Wartezeitverteilungen in Tandem-Warteschlangennetzen mit stückweise approximierenden Dichten maximaler Entropie

Diplomarbeit

im Studiengang Informatik an der

Rheinischen Friedrich-Wilhelms-Universität Bonn

unter Leitung von Prof. Dr. J. Ch. Strelen

eingereicht von

Werner Sandmann

im Januar 1998

Bezeichnungen

Zahlenbereiche

$\mathbb{N}$	$=$	$\{0, 1, 2, \ldots\}$	die Menge der natürlichen Zahlen
$\mathbb{N}^+$	$=$	$\mathbb{N} \setminus \{0\}$	die Menge der natürlichen Zahlen ohne Null
$\mathbb{Z}$	$=$	$\{\ldots, -1, 0, 1, 2, \ldots\}$	die Menge der ganzen Zahlen
$\mathbb{R}$	$=$	$(-\infty, \infty)$	die Menge der reellen Zahlen
$\mathbb{R}_+$	$=$	$[0, \infty)$	die Menge der nichtnegativen reellen Zahlen
$\mathbb{R}^+$	$=$	$(0, \infty)$	die Menge der positiven reellen Zahlen
$\overline{\mathbb{R}}_+$	$=$	$[0, \infty]$	die Menge der nichtnegativen reellen Zahlen einschließlich ∞
$\overline{\mathbb{R}}^+$	$=$	$(0, \infty]$	die Menge der positiven reellen Zahlen einschließlich ∞
$\overline{\mathbb{R}}$	$=$	$[-\infty, \infty]$	die Menge der reellen Zahlen einschließlich $\pm\infty$

Wahrscheinlichkeitsverteilungen

$Exp(\lambda)$	Exponentialverteilung mit Parameter $\lambda \in \mathbb{R}^+$
$E_k(\lambda)$	Erlangverteilung mit $k \in \mathbb{N}^+$ Phasen und Parameter $\lambda \in \mathbb{R}^+$
$H_2(\lambda_1, \lambda_2, p)$	Hyperexponentialverteilung mit Parametern $\lambda_1, \lambda_2 \in \mathbb{R}^+,\ p \in [0, 1]$
$U(a, b)$	Stetige Gleichverteilung auf dem Intervall $[a, b]$
$N(\mu, \sigma^2)$	Normalverteilung mit Erwartungswert $\mu \in \mathbb{R}$ und Varianz $\sigma^2 \in \mathbb{R}_+$
$N(0, 1)$	Standardisierte Normalverteilung

Inhaltsverzeichnis

1 Einleitung

Die vorliegende Arbeit beschäftigt sich mit der Analyse von Tandem-Warteschlangennetzen, deren Knoten aus jeweils einer Warteschlange unendlicher Kapazität und einem Bediener bestehen. Solche Netze werden in dieser Arbeit abkürzend mit *Tandemsystem* bezeichnet. Vereinfacht gesprochen ist ein Tandemsystem eine Hintereinanderschaltung endlich vieler elementarer Warteschlangensysteme bzw. Bedienstationen, wobei ein ankommender Kunde in die erste Bedienstation eintritt, nacheinander alle folgenden Stationen in festgelegter Reihenfolge durchläuft und schließlich das System verläßt. Eine besondere Bedeutung von Tandemsystemen besteht darin, daß sie ein Modell für reale fließbandähnliche Systeme darstellen. Dies beinhaltet sowohl Rechensysteme, als auch konkrete Abläufe des realen Lebens.

Ein Beispiel für die Anwendung der Modellierung mit Tandemsystemen ist die Abfertigung ankommender Flugzeuge an einem Flughafen. Zunächst erfolgt der Ausstieg der Passagiere, sowie der Crewmitglieder, anschließend durchläuft die Maschine nacheinander verschiedene Sicherheitschecks und Reparaturen, und wird aufgetankt, bevor sie zum nächsten Start bereitsteht. Hierbei entsprechen die Flugzeuge den Kunden des Tandemsystems und die in der festgelegten Reihenfolge zu durchlaufenden Maßnahmen bzw. Arbeitsgänge den Bedienern, an denen es eventuell zu Wartezeiten kommen kann.

Ein Rechensystem, das als Einbediener-Tandemsystem modelliert werden kann, ist zum Beispiel ein aus einem Eingabegerät, mehreren für eine bestimmte Art der Datenverarbeitung spezialisierten Rechnern und einem Ausgabegerät bestehendes System, in dem jeweils in fester Reihenfolge die Ausgangsdaten eines Rechners als Eingangsdaten eines bestimmten anderen Rechners von diesem weiter bearbeitet werden. So kann etwa die Berechnung der Komposition mehrerer Funktionen für verschiedene Funktionsargumente als Tandemsystem modelliert werden, wobei die Argumente den Kunden entsprechen und jede Bedienstation genau eine Funktion realisiert. Die Eingabe der Argumente bildet den Ankunftsprozeß.

Allgemeiner kann man sagen, daß alle in einer gewissen Art und Weise fließbandähnlichen Vorgänge, also solche, die nacheinander von verschiedenen Menschen oder Maschinen ausgeführt werden, als Tandemsystem modellierbar sind. Tandemsysteme sind also durchaus in der Praxis relevant und nicht nur von theoretischem Interesse.

Tandemsysteme werden bereits seit den 50er Jahren in zahlreichen Varianten ausführlich in der Literatur behandelt, so zum Beispiel in [Mor55], [Jac56], [Rei57] [Nel58], [Bur64], [Fri65], [Bur72], [New79], [Web79].

Es ist zu beachten, daß in der Literatur der Begriff Tandemsystem, oder englisch *Tandem-Queue*, oftmals auch für aus nur zwei hintereinandergeschalteten Bedienstationen bestehendende Systeme benutzt wird. Desweiteren gibt es zahlreiche Varianten mit Einschränkungen bezüglich der Warteschlangenkapazitäten, Verteilungen der Bedienzeiten und Zwischenankunftszeiten, sowie Varianten mit Rückkopplungen, verschiedenen Bedieneranzahlen pro Station oder verschiedenen Bedienungsdisziplinen. Häufig werden auch die Begriffe *Seriensystem* oder *Etappensystem* für Tandemsysteme benutzt.

In dieser Arbeit ist stets, wenn nicht explizit anders gesagt, die Rede von einem offenen System von hintereinandergeschalteten elementaren FCFS-Bedienstationen mit unendlicher Warteschlangenkapazität, beliebig verteilten unabhängigen Bedienzeiten an den einzelnen Stationen, sowie beliebig verteilten unabhängigen Zwischenankunftszeiten an der ersten Station. Nur an dieser Station seien Ankünfte erlaubt. Kunden, die von einer Station vollständig bedient wurden, treten anschließend in die nächste Station ein, Abgänge gebe es nur nach Fertigstellung der Bedienung in der letzten Station. Für ein solches System wird in dieser Arbeit die Bezeichnung $G(\rightarrow /G/1)^{\alpha}$-System eingeführt, wobei α die Anzahl der Bedienstationen bezeichnet.

Bei der Analyse solcher Tandemsysteme ergibt sich im Gegensatz zu elementaren Warteschlangensystemen, die nur aus einer Bedienstation bestehen, das Problem, daß der Ankunftsprozeß in den einzelnen Bedienstationen mit Ausnahme der zuerst zu durchlaufenden Station abhängig vom Abgangsprozeß der jeweils vorhergehenden Bedienstation ist. Ein Tandemsystem kann also im allgemeinen nicht analysiert werden, indem man die einzelnen Stationen isoliert betrachtet und analysiert, und dann die Einzelergebnisse zusammenfügt. Eine exakte Analyse ist daher bisher lediglich für einige der bereits erwähnten sehr eingeschränkten Varianten möglich, nicht aber für den hier betrachteten allgemeinen Fall.

Dies motiviert den Ansatz der numerischen Analyse durch iterative Berechnung der stationären Verteilungen von Leistungsgrößen in Tandemsystemen, der in dieser Arbeit bis hin zur Entwicklung und Erprobung entsprechender Algorithmen ausgeführt wird. Approximiert werden dabei jeweils die Dichten der Leistungsmaße wie Warte- und Verweilzeit, wobei neben dem auch in vielen anderen Arbeiten untersuchten stationären Verhalten von Tandemsystemen hier auch das transiente Verhalten betrachtet wird. Aus den approximierenden Dichten können zudem beliebige höhere Momente der Leistungsmaße berechnet werden, ohne daß ein zusätzliches Näherungsverfahren benötigt wird.

