMAP, so daß das DBMAP₁+DMAP₂/G/1/N-System mit den in Kapitel 5 bereitgestellten Methoden mit sehr viel geringeren Modifizierungen behandelt werden kann als bei der Methode der Restarbeit. Die zu Abgangszeiten eingebettete MK wird nur von dem aus DBMAP₁+DMAP₂ resultierenden DBMAP bestimmt. Seien DBMAP₁ durch $\mathbf{D}_w^{(1)}$, $w \in \mathbb{N}_0$, $b_{max}^{(1)}$, $\widehat{\mathbf{D}}^{(1)} := \sum_{w=0}^{b_{max}^{(1)}} \mathbf{D}_w^{(1)}$ und DMAP₂ durch $\mathbf{C}^{(2)}$ und $\mathbf{D}^{(2)}$ vorgegeben. Für den Überlagerungs-DBMAP, gegeben durch $\widetilde{\mathbf{D}}_w$, $w \in \mathbb{N}_0$, $\sum_{w=0}^{\infty} \widetilde{\mathbf{D}}_w = \widetilde{\mathbf{D}}$, erhält man mittels des Kroneckerproduktes \otimes

$$\widetilde{\mathbf{D}}_w = \begin{cases} \mathbf{D}_0^{(1)} \otimes \mathbf{C}^{(2)} & \text{für } w = 0 \\ \mathbf{D}_w^{(1)} \otimes \mathbf{C}^{(2)} + \mathbf{D}_{w-1}^{(1)} \otimes \mathbf{D}^{(2)} & \text{für } 1 \leq w \leq b_{max}^{(1)} \\ \mathbf{D}_w^{(1)} \otimes \mathbf{D}^{(2)} & \text{für } w = b_{max}^{(1)} + 1 . \end{cases}$$

Für einen allgemeinen DBMAP₁ ist es schwierig den Maximalwert der zu ziehenden Batchgröße davon abhängig zu machen, ob DBMAP₁ allein oder zusammen mit DMAP₂ Ankünfte erzeugt. Ist DBMAP₁ dagegen ein DMAP₁^[X], so kann diese Kopplung wie in Gl. (6.1) erfolgen durch

$$\widetilde{\mathbf{D}}_w = \begin{cases} \mathbf{D}_0^{(1)} \otimes \mathbf{C}^{(2)} & \text{für } w = 0 \\ \mathbf{D}_w^{(1)} \otimes \mathbf{C}^{(2)} + \mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)} & \text{für } 1 \leq w \leq b_{max}^{(1)} \end{cases} \quad (6.22)$$

mit

$$\mathbf{D}_w^{(1)} = \begin{cases} x(w)\mathbf{D}^{(1)} & w > 0 \\ \mathbf{C}^{(1)} & w = 0 \end{cases} \quad \text{und} \quad \mathbf{D}_w^{(1\bullet)} = \begin{cases} \hat{x}(w)\mathbf{D}^{(1)} & w > 0 \\ \mathbf{C}^{(1)} & w = 0 \end{cases} .$$

Genauso wie in Gl. (6.14) können Belegungen nach Ankünften von Strom 1, Strom 2 oder beiden Strömen getrennt behandelt werden. Für AF erhält man

$$P\{L_n = \mu, X_n = w, \text{Ankunft von Strom 1 allein}\} = \underline{y}_\mu [\mathbf{D}_w^{(1)} \otimes \mathbf{C}^{(2)}]_{\underline{e}} ,$$

$$P\{L_n = \mu, \text{Ankunft von Strom 2 allein}\} = \underline{y}_\mu [\mathbf{D}_0^{(1)} \otimes \mathbf{D}^{(2)}]_{\underline{e}}$$

und

$$P\{L_n = \mu, X_n = w-1, \text{Ankunft von Strom 1 und 2}\} = \underline{y}_\mu \mathbf{1}_{w \geq 2} [\mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} .$$

Ist Strom 1 durch DMAP₁^[X] gegeben, so entsprechen sich die Ausdrücke (vgl. Gl. (6.14))

$$\begin{aligned} & x(w)\mathcal{L}(\mu, 1) \text{ und } \underline{y}_\mu [\mathbf{D}_w^{(1)} \otimes \mathbf{C}^{(2)}]_{\underline{e}} , \\ & \mathcal{L}(\mu, 2) \text{ und } \underline{y}_\mu [\mathbf{D}_0^{(1)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} \text{ sowie} \\ & \hat{x}(w-1)\mathcal{L}(\mu, 3) \text{ und } \underline{y}_\mu \mathbf{1}_{w \geq 2} [\mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} . \end{aligned}$$

Es gelten wieder dieselben Besonderheiten für die Auswahl der Testzelle wie in Abschnitt 6.1.4. Für DF muß \underline{y}_μ gemäß Satz 2 um $\underline{\Delta}_\mu$ geeignet ergänzt werden.

6.4.1 Teilstrom-Verlustwahrscheinlichkeiten (AF²)

Die Ausdrücke für die Verlust-WKTh haben denselben strukturellen Aufbau wie die Gln. (6.15), (6.16) und (6.17).

Teilstrom DMAP₁^[X]

$$P_{\text{Verlust1}} = \frac{1}{\mathcal{A}_1} \sum_{\mu=0}^{N+1} \sum_{w=N+2-\mu}^{\infty} \frac{w - (N+1-\mu)}{w} \underline{y}_\mu [\mathbf{D}_w^{(1)} \otimes \mathbf{C}^{(2)} + \mathbf{1}_{w \geq 2} \mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} \quad (6.23)$$

Teilstrom DMAP₂

$$P_{\text{Verlust2}} = \frac{1}{\mathcal{A}_2} \left\{ \underline{y}_{N+1} [\mathbf{D}_0^{(1)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} + \sum_{\mu=0}^{N+1} \sum_{w=N+2-\mu}^{\infty} \frac{w - (N+1-\mu)}{w} \underline{y}_\mu [\mathbf{1}_{w \geq 2} \mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} \right\} \quad (6.24)$$

Gesamtstrom DMAP₁^[X]+DMAP₂

$$P_{\text{Verlust}\{1+2\}} = \frac{1}{\mathcal{A}_{\{1+2\}}} \left\{ \underline{y}_{N+1} [\mathbf{D}_0^{(1)} \otimes \mathbf{D}^{(2)}]_{\underline{e}} + \mathcal{A}_1 \cdot P_{\text{Verlust1}} \right\} \quad (6.25)$$

Die verschiedenen Divisoren sorgen wieder für die Einschränkung auf Ankunftszeitpunkte des jeweiligen Stromes. Dabei repräsentiert

- $\mathcal{A}_1 := \pi(\widehat{\mathbf{D}}^{(1)} - \mathbf{D}_0^{(1)}) \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)})_{\underline{e}}$ die WKT, daß eine Testzelle von Strom 1 ankommt,
- $\mathcal{A}_2 := \pi(\widehat{\mathbf{D}}^{(1)} \otimes \mathbf{D}^{(2)})_{\underline{e}}$ die WKT, daß eine Zelle von Strom 2 ankommt und
- $\mathcal{A}_{\{1+2\}} := \pi(\widehat{\mathbf{D}}^{(1)} \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)}) - \mathbf{D}_0^{(1)} \otimes \mathbf{C}^{(2)})_{\underline{e}}$ die WKT, daß der Gesamtstrom, d.h. wenigstens einer der Teilströme, eine Ankunft erzeugt.

²DF ergibt sich durch geeignete Ergänzung gemäß Satz 2.

6.4.2 Teilstrom-Wartezeiten

Zwar gibt es (bisher) keinen geschlossenen Ausdruck für die WF der Wartezeit des DBMAP/G/1/ N -Systems. Jedoch kann für den ATM-relevanten deterministischen Bedienprozeß eine Näherung mittels der Belegung angegeben werden, so daß die daraus resultierenden Momente obere Schranken derjenigen der Wartezeit sind. Man erhält dann mit den Normierungskonstanten $N_i = \sum_{\ell=0}^N P\{W_i = \ell \cdot D\}$, $i = 1, 2$ (für AF)

$$P\{W_1 = \ell \cdot D\} = \frac{\mathbf{1}_{\ell=0}^N}{N_1} \left[\sum_{w=1}^{b_{\max}^{(1)}} \frac{1}{w} \sum_{\psi=1}^w \underline{y}_{\ell-\psi+1} \cdot \left(\mathbf{D}_w^{(1)} \otimes \mathbf{C}^{(2)} \right) \cdot \underline{e} + \right. \\ \left. + \sum_{w=2}^{b_{\max}^{(1)}} \frac{1}{w} \sum_{\psi=1}^w \underline{y}_{\ell-\psi+1} \cdot \left(\mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)} \right) \cdot \underline{e} \right], \quad (6.26)$$

$$P\{W_2 = \ell \cdot D\} = \frac{\mathbf{1}_{\ell=0}^N}{N_2} \left[\sum_{w=1}^{b_{\max}^{(1)}} \frac{1}{w} \sum_{\psi=1}^w \underline{y}_{\ell-\psi+1} \cdot \left(\mathbf{D}_{w-1}^{(1\bullet)} \otimes \mathbf{D}^{(2)} \right) \cdot \underline{e} \right]. \quad (6.27)$$

6.4.3 Nachbildung von Prioritäten

Im ATM-Zellkopf ist ein Bit für Prioritäten vorgesehen. Ist das Bit zu 1 gesetzt, so hat die Zelle in Stausituationen Vorrang. Die Nachbildung priorisierter Zellankünfte in den obigen Modellen kann dadurch erreicht werden, daß die Zulassungsposition der Zelle von Strom 2 nicht nach einer Gleichverteilung, sondern gemäß einer geeigneten anderen Verteilung ausgewählt wird. Diese trägt dann dem Anteil der priorisierten Zellen von Strom 2 Rechnung. Im Grenzfall, wenn alle Zellen von Strom 2 ein gesetztes Prioritätsbit haben, wird die Zelle von Strom 2 immer als erste zugelassen. Schwieriger wird es allerdings, ein Verdrängen einer schon wartenden Zelle zu behandeln. In [54] wird dafür eine Näherung angegeben.

7 Modellgestützte Untersuchung von ATM-Verkehrssituationen

7.1 Validierung der verschiedenen Modelle

Für jeden Eingabeprozess, der sowohl eine Darstellung als SMP als auch eine als DBMAP besitzt, können mit den beiden völlig verschiedenen Verfahren aus Kapitel 4 und 5 die WF der Belegung und Verlust-WKTn für AF und DF berechnet werden. Das liefert eine gegenseitige "numerisch-analytische" Validierung der implementierten Algorithmen. Insbesondere für jeden $DMAP^{[X]}$ existiert immer ein äquivalenter $SMP^{[X]}$. Dieselbe Validierung besteht für das CAC-Modell $DMAP_1^{[X]} + DMAP_2/D/1/N$ hinsichtlich der Teilstrom-WFn der Belegung und der Teilstromverlust-WKTn.

Tab. 7.1 stellt die Möglichkeiten beider Verfahren gegenüber. In allen Fällen zeigen die mit beiden Verfahren berechneten WKTn auch bei sehr kleinen Werten um 10^{-9} hervorragende

Methoden der Restarbeit	M/G/1-Typ-Wartesysteme
beliebige deterministische Bedienzeiten $D \geq 1$	
WFn der von jedem Strom gesehenen Belegung	
Teilstrom-Verlustwahrscheinlichkeiten	
WFn der Wartezeiten jedes Teilstromes	—
—	Charakterisierung des Abgangsprozesses als SMP, Rand- und Verbund-WFn der Zwischenabgangszeiten
—	WFn der Belegung zu Abgangszeitpunkten und zu beliebigen Zeitpunkten für den <i>Gesamtabgangsstrom</i>
lokaler Korrelationskoeffizient der Belegung zu Ankunftszeitpunkten	lokaler Korrelationskoeffizient der Belegung zu Abgangszeitpunkten und derjenige der Zwischenabgangszeiten
Rand- und Verbund-WFn der Restarbeit	
bedingte Verlustwahrscheinlichkeit für Strom 2	—
	Der Zustandsraum der Systemmatrix ist um den Faktor 3 kleiner, daher schneller
Teilstrom-Dienstgüteparameter sind auch für den <i>transienten</i> Fall zugänglich	—

Tab. 7.1 Vergleich der Möglichkeiten der Methode der Restarbeit und der M/G/1-Typ-Wartesysteme bezogen auf das $DMAP_1^{[X]} + DMAP_2/D/1/N$ -System.

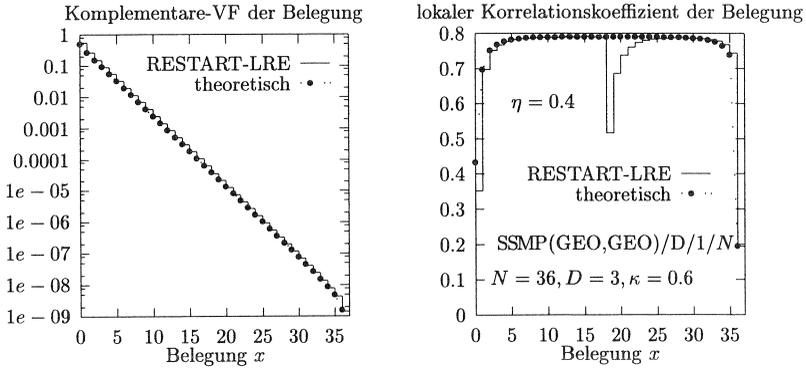


Abb. 7.1 Simulierte und berechnete Verläufe für die Belegung

Übereinstimmung. Für `double-precision`-Zahlen in C-Programmen unterscheiden sich die numerischen Werte in Exponentialdarstellung mit 16 Stellen nur in den letzten 4 der 15 Nachkommastellen.

Alle berechenbaren WF_n (Belegung, Wartezeit, Zwischenabgangszeiten auch für zwei unterscheidbare Ströme) wurden auch mittels Computersimulation im Rahmen von [118] validiert. Dabei wurden allerdings nur solche Fälle erfaßt, bei denen alle auftretenden WKT_n-werte nicht unter 10^{-5} liegen. Andernfalls werden die Simulationsdauern bei weitem zu groß. Es wurde gute Übereinstimmung gefunden.

Auch das RESTART-LRE-Verfahren zur Schätzung sehr seltener Ereignisse aus [109] wurde zum Vergleich der berechneten Belegungs-WF des $\text{SMP}^{(X)}/D/1/N$ mit WKT_n im Bereich von 10^{-9} herangezogen und dafür gute Übereinstimmung gefunden, s. [109] und Abb. 7.1.

7.2 Korrelationseinfluß auf QOS-Parameter eines Gesamtstromes

In diesem Abschnitt wie auch in den folgenden wird als Eingabeprozess ein SSMP(GEO, GEO) verwendet, d.h. die beiden Zustandsverweilzeiten sind geometrisch verteilt.

$$\begin{aligned} f_{11}(k) &= f_{12}(k) = (1 - q_1)q_1^{k-1} \\ f_{21}(k) &= f_{22}(k) = (1 - q_2)q_2^{k-1} \end{aligned} \quad \text{für } k \geq 1.$$

Für diesen SSMP gibt es eine (einfache) Repräsentation als DMAP, so daß eine "numerisch-analytische" Validierung der berechneten WKT_n durch Verwendung der in den Kapiteln 4 und 5 gegebenen Formeln möglich ist. Der DMAP ist gegeben durch

$$\mathbf{C} = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix} \quad \text{und} \quad \mathbf{D} = \begin{pmatrix} 1 - q_1 & 0 \\ 0 & 1 - q_2 \end{pmatrix} \cdot \mathbf{P}.$$

Dabei ist \mathbf{p} die Übergangsmatrix der eingebetteten MK

$$\mathbf{p} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

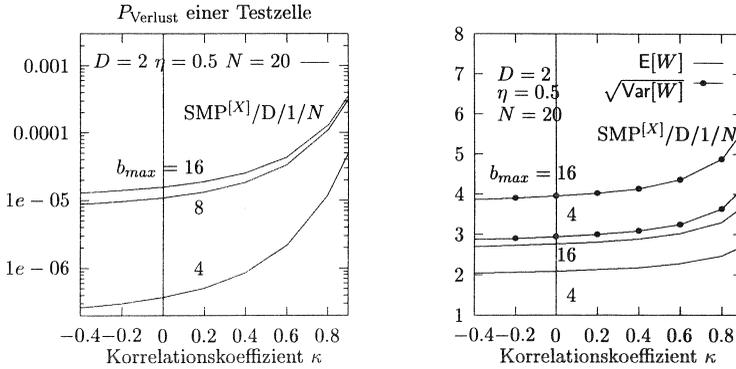


Abb. 7.2 Verlust-WKT sowie Mittelwert und Standardabweichung der Wartezeit einer Testzelle über dem Korrelationskoeffizienten des Eingabeprozesses für verschiedene Werte von b_{max} .

Andererseits ist es bei einem SSMP mit zwei Zuständen leicht möglich, die Korrelation der Zwischenankunftszeiten mittels der Korrelationskoeffizienten der eingebetteten MK vorzugeben. Bei mehr als 2 Zuständen besteht kein einfacher und eindeutiger Zusammenhang mehr; insbesondere läßt sich der Korrelationskoeffizient der eingebetteten MK nicht mehr eindeutig vorgeben, vgl. Gl. (3.24).

Als (monoton fallende) WF der Batchgröße dient eine abgeschnittene geometrische WF

$$x(v) = \frac{1 - q_b}{1 - q_b^{b_{max}}} q_b^{v-1}, \quad 1 \leq v \leq b_{max} .$$

Für die stationäre Lösung der eingebetteten MK wird $\underline{p} = (0.3, 0.7)$ festgelegt. $\underline{P} \cdot \mathbf{p} = \underline{P}$ liefert $P_1 p_{12} = P_2 p_{21}$. Der Korrelationskoeffizient κ wird vorgegeben durch $\kappa = 1 - p_{12} - p_{21} = 1 - p_{12}/P_2$, d.h. er bestimmt p_{12} und damit \mathbf{p} vollständig. Wegen der Wahl von \underline{p} mit $P_1 < P_2$ kann κ in $[-\frac{P_1}{P_2}, 1]$ variiert werden. Nur für $P_1 = P_2 = 0.5$ ist $\kappa = -1$ möglich. Das resultierende Angebot ergibt sich wegen der Batch-ZV zu $\eta = E[X] \cdot D/E[A]$.

Um das Angebot und den Korrelationskoeffizienten von \mathbf{p} vorgeben zu können, werden die Parameter q_i , $i = 1, 2$, die die Zustandsverweilzeiten bestimmen, aus $E[A] = P_1 E[A^{(1)}] + P_2 E[A^{(2)}] = E[X] \cdot D/\eta$ mit der sog. "balanced-mean"-Bedingung [32] $P_1 E[A^{(1)}] = 0.2 \cdot P_2 E[A^{(2)}]$ festgelegt. Dabei bedeutet $A^{(i)}$ die Zustandsverweilzeit im Zustand $i = 1, 2$, und wegen der geometrisch verteilten Verweilzeiten ist $E[A^{(i)}] = 1/(1 - q_i)$ und $\text{Var}[A^{(i)}] = q_i/(1 - q_i)$. $\text{Corr}[A_n, A_{n+1}]$ und κ sind nach Gl. (3.25) wie folgt verknüpft:

$$\text{Corr}[A_{n+1}, A_n] = \kappa \cdot \left(1 + \frac{P_1 \text{Var}[A^{(1)}] + P_2 \text{Var}[A^{(2)}]}{P_1 P_2 (E[A^{(1)}] - E[A^{(2)}])^2} \right)^{-1} ,$$

Der Proportionalitätsfaktor hängt somit auch vom Angebot ab¹, hier aber sehr schwach. In den Diagrammen in Abb. 7.2 wird bei $\eta = 0.5$ und $E[X] \approx 1.6$ (d.h. $q_b = 0.4$) für die

¹ Dies ist auch der Fall, wenn die q_i vorgegeben werden und damit aus $E[A] = P_1 E[A^{(1)}] + P_2 E[A^{(2)}]$ und dem Verkehrsangebot die WKTn P_1 und $P_2 = 1 - P_1$ bestimmt werden. Dann kann sich aber der mögliche Bereich für κ verkleinern.