Genauer ausgedrückt werden die Dichten der bei der Analyse auftretenden Zufallsvariablen, deren Definitionsbereich jeweils eine Teilmenge der reellen Zahlen ist, durch sogenannte *101-Dichten* ([Str89]) stückweise approximiert, das heißt, der Definitionsbereich wird in endlich viele Intervalle aufgeteilt, auf denen die Dichten dann durch Dichten mit maximaler Entropie ([SJ80]) angenähert werden.

Um Iterationsvorschriften zur Berechnung der Leistungsmaße herzuleiten, werden Tandemsysteme als *(max,+)-lineare Systeme*, denen die *(max,+)-Algebra* ([BCOQ92]) zugrundeliegt, dargestellt. Solche (max,+)-lineare Systeme können zur Analyse allgemeiner stochastischer Ereignisgraphen eingesetzt werden. Ihre Benutzung in dieser Arbeit bietet aufgrund ihrer Anwendbarkeit auf zahlreiche weitere Warteschlangennetze den Vorteil, daß die hier beschriebene Entwicklung eines Analysealgorithmus für Tandemsysteme leicht für andere Netze modifiziert werden kann.

In Kapitel 2 dieser Arbeit werden die für die Analyse benötigten mathematischen Grundlagen, das Prinzip der maximalen Entropie, die stückweise Approximation von Dichten und die (max,+)-linearen Systeme, sowie ein geeignetes Analysewerkzeug (*VERENA*) erläutert. Dieses Analysewerkzeug VERENA ([Sie91],[NB93]), eine Programmbibliothek für die Programmiersprache Modula-2, stellt Datenstrukturen und Operationen für Zufallsvariablen zur Verfügung, wobei die Darstellung stetiger Zufallsvariablen gerade durch 101-Dichten erfolgt. Zudem wird in Kapitel 2 eine Erweiterung für 101-Dichten in Form eines neuen Operators für die Maximumbildung von 101-Dichten, sowie ein Modul für VERENA, das Prozeduren zur Berechnung von Maximum und Minimum stetiger Zufallsvariablen enthält, entwickelt und anhand einiger Beispiele erprobt.

In Kapitel 3 werden zunächst formale Definitionen von Tandemsystemen, ihrer Zustände und ihrer Leistungsmaße gegeben und Notationen eingeführt. Weiterhin werden in diesem Kapitel bisher bekannte Ergebnisse für verschiedene näher spezifizierte Tandemsysteme zusammengefaßt, sowie zwei Näherungsverfahren für allgemeine $G(\rightarrow/G/1)^{\alpha}$-Systeme vorgestellt. Schließlich wird bereits hier ein erster Algorithmus zur Analyse von Tandemsystemen unter Verwendung von 101-Dichten entwickelt, bei dem die spezielle Darstellung von Tandemsystemen als (max,+)-lineares System nicht berücksichtigt wird, sondern von der Annahme ausgegangen wird, daß die Abgangsprozesse aus den einzelnen Stationen Erneuerungsprozesse sind und somit die einzelnen Stationen isoliert wie elementare G/G/1-Systeme analysiert werden können.

In Kapitel 4 erfolgt die Darstellung von Tandemsystemen als (max,+)-lineares System, woraus direkt eine Iterationsvorschrift zur Berechnung von Leistungsmaßen und eine entsprechende 101-Approximation resultiert. Die Iteration wird dann so umgeformt, daß Faltungen und Maxima ausschließlich von unabhängigen Zufallsvariablen gebildet werden. Damit erhält man schließlich eine 101-Approximation, für die keine falschen, vereinfachenden Unabhängigkeitsannahmen gemacht werden müssen. Desweiteren wird in Kapitel 4 die Konvergenz transienter Größen gegen stationäre Größen untersucht und bewiesen.

In Kapitel 5 erfolgt für die mit den verschiedenen Algorithmen berechneten Leistungsmaße einiger konkreter Tandemsysteme eine Bewertung der Ergebnisse und, soweit möglich, ein Vergleich mit exakten Resultaten und anderen Approximationsverfahren.

Kapitel 6 enthält eine Zusammenfassung der wichtigsten Ergebnisse der Arbeit, sowie einen Überblick über weiterhin offene Fragen und Ansatzpunkte künftiger Lösungen.

2 Zugrundeliegende Verfahren

Bei der numerischen Analyse stochastischer Modelle ergibt sich allgemein das Problem der Darstellung stetiger Wahrscheinlichkeitsdichten auf den reellen Zahlen in einem Rechner. Die Grundlagen für eine solche approximative Darstellung, Dichten maximaler Entropie und 101-Dichten, sowie Operationen auf diesen Dichten werden in 2.1 bzw. 2.2 beschrieben. In 2.2 wird darüber hinaus das Analysewerkzeug VERENA, das Dichten stetiger Zufallvariablen durch 101-Dichten approximiert, ein neuer Operator für die Maximumbildung von stetigen Zufallsvariablen mit 101-Dichten, sowie ein neues Modul für VERENA, das Prozeduren zur Approximation des Maximums und des Minimums stetiger Zufallsvariablen enthält, vorgestellt und anhand einiger Beispiele erprobt. In 2.3 schließlich wird ein mathematisches Modell zur Analyse allgemeiner stochastischer Ereignisgraphen, das (max,+)-lineare System, definiert und erläutert.

2.1 Das Prinzip der maximalen Entropie

Eine wichtige Grundlage zum Verständnis der später zur Approximation stetiger Dichten benutzten 101-Dichten ist das Prinzip der maximalen Entropie. Informell beschrieben werden Wahrscheinlichkeitsverteilungen und -dichten durch Verteilungen maximaler Entropie (VME) und Dichten maximaler Entropie (DME) angenähert, deren ersten Momente mit den entsprechenden Momenten der jeweils anzunähernden Dichte übereinstimmen.

Genauer gesagt wird zunächst eine Menge von Verteilungen bzw. Dichten durch Festlegung des Ergebnisraums und einiger Momente definiert, und dann schließlich aus dieser Menge eine Verteilung bzw. Dichte, deren Entropie maximal ist, ausgewählt ([SJ80]).

Sei X eine Zufallsvariable mit unbekannter Verteilung P_X und bekannten Erwartungswerten

$$E[g_k(X)] = G_k, \qquad k \in \{0, 1, ...K\},\ K \in \mathbb{N}.$$

Dabei seien die g_k bekannte Funktionen mit $g_0(x) = 1$ und $G_k \in \mathbb{R}$ mit $G_0 = 1$.
Obige Gleichung definiert eine Menge von Verteilungen, aus der schließlich eine Verteilung mit maximaler Entropie gesucht wird. Im folgenden wird nun zunächst der Entropiebegriff definiert.

Definition 1 (Entropie) *Sei $\Omega \subseteq \mathbb{Z}$ ein diskreter Wahrscheinlichkeitsraum. Dann ist die Entropie H einer Verteilung $(p_n, n \in \mathbb{N})$ definiert als*

$$H := -\sum_{n\in\Omega} p_n \ln p_n.$$

Sei $\Omega \subseteq \overline{\mathbb{R}}$. Dann ist die Entropie H einer Dichte f auf Ω definiert als

$$H := -\int_{\Omega} f(x) \ln f(x) dx.$$

Im diskreten Fall muß für die Verteilung

$$p_n \geq 0 \text{ für alle } n \in \Omega \qquad \text{und} \qquad \sum_{n\in\Omega} p_n = 1$$

und im stetigen Fall muß für die Dichte

$$f(x) \geq 0 \text{ für alle } x \in \Omega \qquad \text{und} \qquad \int_{\Omega} f(x)dx = 1$$

gelten. Im Folgenden werden nun nur noch Entropien für Dichten betrachtet.

Sei $\Omega \subseteq \mathbb{R}$ und $f_X(x)$ eine unbekannte Dichte mit den bekannten Momenten

$$M_j = \int_{\Omega} x^j f_X(x)dx, \quad j \in \{0, 1, ..., K\} \qquad \text{mit } M_0 = 1.$$

Eine Approximation $f(x)$ der Dichte $f_X(x)$ mit denselben Momenten wird durch Maximierung der Entropie

$$H = -\int_{\Omega} f(x) \ln f(x)dx$$

gesucht. Solch eine Approximation heißt *von der Ordnung K*. Dieses Maximierungsproblem führt auf die Eulersche Differentialgleichung, die dann folgende *Dichte maximaler Entropie der Ordnung K (DME-K)* liefert ([KLR73]):

$$f(x) = e^{-\lambda(x)}$$

mit

$$\lambda(x) := \sum_{j=0}^{K} \lambda_j x^j, \quad \lambda_0 := 1 - \lambda_0^*, \quad \lambda_j := -\lambda_j^*, \quad j \in \{1, \ldots, K\}$$

wobei die λ_j^* Lagrange-Multiplikatoren sind. Das Hauptproblem ist nun die Berechnung der λ_j aus den nichtlinearen Gleichungen

$$M_k = \mu_k := \int_{\Omega} x^k e^{-\lambda(x)}dx, \quad k \in \{0, 1, ..., K\}.$$

In dieser Arbeit interessieren dabei im Hinblick auf 101-Dichten lediglich Approximationen der Ordnungen 0 und 1, deren spezielle Form über bestimmten Intervallen später untersucht wird. Zunächst sollen zwei Beispiele die beschriebene Vorgehensweise verdeutlichen.

Beispiel 1 (DME-0 über $\mathbb{R}$) *Sei $\Omega \subseteq \mathbb{R}$ und K=0. Dann gilt*

$$\lambda(x) = \lambda_0, \qquad f(x) = e^{-\lambda_0},$$

$$\mu_0 = \int_{\Omega} f(x)dx = \int_{\Omega} e^{-\lambda_0}dx = 1.$$

Somit ist f die Dichte der stetigen Gleichverteilung auf Ω und damit konstant, d.h. jede DME-0 über einer beliebigen Teilmenge von $\mathbb{R}$ ist eine konstante Funktion.

Beispiel 2 (DME-1 über $\overline{\mathbb{R}}_+$) *Sei $\Omega = \overline{\mathbb{R}}_+$ und K=1. Dann gilt*

$$\lambda(x) = \lambda_0 + \lambda_1 x, \qquad f(x) = e^{-\lambda_0 - \lambda_1 x} = e^{-\lambda_0} e^{-\lambda_1 x}$$

und

$$\begin{aligned}
\mu_0 &= \int_\Omega e^{-\lambda_0 - \lambda_1 x} dx \\
&= \left[-\frac{1}{\lambda_1} e^{-\lambda_0 - \lambda_1 x} \right]_0^\infty \\
&= -\frac{1}{\lambda_1} e^{-\infty} + \frac{1}{\lambda_1} e^{-\lambda_0} \\
&= \frac{1}{\lambda_1} e^{-\lambda_0}.
\end{aligned}$$

Wegen $\mu_0 = 1$ folgt

$$\lambda_1 = e^{-\lambda_0}$$

und

$$f(x) = e^{-\lambda_0} e^{-\lambda_1 x} = \lambda_1 e^{-\lambda_1 x},$$

also ist f die Dichte einer Exponentialverteilung mit Parameter λ_1.

Um den Parameter λ_1 zu berechnen, benötigt man

$$\begin{aligned}
\mu_1 &= \int_\Omega x e^{-\lambda_0 - \lambda_1 x} dx \\
&= \left[\frac{1}{\lambda_1^2} e^{-\lambda_0 - \lambda_1 x} (-\lambda_1 x - 1) \right]_0^\infty \\
&= \frac{1}{\lambda_1^2} e^{-\lambda_0} \\
&= \frac{1}{\lambda_1},
\end{aligned}$$

also $\lambda_1 = \frac{1}{\mu_1}$.

Somit ist f die Dichte der Exponentialverteilung mit Parameter $\frac{1}{\mu_1}$, d.h. jede DME-1 über ganz $\overline{\mathbb{R}}_+$ ist immer die Dichtefunktion einer Exponentialverteilung.

Die soeben hergeleiteten speziellen Formen der Dichten werden bei der Definition der 101-Dichten eine wesentliche Rolle spielen.

2.2 Stückweise Approximation von Dichten

In diesem Abschnitt wird zunächst allgemein das Prinzip der stückweisen Approximation von Dichten durch Dichten maximaler Entropie, genauer durch sogenannte 101-Dichten (kurz 101-Approximation) nach [Str89] beschrieben (2.2.1), dann die Wahl der auftretenden Approximationsparameter näher untersucht (2.2.2) und ein Operator für die in dieser Arbeit bedeutsame Maximumbildung zweier stetiger Zufallsvariablen in der entsprechenden approximativen Darstellung hergeleitet (2.2.3).

Weiterhin wird die praktische Realisierung der 101-Approximation durch das Analysewerkzeug VERENA kurz beschrieben (2.2.4) und eine Erweiterung dieses Analysewerkzeugs um die Operationen Minimum und Maximum stetiger Zufallsvariablen, sowie spezieller Prozeduren zur Berechnung von Momenten stetiger Zufallsvariabler erläutert (2.2.5).

Schließlich wird die Erprobung der Implementierung des Maximumoperators erläutert und eine Übersicht über die größten dabei aufgetretenen Fehler gegeben (2.2.6).

2.2.1 Die 101-Approximation von Dichten

Die in [Str89] beschriebene *stückweise Approximation* von Dichten mittels Dichten maximaler Entropie (DME) beruht auf der Zerlegung von $\overline{\mathbb{R}}$ in paarweise disjunkte Intervalle

$$\begin{aligned}
X_{-\infty} &:= (-\infty, -Nd), \\
X_i &:= [id, (i+1)d), \ i \in \mathcal{L} := \{-N, -N+1, ..., L-1\}, \\
X_{\infty} &:= [Ld, \infty)
\end{aligned}$$

mit $d \in \mathbb{R}^+$; $N,\ L \in \mathbb{N}$.

Die Approximation einer gegebenen Dichte erfolgt nun derart, daß diese Dichte über jedem der obigen Intervallen durch DME angenähert wird.

Die DME über $X_{-\infty}$ sei von der Ordnung K_1, die DME über den $X_i,\ i \in \mathcal{L}$ seien von der Ordnung K_2 und die DME über X_∞ von der Ordnung K_3. Solche Dichten heißen $K_1K_2K_3$-*Dichten*. Für $K_2 > 0$ oder $K_1,\ K_3 > 1$ müssen zur Berechnung der Dichteparameter jeweils nichtlineare Gleichungen gelöst werden, was gravierende numerische Probleme zur Folge hat. Im folgenden wird daher der Spezialfall der 101-Dichten betrachtet. Diese bieten den Vorteil, daß keine nichtlinearen Gleichungen bei der Berechnung der Parameter auftreten. Diese stückweise Approximation von Dichten mit 101-Dichten wird auch abkürzend *101-Approximation* genannt.

Um die DME der Ordnung 0 (DME-0) über den Intervallen $X_i,\ i \in \mathcal{L}$ und die DME der Ordnung 1 (DME-1) über $X_{-\infty}$ bzw. X_∞ zu finden, benötigt man neben den Überlegungen aus dem vorhergehenden Abschnitt die Tatsache, daß über diesen Intervallen der Anteil von M_0 nicht 1 beträgt, da jedes dieser Intervalle nur ein Teil des gesamten Ergebnisraumes ist. Ebenso sind die ersten Momente der äußeren Intervalle $X_{-\infty}$ und X_∞ jeweils nur Teile des ersten Momentes über dem gesamten Ergebnisraum.

Sei $f(t)$ die gegebene Dichte,

$$F(t) := \int_{-\infty}^{t} f(x)dx \quad \text{die Verteilungsfunktion von f,}$$

$$M(t) := \int_{-\infty}^{t} xf(x)dx;$$

insbesondere gilt also

$$F(\infty) = M_0 \qquad \text{und} \qquad M(\infty) = M_1.$$

Somit gilt für $i \in \mathcal{L}, d \in \mathbb{R}^+$:

$$\int_{id}^{(i+1)d} f(x)dx = F((i+1)d) - F(id).$$

Für die DME-0 über $X_i,\ i \in \mathcal{L}$ ergibt sich also :

$$f_i := f_i(x) = \frac{F((i+1)d) - F(id)}{d}.$$

Schwieriger gestaltet sich die Berechnung der DME-1 über $X_{-\infty}$ und X_∞.