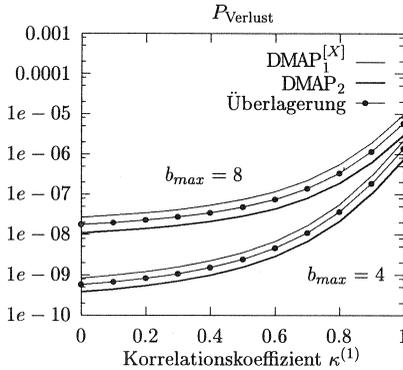


Abb. 7.3 Verlust-WKTn der Teilströme und des Gesamtstromes über dem Korrelationskoeffizienten von Strom 1 bei $\eta_1 = 0.3$, $\eta_2 = 0.2$, $N = 30$, $D = 4$ und unkorreliertem Strom 2.

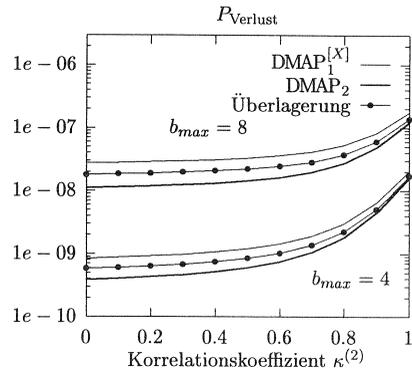


Abb. 7.4 Verlust-WKTn der Teilströme und des Gesamtstromes über dem Korrelationskoeffizienten von Strom 2 bei $\eta_1 = 0.3$, $\eta_2 = 0.2$, $N = 30$, $D = 4$ und unkorreliertem Strom 1.

drei Werte $b_{max} = 4, 8, 16$ die Abhängigkeit der Verlust-WKT einer Testzelle vom Korrelationskoeffizienten dargestellt. b_{max} repräsentiert dabei die Anzahl von Eingangsleitungen, die Zellen produzieren, die auf dieselbe Ausgangsleitung vermittelt werden.

Abb. 7.2 läßt erkennen, daß die positive Korrelation im Eingabeprozess einen erheblichen Einfluß auf alle drei Dienstgüteparameter (Verlust-WKT, mittlere Wartezeit und ihre Standardabweichung als Maß für Verzögerungsschwankungen) hat. Am empfindlichsten ist die Verlust-WKT. Trotz gleichen Angebotes verschlechtert eine größere Maximalzahl von gleichzeitig ankommenden Zellen die Dienstgüte. Je größer diese Maximalzahl, desto mehr wird der Einfluß der Korrelation verdeckt.

Wenn man Einzelquellen annimmt, die für sich genommen unkorreliert sind, in der Überlagerung aber auf einen (positiv) korrelierten Strom führen, so belegen die numerischen Ergebnisse den negativen Einfluß der Korrelation, die zwangsläufig mit der Überlagerung von unabhängigen Quellen entsteht [119].

7.3 Korrelationseinfluß auf die QOS-Parameter der Teilströme

Nun werden zwei Ströme überlagert. Der Korrelationskoeffizient eines der Ströme wird variiert und der andere Strom unkorreliert gehalten. Wiederum wirkt sich eine positive Korrelation in einer Vergrößerung der Verlust-WKTn beider Ströme aus. Dabei ist der Einfluß der Korrelation im $\text{DMAP}_1^{[X]}$ -Strom natürlich erheblich höher, da auch das Teilangebot $\eta_1 = 0.3$ höher gewählt ist (Abb. 7.3). Der Einfluß der Korrelation des zweiten Stromes DMAP_2 ist nur merklich (Abb. 7.4), wenn das Angebot dieses Stromes nahe bei dem des ersten Stromes liegt.

Mit diesen Ergebnissen kann gefolgert werden, daß für die Zulassung einer "großen", positiv korrelierten Quelle (z.B. eine Videoquelle) die Einbeziehung eines Schätzwertes für seine Korrelation erforderlich ist. Gerade am Eingang der Teilnehmeranschlußleitung des

empfangenden Teilnehmers, die nur geringen "existierenden Verkehr" aufweisen dürfte und wo die Leitungsbitrate von 600 auf 150 Mbit/s fällt – hier durch $D = 4$ nachgebildet –, kann also der Korrelationseinfluß erheblich sein. Die Berechnungen beziehen sich auf AF und benutzen den SSMP(GEO,GEO) aus dem letzten Abschnitt für beide Eingabeströme.

7.4 Korrelation des Abgangsstromes

Ein Grund für die Notwendigkeit, korrelierte Eingabeprozesse bei der Modellierung von ATM-Multiplexern zu berücksichtigen, ist die Tatsache, daß der Abgangsstrom eines Multiplexers korreliert ist. Ein solcher Abgangsstrom bzw. ein Teil davon, tritt in einem nachfolgenden Multiplexer wieder als Eingabeprozess auf.

Das M/G/1-Typ-Wartesystem erlaubt auch die Untersuchung des Abgangsstromes. Abb. 7.5 zeigt den globalen Korrelationskoeffizienten der Zwischenabgangszeiten in Abhängigkeit des Angebotes für je zwei Werte des Korrelationskoeffizienten der Zwischenankunftszeiten ($\kappa = 0.2$, d.h. $\text{Corr}[A_n, A_{n+1}] = 0.02$ und $\kappa = 0.8$, d.h. $\text{Corr}[A_n, A_{n+1}] = 0.08$) und der Anzahl b_{max} gleichzeitig ankommender Zellen. Wie zu erwarten ist, stimmt $\text{Corr}[E_n, E_{n+1}]$ bei kleinen Angebotswerten mit $\text{Corr}[A_n, A_{n+1}]$ überein, da das Wartesystem das "Gefüge" der Zeiten zwischen zwei Zellen kaum verändert. Mit wachsendem Angebot wird der Zellenstrom durch das Wartesystem merklich beeinflußt (z.B. Zellenverlust), was sich auch in einem steigenden Korrelationskoeffizienten der Zwischenabgangszeiten ausdrückt. Die Ähnlichkeit der Zwischenabgangszeiten nimmt mit steigendem Angebot zu: Je höher das Angebot, desto häufiger befindet sich das System in der Betriebsperiode, in der mit konstantem Zeitabstand Zellen das System verlassen. Erstaunlich ist allerdings der Abfall von $\text{Corr}[E_n, E_{n+1}]$ für Angebotswerte größer 1. Im unendlichen Wartesystem würde sich ab einem solchen Angebot eine unendliche Warteschlange bilden. Der Abgangsprozeß wäre streng deterministisch. Hier dagegen gibt es weiterhin Freiperioden, allerdings mit wachsendem Angebot immer seltener. Dennoch ist der Abgangsstrom auch hier praktisch "deterministisch". Intuitiv würde man also einen weiteren Anstieg mit 1 als Asymptote erwarten. Aber ist ein deterministischer Prozeß korreliert mit Koeffizient 1 ?² Ein anderer Standpunkt wäre es, "deterministisch" als Grenzfall eines Erneuerungsprozesses aufzufassen, der (immer) unkorreliert ist. Damit würde sich das Abfallen erklären: Je höher das Angebot, desto näher kommt der Abgangsprozeß einer deterministischen Folge.

Wie schon bei anderen Zufallsgrößen beobachtet, wächst auch $\text{Corr}[E_n, E_{n+1}]$ mit κ , während die Batchgröße eine abschwächende Wirkung hat.

In [56] werden auch *lokale* Korrelationseigenschaften untersucht.

7.5 Effektive Bandbreite

Die numerischen Ergebnisse der bisherigen Abschnitte machen deutlich, daß (bei sonst gleichen Stromeigenschaften, z.B. unveränderter Ankunftsrate) positive Korrelation in einem der Ströme die Dienstgüteparameter beider Ströme in einer Überlagerung verschlechtern kann gegenüber dem unkorrelierten Fall. Die Korrelation der Ströme des existierenden Verkehrs wie auch diejenige der neu hinzukommenden Verbindung sollte daher von der Verbindungsannahmesteuerung in der Annahmeentscheidung berücksichtigt werden.

²Eine MK mit 2 Zuständen (0 und 1) hilft hier nicht weiter: Sie liefert bei einem Korrelationskoeffizienten von -1 die (deterministische) Folge $\dots 010101\dots$, bei einem Koeffizienten von 1 dagegen $\dots 000000\dots$ oder $\dots 111111\dots$, je nach Startwert.

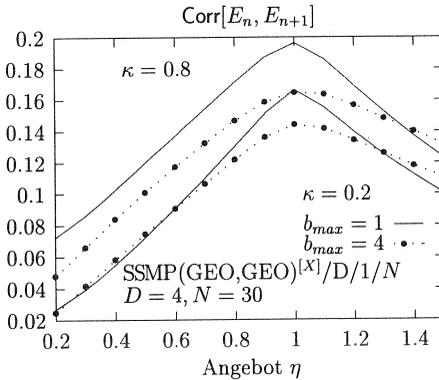


Abb. 7.5 Der Korrelationskoeffizient der Zwischenabgangszeiten in Abhängigkeit des Angebotes für zwei verschiedene κ des Eingabeprozesses und zwei verschiedene Batchgrößen.

Das Konzept der effektiven Bandbreite berücksichtigt implizit Korrelationen. Nach [100] ist die effektive Bandbreite $C_{\text{eff}}(P_{\text{Verlust}}, N)$ eines Ankunftsprozesses definiert als diejenige Abgangsrate eines (fiktiven) Wartesystems mit deterministischer Bedienzeit und Pufferplatzzahl N , bei der der Ankunftsprozess die vorgegebene Verlust-WKT P_{Verlust} erfährt. Betrachtet wird nun der Multiplexer am Eingang eines Teilabschnittes des untersuchten Weges, auf dem die neue Verbindung um Zulassung nachfragt. Die Warteplatzzahl sei N_0 . Mit Hilfe eines (fiktiven) Wartesystems dieser Puffergröße wird zunächst die effektive Bandbreite sowohl des existierenden Verkehrs wie auch der neuen Verbindung bestimmt. Ist die Summe beider Raten nicht größer als die Zellrate der Leitung des betrachteten Teilabschnittes, so kann die neue Verbindung zugelassen werden, andernfalls nicht.

Die Gültigkeit dieses Vorgehens wird mittels eines Modellprozesses verifiziert. Die Ergebnisse dieses Abschnittes wurden für AF gewonnen. Es sind die Ströme $\text{DMAP}_1^{[X]}$ und DMAP_2 vorgegeben. In zwei Wartesystemen $\text{DMAP}_1^{[X]}/D/1/N_0$ und $\text{DMAP}_2/D/1/N_0$ wird für $P_{\text{Verlust}} < 10^{-9}$ die jeweilige Abgangsrate $C_{\text{eff}}^{(i)}$ bestimmt. Dann wird der Summenprozess in ein Wartesystem $\text{DMAP}_1^{[X]} + \text{DMAP}_2/D/1/N_0$ eingespeist, wobei jetzt die Abgangsrate zu $C_{\text{eff}}^{(1)} + C_{\text{eff}}^{(2)}$ gewählt wird. Es zeigt sich, daß die Verlust-WKTn sogar kleiner sind als diejenigen der Teilströme bei Bestimmung der effektiven Bandbreite.

Eine Vergrößerung der Rate des Bedienprozesses im zeitdiskreten Wartesystem ist (hier) nur sinnvoll möglich, wenn die deterministische Bedienzeit groß genug gewählt ist: Für $D = 2$ kann die Bedienrate nicht mehr sinnvoll vergrößert werden, da nur noch die Halbierung bleibt.

Daher wird die Abgangsrate für ein genügend großes D bestimmt durch Variieren des Angebotes durch den Faktor α in beiden DMAPs, der $\text{Corr}[A_n, A_{n+1}]$ nicht tangiert:

$$\mathbf{C} = \begin{pmatrix} 1 - \alpha q_1 & 0 \\ 0 & 1 - \alpha q_2 \end{pmatrix} \quad \mathbf{D} = \begin{pmatrix} \alpha q_1 & 0 \\ 0 & \alpha q_2 \end{pmatrix} \cdot \mathbf{P}$$

Abb. 7.6 zeigt den Verlauf von P_{Verlust} als Funktion von D bei festem Angebot. Ab $D = 80$ ist P_{Verlust} nicht mehr von D abhängig.

Tab. 7.4 zeigt die resultierenden Verlust-WKTn bei Überlagerung der beiden Ströme in ei-

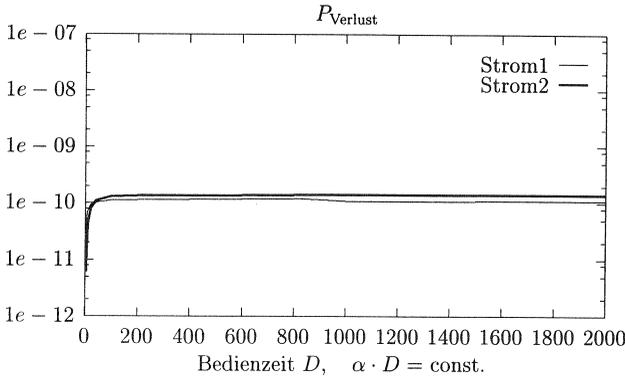


Abb. 7.6 P_{Verlust} als Funktion von D mit $\alpha \cdot D = \text{const.}$

D	α	P_{Verlust}	D	α	P_{Verlust}
100	0.006	1.13e-10	800	0.00075	1.20e-10
200	0.003	1.14e-10	1000	0.0006	1.09e-10
400	0.0015	1.16e-10	2000	0.0003	1.13e-10

Tab. 7.2 Verlust-WKTn bei Abgangsrate $1/D$ für Strom 1. $q_1 = 0.9$, $q_2 = 0.3$, $\text{Corr}[A_n, A_{n+1}] = 0.029$, $\sqrt{\text{Var}[A]}/E[A] = 1.55$, $b_{\text{max}} = 4$, $E[X] = 1.56$, $N = 30$.

D	α	P_{Verlust}	D	α	P_{Verlust}
80	0.025	1.2e-10	800	0.0025	1.39e-10
100	0.02	1.3e-10	1000	0.002	1.39e-10
200	0.01	1.34e-10	2000	0.001	1.40e-10

Tab. 7.3 Verlust-WKTn bei Abgangsrate $1/D$ für Strom 2. $q_1 = 0.9$, $q_2 = 0.3$, $\text{Corr}[A_n, A_{n+1}] = 0.029$, $\sqrt{\text{Var}[A]}/E[A] = 1.55$, $N = 30$.

nem Wartesystem mit einer Abgangsrate, die der Summe der effektiven Bandbreiten beider Ströme entspricht.

Natürlich belegen die obigen Ergebnisse nicht, daß die effektive Bandbreite immer korrekte Annahmeentscheidungen liefert. Eine theoretische Abschätzung wäre hier aussagekräftiger. Bisher wurden theoretische Betrachtungen nur zu *unendlichen zeitkontinuierlichen* Wartesystemen gemacht [100, 66, 102, 35] (*fluid flow*-Modell, Poisson-Eingabeprozess, MMPP), die diese universelle Eigenschaft der effektiven Bandbreite auch für endliche Wartesysteme und damit für den technisch relevanten Fall vermuten lassen.

Auch für größere Verlust-WKTn und kleinere Warteplatzzahl ($N = 10$, $P_{\text{Verlust}} < 10^{-4}$) ergeben sich für die Überlagerung bei effektiver Bandbreite als Abgangsrate Verlust-WKTn der Teilströme im Bereich der Werte, die für jeden Einzelstrom zur Ermittlung der effektiven Bandbreite festgelegt wurden.

7.6 Einfluß der Teilangebote auf Verluste

Zur Untersuchung des Einflusses des Angebotsverhältnisses wird von zwei bis auf das jeweilige Angebot gleichen Strömen ausgegangen ($\kappa^{(1)} = \kappa^{(2)} = 0.2$, kein Batch). Das Gesamtangebot wird konstant bei 0.8 gehalten und das Angebotsverhältnis η_1/η_2 verändert, s. Abb. 7.7. Es zeigt sich, daß

- der schwächere Strom eine leicht höhere Verlust-WKT erfährt, während diejenige des stärkeren Stromes sich schließlich dem Wert annähert, den ein Einzelstrom mit

Strom 1	effektive Bandbreite		P_{Verlust1}	P_{Verlust2}
	Strom 2	Strom1+2		
1/100	1/100	1/50	9.14e-11	3.82e-11
1/200	1/100	1/66	8.16e-11	3.55e-11
		1/67	1.14e-10	5.01e-11
1/100	1/200	1/66	7.22e-11	2.87e-11
		1/67	9.42e-11	3.77e-11
1/400	1/80	1/66	8.87e-11	4.06e-11
		1/67	1.29e-10	6.0e-11
1/400	1/200	1/133	9.99e-11	4.43e-11
		1/134	1.18e-10	5.24e-11
1/400	1/800	1/266	8.49e-11	3.35e-11
		1/267	9.08e-11	3.59e-11
1/800	1/400	1/266	1.01e-10	4.45e-11
		1/267	1.09e-10	4.85e-11
1/1000	1/800	1/444	9.53e-11	3.93e-11
		1/445	9.99e-11	4.13e-11
1/1000	1/1000	1/500	9.39e-11	3.76e-11
1/2000	1/800	1/571	1.09e-10	4.89e-11
		1/572	1.14e-10	5.10e-11

Tab. 7.4 Verlust-WKTn für beide Ströme in der Überlagerung in einem Wartesystem mit $C_{\text{eff}}^{(1)} + C_{\text{eff}}^{(2)}$ als Abgangsrate, $N = 30$.

Angebot 0.8 allein erfährt,

- die Verlust-WKTswerte beider Ströme nicht um Größenordnungen auseinanderliegen,
- die Verlust-WKT des Gesamtstromes sehr nahe an der des stärkeren Stromes liegt.

Dies deutet darauf hin, daß die Mischung von Strömen unterschiedlicher Dienstgütereorderungen (etwa Video mit Daten, d.h. geforderte Verlust-WKTn von 10^{-9} bzw. 10^{-3}) in demselben Multiplexer ohne Priorisierung von Zellen des sensibleren Stromes und bei Garantie der Dienstgütere beider Ströme nur möglich ist, wenn für beide Ströme etwa die Verlust-WKT von Videoverkehr (10^{-9}) eingehalten wird. Ströme unterschiedlicher Dienstgütereorderungen können mittels der VCI und VPI unterschieden werden. Andererseits zeigt die Angebotsabhängigkeit der Teilstrom-Verlust-WKTn, daß zur Dienstgütere garantie in einer Mischung aus einzelnen Strömen derselben Dienstklasse (aber mit unterschiedlichem Angebot) zusätzliche Maßnahmen erforderlich sein können, um für alle beteiligten Ströme dieser Klasse die (gleichen) Verlust-WKTswerte zu garantieren.

7.7 Gleichzeitigkeit von Ankünften und Abgängen

Von Interesse aus Sicht der Wartesystemtheorie ist die Frage, mit welcher WKT Ankünfte und Abgänge im zeitdiskreten Wartesystem gleichzeitig stattfinden. Im zeitkontinuierlichen Fall ist diese WKT immer Null. Mit den Überlegungen aus Abschnitt 4.1.7 bzw. 5.1.6 lassen sich für die bisher verwendeten Beispiele diese WKTn numerisch bestimmen. Sie ergeben sich (bei den berechneten Beispielen) als kaum abhängig von der Warteplatzzahl

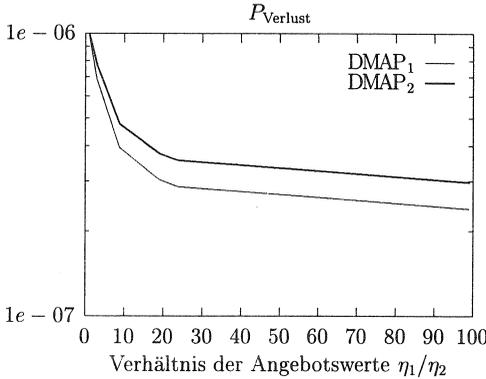


Abb. 7.7 Abhängigkeit der Teilstromverlust-WKTn vom Verhältnis der Angebote. Gesamtangebot $\eta_{ges} = 0.8$, $D = 4$, $\kappa^{(1)} = \kappa^{(2)} = 0.2$, $N = 30$, $\eta_1 > \eta_2$.