Für X_∞ und DME-Ordnung 1 gilt

$$\begin{aligned} \lambda(x) &= \lambda_0 - \lambda_1 x, \\ f(x) &= e^{-\lambda_0 - \lambda_1 x} \end{aligned}$$

und

$$\begin{aligned} \int_{X_\infty} e^{-\lambda_0-\lambda_1 x} dx &= \int_{Ld}^{\infty} e^{-\lambda_0-\lambda_1 x} dx \\ &= \left[-\frac{1}{\lambda_1} e^{-\lambda_0-\lambda_1 x} \right]_{Ld}^{\infty} \\ &= -\frac{1}{\lambda_1} e^{-\lambda_0-\lambda_1 \infty} + \frac{1}{\lambda} e^{-\lambda_0-\lambda_1 Ld} \\ &= \frac{1}{\lambda_1} e^{-\lambda_0-\lambda_1 Ld}. \end{aligned}$$

Mit $\int_{-\infty}^{\infty} f(x)dx = 1$ folgt

$$\begin{aligned} 1 - F(Ld) &= \int_{Ld}^{\infty} f(x)dx \\ &= \frac{1}{\lambda_1} e^{-\lambda_0-\lambda_1 Ld}. \end{aligned}$$

Nun löst man wie folgt nach λ_0 auf :

$$\begin{aligned} e^{-\lambda_0-\lambda_1 Ld} &= \lambda_1(1-F(Ld)), \\ -\lambda_0-\lambda_1 Ld &= \ln(\lambda_1(1-F(Ld))) = \ln\lambda_1 + \ln(1-F(Ld)), \\ \lambda_0 &= -\lambda_1 Ld - \ln\lambda_1 - \ln(1-F(Ld)). \end{aligned}$$

Zudem gilt

$$\begin{aligned} \int_{Ld}^{\infty} x e^{-\lambda_0-\lambda_1 x} dx &= \left[\frac{1}{\lambda_1^2} e^{-\lambda_0-\lambda_1 x}(-\lambda_1 x - 1)\right]_{Ld}^{\infty} \\ &= -\frac{1}{\lambda_1^2} e^{-\lambda_0-\lambda_1 Ld}(-\lambda_1 Ld - 1) \\ &= -\frac{1}{\lambda_1^2}\lambda_1(1-F(Ld))(-\lambda_1 Ld - 1) \\ &= (\frac{1}{\lambda_1} + Ld)(1-F(Ld)). \end{aligned}$$

Mit $\int_{Ld}^{\infty} x f(x) dx = M(\infty) - M(Ld)$ folgt

$$\begin{aligned} M(\infty) - M(Ld) &= (\frac{1}{\lambda_1} + Ld)(1-F(Ld)), \\ \frac{1}{\lambda_1} &= \frac{M(\infty)-M(Ld)}{1-F(Ld)} - Ld, \\ \lambda_1 &= \frac{1}{\frac{M(\infty)-M(Ld)}{1-F(Ld)} - Ld}. \end{aligned}$$

Somit läßt sich die DME-1 von f über X_∞ darstellen als

$$\begin{aligned} f(x) &= e^{\lambda_1 Ld + \ln\lambda_1 + \ln(1-F(Ld)) - \lambda_1 x} \\ &= e^{\lambda_1 Ld} e^{\ln\lambda_1} e^{\ln(1-F(Ld))} e^{-\lambda_1 x} \\ &= e^{\lambda_1 Ld}\lambda_1(1-F(Ld)) e^{-\lambda_1 x}. \end{aligned}$$

Die DME-1 über X_∞ hat also die Form einer Exponentialfunktion

$$f_\infty(x) := ce^{-\lambda x}$$

mit

$$c = \begin{cases} e^{\lambda Ld}\lambda(1-F(Ld)), & \text{falls } F(0) < 1 \\ 0, & \text{falls } F(0) = 1 \end{cases} \tag{1}$$

und

$$\lambda = \begin{cases} \frac{1}{\frac{M(\infty)-M(Ld)}{1-F(Ld)} - Ld}, & \text{falls } F(0) < 1 \\ 1, & \text{falls } F(0) = 1. \end{cases} \tag{2}$$

Analog berechnet man die DME-1 über $X_{-\infty}$ und erhält ebenfalls eine Exponentialfunktion

$$f_{-\infty}(x) := c' e^{\lambda' x}$$

mit

$$c' = \begin{cases} e^{\lambda' N d} \lambda' F(-Nd), & \text{falls } F(0) > 0 \\ 0, & \text{falls } F(0) = 0 \end{cases} \tag{3}$$

und

$$\lambda' = \begin{cases} \frac{1}{-Nd - \frac{M(-Nd)}{F(-Nd)}}, & \text{falls } F(0) > 0 \\ const > 0, & \text{falls } F(0) = 0. \end{cases} \tag{4}$$

Somit ist gezeigt, daß DME der Ordnung 0 bzw. 1 immer eine spezielle Form haben, sodaß sich für 101-Dichten γ ergibt

$$\gamma(t) = \begin{cases} c' e^{\lambda' t} & : \ t \in X_{-\infty}, \\ f_i & : \ t \in X_i, \ i \in \mathcal{L}, \\ c e^{-\lambda t} & : \ t \in X_{\infty} \end{cases}$$

Diese Dichten werden eindeutig durch die Parameter $c', \lambda', c, \lambda, N, L, d$ und $f_i, i \in \mathcal{L}$ beschrieben.

Durch geeignete Wahl von N, L und d läßt sich die Genauigkeit der Approximation beeinflussen. Es ist klar, daß eine feinere Unterteilung der endlichen Intervalle $X_i, i \in \mathcal{L}$ eine höhere Genauigkeit gewährleistet, wobei jedoch für Intervallbreiten $d \approx 0$ die auftretenden Rundungsfehler die Ergebnisse stark beeinflussen. In der Praxis hat sich bisher die Wahl $d = 0.1$ in nahezu allen Fällen als zufriedenstellend erwiesen.

Ein Sonderfall ergibt sich aus der speziellen Form der DME-1 für die 101-Dichte einer exponential verteilten Zufallsvariablen mit Parameter ν. Nach Definition gilt $c' = 0$ und $\lambda' = const$. Mit der Wahl $N := L := 0$ folgt $c = \lambda = \nu$; also ist bei dieser Wahl von N und L die Dichte einer Exponentialverteilung identisch mit ihrer 101-Dichte.

Weiterführende allgemeine Überlegungen zur Wahl dieser 101-Parameter werden im nächsten Abschnitt dieses Kapitels angestellt.

Es werden nun zunächst drei Operatoren auf 101-Dichten definiert, die später zur Formulierung von Algorithmen benötigt werden.

Γ_{101} sei der Operator, der eine gegebene Dichte in eine 101-Dichte transformiert, das heißt, Γ_{101} liefert die oben genannten Parameter $c,\ \lambda,\ c',\ \lambda'$ und f_i zu gegebenen Parametern N, L und d.

$\otimes_{101}$ sei der Operator, der eine 101-Dichte als Approximation der Faltung zweier 101-Dichten liefert. Seien $\gamma_1,\ \gamma_2$ gegebene 101-Dichten und seien die Parameter N, L und d einer 101-Dichte γ gegeben, dann ist $\otimes_{101}$ definiert durch

$$\gamma_1 \otimes_{101} \gamma_2 = \gamma := \Gamma_{101}(\gamma_1 \otimes \gamma_2),$$

wobei $\otimes$ die geläufige Faltung von Zufallsvariablen bezeichnet.

MAX_{101} sei der Operator, der zu einer gegebenen Zufallsvariablen X mit 101-Dichte γ die Verteilung der Zufallsvariablen $\max(X,0)$ liefert, also

$$MAX_{101}\gamma = (\tilde{\gamma}, p_0),$$

wobei $\tilde{\gamma}$ die 'Pseudo-101-Dichte' mit Parametern $c,\ \lambda,\ L$ und $f_i,\ i \in \{0,..L-1\}$ wie für γ ist und mit

$$c' = 0,\ \lambda' = 1,\ N = 0,$$

sowie

$$p_0 = \int_{-\infty}^{0} \gamma(t)dt.$$

Der Begriff 'Pseudo-101-Dichte' soll dabei verdeutlichen, daß die 101-Dichte $\tilde{\gamma}$ an der Stelle 0 unstetig ist, da die Wahrscheinlichkeitsmasse auf dem Intervall $(-\infty,\ 0)$ auf den Wert 0 gesetzt wird.

2.2.2 Zur Wahl der Parameter approximierender 101-Dichten

In [Str89] ist die 101-Approximation allgemein beschrieben, ohne daß auf die Wahl der 101-Parameter N, L und d näher eingegangen wird. In der Praxis stellt jedoch gerade diese Parameterwahl das zentrale Problem der 101-Approximation dar, da die Genauigkeit der Approximation entscheidend von einer geeigneten Wahl von N, L und d abhängt. Daher ist es nun zunächst erforderlich, zu erläutern, wann eine solche Parameterwahl als geeignet angesehen werden kann.

Eine approximierende 101-Dichte konvergiert bei Verwendung potentiell unendlich vieler Intervalle mit beliebig kleinen Intervallbreiten gegen die exakte Dichte. Es erscheint also wünschenswert, bei der 101-Approximation möglichst viele Intervalle mit möglichst kleiner Intervallbreite zu verwenden. Für die 101-Parameter hieße das $d \to 0$; $N, L \to \infty$. Es ist jedoch klar, daß bei der Implementierung auf einem realen Rechner für $d \to 0$ erhebliche Rundungsfehler aufgrund der dann sehr kleinen Wahrscheinlichkeitsmassen und Intervallbreiten und zusätzlich für $N, L \to \infty$ wegen der damit ebenfalls beliebig wachsenden Anzahl von Intervallen Speicherprobleme auftreten können.

Als Faustregeln für die Parameterwahl lassen sich also zunächst informell formulieren :

- Wähle d klein genug, aber nicht zu klein
- Wähle N und L groß genug, aber nicht zu groß

Ein Ansatz zur Erfüllung dieser Faustregeln soll nun näher beschrieben werden. Insbesondere soll die Bedeutung der Formulierungen "klein genug" bzw. "groß genug" und "zu klein" bzw. "zu groß" präzisiert werden.

Die Forderung, d klein zu wählen, resultiert daraus, daß, je größer d, desto größer die mögliche Abweichung der 101-Dichte von der exakten Dichte, da lediglich über allen endlichen Intervallen die Norm der 101-Dichte mit der Norm der exakten Dichte identisch ist. Die einzelnen Funktionswerte der exakten Dichte innerhalb dieser Intervalle können jedoch stark von denen der jeweiligen DME-0, die ja eine konstante Funktion ist, abweichen.

Die Forderung, d nicht zu klein zu wählen, folgt aus der begrenzten Genauigkeit der Zahlendarstellung in realen Rechnern. Die Interpretation von "zu klein" kann also je nach Rechnertyp unterschiedlich sein.

Ähnliches gilt für die Forderung, N und L nicht zu groß zu wählen. Dabei spielt die eingangs erwähnte Speicherproblematik bei heute üblichen Speicherkapazitäten eine untergeordnete Rolle. Vielmehr können auch hier zum Teil erhebliche Rundungsfehler aufgrund sehr kleiner Wahrscheinlichkeitsmassen in weit "außen" liegenden endlichen Intervallen auftreten.

Die Forderung, N und L groß genug zu wählen, beruht darauf, daß in den äußeren Intervallen $X_{-\infty}$ und X_∞, über denen die exakte Dichte durch DME-1 approximiert wird, die Abweichung der Funktionswerte der exakten Dichte und der jeweiligen DME-1 sehr groß werden kann, da diese Intervalle im Gegensatz zu den inneren endlichen Intervallen sehr groß (unendlich) sind und die exakte Dichte lediglich durch eine Exponentialfunktion mit identischer Norm und erstem Moment angenähert wird. Deren Funktionsverlauf kann jedoch stark von dem der exakten Dichte abweichen. In diesem Sinne sind also die Gründe für ein kleines d und für ausreichend große N und L gleichbedeutend.

Die obigen Erläuterungen legen nahe, N, L und d so zu wählen, daß die Zufallsvariable, deren Dichte approximiert werden soll, in den äußeren Intervallen $X_{-\infty}$ und X_∞ Werte nur mit einer Wahrscheinlichkeit annimmt, die eine vorgegebene Schranke $\epsilon \in \mathbb{R}_+$ nicht überschreitet, das heißt, diese Schranke ϵ begrenzt die Wahrscheinlichkeitsmasse, die auf die äußeren Intervalle entfallen darf. Als Hilfsmittel dient dabei die *Tschebyschew-Ungleichung* :

$$P\left[|X - E[X]| \geq c\right] \leq \underbrace{\frac{var(X)}{c^2}}_{Tschebyschew-WK} \qquad \text{für Zufallsvariable } X;\ c \in \mathbb{R}_+.$$

Soll die Tschebyschew-Wahrscheinlichkeit eine Schranke ϵ nicht überschreiten, so muß gelten :

$$\frac{var(X)}{c^2} \leq \epsilon, \qquad c, \epsilon \in \mathbb{R}_+, \tag{5}$$

$$c \geq \sqrt{\frac{var(X)}{\epsilon}}, \qquad c, \epsilon \in \mathbb{R}_+. \tag{6}$$

Wählt man nun für die zu approximierende Dichte einer Zufallsvariablen X die 101-Parameter

N, L und d mit

$$\begin{aligned} -Nd &:= E[X] - c, \\ Ld &:= E[X] + c, \end{aligned}$$

dann folgt für die Wahrscheinlichkeit, daß X Werte in den äußeren Intervallen $X_{-\infty}$ und X_{∞} annimmt

$$\begin{aligned} P[X \in X_{-\infty} \vee X \in X_{\infty}] &= P[X \leq -Nd \ \vee \ X \geq Ld] \\ &= P[X \leq E[X] - c \ \vee \ X \geq E[X] + c] \\ &= P[X - E[X] \leq -c \ \vee \ X - E[X] \geq c] \\ &= P[|X - E[X]| \geq c] \end{aligned}$$

und mit der Tschebyschew-Ungleichung

$$P[X \in X_{-\infty} \vee X \in X_{\infty}] \leq \frac{var(X)}{c^2}.$$

Diese Wahrscheinlichkeit ist kleiner als die Schranke ϵ, wenn Gleichung (6) gilt, insbesondere auch dann, wenn in (6) Gleichheit gilt. Damit folgt

$$\begin{aligned} -Nd &\leq E[X] - \sqrt{\frac{var(X)}{\epsilon}}, \\ Ld &\geq E[X] + \sqrt{\frac{var(X)}{\epsilon}}. \end{aligned}$$

Wählt man nun $d > 0$ beliebig aber fest, so folgen für N und L in Abhängigkeit von d die Bedingungen

$$N \geq \frac{1}{d}\left(\sqrt{\frac{var(X)}{\epsilon}} - E[X]\right), \tag{7}$$

$$L \geq \frac{1}{d}\left(\sqrt{\frac{var(X)}{\epsilon}} + E[X]\right). \tag{8}$$

Beachtet man zudem, daß N und L natürliche Zahlen sind und die Forderung, N und L nicht zu groß zu wählen, so erhält man die Parameterwahl

$$N = \left\lceil \max\left(\frac{1}{d}\left(\sqrt{\frac{var(X)}{\epsilon}} - E[X]\right);\ 0\right)\right\rceil, \tag{9}$$

$$L = \left\lceil \max\left(\frac{1}{d}\left(\sqrt{\frac{var(X)}{\epsilon}} + E[X]\right);\ 0\right)\right\rceil. \tag{10}$$

Beispiel 3 (Erlang-k-Verteilung) *Sei $X \sim E_k(\mu)$ und seien $d, \epsilon > 0$ fest gewählt. Dann gilt*

$$E[X] = \frac{1}{\mu}, \qquad var(X) = \frac{1}{k\mu^2},$$

und damit für N und L

$$N = \left\lceil \max\left(\frac{1}{d}\sqrt{\frac{1}{k\mu^2\epsilon}} - \frac{1}{\mu};\ 0\right)\right\rceil = \left\lceil \frac{1}{d\mu}\left(\frac{1}{\sqrt{k\epsilon}} - 1\right)\right\rceil,$$

$$L = \left\lceil \max\left(\frac{1}{d}\sqrt{\frac{1}{k\mu^2\epsilon}} + \frac{1}{\mu};\ 0\right)\right\rceil = \left\lceil \frac{1}{d\mu}\left(\frac{1}{\sqrt{k\epsilon}} + 1\right)\right\rceil.$$

Beispiel 4 (Erlang-5-Verteilung mit Parameter 0.5) *Sei $X \sim E_5(0.5)$, $d = 0.1$ und $\epsilon = 0.05$. Dann gilt*

$$N = 20,\ L = 60,\ -Nd = -2,\ Ld = 6,\ \textit{also}$$

$$X_{-\infty} = (-\infty,\ -2), \qquad X_{\infty} = [6,\ \infty).$$

Es ergeben sich also 80 innere Intervalle der Breite 0.1.