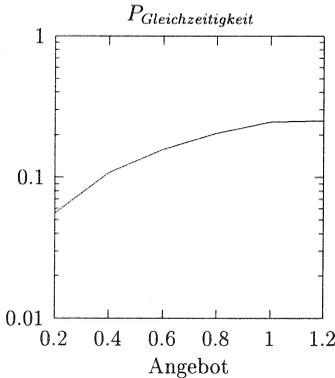


Abb. 7.8 WKTn für gleichzeitiges Auftreten von Ankünften und Abgängen in Abhängigkeit vom Angebot, DF. $D = 4$, $\kappa = 0.2$, $b_{max} = 1$, $N = 20$.

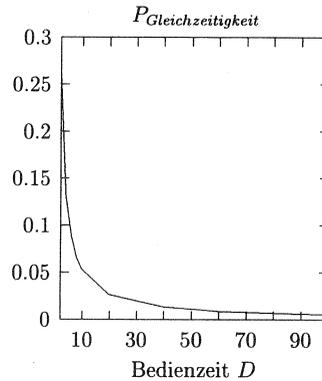


Abb. 7.9 WKTn für gleichzeitiges Auftreten von Ankünften und Abgängen in Abhängigkeit von der Bedienzeit, DF. $D = 4$, $\kappa = 0.2$, $b_{max} = 1$, $N = 20$.

($N \geq 10$). Sie sind etwa umgekehrt proportional zur Bedienzeit (bei festem Angebot) (Abb. 7.9) und steigen mit wachsendem Angebot an (Abb. 7.8). Die näherungsweise bestehende Unabhängigkeit von der Warteplatzzahl erklärt sich bei hohem Angebot dadurch, daß Abgänge fast immer im Abstand der deterministischen Bedienzeit erfolgen und die Anzahl der Ankünfte natürlich nicht von der Warteplatzzahl abhängt. Bei niedrigem Angebot, d.h. langen Freiperioden, können Abgänge und Ankünfte nur gleichzeitig auftreten, wenn sich eine (kleine) Warteschlange bildet. Diese wird aber von der Zahl der Warteplätze kaum beeinflusst. Die WKT gleichzeitiger Ankünfte und Abgänge sinkt, je näher man dem reinen Verlustsystem kommt ($N < 8$).

7.8 Schnelle Algorithmen für Markov-Ketten

Alle Modelle dieser Arbeit führen schließlich auf MKn. Ihre Zustands-WKTn werden benötigt, um Dienstgüteparameter zu berechnen. Für die Analyse des stationären Falls wird insbesondere die stationäre Lösung der MK benötigt. Der Vektor \underline{P} der stationären Zustands-WKTn ist bestimmt durch die Übergangsmatrix \mathbf{p} mit

$$\underline{P} \cdot \mathbf{p} = \underline{P}, \quad \underline{P} \cdot \underline{e} = 1 \quad (7.1)$$

Wenn die Lösung existiert, kann sie (nach Erzeugung der Koeffizientenmatrix des resultierenden linearen Gleichungssystems) mittels Gauß'scher Elimination oder sog. LU-Zerlegung³ [39] gewonnen werden. Dabei ist zu beachten, daß Gl. (7.1) (redundant) überbestimmt ist.

Sehr viel einfacher ist der folgende Algorithmus, der unmittelbar mit der Übergangsmatrix arbeitet. Außerdem hat er die wichtige Eigenschaft, daß in jedem Iterationsschritt ausschließlich positive Zahlen erzeugt werden, da keine Subtraktionen auftreten. Dadurch sind auch die Lösungen positiv. Beim Gauß-Algorithmus oder der LU-Zerlegung können durch Subtraktionen dem Betrag nach sehr kleine negative Zahlen entstehen, die als WKTn keinen Sinn ergeben. Für die Bestimmung von WKTn in der Größenordnung von 10^{-9} sind solche Resultate unbrauchbar, auch wenn die gefundene Lösung mit einigen negativen Zahlen die Gleichungen sehr gut erfüllt.

7.8.1 Der Algorithmus von Grassmann et al. (1985)

Sei $\mathbf{p} = (p_{ij})_{0 \leq i, j \leq m}$ die Übergangsmatrix einer irreduziblen MK. Dann erhält man die stationäre Lösung $\underline{P} = \underline{P} \cdot \mathbf{p}$ durch folgenden

Algorithmus

1. Für alle $0 \leq \mu, \nu \leq m$ setze $a_{\mu\nu} := p_{\mu\nu}$.
2. Für $\nu = m, m-1, \dots, 1$ berechne folgendes:
 $S := \sum_{j=0}^{\nu-1} a_{\nu j}$.
 $a_{i\nu} := a_{i\nu}/S$ für $0 \leq i \leq \nu-1$.
 $a_{ij} := a_{ij} + a_{i\nu}a_{\nu j}$ für $0 \leq i, j \leq \nu-1$.
3. $TOT := 1, r_0 := 1$.
4. Für $j = 1, 2, \dots, m$ berechne:
 $r_j := \sum_{k=0}^{j-1} r_k a_{kj}$.⁴
 Setze $TOT := TOT + r_j$.
5. Dann ist $P_j = r_j/TOT$ für $0 \leq j \leq m$.

Der Algorithmus verändert die Einträge der Matrix (a_{ij}). Um die gefundene Lösung nochmals durch Einsetzen überprüfen zu können, ist Schritt 1. erforderlich. Eine detaillierte Begründung des Algorithmus findet sich in [43].

³Die LU-Zerlegung einer Matrix bezeichnet ihre Darstellung als Produkt aus einer unteren und einer oberen Dreiecksmatrix.

⁴In [43], p. 1112 findet sich ein kleiner Druckfehler: Anstelle von $r_j = a_{0j} + \sum_{k=0}^{j-1} r_k a_{kj}$ muß $r_j = \sum_{k=0}^{j-1} r_k a_{kj}$ erscheinen. Der Summand a_{0j} ist schon in der Summe über k berücksichtigt wegen $r_0 = 1$. Nur mit dieser Korrektur arbeitet der Algorithmus korrekt.

7.8.2 Erweiterung des Grassmann-Algorithmus auf Blockmatrizen

In [70] wird der obige Algorithmus verallgemeinert auf den Fall von Blockmatrizen. Solche treten bei allen in dieser Arbeit behandelten Modellen auf. Durch Verwendung der Blockmatrizen kann sich eine erhebliche Beschleunigung ergeben. Ungünstig sind jedoch die an die Stelle der Zahlendivisionen tretenden Multiplikationen mit invertierten Matrizen. Zum einen ist die Matrixinversion selbst rechenaufwendig⁵. Andererseits führen Standardverfahren zur Invertierung Subtraktionen aus⁶, so daß die im Verlauf des Algorithmus invertierten Matrizen, obwohl sie theoretisch nur positive Einträge haben sollten [70], p. 166, häufig diese Eigenschaft durch Rundungsfehler verlieren. Dieses Problem tritt beispielsweise bei den Berechnungen in Kapitel 8 für große Warteplatzzahlen auf, wenn die Erweiterung von Kramer auf die folgende Blockmatrix vom M/G/1-Typ angewendet wird.

7.8.3 Anpassung des Kramer-Algorithmus an M/G/1-Typ-Matrizen

Die folgende Übergangsmatrix mit Blockstruktur tritt in Kapitel 8 auf:

$$\mathbf{Q} = \begin{pmatrix} \mathbf{D}_0 & \mathbf{D}_1 & \mathbf{D}_2 & \cdots & \mathbf{D}_N & \widehat{\mathbf{D}}_{N+1} \\ \mathbf{D}_0 & \mathbf{D}_1 & \mathbf{D}_2 & \cdots & \mathbf{D}_N & \widehat{\mathbf{D}}_{N+1} \\ \mathbf{0} & \mathbf{D}_0 & \mathbf{D}_1 & \cdots & \mathbf{D}_{N-1} & \widehat{\mathbf{D}}_N \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_0 & \cdots & \mathbf{D}_{N-2} & \widehat{\mathbf{D}}_{N-1} \\ \vdots & & & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{D}_0 & \widehat{\mathbf{D}}_1 \end{pmatrix} \quad (7.2)$$

mit $\widehat{\mathbf{D}}_k := \sum_{i=k}^{\infty} \mathbf{D}_i$ und den $(m \times m)$ -Matrizen \mathbf{D}_i , $i \in \mathbb{N}_0$. Die stationäre Lösung $(\underline{x}_0, \underline{x}_1, \dots, \underline{x}_{N+1})$ erfüllt die folgenden Gleichungen:

$$\text{Gl. } \boxed{k} \text{ für } 0 \leq k \leq N: \quad \underline{x}_k = \underline{x}_0 \mathbf{D}_k + \sum_{i=1}^{k+1} \underline{x}_i \mathbf{D}_{k+1-i} \quad (7.3)$$

$$\text{Gl. } \boxed{N+1}: \quad \underline{x}_{N+1} = \underline{x}_0 \widehat{\mathbf{D}}_{N+1} + \sum_{i=1}^{N+1} \underline{x}_i \widehat{\mathbf{D}}_{N+2-i} \quad (7.4)$$

Löst man die Gl. $\boxed{0}$ nach \underline{x}_0 auf und ersetzt damit in den verbleibenden Gleichungen \underline{x}_0 , und wird so mit Gl. \boxed{k} , $k = 1, 2, \dots, N+1$ weiterverfahren, d.h. nach \underline{x}_k aufgelöst und in den verbleibenden Gln. $\boxed{k+1}$, ..., $\boxed{N+1}$ \underline{x}_k ersetzt, so erhält man schließlich ein Gleichungssystem folgender Form:

$$\underline{x}_0 = \underline{x}_1 \mathbf{D}_0 (\mathbf{I} - \mathbf{D}_0)^{-1} \quad (7.5)$$

$$\underline{x}_k = \underline{x}_{k+1} \mathbf{D}_0 (\mathbf{I} - \mathbf{M}_1^{(N+2-k)})^{-1}, \quad 1 \leq k \leq N \quad (7.6)$$

$$\underline{x}_{N+1} = \underline{x}_{N+1} \mathbf{M}_1^{(1)} \quad (7.7)$$

Darin sind die Matrizen $\mathbf{M}_1^{(N+2-k)}$, $k = 1, \dots, N+1$ durch die folgende Prozedur zu gewinnen, die für jedes k weitere $k-1$ Matrizen $\mathbf{M}_i^{(N+2-k)}$, $i = 2, \dots, k$ benötigt:

⁵Die Anzahl der Operationen wächst kubisch mit dem Matrixrang.

⁶In [93] wird die LU-Zerlegung als die günstigste Methode zur Invertierung empfohlen.

Prozedur: Beginne mit den $N + 1$ Matrizen

$$\begin{cases} \mathbf{M}_i^{(N+1)} = \mathbf{D}_i + \mathbf{D}_0(\mathbf{I} - \mathbf{D}_0)^{-1}\mathbf{D}_i & , \quad 1 \leq i \leq N \\ \mathbf{M}_{N+1}^{(N+1)} = \widehat{\mathbf{D}}_{N+1} + \mathbf{D}_0(\mathbf{I} - \mathbf{D}_0)^{-1}\widehat{\mathbf{D}}_{N+1} \end{cases} \quad (7.8)$$

und speichere $\mathbf{M}_i^{(N+1)}$ in dem zwei-dimensionalen Feld \mathcal{F}_i .

Berechne dann für $k = N, N - 1, \dots, 2$ die k Matrizen

$$\begin{cases} \mathbf{M}_i^{(k)} = \mathbf{D}_i + \mathbf{D}_0(\mathbf{I} - \mathbf{M}_1^{(k+1)})^{-1}\mathbf{M}_{i+1}^{(k+1)} & , \quad 1 \leq i \leq k - 1 \\ \mathbf{M}_k^{(k)} = \widehat{\mathbf{D}}_k + \mathbf{D}_0(\mathbf{I} - \mathbf{M}_1^{(k+1)})^{-1}\mathbf{M}_{k+1}^{(k+1)} \end{cases} \quad (7.9)$$

und speichere $\mathbf{M}_i^{(k)}$ in $\mathcal{F}_{i+N+1-k}$, wodurch die Matrizen des vorherigen Schrittes bis auf $\mathbf{M}_1^{(k+1)}$ alle überschrieben werden. $\mathbf{M}_1^{(k+1)}$ bleibt dagegen in \mathcal{F}_{N+1-k} gespeichert.

Für $k = 1$ berechne $\mathbf{M}_1^{(1)} = \widehat{\mathbf{D}}_1 + \mathbf{D}_0(\mathbf{I} - \mathbf{M}_1^{(2)})^{-1}\mathbf{M}_2^{(2)}$.

Nach [70] ist $\mathbf{M}_1^{(1)}$ stochastisch und irreduzibel, wenn dieses auch für \mathbf{Q} gilt.

Der Grassmann-Algorithmus kann auf Gl. (7.7) angewendet werden und liefert eine stochastische Lösung für $\mathbf{M}_1^{(1)}$, die ein Vielfaches von \underline{x}_{N+1} ist. Sukzessives Berechnen der Gln. (7.6) und (7.5) ergibt ein Vielfaches der stationären Lösung $(\underline{x}_0, \dots, \underline{x}_{N+1})$, die schließlich durch Normierung gewonnen wird.

Mit der oben beschriebenen Speichermethode benötigt man bis auf geringen Speicherplatz zur Berechnung der Matrixprodukte und der Invertierten nur Speicher für $N + 1$ ($m \times m$)-Matrizen. Nach Abschluß der Prozedur befinden sich die $\mathbf{M}_1^{(k+1)}$ in den Feldern \mathcal{F}_{N+1-k} , $1 \leq k \leq N$, von wo sie für die Berechnungen in (7.6) und (7.5) abgerufen werden können. Für die Berechnungen in Kapitel 8 ist ein Haushalten mit den Speicherressourcen sehr wichtig, da die Phasenzahl der verwendeten DMAPs gerade mit der untersuchten Periodenlänge übereinstimmt. Bei dem in [9] vorgeschlagenen Verfahren, das etwa ebensoviel Speicher benötigt und nicht schneller ist als das oben beschriebene, wird die Situation größten Speicherbedarfs erst am Ende der gesamten Berechnungen erreicht, während dies hier gleich zu Beginn der Fall ist. Das erleichtert abzuschätzen, ob der zur Verfügung stehende Speicher für die Berechnung ausreicht.

Der so gewonnene Algorithmus war bei den Berechnungen in Kapitel 8 für $N > 30$ und zwei überlagerte Videoquellen nach Tabelle 8.1 numerisch nicht mehr stabil. Mit kleinen Änderungen kann die obige Prozedur auch auf die Übergangsmatrizen, die aus den SM-Kernen gemäß den Abbn. 5.2 und 5.3 für das DBMAP/G/1/N-Wartesystem folgen, angewendet werden. Dort ist die Beschleunigung anteilmäßig geringer, weil bei großen Phasenzahlen die Erzeugung der Blockmatrixeinträge \mathbf{A}_k und \mathbf{B}_k schon recht zeitintensiv ist.

8 Analytische Modelle für VBR-Videoquellen mit periodischer Korrelationsfunktion

8.1 Messungen an Videoquellen

Vom ATM-Konzept erwartet man für die Dauer der Übertragung von Videosignalen eine gleichbleibend gute Bildqualität, da das statistische Multiplexen *variable* Bitraten (VBR) zuläßt. Bei Video-Codern mit *konstanter* Ausgangsbitrate, wie sie in STM-Netzen gefordert wird, muß auf einen gröberen Quantisierer umgeschaltet werden, wenn bei einem Szenenwechsel die Differenzinformation zu groß wird. In dieser Übertragungsphase verringert sich die Bildqualität [68], p.834 .

Die meisten Untersuchungen zur statistischen Natur von VBR-Videoquellen gehen von der Annahme aus, daß ein ganzer Videorahmen (d.h. ein Bild) erst zwischengespeichert wird und dann Zellen mit konstanter Rate gesendet werden. Damit ergibt sich für die Dauer eines Rahmens eine konstante Zellrate [10, 110, 89, 78]. Außerdem resultiert dann ein Verlauf der Autokorrelationsfunktion der Bitrate, bzw. der Anzahl der in einem genügend kleinen Meßintervall erzeugten ATM-Zellen, der recht gut als negativ-exponentiell angenähert werden kann. Dies erklärt sich dadurch, daß periodische Schwankungen der Korrelationsfunktion, wie sie infolge von Ähnlichkeiten zwischen benachbarten Bildzeilen (sog. "spatial correlation") und Ähnlichkeiten zwischen entsprechenden Bildbereichen in zeitlich aufeinanderfolgenden Rahmen (sog. "temporal correlation") zu erwarten sind, durch die Zwischenspeicherung ("*Frame-Buffering*") absorbiert werden.

Ohne *Frame-Buffering* ergeben sich *meßbare Periodizitäten* in der Korrelationsfunktion der Anzahl von Zellen innerhalb des Meßintervalls [90, 46]. Bei der in [90, 46] gewählten Länge des Meßintervalls geht diese Periodizität auf Ähnlichkeiten zwischen benachbarten Rahmen zurück. Dagegen verschwindet die periodische Struktur in der Korrelationsfunktion der *Zwischenankunftszeiten* ohne *Frame-Buffering* fast vollständig [90].

Zur Nachbildung der Quelleneigenschaften wird in [46] ein (modifizierter) ARMA-Prozeß (Auto-Regressive Moving Average) eingesetzt. Es können die Modellparameter so angepaßt werden, daß Mittelwert, Varianz und Korrelationsfunktion des SP der Anzahl von Zellen innerhalb eines Meßintervalls mit den Werten der realen Quelle gut übereinstimmen. Der Meßaufbau besteht dabei aus einem Videocodern, der seine Zellen auf eine 150-Mbit/s-Leitung einspeist. Das Meßintervall entspricht der Länge von 64 ATM-Zellen auf dieser Leitung. Für drei verschiedene Szenen wurde die Korrelationsfunktion in guter Übereinstimmung rekonstruiert. Zwei davon sind in Abb. 8.1 und 8.2 dargestellt. Leider sind Wartesysteme mit ARMA-Eingabeprozess analytisch bisher nicht faßbar. Man muß also zur Leistungsbewertung von Videoverkehr in ATM-Multiplexern auch mit diesem Modell auf die Computersimulation zurückgreifen. Ferner sind die Verlust-WKTn für Videoquellen (10^{-9}) simulativ ohne spezielle Verfahren der Schätzung sehr seltener Ereignisse nicht bestimmbar. Daher beschränkt sich [46] auch auf die simulative Untersuchung der Wartezeiten. Details über ARMA-Modelle finden sich in [11, 24].

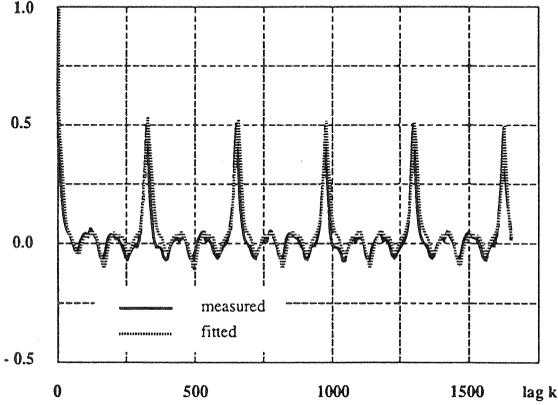


Abb. 8.1 Gemessene (und mittels ARMA-Prozeß nachgebildete) Korrelationsfunktion der Anzahl von Zellen der VBR-Video-Szene "head and shoulder scene" zitiert nach [46].

In [23] wird der Versuch unternommen, Videoquellen mittels eines GMDP (Geometrically Modulated Deterministic Process) nachzubilden. Ein Wartesystem mit GMDP-Eingabeprozess ist der analytischen Betrachtung zugänglich. Die Autoren berichten aber, daß der GMDP sich als ungeeignet herausstellt, weil er das periodische Verhalten der Korrelationsfunktion nicht nachbilden kann, sondern wieder nur den negativ-exponentiellen Abfall wie in Modellen für Videoquellen *mit* Frame-Buffering.

Im vorliegenden Kapitel wird die Eignung des DMAP^[X] zur Nachbildung der in [90, 46] gemessenen Eigenschaften von Videoquellen untersucht. Obwohl es hier um eine Einzelquelle geht, wird die Batch-ZV mit einbezogen. In [90] findet sich eine gemessene WF für die Anzahl von Zellen im Meßintervall, die mittels $x(v)$ direkt ins DMAP^[X]-Modell eingehen kann.