Bemerkung : Erlang-verteilte Zufallsvariablen nehmen nur positive Werte an. Somit läßt sich die Parameterwahl noch verbessern, indem man $N := 0$ setzt, da keine Wahrscheinlichkeitsmasse auf das Intervall $(-\infty,\ 0)$ entfällt. Damit spart man im obigen Beispiel 20 Intervalle. Ähnliches gilt für viele weitere Verteilungen. Dies zeigt, daß die mit Hilfe der Tschebyschew-Ungleichung gefundenen Schranken nicht scharf sind; sie garantieren lediglich, daß N und L groß genug sind, um die Vorgabe einer Wahrscheinlichkeitsmasse von weniger als ϵ in den äußeren Intervallen zu erfüllen.

Beispiel 5 (Standardisierte Normalverteilung) *Sei $X \sim N(0,1)$, das heißt, X ist normalverteilt mit Erwartungswert 0 und Varianz 1, und seien $d, \epsilon > 0$ fest gewählt. Dann ergibt sich die Parameterwahl*

$$N := L := \frac{1}{d\sqrt{\epsilon}}.$$

und damit

$$X_{-\infty} = (-\infty,\ -\frac{1}{\sqrt{\epsilon}}), \qquad X_{\infty} = [\frac{1}{\sqrt{\epsilon}},\ \infty).$$

Die hier beschriebene Art der Parameterwahl stellt einen ersten Ansatz zur Systematisierung dieser Parameterwahl dar, sollte jedoch in einigen Punkten noch verbessert und konkretisiert werden.

So läßt sich zum Beispiel keine allgemeine Aussage über die geeignete Wahl der Schranke ϵ machen, insbesondere ist diese aus der Tschebyschew-Ungleichung resultierende Schranke nicht scharf, das heißt, die Tschebyschew-Wahrscheinlichkeit kann zum Teil deutlich größer sein als die tatsächliche Wahrscheinlichkeit. Dies führt dazu, daß die Wahrscheinlichkeitsmasse, die tatsächlich auf die äußeren Intervalle entfällt, unter Umständen sehr viel kleiner ist als ϵ. Dies wiederum führt dann zu Konflikten mit der Forderung, N und L nicht zu groß zu wählen.

Ein weiteres Problem ergibt sich bei der Darstellung von Zufallsvariablen, deren Erwartungswert und Varianz unbekannt sind. In der Praxis will man jedoch häufig gerade solche Zufallsvariablen darstellen, so etwa Zufallsvariablen, die aus Operationen auf anderen Zufallsvariablen resultieren. Es stellt sich also das Problem, diese Werte zu approximieren, ohne die Bedingungen (7) und (8) zu verletzen. Dazu wähle man analog zu (9) und (10)

$$N := \left\lceil \max\left(\frac{1}{d}\left(\sqrt{\frac{\widetilde{var}(X)}{\epsilon}} - \widetilde{E_N}[X]\right) ; \ 0\right)\right\rceil \tag{11}$$

$$L := \left\lceil \max\left(\frac{1}{d}\left(\sqrt{\frac{\widetilde{var}(X)}{\epsilon}} + \widetilde{E_L}[X]\right) ; \ 0\right)\right\rceil \tag{12}$$

mit $\widetilde{var}(X) = var(X) + \alpha_1$, $\widetilde{E_N}[X] = E[X] - \alpha_2$, $\widetilde{E_L}[X] = E[X] + \alpha_3$, $\alpha_i \in \mathbb{R}_+$, das heißt, es gilt

$$\widetilde{var}(X) \geq var(X), \ \widetilde{E_N}[X] \leq E[X], \ \widetilde{E_L}[X] \geq E[X].$$

Um zusätzlich die Forderung nach möglichst kleinen N und L zu erfüllen, sollten diese Näherungen möglichst scharf sein, das heißt, die α_i sollten möglichst klein sein.

Das oben beschriebene Problem der geeigneten Wahl von ϵ besteht hierbei natürlich ebenso.

Zusammenfassend kann man also sagen, daß mit der beschriebenen Art der Parameterwahl eine erste Richtlinie für eine weitere Systematisierung gegeben ist, was insofern ein deutlicher Fortschritt ist, daß bisher keinerlei Aussagen zu dieser Parameterwahl gemacht werden konnten. Weiterführende Untersuchungen, deren Notwendigkeit erläutert wurde, sollen im Rahmen dieser Arbeit nicht gemacht werden, denn die Wahl der 101-Parameter stellt kein spezielles Problem bei der Analyse von Tandemsystemen, sondern ein allgemeines Problem der 101-Approximation dar.

2.2.3 Die 101-Dichte des Maximums unabhängiger Zufallsvariablen

Der Operator MAX_{101} ermöglicht die Berechnung des Maximums einer Zufallsvariablen mit 101-Dichte und 0; die Berechnung des Maximums zweier Zufallsvariablen mit 101-Dichten wird nicht berücksichtigt. Im weiteren Verlauf dieser Arbeit wird die Bildung des Maximums von Zufallsvariablen jedoch eine zentrale Rolle spielen. Daher ist es erforderlich, daß auch die Maximumberechnung von Zufallsvariablen mit 101-Dichten möglich ist.

Seien X, Y unabhängige stetige Zufallsvariablen. Dann ist das Maximum $\max(X, Y)$ von X und Y ebenfalls eine stetige Zufallsvariable und hat die Verteilungsfunktion

$$F_{\max}(x) = F_X(x)F_Y(x), \ x \in \overline{\mathbb{R}}.$$

Somit folgt für die Dichte des Maximums

$$f_{\max}(x) = F'_{\max}(x) = F_X(x)f_Y(x) + f_X(x)F_Y(x), \ x \in \overline{\mathbb{R}}. \tag{13}$$

Bemerkung : Die Zufallsvariable $\min(X, Y)$ hat die Verteilungsfunktion

$$\begin{aligned} F_{\min}(x) &= 1-(1-F_X(x))(1-F_Y(x)),\ x \in \overline{\mathbb{R}} \\ &= F_X(x)+F_Y(x)-F_X(x)F_Y(x) = F_X(x)+F_Y(x)-F_{\max}(x),\ x \in \overline{\mathbb{R}}. \end{aligned}$$

und damit die Dichte

$$f_{\min}(x) = f_X(x)+f_Y(x)-f_{\max}(x),\ x \in \overline{\mathbb{R}}.$$

In diesem Abschnitt wird aus den 101-Dichten zweier unabhängiger stetiger Zufallsvariablen X und Y die 101-Dichte der Zufallsvariablen $\max(X, Y)$ berechnet. Dabei sei vorausgesetzt, daß die Parameter N, L und d der 101-Dichten von X, Y und $\max(X, Y)$ jeweils identisch sind, das heißt, allen drei 101-Approximationen liegt die gleiche Zerlegung von $\overline{\mathbb{R}}$ zugrunde. Weiterhin sei vorausgesetzt, daß die Parameter c, c', λ, λ' und $f_i, i \in \mathcal{L}$ für X und Y bekannt sind.

Die 101-Parameter von X seien im folgenden mit c_1, c_1', λ_1, λ_1' und f_i^X, die 101-Parameter von Y entsprechend mit c_2, c_2', λ_2, λ_2' und f_i^Y bezeichnet. Die Verteilungsfunktionen von X bzw. Y seien mit F_X bzw. F_Y und die 101-Dichten mit γ_X und γ_Y bezeichnet.

Die Dichte des Maximums muß nun über den Intervallen $X_{-\infty}$, $X_i, i \in \mathcal{L}$ und X_∞ berechnet werden. Dabei geht man jeweils von den 101-Dichten von X und Y aus.