Die Periodizität der Korrelationsfunktion wird erreicht, indem der MK der Phasen des DMAP periodische Zustände gegeben werden.

Eine analytisch einsetzbare Repräsentation der Videoszenen, die *gleichzeitig* ein Meßintervall der Größe von 64 ATM-Zellen, die gemessene WF der im Meßintervall erzeugten Zellen und die gemessenen Spitzenwerte der Korrelationsfunktion bei einer Periode von 325 Meßintervallängen berücksichtigt, ergibt sich leider nicht. Dies liegt in der Hauptsache daran, daß für eine Periode der Länge 325 mindestens 325 Phasen vorgesehen werden müssen. Bei einer Überlagerung von zwei DMAPs dieser Phasenzahl erhielte man schon eine MK mit wenigstens 105625 Phasen. So große Zustandsräume sind numerisch kaum behandelbar. Dennoch kann für sehr viel kleinere Perioden der Einfluß der periodischen Korrelationsfunktion *im Vergleich* zum nicht-periodischen Verlauf in Modellen mit *Frame-Buffering* studiert werden. Für diesen Vergleich wird das DBMAP/D/1/N-Wartesystem für $D = 1$ und DF-Regel verwendet, wie in Abschnitt 5.2 beschrieben.

Die numerischen Ergebnisse deuten an, daß gerade negative Werte der Korrelationsfunktion zu Beginn der ersten Periode die Dienstgüteparameter gegenüber dem Fall mit *Frame-Buffering* erheblich verringern. Daraus ist zu schließen, daß die *periodischen Schwankungen* in der Korrelationsfunktion, die durch *Frame-Buffering* unterdrückt werden, durchaus von Vorteil sein können. Ein anderer Nachteil des *Frame-Buffering* ist bekanntlich die zusätzli-

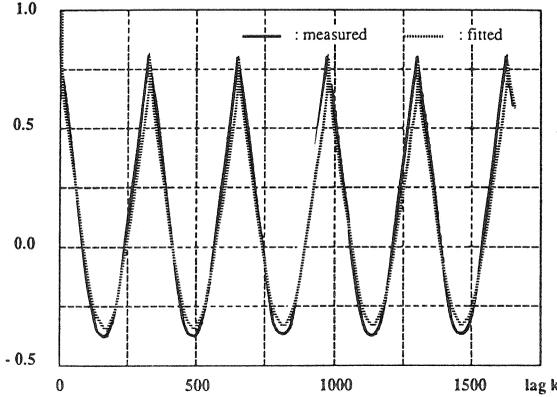


Abb. 8.2 Gemessene (und mittels ARMA-Prozeß nachgebildete) Korrelationsfunktion der Anzahl von Zellen der VBR-Video-Szene "low-flying aircraft scene" zitiert nach [46].

che Verzögerung des Signals infolge der Zwischenspeicherung eines ganzen Videorahmens. Es werden erst Zellen abgeschickt, wenn der Rahmen komplett gespeichert ist. Diese zusätzliche Verzögerung würde für den in [90, 46] verwendeten Codec 40 ms (etwa 1365 Zellen bei einer Informationsfeldlänge von 48 byte) ausmachen. Ein Grund für die Wahl relativ kleiner Zellen bei ATM war die Begrenzung der Verzögerung bei Paketierung [26].

Besonders günstig wäre somit eine Kodierung, die negative Werte der Korrelationsfunktion zu Beginn der ersten Periode erzeugt. Solche wurden in [90, 46] gemessen.

8.2 Die Korrelationsfunktion der Anzahl von Ankünften eines DMAP^[X]

Mit der Korrelationsfunktion $\text{Corr}(h)$ ist der Korrelationskoeffizient als Funktion seiner Ordnung gemeint. Bedeutet \mathcal{N}_t die Anzahl der Zellen, die im Intervall $(t; t + \tau]$ erzeugt werden, so ist

$$\text{Corr}(h) = \frac{E[\tau\mathcal{N}_t \cdot \tau\mathcal{N}_{t+h\tau}] - E^2[\tau\mathcal{N}_t]}{E[\tau\mathcal{N}_t^2] - E^2[\tau\mathcal{N}_t]} \quad (8.1)$$

Es wird also die Anzahl der Ankünfte im Intervall $(t; t + \tau]$ zu derjenigen im um h Intervalllängen τ verschobenen Intervall $(t + h\tau; t + (h + 1)\tau]$ in Beziehung gesetzt. Die drei bestimmenden Erwartungswerte in Gl. (8.1) ergeben sich nach Abschnitt 3.5.1 zu

$$\begin{aligned} E[\tau\mathcal{N}_t \cdot \tau\mathcal{N}_{t+h\tau}] &= \pi \left[\sum_{v=1}^{\infty} v\mathbf{P}(v, \tau) \right] (\mathbf{C} + \mathbf{D})^{(h-1)\tau} \left[\sum_{w=1}^{\infty} w\mathbf{P}(w, \tau) \right] \underline{\mathbf{e}} = \\ &= [\bar{x}^{(1)}]^2 \pi \left[\mathbf{D} \sum_{i=1}^{\tau} (\mathbf{C} + \mathbf{D})^{\tau-i} \right] (\mathbf{C} + \mathbf{D})^{(h-1)\tau} \left[\sum_{i=1}^{\tau} (\mathbf{C} + \mathbf{D})^{i-1} \mathbf{D} \right] \underline{\mathbf{e}} \quad , \end{aligned} \quad (8.2)$$

$$E[\tau\mathcal{N}_t^2] = 2[\bar{x}^{(1)}]^2 \pi \mathbf{D} \sum_{i=0}^{\tau-2} \sum_{\xi=0}^i (\mathbf{C} + \mathbf{D})^{\xi} \mathbf{D} \underline{\mathbf{e}} + \bar{x}^{(2)} \tau \pi \mathbf{D} \underline{\mathbf{e}} \quad , \quad (8.3)$$

$$E[\tau \mathcal{N}_t] = \bar{x}^{(1)} \tau \underline{\pi} \mathbf{D} \underline{e} \quad . \quad (8.4)$$

Mit $\underline{\pi}(\mathbf{C} + \mathbf{D}) = \underline{\pi}$ und $(\mathbf{C} + \mathbf{D})\underline{e} = \underline{e}$ ergibt sich leicht, daß (8.1) invariant ist gegen eine Vertauschung von \mathbf{C} und \mathbf{D} . Daher gibt es immer zwei DMAPs mit derselben Korrelationsfunktion, aber unterschiedlichen Ankunftsrate $\underline{\pi} \mathbf{D} \underline{e}$ und $1 - \underline{\pi} \mathbf{D} \underline{e}$, wenn nicht $\underline{\pi} \mathbf{D} \underline{e} = 0.5$.

8.3 Wahl der Matrizen \mathbf{C} und \mathbf{D} zur Erzeugung des periodischen Verlaufs

Für $\tau = 1$ reduziert sich Gl. (8.2) zu

$$E[\mathcal{N}_t \cdot \mathcal{N}_{t+h}] = [\bar{x}^{(1)}]^2 \underline{\pi} \mathbf{D} (\mathbf{C} + \mathbf{D})^{h-1} \mathbf{D} \underline{e} \quad . \quad (8.5)$$

Die übrigen Ausdrücke in Gl. (8.1) hängen nicht von h ab. Somit wird $\text{Corr}(h)$ periodisch in h , wenn dieses für $E[\mathcal{N}_t \cdot \mathcal{N}_{t+h}]$ gilt. Erreicht wird das, indem die Übergangsmatrix $\mathbf{C} + \mathbf{D}$ mit periodischen Zuständen¹ vorgegeben wird. Die einfachste Wahl, die bei gegebener Zustandszahl die größte Periode erzeugt, ist die Matrix der zyklischen Verschiebung

$$\mathbf{C} + \mathbf{D} = \mathbf{M}_{cycl} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 0 \\ & & & & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix} \quad . \quad (8.6)$$

Diese Wahl impliziert $\mathbf{D} = \text{Diag}(d_1, d_2, \dots, d_m) \cdot \mathbf{M}_{cycl}$ mit $0 \leq d_i \leq 1$, $1 \leq i \leq m$. Nachfolgend wird zur Vereinfachung der Notation der Vektor $\underline{d} := (d_1, d_2, \dots, d_m)$ verwendet. Die eindeutige stationäre Lösung der MK ergibt sich zu $\underline{\pi} = (\frac{1}{m}, \frac{1}{m}, \dots, \frac{1}{m})$. Weiter erhält man

$$E[\mathcal{N}_t \cdot \mathcal{N}_{t+h}] = \frac{[\bar{x}^{(1)}]^2}{m} \underline{d} \cdot [\mathbf{M}_{cycl}]^h \cdot \underline{d}^T$$

und

$$\text{Corr}(h) = \frac{m \underline{d} \cdot \mathbf{M}_{cycl}^h \cdot \underline{d}^T - (\underline{d} \cdot \underline{e})^2}{[\bar{x}^{(2)}]} \quad . \quad (8.7)$$

Die so erzeugte Korrelationsfunktion hat von h unabhängige Spitzenwerte. Ein Abfall der Spitzenwerte, wie er auch gemessen wurde, kann dadurch erreicht werden, daß z.B. die erste Zeile von \mathbf{M}_{cycle} durch $(0, a, b, 0, \dots, 0)$, $a + b = 1$ ersetzt wird². Mit $a = 0.99$ fallen die Spitzenwerte etwa geometrisch jeweils um den Faktor 0.999 entsprechend den Meßergebnissen in [90].

Für $\tau > 1$ erhält man meistens zusätzliche Oszillationen der Korrelationsfunktion innerhalb

¹Für MKn mit periodischen Zuständen s. z.B. [20].

²Mit dieser Ersetzung benötigt man aber wieder die allgemeine Formel (8.1) zur Berechnung der Korrelationsfunktion.

der Grundperiode. Der (anscheinend) einzige Fall, bei dem trotz $\tau > 1$ keine verfälschenden Oszillationen auftreten, ist für die Wahl

$$\mathbf{C}' := \begin{pmatrix} \mathbf{0} & \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & & \mathbf{0} \\ \vdots & & & \ddots & \vdots \\ \mathbf{0} & & & & \mathbf{I} \\ \mathbf{C} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}, \quad \mathbf{D}' := \begin{pmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & & \vdots \\ \mathbf{0} & \mathbf{0} & & \mathbf{0} \\ \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}$$

gegeben, wenn $\mathbf{C} + \mathbf{D} = \mathbf{M}_{cycl}$ und die Zahl der Hauptzeilen und -spalten gerade τ ist. Dann sind die von \mathbf{C}' und \mathbf{D}' für $\tau > 1$ bzw. \mathbf{C} und \mathbf{D} für $\tau = 1$ erzeugten Korrelationsfunktionen gleich. Dies ergibt sich daraus, daß für \mathbf{C}' und \mathbf{D}' der erste Summand in Gl. (8.3) verschwindet und Gl. (8.2) sich zu $[\bar{x}^{(1)}]^2 \tau \pi \mathbf{D} \mathbf{M}_{cycle}^{h-1} \mathbf{D} \underline{e}$ berechnet mit $\pi = (\frac{1}{m\tau}, \dots, \frac{1}{m\tau})$. Der von \mathbf{C}' und \mathbf{D}' erzeugte DMAP hat gegenüber dem von \mathbf{C} und \mathbf{D} erzeugtem um den Faktor τ größere Zwischenankunftszeiten, s. Abschnitt 3.4.1. Wichtig ist auch, daß die Spitzenwerte von $\text{Corr}(h)$ für $b_{max} > 1$ häufig stark verkleinert werden wegen $\bar{x}^{(2)} \geq [\bar{x}^{(1)}]^2$ [80].

8.4 Korrelationsfunktionen, die den gemessenen Verläufen ähnlich sind

Das Meßintervall in [90, 46] beträgt 64 ATM-Zellen, so daß infolge des kleinsten gemessenen Wertes der Zwischenankunftszeiten (7.63 Zellenlängen) die Maximalzahl von Zellen innerhalb des Meßintervalls sich zu 8 ergibt, [46], p. 286. Wählt man die Matrizen \mathbf{C}' und \mathbf{D}' mit $\tau = 64$, für die Untermatrizen \mathbf{C} und \mathbf{D} $m = 325$ (für eine Periode³ der Größe 325) sowie $b_{max} = 8$, so hätte man schon viele Eigenschaften der gemessenen Quellen erfaßt. Das bedeutet jedoch eine Anzahl von $325 \cdot 64 = 20800$ Phasen für eine einzige Quelle. Das sind für numerisch-analytische Untersuchungen bei weitem zu viele.

Somit muß darauf verzichtet werden, eine so große Periode im Modell nachzubilden. Dennoch kann die Bedeutung der periodischen Korrelationsfunktion im Vergleich mit der nicht-periodischen bei *Frame-Buffering* modellhaft studiert werden. Tabelle 8.1 zeigt die Komponenten des Vektors \underline{d} , die auf die in Abb. 8.3 bis 8.6 dargestellten Korrelationsfunktionen führen. Szene 1 ähnelt der "head and shoulders scene" und Szene 2 der "low-flying aircraft scene". Zur Repräsentation einer Quelle mit *Frame-Buffering* wird das Model für "video source with uniform activity level" in [10] verwendet. Als Parameter in der dort verwendeten Notation wird gewählt: $M = 10$ (Anzahl der Phasen), $\alpha = 0.07926$, $\beta = 0.0424$, $d = 10.1$. Die Ankunftsrate beträgt ebenfalls 0.345.

8.4.1 Korrelationsfunktion der Überlagerung zweier Quellen

Da die Überlagerung zweier $\text{DMAP}^{[X]}$ einen DBMAP ergibt, kann mit den Methoden in Abschnitt 3.5 auch die Korrelationsfunktion der Überlagerung berechnet werden. Sind $\text{DMAP}_1^{[X_1]}$ und $\text{DMAP}_2^{[X_2]}$ durch $(\mathbf{C}^{(1)}, \mathbf{D}^{(1)}, x_1(v))$ und $(\mathbf{C}^{(2)}, \mathbf{D}^{(2)}, x_2(v))$ gegeben, so erhält man für $\tau = 1$ aus den Gln. (3.52), (3.53) und (3.54) zusammen mit Gl. (3.40)

$$\sum_{v=1}^{\infty} v \mathbf{P}(v, 1) = (\mathbf{C}^{(1)} + \mathbf{D}^{(1)}) \otimes \mathbf{D}^{(2)} \bar{x}_2^{(1)} + \bar{x}_1^{(1)} \mathbf{D}^{(1)} \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)}),$$

³Die Messungen in [90, 46] wurden vor der Standardisierung der Zellgröße abgeschlossen und beziehen sich daher auf Zellen der Länge 32+4 Bytes; bei 48+5 Bytes dürfte sich die Periode aber nicht wesentlich verkleinern.

Szene	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}
1	0.2	0.15	0.1	0.05	0.5	0.1	0.3	0.05	0.05	0.05
2	0.2	0.18	0.16	0.14	0.11	0.05	0.01	0.01	0.01	0.01
3	0.6	0.4	0.4	0.3	0.3	0.2	0.15	0.15	0.15	0.1
Szene	d_{11}	d_{12}	d_{13}	d_{14}	d_{15}	d_{16}	d_{17}	d_{18}	d_{19}	d_{20}
1	0.7	0.8	0.7	0.15	0.2	0.3	0.3	0.4	0.9	0.9
2	0.01	0.01	0.4	0.4	0.7	0.9	0.9	0.9	0.9	0.9
3	0.1	0.1	0.15	0.15	0.15	0.4	0.6	0.8	0.8	0.9

Tab. 8.1 Die Werte für d_i , $1 \leq i \leq m$ der drei Szenen mit periodischer Korrelationsfunktion (Periode $m = 20$, Ankunftsrate 0.345).

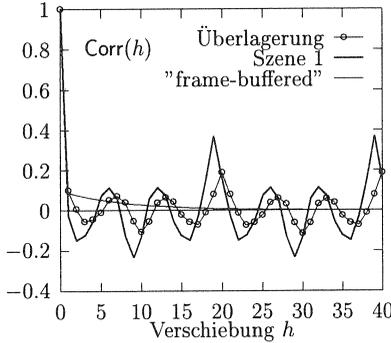


Abb. 8.3 Korrelationsfunktionen der Anzahl von Zellen: Szene 1, Modell für Frame-Buffering und die Überlagerung von beiden.

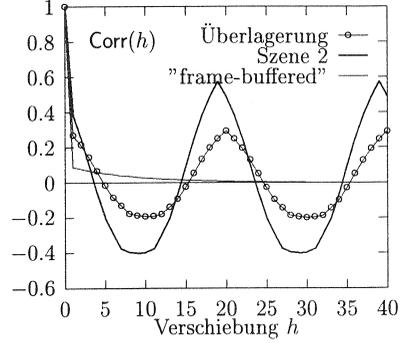


Abb. 8.4 Korrelationsfunktionen der Anzahl von Zellen: Szene 2, Modell für Frame-Buffering und die Überlagerung beider.

$$E[\mathcal{N}_t \mathcal{N}_{t+h}] = \pi \sum_{v=1}^{\infty} v \mathbf{P}(v, 1) \cdot \left[(\mathbf{C}^{(1)} + \mathbf{D}^{(1)}) \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)}) \right]^{h-1} \cdot \sum_{v=1}^{\infty} v \mathbf{P}(v, 1) \mathbf{e}, \quad (8.8)$$

$$E[\mathcal{N}_t^2] = \pi \left[(\mathbf{C}^{(1)} \otimes \mathbf{D}^{(2)}) \bar{x}_2^{(2)} + (\mathbf{D}^{(1)} \otimes \mathbf{C}^{(2)}) \bar{x}_1^{(2)} + (\mathbf{D}^{(1)} \otimes \mathbf{D}^{(2)}) (\bar{x}_2^{(2)} + 2\bar{x}_1^{(1)} \bar{x}_2^{(1)} + \bar{x}_1^{(2)}) \right] \mathbf{e}, \quad (8.9)$$

$$E[\mathcal{N}_t] = \pi \left(\left[(\mathbf{C}^{(1)} + \mathbf{D}^{(1)}) \otimes \mathbf{D}^{(2)} \right] \bar{x}_2^{(1)} + \left[\mathbf{D}^{(1)} \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)}) \right] \bar{x}_1^{(1)} \right) \mathbf{e}. \quad (8.10)$$

Die Abbn. 8.3 bis 8.6 enthalten auch Korrelationsfunktionen der Überlagerung zweier Szenen. Die Spitzenwerte der Überlagerung liegen zwischen denen der Einzelquellen. Der Wert h , bei dem zum ersten Mal ein negativer Wert auftritt, wächst in der Überlagerung.

8.4.2 Korrelationsfunktion der Zwischenankunftszeiten

Die Messungen in [46] beziehen sich auf die Anzahl der Ankünfte im Meßintervall, nicht aber das Korrelationsverhalten der Zwischenankunftszeiten. Für den hier betrachteten DMAP

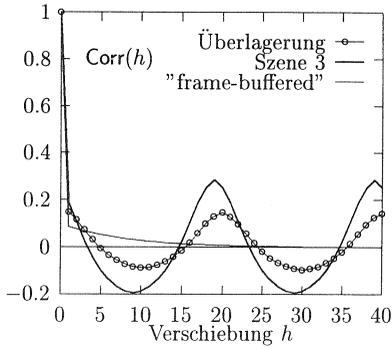


Abb. 8.5 Korrelationsfunktionen der Anzahl von Zellen: Szene 3, Modell für Frame-Buffering und die Überlagerung von beiden.