<u>**1. Schritt :**</u> Berechnung der DME-1 von $\max(X, Y)$ auf $X_{-\infty}$

Sei $x \in X_{-\infty}$. Dann gilt für die 101-Dichten von X und Y

$$\gamma_X(x) = c_1' e^{\lambda_1' x}, \qquad \gamma_Y(x) = c_2' e^{\lambda_2' x}.$$

Damit erhält man für die Verteilungsfunktionen

$$F_X(x) = \frac{c_1'}{\lambda_1'} e^{\lambda_1' x}, \qquad F_Y(x) = \frac{c_2'}{\lambda_2'} e^{\lambda_2' x}.$$

Für die Dichte des Maximums folgt also

$$\begin{aligned} f_{\max}(x) &= \frac{c_1'}{\lambda_1'} e^{\lambda_1' x} c_2' e^{\lambda_2' x} + c_1' e^{\lambda_1' x} \frac{c_2'}{\lambda_2'} e^{\lambda_2 x} \\ &= \frac{c_1' c_2'}{\lambda_1'} e^{(\lambda_1'+\lambda_2')x} + \frac{c_1' c_2'}{\lambda_2'} e^{(\lambda_1'+\lambda_2')x} \\ &= \left(\frac{c_1' c_2'}{\lambda_1'} + \frac{c_1' c_2'}{\lambda_2'} \right) e^{(\lambda_1'+\lambda_2')x}. \end{aligned}$$

Diese Dichte hat bereits die Form einer DME-1 über $X_{-\infty}$, die Parameter erfüllen die Gleichungen (3) und (4). Es gilt also für die DME-1-Parameter von $\max(X, Y)$ auf $X_{-\infty}$

$$c'_{\max} = \frac{c_1' c_2'}{\lambda_1'} + \frac{c_1' c_2'}{\lambda_2'} \qquad \text{und} \qquad \lambda'_{\max} = \lambda_1' + \lambda_2'.$$

2. Schritt : Berechnung der DME-0 ($f_i^{\max}$) von $\max(X,Y)$ auf X_i für $i \in \mathcal{L}$

Sei $x \in X_i = [id,(i+1)d), i \in \mathcal{L}$. Dann gilt für die Verteilungsfunktionen von X und Y

$$\begin{aligned} F_X(x) &= F_X(-Nd) + df_{-N}^X + df_{-N+1}^X + \cdots + df_{i-1}^X + (x-id)f_i^X \\ &= F_X(-Nd) + d\sum_{j=-N}^{i-1} f_j^X + (x-id)f_i^X, \\ F_Y(x) &= F_Y(-Nd) + df_{-N}^Y + df_{-N+1}^Y + \cdots + df_{i-1}^Y + (x-id)f_i^Y \\ &= F_Y(-Nd) + d\sum_{j=-N}^{i-1} f_j^Y + (x-id)f_i^Y. \end{aligned}$$

Für die Dichte des Maximums auf X_i gilt also

$$\begin{aligned} f_{\max}(x) &= (F_X(-Nd) + d\sum_{j=-N}^{i-1} f_j^X + (x-id)f_i^X)f_i^Y \\ &\quad + f_i^X(F_Y(-Nd) + d\sum_{j=-N}^{i-1} f_j^Y + (x-id)f_i^Y). \end{aligned}$$

Damit ergibt sich das Problem, daß $f_i^{\max}$ keine konstante Funktion auf X_i, also keine DME-0 auf X_i ist. Betrachtet man die von x abhängigen Terme, so fällt jedoch auf, daß diese Terme lineare Funktionen mit den Randwerten 0 und df_i^X bzw. 0 und df_i^Y sind. Da eine DME-0 über X_i konstant sein muß und zudem das nullte Moment mit dem der exakten Dichte übereinstimmen muß, ergeben sich für die Wahrscheinlichkeitsmassen von X und Y auf X_i die konstanten Werte $\frac{d}{2}f_i^X$ bzw. $\frac{d}{2}f_i^Y$ und man erhält für die DME-0 des Maximums auf X_i

$$\begin{aligned} f_i^{\max} &= (F_X(-Nd) + d\sum_{j=-N}^{i-1} f_j^X + \frac{d}{2}f_i^X)f_i^Y + f_i^X(F_Y(-Nd) + d\sum_{j=-N}^{i-1} f_j^Y + \frac{d}{2}f_i^Y) \\ &= (F_X(-Nd) + d\sum_{j=-N}^{i-1} f_j^X)f_i^Y + f_i^X(F_Y(-Nd) + d\sum_{j=-N}^{i-1} f_j^Y) + df_i^X f_i^Y \\ &= (\frac{c_1'}{\lambda_1'}e^{-\lambda_1' Nd} + d\sum_{j=-N}^{i-1} f_j^X)f_i^Y + f_i^X(\frac{c_2'}{\lambda_2'}e^{-\lambda_2' Nd} + d\sum_{j=-N}^{i-1} f_j^Y) + df_i^X f_i^Y. \end{aligned}$$

3. Schritt : Berechnung der DME-1 von $\max(X,Y)$ auf X_∞

Sei $x \in X_\infty$. Dann gilt für die 101-Dichten von X und Y

$$\gamma_X(x) = c_1 e^{-\lambda_1 x}, \qquad \gamma_Y(x) = c_2 e^{-\lambda_2 x}.$$

Somit gilt für die Verteilungsfunktionen von X und Y

$$
\begin{aligned}
F_X(x) &= F_X(Ld) + \int_{Ld}^{x} c_1 e^{-\lambda_1 t} dt \\
&= F_X(Ld) + \int_{Ld}^{\infty} c_1 e^{-\lambda_1 t} dt - \int_{x}^{\infty} c_1 e^{-\lambda_1 t} dt \\
&= 1 - \frac{c_1}{\lambda_1} e^{-\lambda_1 x}, \\
F_Y(x) &= F_Y(Ld) + \int_{Ld}^{x} c_2 e^{-\lambda_2 t} dt \\
&= F_Y(Ld) + \int_{Ld}^{\infty} c_2 e^{-\lambda_2 t} dt - \int_{x}^{\infty} c_2 e^{-\lambda_2 t} dt \\
&= 1 - \frac{c_2}{\lambda_2} e^{-\lambda_2 x}.
\end{aligned}
$$

Für die Dichte des Maximums auf X_∞ folgt also

$$
\begin{aligned}
f_{\max}(x) &= (1 - \frac{c_1}{\lambda_1} e^{-\lambda_1 x}) c_2 e^{-\lambda_2 x} + c_1 e^{-\lambda_1 x} (1 - \frac{c_2}{\lambda_2} e^{-\lambda_2 x}) \\
&= c_1 e^{-\lambda_1 x} + c_2 e^{-\lambda_2 x} - \frac{c_1 c_2 (\lambda_1 + \lambda_2)}{\lambda_1 \lambda_2} e^{-(\lambda_1 + \lambda_2) x}.
\end{aligned}
$$

Diese Dichte ist keine DME-1 auf X_∞, es ist also erforderlich, die Wahrscheinlichkeitsmasse und den Anteil des 1. Momentes von $\max(X, Y)$ auf X_∞ zu berechnen, um die DME-1 zu finden.

Es gilt

$$
\begin{aligned}
1 - F_{\max}(Ld) &= \int_{Ld}^{\infty} f_{\max}(x) dx \\
&= \frac{c_1}{\lambda_1} e^{-\lambda_1 Ld} + \frac{c_2}{\lambda_2} e^{-\lambda_2 Ld} - \frac{c_1 c_2}{\lambda_1 \lambda_2} e^{-\lambda_1 + \lambda_2} Ld
\end{aligned}
$$

und

$$
\begin{aligned}
M_{\max}(\infty) - M_{\max}(Ld) &= \int_{Ld}^{\infty} x f_{\max}(x) dx \\
&= (\lambda_1 Ld + 1) \frac{c_1}{\lambda_1^2} e^{-\lambda_1 Ld} + (\lambda_2 Ld + 1) \frac{c_2}{\lambda_2^2} e^{-\lambda_2 Ld} \\
&\quad + ((\lambda_1 + \lambda_2) Ld + 1) \frac{c_1 c_2}{\lambda_1 \lambda_2 (\lambda_1 + \lambda_2)} e^{-(\lambda_1 + \lambda_2) Ld}.
\end{aligned}
$$

Daraus erhält man gemäß (1) und (2) die DME-1-Parameter von $\max(X, Y)$ auf X_∞ zu gegebenen DME-1-Parameter von X und Y.

Analog zu den Operatoren Γ_{101}, $\otimes_{101}$ und MAX_{101} wird nun ein Operator für das Maximum zweier Zufallsvariablen mit 101-Dichten definiert.