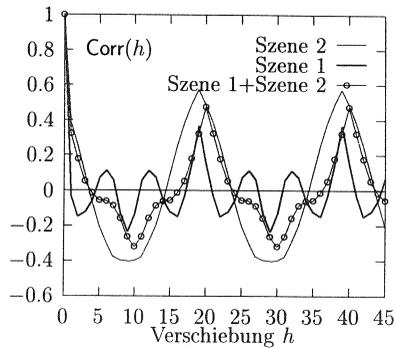


Abb. 8.6 Korrelationsfunktionen der Anzahl von Zellen: Szene 1, Szene 2 und die Überlagerung von beiden.

oder $\text{DMAP}^{[X]}$ kann auch die Korrelationsfunktion der Zwischenankunftszeiten angegeben werden. Mit $\mathbf{q}(t) = \mathbf{C}^{t-1} \mathbf{D}$ erhält man nach Abschnitt 3.5.2 $P\{A_n = t, A_{n+h} = t'\} = P\mathbf{q}(t)[(\mathbf{I} - \mathbf{C})^{-1} \mathbf{D}]^{h-1} \mathbf{q}(t') \underline{\mathbf{e}}$ und daraus

$$E[A_n \cdot A_{n+h}] = \frac{1}{\pi \mathbf{D} \underline{\mathbf{e}}} \cdot \pi [(\mathbf{I} - \mathbf{C})^{-1} \mathbf{D}]^h (\mathbf{I} - \mathbf{C})^{-1} \underline{\mathbf{e}}, \quad h \geq 1,$$

$$E[A_n^2] = \frac{1}{\pi \mathbf{D} \underline{\mathbf{e}}} [2\pi (\mathbf{I} - \mathbf{C})^{-1} \underline{\mathbf{e}} - 1], \quad E[A_n] = \frac{1}{\pi \mathbf{D} \underline{\mathbf{e}}}.$$

Abb. 8.7 zeigt beide Korrelationsfunktionen für Szene 2. Die Periode der Zwischenankunftszeiten beträgt nur 7, gegenüber 20 bei der Anzahl der Ankünfte. Wählt man \mathbf{C} und \mathbf{D} so, daß die zu Ankunftszeitpunkten eingebettete MK \mathbf{p} periodische Zustände hat, d. h.

$$\mathbf{C} := \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \ddots & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{C}_T \end{pmatrix}, \quad \mathbf{D} := \begin{pmatrix} \mathbf{0} & \mathbf{D}_1 & \cdots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \\ \mathbf{0} & \vdots & \ddots & \mathbf{D}_{T-1} \\ \mathbf{D}_T & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}, \quad (8.11)$$

so daß sich als Periode T ergibt und \mathbf{p} wie folgt lautet

$$\mathbf{p} = (\mathbf{I} - \mathbf{C})^{-1} \mathbf{D} = \begin{pmatrix} \mathbf{0} & (\mathbf{I} - \mathbf{C}_1)^{-1} \mathbf{D}_1 & \cdots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & & \ddots & (\mathbf{I} - \mathbf{C}_{T-1})^{-1} \mathbf{D}_{T-1} \\ (\mathbf{I} - \mathbf{C}_T)^{-1} \mathbf{D}_T & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}, \quad (8.12)$$

so erhält man oft keine Periodizität mehr für die Korrelationsfunktion der Anzahl der Ankünfte, s. Abb. 8.8. Daher benötigt man für eine vollständige Charakterisierung beide Korrelationsfunktionen.

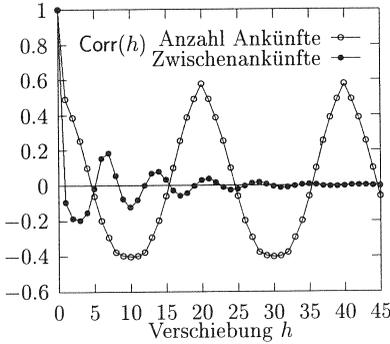


Abb. 8.7 Korrelationsfunktionen der Anzahl von Ankünften und der Zwischenankunftszeiten.

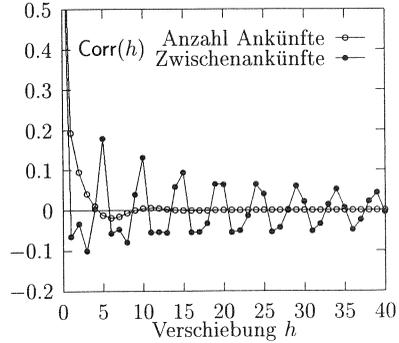


Abb. 8.8 Verschwindende Periode für die Anzahl der Ankünfte, aber Periodizität in der Korrelationsfunktion der Zwischenankunftszeiten. Periode $T = 5$.

8.5 Das Modell zum Szenenvergleich

Mittels eines Wartesystems, das einen einzelnen Multiplexer am Eingang eines Teilabschnittes nachbildet, werden Quellenmodelle mit periodischer Korrelationsfunktion, wie sie ohne *Frame-Buffering* auftritt, und solche ohne diese Periodizität in der Korrelationsfunktion verglichen. Dazu werden zwei Ströme überlagert und für beide getrennt die Dienstgüteparameter Verlust-WKTn, Mittelwert und Standardabweichung der Wartezeit berechnet. Dabei wird das DEMAP/D/1/N-Wartesystem $D = 1$ aus Abschnitt 5.2 für DF eingesetzt mit dem Überlagerungseingabeprozess $\text{DMAP}_1^{[X]} + \text{DMAP}_2$. Entsprechend den Ausführungen in Abschnitt 6.4 müssen die Teilstrom-Dienstgüteparameter berechnet werden.

Verlustwahrscheinlichkeiten

a) für Strom 1: $P_{\text{Verlust},1} := P\{\text{Strom 1 verliert eine Testzelle unter der Bedingung, daß der ankommende Batch eine Testzelle von Strom 1 enthält}\}$.

b) für Strom 2: $P_{\text{Verlust},2} := P\{\text{Strom 2 verliert eine Zelle unter der Bedingung, daß der ankommende Batch eine Zelle von Strom 2 enthält}\}$.

Mit diesen Definitionen erhält man für beide Ströme

$$P_{\text{Verlust},1} = \frac{\mathcal{T}_1 + \mathcal{T}_{12}}{\underline{\pi} [\mathbf{D}^{(1)} \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)})] \underline{e}}, \quad (8.13)$$

$$P_{\text{Verlust},2} = \frac{\mathcal{T}_{12}}{\underline{\pi} [(\mathbf{C}^{(1)} + \mathbf{D}^{(1)}) \otimes \mathbf{D}^{(2)}] \underline{e}} \quad (8.14)$$

mit $\underline{\pi} [(\mathbf{C}^{(1)} + \mathbf{D}^{(1)}) \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)})] = \underline{\pi}$, und (s. auch Gl. (5.41))

$$\begin{aligned}
\mathcal{T}_1 &:= \sum_{\mu=1}^{N+1} \sum_{v=N+2-(\mu-1)}^{b_{\max}^{(1)}} \frac{v - (N + 1 - (\mu - 1))}{v} \underline{x}_\mu \cdot (\mathbf{D}_v^{(1)} \otimes \mathbf{C}^{(2)}) \cdot \underline{e} + \\
&+ \sum_{v=N+2}^{b_{\max}^{(1)}} \frac{v - (N + 1)}{v} \underline{x}_0 \cdot (\mathbf{D}_v^{(1)} \otimes \mathbf{C}^{(2)}) \cdot \underline{e} \quad , \\
\mathcal{T}_{12} &:= \sum_{\mu=1}^{N+1} \sum_{v=N+1-(\mu-1)}^{b_{\max}^{(1)}} \frac{v + 1 - (N + 1 - (\mu - 1))}{v + 1} \underline{x}_\mu \cdot (\mathbf{D}_v^{(1)} \otimes \mathbf{D}^{(2)}) \cdot \underline{e} + \\
&+ \sum_{v=N+1}^{b_{\max}^{(1)}} \frac{v + 1 - (N + 1)}{v + 1} \underline{x}_0 \cdot (\mathbf{D}_v^{(1)} \otimes \mathbf{D}^{(2)}) \cdot \underline{e} \quad ,
\end{aligned}$$

wobei $\mathbf{D}_v^{(1)} = x_1(v)\mathbf{D}^{(1)}$ ist. Die Nenner repräsentieren die WKTn für die Bedingungsereignisse

$$\begin{aligned}
\pi \left[\mathbf{D}^{(1)} \otimes (\mathbf{C}^{(2)} + \mathbf{D}^{(2)}) \right] \underline{e} &= P\{\text{Strom 1 erzeugt eine Testzelle}\} \text{ und} \\
\pi \left[(\mathbf{C}^{(1)} + \mathbf{D}^{(1)}) \otimes \mathbf{D}^{(2)} \right] \underline{e} &= P\{\text{Strom 2 erzeugt eine Zelle}\}.
\end{aligned}$$

WFn der Teilstrom-Wartezeit

Wegen $D = 1$ ergibt sich die Wartezeit direkt aus der Belegung. Eine Zelle eines ankommenden Batch, die als ψ -te zugelassen wird, also nicht verlorenght, sieht eine Wartezeit, die sich additiv aus dem Belegungswert und der Anzahl $\psi - 1$ der Zellen desselben Batch zusammensetzt, die vor dieser Zelle ins System gelangt sind. Mit den Teilausdrücken $\mathcal{W}^{(\nu)}(\ell)$, $\nu = 1, 2$ und $\mathcal{W}^{(12)}(\ell)$ folgt

$$P\{W_1 = \ell\} = \frac{\mathbf{1}_{\ell=0}^N}{N_1} \left[\mathcal{W}^{(1)}(\ell) + \mathcal{W}^{(12)}(\ell) \right] \quad , \quad N_1 = \sum_{\ell=0}^N P\{W_1 = \ell\} \quad , \quad (8.15)$$

$$P\{W_2 = \ell\} = \frac{\mathbf{1}_{\ell=0}^N}{N_2} \left[\mathcal{W}^{(12)}(\ell) + \mathcal{W}^{(2)}(\ell) \right] \quad , \quad N_2 = \sum_{\ell=0}^N P\{W_2 = \ell\} \quad (8.16)$$

mit

$$\mathcal{W}^{(1)}(\ell) := \sum_{v=1}^{b_{\max}^{(1)}} \frac{1}{v} \sum_{\psi=1}^v \left[\mathbf{1}_{\ell=\psi-1} (\underline{x}_0 + \underline{x}_1) + \mathbf{1}_{\ell=\psi}^N \underline{x}_{\ell-\psi+2} \right] \cdot (\mathbf{D}_v^{(1)} \otimes \mathbf{C}^{(2)}) \cdot \underline{e} \quad ,$$

$$\mathcal{W}^{(2)}(\ell) := \left[\mathbf{1}_{\ell=0} (\underline{x}_0 + \underline{x}_1) + \mathbf{1}_{\ell=1}^N \underline{x}_{\ell+1} \right] \cdot (\mathbf{D}_0^{(1)} \otimes \mathbf{D}^{(2)}) \cdot \underline{e} \quad ,$$

$$\mathcal{W}^{(12)}(\ell) := \sum_{v=1}^{b_{\max}^{(1)}} \frac{1}{v+1} \sum_{\psi=1}^{v+1} \left[\mathbf{1}_{\ell=\psi-1} (\underline{x}_0 + \underline{x}_1) + \mathbf{1}_{\ell=\psi}^N \underline{x}_{\ell-\psi+2} \right] \cdot (\mathbf{D}_v^{(1)} \otimes \mathbf{D}^{(2)}) \cdot \underline{e} \quad .$$

8.6 Numerische Ergebnisse für periodische und nicht-periodische Korrelationsfunktionen

Die nachfolgenden numerischen Ergebnisse beziehen sich auf die in Tabelle 8.1 aufgeführten Szenen. Als Vergleichsszene für eine Quelle mit *Frame-Buffering* wird eine Videoquelle "with uniform activity level" gemäß [10] verwendet mit den Parametern (in der dort verwendeten Notation): Zahl der Phasen $M = 10$, $\alpha = 0.07926$, $\beta = 0.0424$, $d = 10.1$. Die Abbn. 8.3, 8.4 und 8.5 zeigen die Korrelationsfunktion des Modells mit *Frame-Buffering* unter der Bezeichnung "frame-buffered" jeweils mit den Korrelationsfunktionen der Überlagerung mit einem Modell ohne *Frame-Buffering*. Die Ankunftsrate beider Ströme in der Überlagerung ist 0.345.

1. GEO^[X]+DMAP/D/1/N (Tabelle 8.2). Der GEO^[X]-Strom repräsentiert eine Überlagerung von vielen Strömen auf einem Teilabschnitt, der DMAP eine der oben betrachteten Modelle für Videoszenen mit und ohne *Frame-Buffering*. Ziel ist der Vergleich der für die vier Modelle resultierenden Dienstgüteparameter. Tabelle 8.2 führt die Ergebnisse auf. $b_{max}^{(1)}$ repräsentiert die Maximalzahl der gleichzeitig ankommenden Zellen des GEO^[X]-Stromes. Solange $b_{max}^{(1)}$ im Vergleich zur Warteplatzzahl klein ist, ergibt sich ein beträchtlicher Unterschied in den Verlust-WKTn: Sie liegen für das Modell ohne *Frame-Buffering* (mit periodischer Korrelationsfunktion) weit unter denen des Modells mit *Frame-Buffering*. Bei den Mittelwerten der Wartezeit (und ähnlich bei den Standardabweichungen) ist der Unterschied wesentlich kleiner. Hierbei ist zu beachten, daß die Wartezeit wegen der Mittelung über die Zulassungsposition einer Testzelle von Strom 1 bzw. der Zelle von Strom 2 nicht direkt mit der zu Ankunftszeitpunkten gesehenen Belegung zusammenhängt. Für die Szenen 2 und 3 erhält man sehr ähnliche Ergebnisse wie die hier für Szene 1 aufgeführten, wobei diejenigen von Szene 2 und 3 etwas größer sind als diejenigen von Szene 1, aber immer noch weit unter denen des Modells mit *Frame-Buffering* liegen. Wie schon in den Abschnitten 7.3 und 7.6 gezeigt, liegen die Teilstrom-Verlust-WKTn der Ströme derselben Überlagerung nicht um Größenordnungen auseinander.

Verwendete Parameter

- Periode der Korrelationsfunktion: 20,
- Ankunftsrate beider Ströme: 0.345,
- Ankunftsrate der Überlagerung: 0.730,
- WF der Zwischenankunftszeiten des GEO-Stromes: $a(k) = (1 - q)q^{k-1}$ $q = 0.73$ für $b_{max}^{(1)} = 2$ und $q = 0.8$ für $b_{max}^{(1)} = 10$,
- WF der Mehrfachankünfte des GEO^[X]-Stromes:
 $x(v) = (1 - q_b)q_b^{v-1} / (1 - q_b^{b_{max}^{(1)}})$, $1 \leq v \leq b_{max}^{(1)}$, $q_b = 0.4$.

2. Einfluß des Abfalls innerhalb der ersten Periode (Abb. 8.9). Geht man bei der Erzeugung der periodischen Korrelationsfunktion von (d_1, d_2, \dots, d_m) über zum Vektor $(d_1, d_1, d_2, d_2, \dots, d_m, d_m)$, so bleibt die Ankunftsrate erhalten, und die Korrelationsfunktion wird in Richtung der h -Achse um den Faktor 2 gestreckt. Man erhält also eine Kurve mit um den Faktor 2 größerer Periode, aber auch um den Faktor 2 größeren Nulldurchgängen. Abb.

a) Teilstrom-Verlustwahrscheinlichkeiten

N	$b_{max}^{(1)} = 10$				$b_{max}^{(1)} = 2$			
	Szene1	GEO ^[X]	frame-buffered	GEO ^[X]	Szene1	GEO ^[X]	frame-buffered	GEO ^[X]
10	4.45e-4	8.32e-4	9.30e-4	1.48e-3	8.66e-6	1.23e-5	1.41e-4	1.75e-4
15	3.83e-5	7.22e-5	1.26e-4	1.98e-4	6.92e-8	9.78e-8	6.73e-6	8.32e-6
20	3.34e-6	6.30e-6	1.73e-5	2.70e-5	5.51e-10	7.80e-10	3.24e-7	4.00e-7
30	2.56e-8	4.82e-8	3.25e-7	5.09e-7	3.48e-14	4.92e-14	9.31e-10	7.53e-10

b) Mittlere Teilstrom-Wartezeit (gleiche Tendenz für die Standardabweichungen)

N	$b_{max}^{(1)} = 10$				$b_{max}^{(1)} = 2$			
	Szene1	GEO ^[X]	frame-buffered	GEO ^[X]	Szene1	GEO ^[X]	frame-buffered	GEO ^[X]
10	0.751	1.103	0.962	1.207	0.417	0.573	0.612	0.683
15	0.783	1.139	1.026	1.269	0.417	0.573	0.619	0.689
20	0.788	1.143	1.038	1.281	0.417	0.573	0.619	0.689
30	0.788	1.144	1.040	1.283	0.418	0.573	0.619	0.689

Tab. 8.2 Teilstrom-Dienstgüteparameter für beide Ströme des GEO^[X]+DMAP/D/1/N-Systems. DMAP repräsentiert Szene 1 (non-frame-buffered) nach Tabelle 8.1 oder das Modell mit Frame-Buffering aus [10].

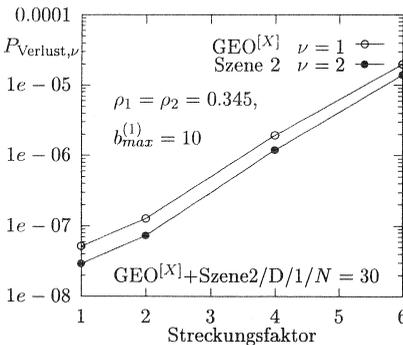


Abb. 8.9 Einfluß des Streckungsfaktors der Nulldurchgänge der Korrelationsfunktion auf die Teilstrom-Verlust-WKTn (am Beispiel von Szene 2).

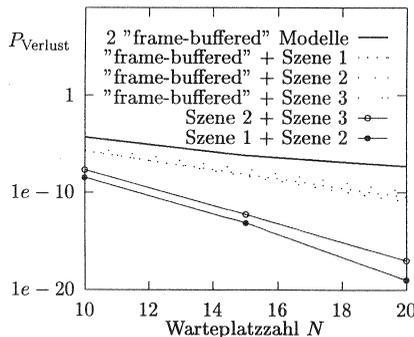


Abb. 8.10 Teilstrom-Verlust-WKTn für verschiedene Überlagerungen. Diese sind für beide Ströme in der Überlagerung gleich.

8.9 zeigt die Abhängigkeit der Verlust-WKTn beider Ströme vom Streckungsfaktor. Die Teilstrom-Verlust-WKTn wachsen mit diesem Streckungsfaktor. Je größer der Streckungsfaktor, desto näher kommt man der nicht-periodischen Korrelationsfunktion des Frame-

Buffering-Modells, die *nur positive* Werte hat. Damit liegt die Wirkung der Streckung auf die Verlust-WKTn also nicht in der Vergrößerung der Periode – diese sollte den Verlust eher vermindern, da die Spitzenwerte weiter auseinanderliegen –, sondern vielmehr in dem größeren ersten Nulldurchgang: Er führt dazu, daß negative Werte der Korrelationsfunktion erst für größere h erreicht werden. Weiterhin sieht man, daß die Verlust-WKTn für den Streckungsfaktor 6 diejenigen des *Frame-Buffering*-Modells überschreitet. Dies geht offenbar darauf zurück, daß die positiven Werte der um den Faktor 6 gestreckten periodischen Korrelationsfunktion der Szene 2 in Abb. 8.4 in der ersten Periode dann erheblich größer sind als diejenigen des *Frame-Buffering*-Modells.

3. Überlagerung zweier Modelle für Video (Abb. 8.10) Nun werden je zwei der Modelle für Video überlagert. Die Teilstromangebote sind wieder 0.345. Abb. 8.10 zeigt nun einen noch stärkeren Unterschied der Verlust-WKTn für die Modelle mit bzw. ohne *Frame-Buffering*. Immer dann, wenn in der Überlagerung ein Strom eine nicht-periodische Korrelationsfunktion hat, sind die Teilstrom-Verlust-WKTn erheblich höher.

Bemerkung: Eine Überlagerung zweier Szenen mit periodischer Korrelationsfunktion *ohne* Ersetzung der ersten Zeile von \mathbf{M}_{cycle} durch $(0, a, b, 0, \dots, 0)$, so daß ein geometrischer Abfall der Spitzenwerte zustande kommt, ergibt eine reduzierbare MK für die Belegung am Ende eines Slot. Diese ist dann nicht mehr mit den Algorithmen [70, 43] zu lösen. Die hier vorgestellten Ergebnisse wurden mit der Ersetzung berechnet.

8.6.1 Diskussion der Ergebnisse

Mit *Frame-Buffering* wird häufig die Vorstellung einer "Glättung" des Zellenstromes assoziiert. Dies ist in dem Sinne richtig, daß Zellratenschwankungen für die Dauer eines Videorahmens verschwinden. Dies führt nach den obigen Erkenntnissen jedoch nicht zu einer Glättung des Verkehrs in dem Sinne, daß die Dienstgüteparameter günstiger werden. Das ist auch nicht verwunderlich: Der Grenzfall der Übertragung mit konstanter Bitrate entspricht einem Pre-Buffering des Gesamtsignals und einer Übertragung mit konstanter Zellrate. Es ist eine allgemein anerkannte Vorstellung, daß der Multiplexgewinn durch statistisches Multiplexen, also Ausnutzen der statistischen Schwankungen, zunimmt. Aus Sicht der stochastischen Modellierung steht die hier gezogene Folgerung, daß negative Werte der Korrelationsfunktion zu Beginn der ersten Periode von wesentlicher Bedeutung sind, im Einklang mit den Ergebnissen von [51], wo ein negativer Korrelationskoeffizient erster Ordnung ($h = 1$) der Zwischenankunftszeiten niedrigere Verlust-WKTn als im unkorrelierten Fall ergibt. Diese Folgerung mißt der Periodizität an sich nur geringe Bedeutung zu. In [90] wurden für dieselben Szenen wie in [46] auch die Korrelationsfunktionen der Zwischenankunftszeiten gemessen. Diese zeigen *keine* ausgeprägte Periodizität, sondern eher ein "Zappeln" um die Nulllinie. Dies wird in dem hier entwickelten Modell ebenfalls recht gut nachgebildet, s. Abb. 8.7: die Periode für *Corr* der Zwischenankunftszeiten ist erheblich kleiner.

Schlußbemerkung und Ausblick

In der vorliegenden Arbeit wurden Verfahren entwickelt, mit denen für die ATM-Thematik wichtige Kenngrößen (insbesondere Verlustwahrscheinlichkeiten) zeitdiskreter Wartesysteme numerisch berechnet werden können. Es zeigt sich, daß mittels dieser Wartesysteme prinzipielle Aussagen hinsichtlich des Verhaltens von ATM-Multiplexern und an diese geknüpfte Konzepte, wie die Verbindungsannahmesteuerung, möglich sind.

Allerdings ist es infolge des sonst enormen Rechenaufwands meist notwendig, sich auf wenige Eigenschaften einer realen Quelle im Modell zu beschränken. Dies wird deutlich insbesondere bei der Nachbildung der periodischen Korrelationsfunktion für Videoquellen in Kapitel 8. Zwar kann man dem DMAP so viele Phasen geben, daß er die gemessene Periode nachbildet; jedoch lassen sich dann Dienstgüteparameter nicht mehr in vertretbarer Rechenzeit ermitteln. Das heißt aber nicht, daß die Modellierung wertlos ist. Prinzipielle Erkenntnisse sind auch für kleinere Perioden zu gewinnen, und das erheblich schneller als mittels der Computersimulation, die zudem die Schätzung von sehr seltenen Ereignissen noch nicht für jede denkbare Konstellation beherrscht (Wie wäre das RESTART-Verfahren [114, 109] etwa für die Behandlung zweier verschiedenener Nachrichtenströme zu modifizieren, wenn das Angebot einer Quelle im Verhältnis zur anderen klein ist ?).

Soll bei gegebener Periode auch die kleinste Zwischenankunftszeit der Zellen dieser Videoquelle nachgebildet werden, so gelingt das mit dem zeitverschobenen DMAP. Allerdings vergrößert sich dann der Zustandsraum und beeinflusst auch die Korrelationsfunktion des Zählprozesses. Diese Abhängigkeit dürfte aber auch das ARMA-Modell in [46] beinhalten, das infolge der Festlegung des Meßintervalls auf 64 ATM-Zellen die Quelleneigenschaften verfälscht [90].

Eine Schwierigkeit stellt der stochastische Begriff der Korrelation dar. Dieser mißt den Grad der Nicht-Unabhängigkeit, allerdings ohne allzusehr zu differenzieren. Auf das "Wesen" einer Verkehrsquelle unter Einbeziehung der globalen Korrelation (*eine* Zahl für den Grad der Nicht-Unabhängigkeit) zu schließen, ist schwierig. Das zeigt sich umgekehrt auch daran, daß die Parametrisierung eines SSMP durch Messungen (bisher) nur möglich ist, wenn lediglich zwei Zustände angenommen werden [30]. Ob ein (eindeutiges) Verfahren gefunden werden kann, wenn ein größerer Zustandsraum angenommen wird, ist fraglich; schon die Bestimmung der Übergangsmatrix einer dreigliedrigen MK aus ihrem Korrelationskoeffizienten führt nicht mehr auf einen eindeutigen Zusammenhang zwischen Matrixeinträgen und globalem Korrelationskoeffizienten (bei vorgegebener stationärer Lösung). Schließlich benötigt das Parametrisierungsverfahren zusätzlich den (gemessenen) lokalen Korrelationskoeffizienten, der stärker differenziert als der globale.

Für das DBMAP/G/1/N-Wartesystem existiert bisher noch keine geschlossene Darstellung der z -Transformierten der Wartezeit-WF genauso wenig wie ein geschlossener Ausdruck der Laplace-Transformierten (der Dichtefunktion) des zeitkontinuierlichen Gegenstücks. Weiterhin wäre von großem Interesse ein Verfahren zur Charakterisierung des Abgangsstromes *nach Teilströmen getrennt*, um den Einfluß eines Multiplexers auf einen einzelnen Strom

nicht nur *eingangsseitig* wie bei den hier entwickelten Modellen nachvollziehen zu können. Bisher ist dies nur simulativ möglich. Das führt schließlich auf die Problemstellung, eine Wartesystemkaskade zu behandeln, um so eine virtuelle Verbindung auf ihrem Weg durch ein Netz verfolgen zu können. Hier scheint die Computersimulation der einzige Weg zu sein, überhaupt zu Ergebnissen zu gelangen. Zur Verifizierung von Verlustwahrscheinlichkeiten sollte das RESTART-LRE-Verfahren dafür ausgebaut werden. Bisher wurde nur das Verzögerungsverhalten simulativ untersucht [118], das mit Blick auf Realzeitdienste allerdings auch von erheblicher Bedeutung für die Funktionsfähigkeit eines ATM-Netzes ist [77].

Anhang

A Beweise

A.1 Zum Beweis von Hilfssatz 6 (DF)

Im Falle $\ell > 0$:

a) Für $1 \leq v \leq N$ (vgl. Gl. (4.11)):

$$\begin{aligned}
 P_{ijkv}\{U_{n-1}^+ - A_{n-1} = \ell\} &= \sum_{\ell'=vD}^{ND+D} P_{ijkv}\{U_{n-1}^+ = \ell', A_{n-1} = \ell' - \ell\} = \\
 &= \sum_{\ell'=vD}^{ND} P_{ijkv}\{U_{n-1} = \ell' - vD, A_{n-1} = \ell' - \ell\} + \\
 &+ \sum_{\ell'=1+ND}^{ND+D} \left[P_{ijkv}\{U_{n-1} = \ell' - vD, A_{n-1} = \ell' - \ell\} + \sum_{\xi=0}^{v-1} P_{ijkv}\{U_{n-1} = \ell' - \xi D, A_{n-1} = \ell' - \ell\} \right] = \\
 &= \sum_{\ell'=0}^{ND-vD+D} \underbrace{P_{ijkv}\{U_{n-1} = \ell', A_{n-1} = \ell' - \ell + vD\}}_{=0 \text{ für } k \neq \ell' \text{ (} U_{n-1} = k \text{ steht in der Bedingung!)}} + \\
 &\quad + \sum_{\xi=0}^{v-1} \sum_{\ell'=1+ND-\xi D}^{D+ND-\xi D} \underbrace{P_{ijkv}\{U_{n-1} = \ell', A_{n-1} = \ell' - \ell + \xi D\}}_{=0 \text{ für } k \neq \ell'} = \\
 &= \mathbf{1}_{k=0}^{ND-vD+D} P_{ijkv}\{A_{n-1} = k - \ell + vD\} + \sum_{\xi=0}^{v-1} \mathbf{1}_{k=1+ND-\xi D}^{D+ND-\xi D} P_{ijkv}\{A_{n-1} = k - \ell + \xi D\} = \\
 &\quad \stackrel{\text{mit (4.8)}}{=} \mathbf{1}_{k=0}^{ND-vD+D} f_{ij}(k - \ell + vD) + \sum_{\xi=0}^{v-1} \mathbf{1}_{k=1+ND-\xi D}^{D+ND-\xi D} f_{ij}(k - \ell + \xi D) .
 \end{aligned}$$

b) Für $1 + N \leq v < \infty$ (vgl. Gl. (4.11)):

$$\begin{aligned}
 P_{ijkv}\{U_{n-1}^+ - A_{n-1} = \ell\} &= \sum_{\ell'=1+ND}^{D+ND} P_{ijkv}\{U_{n-1}^+ = \ell', A_{n-1} = \ell' - \ell\} = \\
 &= \sum_{\ell'=1+ND}^{D+ND-1} \sum_{\xi=0}^N P_{ijkv}\{U_{n-1} = \ell' - \xi D, A_{n-1} = \ell' - \ell\} + \\
 &+ \underbrace{\sum_{\xi=0}^N P_{ijkv}\{U_{n-1} = (N+1-\xi)D, A_{n-1} = (N+1)D - \ell\} + P_{ijkv}\{U_{n-1} = 0, A_{n-1} = (N+1)D - \ell\}}_{\ell'=ND+D} \\
 &= \sum_{\xi=0}^N \sum_{\ell'=1+(N-\xi)D}^{(N-\xi+1)D} P_{ijkv}\{U_{n-1} = \ell', A_{n-1} = \ell' - \ell + \xi D\} + P_{ijkv}\{U_{n-1} = 0, A_{n-1} = (N+1)D - \ell\} = \\
 &= \sum_{\xi=0}^N \mathbf{1}_{k=1+(N-\xi)D}^{(N+1-\xi)D} P_{ijkv}\{A_{n-1} = k - \ell + \xi D\} + \mathbf{1}_{k=0} P_{ijkv}\{A_{n-1} = ND + D - \ell\} = \\
 &\quad \stackrel{\text{mit (4.8)}}{=} \sum_{\xi=0}^N \mathbf{1}_{k=1+(N-\xi)D}^{(N+1-\xi)D} f_{ij}(k - \ell + \xi D) + \mathbf{1}_{k=0} f_{ij}(ND + D - \ell) .
 \end{aligned}$$

A.2 Formaler Beweis von Satz 2 (AF und DF)

Wenn ein w -Batch ankommt und kein Abgang gleichzeitig stattfindet, unterscheiden sich AF und DF nur darin, daß die gesehene Belegung bei u (das ist der letzte Abgangszeitpunkt vor t_A) für AF den Wert N und für DF $N + 1$ nicht überschreitet.

Es bezeichnet $w_{z_{k,r}^{i,j}}(t_A) := P\{\text{Ein } w\text{-Batch kommt zu } t_A \text{ an und sieht die Belegung } r, \text{ und der Eingabeprozess ist in Phase } j \text{ unter der Bedingung, daß bei } 0 \text{ ein Abgang } k \text{ Kunden im System zurückläßt und der Eingabeprozess sich in Phase } i \text{ befindet}\}$.

$w_{\Delta_{k,r}^{i,j}}(t_A)$ ist in Abschnitt 5.1.6 definiert und dort explizit angegeben.

$r = 0$:

$$w_{z_{k,r}^{i,j}}(t_A) = \sum_{j' \in \mathbb{E}} \left[\sum_{u=0}^{t_A-1} \hat{r}_{k,0}^{i,j'}(u) [\mathbf{P}(0, t_A - u - 1) \mathbf{D}_w]_{j',j} \right] + w_{\Delta_{k,r}^{i,j}}(t_A)$$

$1 \leq r \leq N$:

$$w_{z_{k,r}^{i,j}}(t_A) = \sum_{j' \in \mathbb{E}} \left[\sum_{u=0}^{t_A-2} \hat{r}_{k,0}^{i,j'}(u) \sum_{v=1}^{t_A-u-1} \sum_{w'=1}^r [\mathbf{D}_0^{v-1} \mathbf{D}_{w'} \mathbf{P}(r - w', t_A - u - v - 1) \cdot \mathbf{D}_w]_{j',j} \cdot H^c(t_A - u - v) + \sum_{u=0}^{t_A-1} \sum_{\ell=1}^r \hat{r}_{k,\ell}^{i,j'}(u) [\mathbf{P}(r - \ell, t_A - u - 1) \mathbf{D}_w]_{j',j} H^c(t_A - u) \right] + w_{\Delta_{k,r}^{i,j}}(t_A)$$

$r = N + 1$:

$$w_{z_{k,r}^{i,j}}(t_A) = \sum_{j' \in \mathbb{E}} \sum_{r'=N+1}^{\infty} \left[\sum_{u=0}^{t_A-2} \hat{r}_{k,0}^{i,j'}(u) \sum_{v=1}^{t_A-u-1} \sum_{w'=1}^{r'} [\mathbf{D}_0^{v-1} \mathbf{D}_{w'} \mathbf{P}(r' - w', t_A - u - v - 1) \cdot \mathbf{D}_w]_{j',j} \cdot H^c(t_A - u - v) + \sum_{u=0}^{t_A-1} \sum_{\ell=1}^{r'} \hat{r}_{k,\ell}^{i,j'}(u) [\mathbf{P}(r' - \ell, t_A - u - 1) \mathbf{D}_w]_{j',j} H^c(t_A - u) \right] + w_{\Delta_{k,r}^{i,j}}(t_A)$$

Grenzübergang $t_A \rightarrow \infty$ liefert $(\underline{x}_\ell = \underline{0}$ für $\ell > \mathcal{M}$)

$r = 0$:

$$z_r(w) = \underline{\Delta}_r \mathbf{D}_w + \frac{1}{E^*} \underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_w \quad (\text{A.1})$$

$1 \leq r \leq N$:

$$z_r(w) = \underline{\Delta}_r \mathbf{D}_w + \frac{1}{E^*} \sum_{\ell=1}^r [\underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell] \sum_{\sigma=0}^{\infty} \mathbf{P}(r - \ell, \sigma) \mathbf{D}_w H^c(\sigma + 1) \quad (\text{A.2})$$

$r = N + 1$:

$$z_r(w) = \underline{\Delta}_r \mathbf{D}_w + \frac{1}{E^*} \sum_{r'=N+1}^{\infty} \sum_{\ell=1}^{r'} [\underline{x}_0 (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_\ell + \underline{x}_\ell] \sum_{\sigma=0}^{\infty} \mathbf{P}(r' - \ell, \sigma) \mathbf{D}_w H^c(\sigma + 1) \quad (\text{A.3})$$

Unter Beachtung von $H^c(\sigma) = H^c(\sigma + 1) + h(\sigma + 1)$ und Hilfssatz 7 kann $z_r(w)$ für AF durch \underline{y}_r und für DF durch \underline{y}_r und $\underline{\Delta}_r$ ausgedrückt werden, wie in Satz 2 behauptet.

A.3 Der Abgangsprozeß des DBMAP/G/1/N-Systems (AF und DF)

$$\mathbf{A} := \sum_{k=0}^{\infty} \sum_{t=1}^{\infty} \mathbf{A}_k(t), \quad \mathbf{B} := \sum_{k=0}^{\infty} \sum_{t=1}^{\infty} \mathbf{B}_k(t), \quad \mathbf{A}\mathbf{e} = \mathbf{B}\mathbf{e} = \mathbf{e} \quad (\text{Zeilensummen von } \tilde{\mathbf{q}}),$$

$$\mathbf{A}(t)\mathbf{e} := \sum_{k=0}^{\infty} \mathbf{A}_k(t)\mathbf{e} = \widehat{\mathbf{D}}^t h(t)\mathbf{e} = h(t)\mathbf{e},$$

$$\mathbf{B}(t)\mathbf{e} := \sum_{k=0}^{\infty} \mathbf{B}_k(t)\mathbf{e} = \sum_{v=1}^t \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t-v),$$

$$\mathbf{B}_0 = (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_1 \mathbf{A}_0,$$

$$\mathbf{A}_0(t) = \mathbf{P}(0, t) h(t) = \mathbf{D}_0^t h(t),$$

$$\mathbf{B}_0(t) = \sum_{v=1}^t \mathbf{D}_0^{v-1} \mathbf{D}_1 \mathbf{A}_0(t-v) = \sum_{v=1}^t \mathbf{D}_0^{v-1} \mathbf{D}_1 \mathbf{D}_0^{t-v} h(t-v),$$

$$\mathbf{B}_1(t) = \sum_{v=1}^t \mathbf{D}_0^{v-1} [\mathbf{D}_1 \mathbf{A}_1(t-v) + \mathbf{D}_2 \mathbf{A}_0(t-v)] .$$

$$\tilde{\mathbf{q}}(t_2)\mathbf{e} = \begin{pmatrix} \sum_{k=0}^{\infty} \mathbf{B}_k(t_2)\mathbf{e} \\ \sum_{k=0}^{\infty} \mathbf{A}_k(t_2)\mathbf{e} \\ \vdots \\ \sum_{k=0}^{\infty} \mathbf{A}_k(t_2)\mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{B}(t_2)\mathbf{e} \\ \mathbf{A}(t_2)\mathbf{e} \\ \vdots \\ \mathbf{A}(t_2)\mathbf{e} \end{pmatrix} = \begin{pmatrix} \sum_{v=1}^{t_2} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t_2-v) \\ h(t_2)\mathbf{e} \\ \vdots \\ h(t_2)\mathbf{e} \end{pmatrix} .$$

$$\tilde{\mathbf{q}}\tilde{\mathbf{q}}(t_2)\mathbf{e} = \begin{pmatrix} \mathbf{B}_0\mathbf{B}(t_2)\mathbf{e} + (\mathbf{e} - \mathbf{B}_0\mathbf{e})h(t_2) \\ \mathbf{A}_0\mathbf{B}(t_2)\mathbf{e} + (\mathbf{e} - \mathbf{A}_0\mathbf{e})h(t_2) \\ \mathbf{A}h(t_2)\mathbf{e} \\ \vdots \\ \mathbf{A}h(t_2)\mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_0[\mathbf{B}(t_2)\mathbf{e} - h(t_2)\mathbf{e}] + h(t_2)\mathbf{e} \\ \mathbf{A}_0[\mathbf{B}(t_2)\mathbf{e} - h(t_2)\mathbf{e}] + h(t_2)\mathbf{e} \\ h(t_2)\mathbf{e} \\ \vdots \\ h(t_2)\mathbf{e} \end{pmatrix}$$

$$= \begin{pmatrix} (\mathbf{I} - \mathbf{D}_0)^{-1} \mathbf{D}_1 \mathbf{A}_0 \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t_2-v) - h(t_2)\mathbf{e} \right] \\ \mathbf{A}_0 \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t_2-v) - h(t_2)\mathbf{e} \right] \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} h(t_2)\mathbf{e} \\ h(t_2)\mathbf{e} \\ h(t_2)\mathbf{e} \\ \vdots \\ h(t_2)\mathbf{e} \end{pmatrix} .$$

$$\tilde{\mathbf{q}}(t_1)\tilde{\mathbf{q}}(t_2)\mathbf{e} = \begin{pmatrix} \mathbf{B}_0(t_1)[\mathbf{B}(t_2) - \mathbf{A}(t_2)]\mathbf{e} + \mathbf{B}(t_1)\mathbf{A}(t_2)\mathbf{e} \\ \mathbf{A}_0(t_1)[\mathbf{B}(t_2) - \mathbf{A}(t_2)]\mathbf{e} + \mathbf{A}(t_1)\mathbf{A}(t_2)\mathbf{e} \\ \sum_{k=0}^{\infty} \mathbf{A}_k(t_1)h(t_2)\mathbf{e} \\ \vdots \\ \sum_{k=0}^{\infty} \mathbf{A}_k(t_1)h(t_2)\mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathcal{T}_I(t_1, t_2) \\ \mathcal{T}_{II}(t_1, t_2) \\ h(t_1)h(t_2)\mathbf{e} \\ \vdots \\ h(t_1)h(t_2)\mathbf{e} \end{pmatrix} ,$$

$$\begin{aligned} \mathcal{T}_I(t_1, t_2) &= \mathbf{B}_0(t_1) \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t_2-v) - h(t_2)\mathbf{e} \right] + \\ &+ \sum_{v=1}^{t_1} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t_1-v)h(t_2) \quad , \end{aligned}$$

$$\mathcal{T}_{II}(t_1, t_2) = \mathbf{A}_0(t_1) \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1} (\mathbf{I} - \mathbf{D}_0)\mathbf{e} h(t_2-v) - h(t_2)\mathbf{e} \right] + h(t_1)h(t_2)\mathbf{e} .$$

$$\tilde{\mathbf{q}}(t_1)\tilde{\mathbf{q}}(t_2)\mathbf{e} = \begin{pmatrix} \mathbf{B}(t_1)h(t_2)\underline{\mathbf{e}} \\ \mathbf{A}(t_1)h(t_2)\underline{\mathbf{e}} \\ \vdots \\ \mathbf{A}(t_1)h(t_2)\underline{\mathbf{e}} \end{pmatrix} + \begin{pmatrix} \left[\mathbf{B}_0(t_1)(\mathbf{I} - \mathbf{D}_0)^{-1}\mathbf{D}_1\mathbf{A}_0 + \mathbf{B}_1(t_1)\mathbf{A}_0 \right] \cdot \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1}(\mathbf{I} - \mathbf{D}_0)\underline{\mathbf{e}}h(t_2 - v) - h(t_2)\underline{\mathbf{e}} \right] \\ \left[\mathbf{A}_0(t_1)(\mathbf{I} - \mathbf{D}_0)^{-1}\mathbf{D}_1\mathbf{A}_0 + \mathbf{A}_1(t_1)\mathbf{A}_0 \right] \cdot \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1}(\mathbf{I} - \mathbf{D}_0)\underline{\mathbf{e}}h(t_2 - v) - h(t_2)\underline{\mathbf{e}} \right] \\ \mathbf{A}_0(t_1)\mathbf{A}_0 \cdot \left[\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1}(\mathbf{I} - \mathbf{D}_0)\underline{\mathbf{e}}h(t_2 - v) - h(t_2)\underline{\mathbf{e}} \right] \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}.$$

Damit erhält man im allgemeinen Fall die WF bzw. Verbund-WFn (erster bzw. zweiter Ordnung $f_1(t_1, t_2) = P\{E_n = t_1, E_{n+1} = t_2\}$ bzw. $f_2(t_1, t_2) = P\{E_n = t_1, E_{n+2} = t_2\}$) der Zwischenabgangszeiten

$$\begin{aligned} f(t) &= \underline{x}_0(\mathbf{B}(t) - \mathbf{A}(t))\underline{\mathbf{e}} + h(t) = \\ &= \underline{x}_0 \sum_{v=1}^t \mathbf{D}_0^{v-1}(\mathbf{I} - \mathbf{D}_0)\underline{\mathbf{e}}h(t - v) + (1 - \underline{x}_0\underline{\mathbf{e}})h(t) \quad , \end{aligned} \quad (\text{A.4})$$

$$\begin{aligned} f_1(t_1, t_2) &= \left(\underline{x}_0 \sum_{v=1}^{t_1} \mathbf{D}_0^{v-1}\mathbf{D}_1\mathbf{D}_0^{t_1-v}h(t_1 - v) + \underline{x}_1\mathbf{D}_0^{t_1}h(t_1) \right) \cdot \\ &\quad \cdot \left(\sum_{v=1}^{t_2} \mathbf{D}_0^{v-1}(\mathbf{I} - \mathbf{D}_0)\underline{\mathbf{e}}h(t_2 - v) - h(t_2)\underline{\mathbf{e}} \right) + \\ &\quad + \underline{x}_0 \sum_{v=1}^{t_1} \mathbf{D}_0^{v-1}(\mathbf{I} - \mathbf{D}_0)\underline{\mathbf{e}}h(t_1 - v)h(t_2) + (1 - \underline{x}_0\underline{\mathbf{e}})h(t_1)h(t_2) \quad , \end{aligned} \quad (\text{A.5})$$

$$\begin{aligned} f_2(t_1, t_2) &= \underline{\mathbf{x}}\tilde{\mathbf{q}}(t_1)\tilde{\mathbf{q}}(t_2)\mathbf{e} = \\ &\quad \underline{x}_0\mathbf{B}(t_1)h(t_2)\underline{\mathbf{e}} + (1 - \underline{x}_0\underline{\mathbf{e}})h(t_1)h(t_2) + \\ &\quad \left(\underline{x}_0[\mathbf{B}_0(t_1)(\mathbf{I} - \mathbf{D}_0)^{-1}\mathbf{D}_1 + \mathbf{B}_1(t_1)] + \right. \\ &\quad \left. \underline{x}_1[\mathbf{A}_0(t_1)(\mathbf{I} - \mathbf{D}_0)^{-1}\mathbf{D}_1 + \mathbf{A}_1(t_1)] + \underline{x}_2\mathbf{A}_0(t_1) \right) \cdot \mathbf{A}_0 \cdot (\mathbf{B}(t_2)\underline{\mathbf{e}} - h(t_2)\underline{\mathbf{e}}) \quad . \end{aligned} \quad (\text{A.6})$$

B Abkürzungen

ATM	Asynchronous Transfer Mode
STM	Synchronous Transfer Mode
QOS	Quality of Service
VBR	Variable Bit Rate
CBR	Constant Bit Rate
FIFO	First In First Out
LRE	Limited Relative Error
LGS	lineares Gleichungssystem
SP	stochastischer Prozeß
i.i.d.	independent identically distributed
SSS	strict sense stationary
[S]SMP	[Special] Semi-Markov-Process
MEP	Markov-Erneuerungsprozeß
[B]MAP	[Batch] Markovian Arrival Process
MMPP	Markov Modulated Poisson Process
N	Neuts Process (nicht zu verwechseln mit der Puffergröße N)
DBMAP	Discrete Batch Markovian Arrival Process
DMAP	Discrete Markovian Arrival Process
GMDP	Geometrically Modulated Deterministic Process
GEO	Bernoulli-Prozeß (geometrisch verteilte Zwischenankunftszeiten)
D	Deterministic (Bezeichnung für den stochastischen Prozeß, dagegen meint D den jeweiligen deterministischen Wert)
GI	General independent
G	General (d. h. auch korreliert)
MK	Markov-Kette
AF	"Arrival First" (Zulassungsregel bei zeitdiskreten Wartesystemen: Bei Gleichzeitigkeit von Ankunft und Abgang findet erst die Ankunft statt.)
DF	"Departure First" (Gegenstück zu AF)
WKT	Wahrscheinlichkeit
ZV	Zufallsvariable
WF	Wahrscheinlichkeitsfunktion
VF	Verteilungsfunktion
w -Batch	Mehrfachankunft von w Kunden (Zellen) gleichzeitig
VCI	Virtual Channel Identifier
VPI	Virtual Path Identifier

C Symbole

$\ell = \lfloor x \rfloor$	größte ganze Zahl ℓ , die kleiner oder gleich x ist
$\ell = \lceil x \rceil$	kleinste ganze Zahl ℓ , die größer oder gleich x ist
$\mathbf{1}_{\ell=a}^b$; $\mathbf{1}_{\text{[AUSDRUCK]}}$	Indikatorfunktion: $\mathbf{1}_{\ell=a}^b = 1$ für $a \leq \ell \leq b$ und $= 0$ sonst; $\mathbf{1}_{\text{[AUSDRUCK]}} = 1$, falls AUSDRUCK wahr ist, und $= 0$ sonst.
$[matrix]_{ij}$	Matrizelement in der i -ten Zeile und der j -ten Spalte
$[vector]_i$	i -te Vektorkomponente
$\mathbf{Diag}(f_1, \dots, f_m)$	Diagonalmatrix mit m Zeilen und Spalten
\underline{e} bzw. \mathbf{e}	$(1, 1, \dots, 1)^T$, m Komponenten bzw. soviele Komponenten, wie aus dem Zusammenhang nötig
$E[\cdot]$	Erwartungswert
$\text{Var}[\cdot]$	Varianz
$\text{Cov}[\cdot]$	Kovarianz
$\text{Corr}[\cdot]$	(globaler) Korrelationskoeffizient
\otimes	Kronecker Produkt, verknüpft die $(m_1 \times m_2)$ -Matrix \mathbf{A} mit der Matrix \mathbf{B} nach der Vorschrift
	$\mathbf{A} \otimes \mathbf{B} := \begin{pmatrix} [\mathbf{A}]_{11} \cdot \mathbf{B} & \dots & [\mathbf{A}]_{1m_2} \cdot \mathbf{B} \\ \vdots & & \vdots \\ [\mathbf{A}]_{m_1 1} \cdot \mathbf{B} & \dots & [\mathbf{A}]_{m_1 m_2} \cdot \mathbf{B} \end{pmatrix}$
T_n	n -ter Ankunftszeitpunkt
A_n	$A_n := T_{n+1} - T_n$, n -te Zwischenankunftszeit
A	$A = \{A_n; n \in \mathbb{N}_0\}$
$\mathbf{Q}(t)$	SM-Kern $[\mathbf{Q}(t)]_{ij} := Q_{ij}(t) = P\{S_{n+1} = j, T_{n+1} - T_n \leq t S_n = i\}$
$\mathbf{q}(t)$	$[\mathbf{q}(t)]_{ij} = q_{ij}(t) := P\{S_{n+1} = j, T_{n+1} - T_n = t S_n = i\} = p_{ij} f_{ij}(t)$
$f_{ij}(t)$ bzw. $F_{ij}(t)$	WF bzw. VF der Verweilzeit im Zustand i bis zum Übergang nach j
\mathbf{P}	$= \sum_{t=0}^{\infty} \mathbf{q}(t)$, meistens die zu Ankunftszeiten eingebettete MK
\underline{P}	Vektor der stationären Zustands-WKTn zu Ankunftszeitpunkten
S_n	$= S(T_n)$ Zustands-ZV eines SMP beim n -ten Ankunftszeitpunkt
S	$S = \{S_n; n \in \mathbb{N}_0\}$
U_n	Restarbeit beim n -ten Ankunftszeitpunkt <i>ohne</i> Berücksichtigung dieser Ankunft
$P_{ijk\ell}\{U_n = \ell\}$	$= P\{U_n = \ell S_n = j, U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\}$
$P_{ikv}\{U_n = \ell, S_n = j\}$	$= P\{U_n = \ell, S_n = j U_{n-1} = k, S_{n-1} = i, X_{n-1} = v\}$
$u_n(k, i)$	$= P\{U_n = k, S_n = i\}$
$u_n(k, i, v)$	$= P\{U_n = k, S_n = i, X_n = v\}$
$[\underline{u}_n(k)]_i$; $\underline{u}(k)$	$[\underline{u}_n(k)]_i = u_n(k, i)$; $\underline{u}(k) = \lim_{n \rightarrow \infty} u_n(k)$, stationäre Lösung
\underline{u}_n ; \underline{u}	$\underline{u}_n = (\underline{u}_n(0), \dots, \underline{u}_n(ND + D))$; $\underline{u} = \lim_{n \rightarrow \infty} \underline{u}_n$
$\mathbf{C}, \mathbf{D}, \mathbf{D}_v, \widehat{\mathbf{D}}$; $\boldsymbol{\pi}$	$(m \times m)$ -Matrizen, die den DMAP, bzw. DBMAP bestimmen. $\widehat{\mathbf{D}} = \sum_{w=0}^{b_{\max}} \mathbf{D}_w$; $\boldsymbol{\pi} \widehat{\mathbf{D}} = \boldsymbol{\pi}$ (stationäre Lösung)
$\mathbf{P}(w, t)$	Zählfunktion des DBMAP

$\tilde{\mathbf{q}}(t)$	SM-Kern des zu Abgangzeitpunkten eingebetteten MEP im DBMAP/G/1/N-Wartesystem
$\tilde{\mathbf{q}}$	$= \sum_{t=1}^{\infty} \tilde{\mathbf{q}}(t)$
$h(t), H(t)$ bzw. $H^c(t)$	WF, VF bzw. komplementäre VF der Bedienzeit
$\mathbf{A}_k^{[j l l]}, \mathbf{B}_k^{[j l l]},$ $\mathbf{A}_k^{[j l l]}, \mathbf{B}_k^{[j l l]}$	$(m \times m)$ -Matrizen, die das DBMAP/G/1/N-System stochastisch beschreiben; $\mathbf{A}_k^{[j l l]} = \sum_{t=1}^{\infty} \mathbf{A}_k^{[j l l]}(t)$, $\mathbf{B}_k^{[j l l]} = \sum_{t=1}^{\infty} \mathbf{B}_k^{[j l l]}(t)$
$\mathbf{A}_k''', \mathbf{B}_k'''$	Matrizen zur Formalisierung der Gleichzeitigkeit von Ankünften und Abgängen: $\mathbf{A}_k''' \mathbf{D}_0 = \mathbf{A}'_k$, $\mathbf{B}_k''' \mathbf{D}_0 = \mathbf{B}'_k$
$[\underline{x}_k]_i$	$= P\{\text{ein Abgang hinterläßt } k \text{ Kunden im System, und der Eingabeprozess befindet sich in Phase } i\}$
\underline{x}	AF: $= (\underline{x}_0, \dots, \underline{x}_N)$, DF: $= (\underline{x}_0, \dots, \underline{x}_{N+1})$
$E_n; E$	n -te Zwischenabgangszeit; $E = \{E_n; n \in \mathbb{N}_0\}$
E^*	mittlere Zwischenabgangszeit im DBMAP/G/1/N-System (<i>fundamental mean</i> des MEP zu Abgangzeitpunkten)
Phase	im Fall des DBMAP eine andere Bezeichnung für die Zustände der ihn erzeugenden MK mit verschiedenen Übergangstypen
IE	Zustandsraum eines SMP oder Menge der Phasen eines DBMAP
m	Anzahl der Zustände eines SMP oder Phasen eines DBMAP
L_n	Belegung zum n -ten Ankunftszeitpunkt (ohne Berücksichtigung der Ankunft)
W	Wartezeit
$\rho_h^{[Z]}(x)$	lokaler Korrelationskoeffizient h -ter Ordnung des SP $Z = \{Z_n; n \in \mathbb{N}_0\}$
$\kappa_h^{[Z]}$	globaler Korrelationskoeffizient h -ter Ordnung des SP $Z = \{Z_n; n \in \mathbb{N}_0\}$
N	Warteplatzzahl ohne Bediener
\mathcal{M}	$\mathcal{M} = N$ für AF und $\mathcal{M} = N + 1$ für DF
b_{max}	Maximalzahl von gleichzeitig ankommenden Kunden (Zellen)
X_n	ZV der Batchgröße
$x(v), X(v)$ bzw. $X^c(v)$	WF, VF bzw. komplementäre VF der Batchgröße
$D \geq 1$	deterministische Bedienzeit in Zeiteinheiten

Faltung

Gemäß [37], p. 299 wird für eine rechtsseitig-stetige, monoton wachsende Funktion $F: \mathbb{R} \rightarrow \mathbb{R}$, die auf $(-\infty, 0)$ verschwindet, als Faltung definiert

$$F * g(t) := \begin{cases} \int_{x=0}^t g(t-x) dF(x) & \text{für } t \geq 0 \\ 0 & \text{für } t < 0. \end{cases} \quad (\text{C.1})$$

Hat F eine Dichtefunktion f , so ist

$$F * g(t) = \int_{x=0}^t g(t-x) f(x) dx \quad \text{für } t \geq 0. \quad (\text{C.2})$$

Ist F eine Treppenfunktion mit Sprüngen in x_j und Sprunghöhen p_j , so gilt

$$F * g(t) = \sum_{j \in \mathcal{J}(t)} g(t-x_j) p_j \quad \text{für } t \geq 0, \quad \mathcal{J}(t) = \{j \in \mathbb{N} : x_j \leq t\}. \quad (\text{C.3})$$

Summationskonvention: $\sum_{k=a}^b f(k) = 0$ für $b < a$.

Literaturverzeichnis

- [1] R. G. Addie, M. Zukerman. *An Approximation for Performance Evaluation of stationary single Server Queues*. In *IEEE INFOCOM'93*, pp. 835–842, San Francisco, March, 28 - April, 1 1993.
- [2] D. Anick, D. Mitra, M.M. Sondhi. *Stochastic theory of a data-handling system with multiple sources*. Bell System Techn. J., Vol. 61, No. 8, pp. 1871–1894, 1982.
- [3] S. Asmussen. *Applied Probability and Queues*. J. Wiley, New York, 1987.
- [4] A. J. Bayes. *Statistical techniques for simulation models*. Australian Computer Journal, Vol. 2, No. 4, pp. 180–184, Nov. 1970.
- [5] N. Blomqvist. *The covariance function of the M/G/1 queueing system*. Aktuarietidskr., Vol. 50, pp. 157–174, 1967.
- [6] C. Blondia. *The N/G/1 finite capacity queue*. Communications in Statistics - Stochastic Models, Vol. 5, No. 2, pp. 273–294, 1989.
- [7] C. Blondia. *Finite Capacity Vacation Models with Non-Renewal Input*. J. Appl. Prob., Vol. 28, pp. 174–197, 1991.
- [8] C. Blondia. *A discrete-time batch Markovian arrival process as B-ISDN traffic model*. Belgian Journal of Operations Research, Statistics and Computer Science, Vol. 32, No. 3,4, pp. 3–23, 1993.
- [9] C. Blondia and O. Casals. *Performance Analysis of Statistical Multiplexing of VBR Sources*. In *IEEE INFOCOM'92*, pp. 828–838, Florence, Italy, May, 4-8 1992.
- [10] C. Blondia and O. Casals. *Statistical multiplexing of VBR sources: A matrix-analytic approach*. Perf. Eval., Vol. 16, No. 1-3, pp. 5–20, Nov. 1992.
- [11] G.E.P. Box, G.M. Jenkins. *Time Series Analysis (forecasting and control)*. Holden Day, Oakland, 1976.
- [12] P. Brémaud. *Point Processes and Queues*. Springer-Verlag, New York, 1981.
- [13] U. Briem, T.H. Theimer and H. Kröner. *A General Discrete-Time Queueing Model: Analysis and Applications*. In *ITC-13, Vol. 14, Copenhagen*, pp. 13–19, June 19-26 1991.
- [14] T. Bröcker. *Analysis II*. BI Wissenschaftsverlag, Mannheim, 1992.
- [15] CCITT Recommendation I. 311. *B-ISDN General Network Aspects*. 1991.
- [16] CCITT Recommendation I. 371. (Draft). *Traffic Control and Congestion Control in B-ISDN*. June 1992.
- [17] K. L. Chung. *Markov Chains with Stationary Transition Probabilities*. Springer Verlag, Berlin, 1967.
- [18] K. L. Chung. *A Course in Probability Theory*. Harcourt, Brace & World, Inc., New York, 1968.
- [19] E. Cinlar. *Markov Renewal Theory*. Adv. Appl. Prob., Vol. 1, pp. 123–187, 1969.
- [20] E. Cinlar. *Introduction to Stochastic Processes*. Prentice Hall, New Jersey, 1st edition, 1975.
- [21] E. Cinlar. *Markov Renewal Theory: A Survey*. Management Science, Vol. 21, No. 7, March 1975.

- [22] J.W. Cohen. *The Single Server Queue*. Applied Mathematics and Mechanics. North-Holland, Amsterdam, 1969.
- [23] J. Cosmas, A. Odinnma-Okafor. *Characterization of Variable Rate Video Codecs in ATM to a Geometrically Modulated Deterministic Process Model*. In *ITC-13, Vol. 14, Copenhagen*, pp. 773–780, June 19–26 1991.
- [24] D.R. Cox, H.D. Miller. *The Theory of Stochastic Processes*. Chapman & Hall, London, 3rd reprint, 1972.
- [25] D.R. Cox, P.A.W. Lewis. *The Statistical Analysis of Series of Events*. Methuen & Co. Ltd., London, 1966.
- [26] M. de Prycker. *Asynchronous Transfer Mode*. Ellis Horwood, New York, 1991.
- [27] L. Deng, J. W. Mark. *Parameter estimation for Markov modulated Poisson processes via the EM algorithm with time discretization*. Telecommunication Systems, Vol. 1, No. 4, pp. 321–338, 1993.
- [28] W. Ding. *A generator for correlated random sequences with prescribed properties*. 1990. In: F. Breitenecker, I. Troch (Hrg.), Fortschritte in der Simulationstechnik, Band 1, 6. Symposium Simulationstechnik, Wien, 25.-27. Sept. 1990, Vieweg, September 1990.
- [29] W. Ding. *Korrelierte Zufallsprozesse in Wartesystemen von Kommunikationsnetzen*. Dissertation, RWTH Aachen, Aachen, 1991.
- [30] W. Ding. *A unified correlated input process model for telecommunication networks*. In *ITC-13, Vol. 14, Copenhagen*, pp. 539–544, June 19–26 1991.
- [31] W. Ding, F. Schreiber. *Local Correlation Properties of Random Sequences Generated by Queueing System M/M/1*. Archiv f. Elektronik und Übertragungstechnik, Vol. 44, No. 5, pp. 384–389, 1990.
- [32] W. Ding and P. Decker. *Waiting time distribution of a discrete SSMP/G/1 queue and its implications in ATM systems*. In *7th ITC Spec. Sem.*, paper 9.2, Morristown, USA, Oct. 9–11 1990.
- [33] Z. Dziong, J. Choquette, K.-Q. Liao, L. Mason. *Admission Control and Routing in ATM Networks*. Computer Networks and ISDN Systems, Vol. 20, 1990.
- [34] J. Eberspächer. *Hochgeschwindigkeitsnetze (Editorial)*. Informationstechnik und Technische Informatik, No. 4, pp. 3–8, 1993.
- [35] A. I. Elwalid, D. Mitra. *Effective Bandwidth of General Markovian Traffic Sources and Admission Control of High Speed Networks*. In *IEEE INFOCOM93*, San Francisco, pp. 256–265, March 31 - April 2, 1993.
- [36] M.R. Frater et al. *Fast Simulation of Rare Events Using Reverse Time Models*. In *Proceedings of the ITC Spec. Sem.*, Adelaide, 1989.
- [37] L. Fahrmeir, H. Kaufmann, F. Ost. *Stochastische Prozesse*. Hanser, München, 1980.
- [38] W. Fischer, K. Meier-Hellstern. *The MMPP Cookbook*. Perf. Eval., Vol. 18, No. 2, pp. 149–171, 1993.
- [39] G.H. Golub, C.F. van Loan. *Matrix Computations*. John Hopkins University Press, Baltimore and London, 2nd edition, 1991.
- [40] B. Gopinath, J.A. Morrison. *Discrete-Time Single Server Queue with Correlated Inputs*. BSTJ, Vol. 56, No. 9, pp. 1743–1768, Nov. 1977.
- [41] A. Graham. *Non-negative matrices and applicable topics in linear algebra*. Ellis Horwood, Chichester, 1987.
- [42] J. Grandell. *Doubly stochastic processes*. Springer Verlag, Berlin, 1976.

- [43] W. K. Grassmann, M. Taksar, D. Heyman. *Regenerative Analysis and Steady State Distributions for Markov Chains*. Opns. Res., Vol. 33, No. 5, pp. 1107–1116, Sept.–Okt. 1985.
- [44] A. Gravey, J. Louvion and P. Boyer. *On the Geo/D/1 and Geo/D/1/n Queues*. Perf. Eval., Vol. 11, No. 2, pp. 117–125, July 1990.
- [45] A. Gravey, J. Louvion and P. Boyer. *Simultaneity in discrete-time single server queues with Bernoulli inputs*. Perf. Eval., Vol. 14, No. 2, pp. 123–131, Jan 1992.
- [46] R. Grünfelder J.P. Cosmas, S. Mantorpe, A. Odinma-Okafor. *Characterization of Video Codecs as Autoregressive Moving Average Processes and Related Queueing System Performance*. IEEE J. Sel. Areas Com, Vol. 9, No. 3, pp. 284–293, April 1991.
- [47] F. Guillemin, J. Boyer and A. Dupuis. *Burstiness in broadband integrated networks*. Perf. Eval., Vol. 15, pp. 163–176, 1992.
- [48] R. Händel, M.N. Huber. *Integrated Broadband Networks*. Addison-Wesley, 1991.
- [49] H. Heeke. *Statistical Multiplexing Gain for VBR Video Codecs in ATM Networks*. Int. J. Digital and Analog Com. Systems, Vol. 4, pp. 261–268, Jan.-March 1991.
- [50] H. Heffes, D. M. Lucantoni. *A Markov-modulated characterization of packetized voice and data traffic and related statistical multiplexer performance*. IEEE J. Sel. Areas Com., Vol. 4, pp. 856–868, 1986.
- [51] C. Herrmann. *Analysis of the Discrete-time SMP/D/1/s Finite Buffer Queue with Applications in ATM*. In *IEEE INFOCOM'93*, pp. 160–167, Session 2a.3, San Francisco, March 28 - April 1, 1993.
- [52] C. Herrmann. *Correlation Effect on Per-stream QOS Parameters of ATM Traffic Superpositions relevant to Connection Admission Control*. In *IEEE ICC'93*, pp. 1027 – 1031, Geneva, May 23-26, 1993.
- [53] C. Herrmann. *A Performance Model for Statistical Multiplexing of correlated ATM Traffic Superpositions*. In *7. ITG/GI Fachtagung: Messung, Modellierung und Bewertung von Rechen- und Kommunikationssystemen*, paper 4A.2, Aachen, Sept. 21.-23., 1993, Springer-Verlag.
- [54] C. Herrmann. *The discrete-time DBMAP/G/1/s finite buffer queue with priorities relevant to Connection Admission Control in ATM*. In *IEEE Workshop on Visual Processing and Communications*, Rutgers Univ., New Brunswick, NJ, USA, Sept. 19-20, 1994.
- [55] C. Herrmann. *Discrete-time Models for Connection Admission Control (CAC) in ATM*. In *2nd IFIP Workshop on Performance Modelling and Evaluation of ATM Networks*, Bradford, UK, July 4-6, 1994.
- [56] C. Herrmann. *Local Correlation Properties of some Discrete-time Finite Queues with Semi-Markovian Input*. Archiv f. Elektronik und Übertragungstechnik, Vol. 48, No. 5, pp. 244–255, Sept. 1994.
- [57] C. Herrmann. *On the Analytical Model with the Periodic Correlation Function of "non-frame-buffered" VBR Video in ATM*. In *IEEE International Telecommunications Symposium*, Rio de Janeiro, Aug. 23-25, 1994.
- [58] C. Herrmann. *VBR Video in ATM without Frame-Buffering: Influence of a Periodic Correlation Function on QOS Parameters*. In *2nd IFIP Workshop on Performance Modelling and Evaluation of ATM Networks*, Bradford, July 4-6, 1994.
- [59] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Camb. Univ. Press, Cambridge, 1991.
- [60] R.A. Howard. *Dynamic Probabilistic Systems, Vol. II: Semi-Markov and Decision Processes*, Vol. II. J. Wiley, New York, 1971.
- [61] J. J. Hunter. *Mathematical Techniques of Applied Probability, Vol. 1, Discrete Time Models: Basic Theory*. Academic Press, New York, 1983.

- [62] J. J. Hunter. *Mathematical Techniques of Applied Probability, Vol. 2, Discrete Time Models: Techniques and Applications*. Academic Press, New York, 1983.
- [63] J.J. Hunter. *On the Moments of Markov Renewal Processes*. Adv. Appl. Prob., No. 1, pp. 188–210, 1969.
- [64] Y.H. Jeon, I. Viniotis. *Achievable Loss Probabilities and Buffer Allocation Policies in ATM Nodes with Correlated arrivals*. In *IEEE ICC'93*, pp. 365–369, Geneva, Switzerland, 1993.
- [65] M. J. Karol, M. G. Hluchyj, S. P. Morgan. *Input Versus Output Queueing on a Space-Division Packet Switch*. IEEE Trans. Com., Vol. 35, No. 12, pp. 1347–1356, 1987.
- [66] F.P. Kelly. *Effective bandwidths at multi-class queues*. Queueing Systems, Vol. 9, pp. 5–15, 1991.
- [67] J.F.C. Kingman. *On doubly stochastic Poisson processes*. Camb. Phil. Soc., Vol. 60, pp. 923–930, 1964.
- [68] R. Kishimoto, Y. Ogata, F. Inumaru. *Generation Interval Distribution Characteristics of Packetized Variable Rate Video Coding Data Streams in an ATM Network*. IEEE J. Sel. Areas in Com, Vol. 7, No. 5, pp. 833–841, June 1989.
- [69] D. D. Kouvatsos. *Maximum Entropy and the G/G/1/N Queue*. Acta Informatica, Vol. 23, pp. 545–565, 1986.
- [70] M. Kramer. *Computational Methods for Markov Chains occuring in Queueing Theory*. In *Messung, Modellierung und Bewertung von Rechnersystemen, Informatik Fachberichte 154*, Springer, pp.164–175, 1987.
- [71] H. Kröner. *Comparative Performance Study of Space Priority Mechanisms for ATM Networks*. In *Proceedings of the INFOCOM'90*, pp. 1136–1143, San Francisco, 1990.
- [72] W. P. Lidinsky. *Data communications needs*. IEEE Network Magazine, pp. 28–33, March 1990.
- [73] D. V. Lindley. *The theory of queues with a single server*. Proc. Cambridge Phil. Soc., Vol. 48, pp. 277–289, 1952.
- [74] D. Liu, M.F. Neuts. *Counter-examples involving Markovian arrival processes*. Comm. Stat.-Stoch. Mod., Vol. 7, No. 3, pp. 499–509, 1991.
- [75] D.M. Lucantoni. *New Results on the single Server Queue with a Batch Markovian Arrival Process*. Stochastic Models, Communications in Statistics, Vol. 7, No. 1, pp. 1–46, 1991.
- [76] D.M. Lucantoni, M.F. Neuts and K.S. Meier-Hellstern. *A single server queue with server vacations and a class of non-renewal arrival processes*. Adv. Appl. Prob., Vol. 22, pp. 676–705, Sept. 1990.
- [77] K. A. Lutz. *Sprachübertragungsqualität in Kommunikationsnetzen von morgen sichern*. SIE-MENS telecom report, Vol. 3, No. 17, pp. 112–114, 1994.
- [78] B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, J.D. Robbins. *Performance Models of Statistical Multiplexing in Packet Video Communications*. IEEE Trans Com., Vol. 36, No. 7, pp. 834–843, July 1988.
- [79] B.A. Makrucki. *A study of source traffic management and buffer allocation in ATM networks*. In *7th ITC Spec. Sem.*, paper 5.5, Morristown, USA, October 1990.
- [80] R. Mathar and D. Pfeifer. *Stochastik für Informatiker*. B.G. Teubner, Stuttgart, 1. Auflage, 1990.
- [81] K. S. Meier. *A statistical procedure for fitting Markov modulated Poisson processes*. PhD thesis, Univ Delaware, Newark, DE, 1984.
- [82] K. S. Meier-Hellstern. *The analysis of a queue arising in overflow models*. IEEE Trans. Comm., Vol. 37, pp. 367–372, 1989.

- [83] M. Murata et al. *Analysis of a Discrete-Time Single-Server Queue with Bursty Inputs for Traffic Control in ATM Networks*. IEEE J-SAC, Vol. 8, No. 3, pp. 447–458, April 1990.
- [84] S. Narayana, M.F. Neuts. *The first two moment matrices of the counts for the Markovian Arrival Process*. Comm. Stat.-Stoch. Mod., Vol. 8, No. 3, pp. 459–477, 1992.
- [85] Marcel F. Neuts. *Matrix-Geometric Solutions in Stochastic Models*. John Hopkins, Baltimore and London, 1981.
- [86] Marcel F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and their Applications*. Marcel Dekker, New York, 1989.
- [87] M.F. Neuts. *A versatile Markovian point process*. J. Appl. Prob., Vol. 16, pp. 764–779, 1978.
- [88] V. Nollau. *Semi-Markovsche Prozesse*. Verlag Harri Deutsch, Frankfurt am Main, 1980.
- [89] M. Nomura, T. Fujii, N. Ohta. *Basic Characteristics of Variable Video Coding in ATM Environment*. IEEE J-SAC, Vol. 7, No. 5, pp. 752–760, June 1989.
- [90] A. Odinma-Okafor. *The Characterization of Variable Rate Video Signals*. PhD thesis, Dep. of Electronic Eng., Queen Mary & Westfield College, London, Oct. 1991.
- [91] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill, Tokyo, 1965.
- [92] J. F. Patterson, C. Egidio. *Three Keys to the Broadband Future: A View of Applications*. IEEE Network Magazine, pp. 41–47, March 1990.
- [93] W. H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 1988.
- [94] R. Pyke. *Markov Renewal processes: definitions and preliminary properties*. Ann. Math. Stat., Vol. 32, pp. 1231–1242, 1961.
- [95] R. Pyke. *Markov Renewal processes with finitely many states*. Ann. Math. Stat., Vol. 32, pp. 1243–1259, 1961.
- [96] V. Ramaswami. *The N/G/1 queue and its detailed analysis*. Adv. Appl. Prob., No. 12, pp. 222–261, 1980.
- [97] V. Ramaswami. *A stable recursion for the steady state vector in Markov Chains of M/G/1-type*. Stoch. Models, No. 4, pp. 183–188, 1988.
- [98] V. Ramaswami, W. Willinger. *Efficient Traffic Performance Strategies for Packet Multiplexers*. In *ITC Spec. Sem.*, pp. 401–407, Adelaide, Australia, Sept. 25–29 1989.
- [99] E. P. Rathgeb. *Modeling and Performance Comparison of Policing Mechanisms for ATM Networks*. J. on Sel. Areas in Com., Vol. 9, No. 3, pp. 325–334, April 1991.
- [100] R.Guerin, H. Ahmadi, M. Naghshineh. *Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks*. IEEE J. Sel. Areas Comm., Vol. 9, No. 7, pp. 968–981, 1991.
- [101] I.E.G. Richardson, M.J. Riley. *FEC and Multi-Level Coding for ATM Networks*. In *2nd IFIP Workshop on Performance Modelling and Evaluation of ATM Networks*, Bradford, UK, July 4-7, 1994.
- [102] R.J.Gibbens, P.J. Hunt. *Effective bandwidth for the multi-type UAS channel*. Queueing Systems, Vol. 9, pp. 17–28, 1991.
- [103] J.W. Roberts, editor. *COST 224, Performance Evaluation and design of multiservice networks*. Information technologies and sciences. France Telecom, 1991.
- [104] F. Schreiber. *The Extended Bayes-Postulate, Computer Random Experiments, and the Unique Solution of an Elementary Problem in Statistics*. Archiv f. Elektronik und Übertragungstechnik, Vol. 35, No. 1, pp. 34–43, 1981.

- [105] F. Schreiber. *The Bayes-Laplace-Statistic of the Multinomial Distribution*. Archiv f. Elektronik und Übertragungstechnik, Vol. 39, No. 5, pp. 293–298, 1985.
- [106] F. Schreiber. *Measurement of the Correlation Coefficient of the 2-Node Markov Chain*. Archiv f. Elektronik und Übertragungstechnik, Vol. 40, No. 6, pp. 402–405, 1986.
- [107] F. Schreiber. *Effective Control of Simulation Runs by a New Evaluation algorithm for Correlated Random Sequences*. Archiv f. Elektronik und Übertragungstechnik, Vol. 42, No. 6, pp. 347–354, 1988.
- [108] F. Schreiber. *Local Correlation of Occupancy of the Finite Buffer Queue M/M/1/N*. Archiv f. Elektronik und Übertragungstechnik, Vol. 48, No. 2, pp. 63–67, March 1994.
- [109] F. Schreiber, C. Goerg. *Rare Event Simulation: a modified RESTART-Method using the LRE-Algorithm*. In *ITC-14, Antibes Juan-les-Pins, France*, pp. 787–796, June 6-10 1994.
- [110] P. Sen, B. Maglaris, N. Rikli, D. Anastassiou. *Models for Packet Switching of Variable Rate Video Sources*. IEEE J. on Sel. Areas in Com., Vol. 7, No. 5, pp. 865–869, 1989.
- [111] K. Siriam, W. Whitt. *Characterizing superposition arrival processes in packet multiplexers for voice and data*. IEEE J. on Sel. Areas in Com., Vol. 4, No. 6, pp. 833–846, Sept. 1986.
- [112] T. Takine, T. Suda, T. Hasegawa. *Cell Loss and Output Process Analysis of a Finite-Buffer Discrete-time ATM Queueing System with Correlated Arrivals*. In *IEEE INFOCOM'93*, pp. 1259–1269, paper 10C.3.1, San Francisco, March 30-April 1, 1993.
- [113] P. Tran-Gia, Hamid Ahmadi. *Analysis of a discrete-time $G^{[X]}/D/1-S$ queueing system with applications in packet-switching systems*. In *IEEE INFOCOM'88*, paper 9A.1.10, March 27-31, 1988.
- [114] M. Villen-Altamirano, J. Villen-Altamirano. *RESTART: A Method For Accelerating Rare Event Simulation*. In *Queueing, Performance and Control in ATM*, pp. 71–76, ITC-13, Copenhagen, June 19-26 1991.
- [115] B.H. Walke. *Realzeitrechner-Modelle: Theorie und Anwendung*. Oldenbourg Verlag, München, Wien, 1978.
- [116] B.H. Walke. *Datenkommunikation I, Teil 1*. Dr. Alfred Hüthig Verlag, Heidelberg, 1987.
- [117] E. Wallmeier. *A connection acceptance algorithm of mixed traffic streams*. Int. J. Digital and Analog Communication Systems, Vol. 3, pp. 143–153, 1990.
- [118] C. Westermann. *Entwicklung und Bewertung eines Modells für die Simulation einer virtuellen Verbindung innerhalb eines ATM-Netzes*. Master's thesis, RWTH Aachen, Germany, Dec 1993.
- [119] W. Whitt. *Approximating a point process by a renewal process, I.: two basic methods*. Operations Research, Vol. 30, No. 1, pp. 125–147, 1982.
- [120] W. Whitt. *Tail probabilities with statistical multiplexing and effective bandwidth in multi-class queues*. Telecommunication Systems, Vol. 2, No. 1, pp. 71–107, 1993.
- [121] W.P.Cherry, R.L. Disney. *The Superposition of two independent Markov Renewal Processes*. ZASTOSOWANIA MATEMATYKI, APPLICATIONES MATHEMATICAE, Vol. XVII, No. 4, pp. 567–602, 1983.
- [122] D. J. Wright, M. To. *Telecommunication Applications of the 1990s and their Transport Requirements*. IEEE Network Magazine, pp. 34–40, March 1990.
- [123] J. Zhang. *Performance Study of Markov Modulated Fluid Flow Models with Priority Traffic*. In *IEEE INFOCOM'93*, San Francisco, March, 28 - April, 1 1993.

Lebenslauf

Name: Christoph Herrmann
geboren: am 2. Dezember 1964 in Bensberg
Eltern: Wolfgang Herrmann, Dipl.-Ing.(Geodäsie),
Kreisvermessungsdirektor a. D.
Christel Herrmann, Hausfrau
Geschwister: Wolfram und Ursula

Schulbildung

1971-1975 Grundschule "Am Broich" in Bergisch Gladbach
1975-1984 Nicolaus-Cusanus-Gymnasium in Bergisch Gladbach
Abitur Juni 1984

Wehrdienst 1. Juli 1984 bis 30. September 1985

Studium

WS 1985-SS 1991 Elektrotechnik (Fachrichtung Nachrichtentechnik) an der RWTH Aachen
Diplomarbeit "Gewichtsverteilung zyklischer Codes" (Institut für Elektrische Nachrichtentechnik, Prof. Lüke)
Werkstudent August u. September 1990 Fa. Siemens, Medizintechnik, Erlangen: Studie zur Sigma-Delta-Modulation für die Analog-Digital-Wandlung in Computertomographiegeräten.
seit 1991 DFG-Stipendiat im Graduiertenkolleg "Informatik & Technik" (Sprecher: Prof. Spaniol) und wissenschaftlicher Mitarbeiter am Institut für Kommunikationsnetze der RWTH Aachen (Prof. Walke).