$MAXZV_{101}$ sei der Operator, der zu gegebenen Zufallsvariablen X und Y mit 101-Dichten eine 101-Dichte als Approximation des Maximums von X und Y liefert. Seien γ_1, γ_2 die 101-Dichten von X bzw. Y mit identischen Parametern N, L und d und sei γ eine 101-Dichte mit den gleichen Parametern N, L, d. Dann ist $MAXZV_{101}$ definiert durch

$$MAXZV_{101}(\gamma_1, \gamma_2) := \Gamma_{101}(\max(X, Y)),$$

wobei 'max' die geläufige Maximumbildung für Zufallsvariablen bezeichnet und sich die 101-Parameter durch Anwendung des Operators Γ_{101} wie oben beschrieben ergeben.

Ein entsprechender Operator für das Minimum zweier Zufallsvariablen mit 101-Dichten könnte in analoger Weise definiert werden, er wird jedoch nicht benötigt, da sich das Minimum wie zu Beginn dieses Abschnitts bemerkt, direkt aus dem Maximum und den beiden Zufallsvariablen ergibt.

2.2.4 Realisierung von 101-Dichten in VERENA

Wie in der Einleitung bereits erwähnt, stellt das Analysewerkzeug VERENA (VErteilungsREchensystem zur Numerischen Analyse) - eine Modula-2-Programmbibliothek - zahlreiche Datenstrukturen und Operationen auf Zufallsvariablen zur Verfügung. Insbesondere erfolgt die Darstellung von Dichten stetiger Zufallsvariablen durch die beschriebene 101-Approximation. Ebenso sind neben zahlreichen weiteren Operationen die oben beschriebenen Operatoren MAX_{101} und $\otimes_{101}$ implementiert. Ein direkter Zugriff auf die Parameter c, c', λ und λ' ist nicht möglich. Diese Parameter können jedoch aus den Größen

$$W_{links} := F(-Nd) = \frac{c'}{\lambda'} e^{-\lambda' Nd},$$

$$M_{links} := M(-Nd) = (-Nd - \frac{1}{\lambda'}) W_{links},$$

$$W_{rechts} := 1 - F(Ld) = \frac{c}{\lambda} e^{-\lambda Ld},$$

$$M_{rechts} := M(\infty) - M(Ld) = (Ld + \frac{1}{\lambda}) W_{rechts}$$

mittels (1)-(4) berechnet werden.

Diese Größen sind Komponenten des Verbundes 'ZVStrukt', der eine Zufallsvariable beschreibt. Auch der Zugriff auf die Wahrscheinlichkeitsmassen df_i ist möglich. Außerdem ist in VERENA der Typ 'ZufallsVar' als Zeiger auf ein Objekt vom Typ 'ZVStrukt' definiert. Er wird für die Deklaration aller Variablen benutzt, die Zufallsvariablen repräsentieren. Weiterhin stellt VERENA die Prozedur 'KonvertExpM' zur Berechnung der DME-1-Parameter aus obigen Größen, sowie die Prozedur 'KonvertMExp' zur Berechnung obiger Größen aus den DME-1-Parametern zur Verfügung.

VERENA wurde von Dietmar Siepmann ([Sie91]) entwickelt und später mehrfach erweitert. Eine ausführliche Beschreibung findet man in [NB93].

2.2.5 Das Modul 'RV2Plus' - Eine Erweiterung von VERENA

Im Modul 'RV2Plus' sind Prozeduren zur Minimum- und Maximumbildung je zweier unabhängiger stetiger Zufallsvariablen, die in VERENA-Datenstrukturen vorliegen, implementiert. Die Berechnung des Maximums erfolgt dabei gemäß dem beschriebenen Operator $MAXZV_{101}$. Die Berechnung des Minimums spielt in dieser Arbeit zwar keine wesentliche Rolle; die entsprechende Prozedur wurde aber dennoch in 'RV2Plus' implementiert, um eine Art Vollständigkeit dieses Erweiterungsmoduls zu erhalten.

Der Name des Moduls soll dabei die logische Verbundenheit mit dem bestehenden Modul 'RndVar2', welches die Minimum- und Maximumberechnung beliebiger Zufallsvariablen mit Konstanten enthält, deutlich machen.

PROCEDURE MaxStetZV(X, Y: ZufallsVar; VAR M: ZufallsVar);

Die Prozedur 'MaxStetZV' berechnet aus gegebenen unabhängigen Zufallsvariablen X und Y das Maximum 'M' dieser Zufallsvariablen. Dieses Maximum ist nach den Ausführungen der vorausgegangenen Abschnitte eine 101-Approximation des korrekten Ergebnisses. Voraussetzung für die Benutzung dieser Prozedur ist, daß bei der Initialisierung von X, Y und M durch die Prozedur 'InitStetZV' die Grenzen des DME-0 Bereiches und die Intervallbreite d auf diesem Bereich jeweils identisch gewählt werden.

PROCEDURE MinStetZV(X, Y: ZufallsVar; VAR M: ZufallsVar);

Die Prozedur 'MinStetZV' arbeitet analog zu 'MaxStetZV', liefert jedoch in 'M' das Minimum von X und Y zurück.

Weiterhin sind in dem Modul spezielle Prozeduren zur Momentberechnung stetiger Zufallsvariablen enthalten, die in Bezug auf Effizienz und Genauigkeit eine Verbesserung der entsprechenden, bereits bestehenden Prozeduren für allgemeine Zufallsvariablen darstellen. Da sie jedoch nur Hilfsmittel zur Auswertung der in dieser Arbeit erzielten Ergebnisse sind, soll hier nicht näher darauf eingegangen werden.

Der dokumentierte Quelltext, das heißt, das Implementationsmodul, sowie das Definitionsmodul von 'RV2Plus' ist im Anhang enthalten.

2.2.6 Vergleich exakter und approximierter Maxima ausgewählter Zufallsvariablen

Um die Güte des Operators $MAXZV_{101}$ bzw. seiner Implementierung zu untersuchen, wurden die Erwartungswerte, zweiten Momente und Varianzen exakter und approximierter Maxima je zweier unabhängiger Zufallsvariablen verglichen.

Dazu wurden für mehrere ausgewählte Verteilungen von X und Y zunächst die Formeln zur exakten Berechnung der ersten beiden Momente - aus denen direkt die Varianz berechenbar ist - des jeweiligen Maximums hergeleitet.

Die Auswahl der Verteilungen, anhand derer die Implementierung des Operators $MAXZV_{101}$ erprobt wird, erfolgte dabei nach allgemeinen Gesichtspunkten der Modellierung von Warteschlangennetzen.

So verwendet man dabei - auch um die Behandlung mit analytischen Techniken nicht gänzlich auszuschließen - wegen ihrer guten Handhabbarkeit zumeist Exponentialverteilungen oder sogenannte Phasenverteilungen, die durch mehrere exponential verteilte Phasen charakterisiert sind. Beispiele hierfür sind Cox-, Erlang- oder Hyperexponentialverteilungen. Das Prinzip der Kombination von Exponentialverteilungen kann dabei so verallgemeinert werden, daß jede Verteilung mit rationaler Laplace-Tranformierten durch eine Folge exponentieller Phasen dargestellt werden kann.

Ein weiterer Gesichtspunkt war die Übersichtlichkeit der exakten Formeln für die Momente der Maxima jeweils zweier Zufallsvariablen, um den Vergleich nicht durch unnötig viele verwirrende Einzelheiten zu verkomplizieren.

Daher wurden exponential-, hyperexponential- und Erlang-2-verteilte Zufallsvariablen zur Erprobung des Maximumoperators herangezogen, denn sie erfüllen den Zweck, die Approximationsgüte zu demonstrieren, hinreichend. Der Operator $MAXZV_{101}$ und insbesondere seine Implementierung als Prozedur 'MaxStetZV' ist aber selbstverständlich allgemein auf alle stetigen Zufallsvariablen anwendbar.

Zur exakten Berechnung der Maxima und deren Momente ist wiederum anzumerken, daß sich das Maximum von X und Y aus der Beziehung

$$X + Y = \min(X,Y) + \max(X,Y)$$

ergibt.

Somit folgt für die ersten beiden Momente

$$\begin{aligned} E[\max(X,Y)] &= E[X] + E[Y] - E[\min(X,Y)], \\ E[(\max(X,Y))^2] &= E[X^2] + E[Y^2] - E[(\min(X,Y))^2]. \end{aligned}$$

Für einige Verteilungen ist es einfacher, zuerst die Momente des Minimums und dann daraus die Momente des Maximums zu berechnen. Dies setzt natürlich wieder voraus, daß die Verteilungsfunktionen und die Momente von X und Y bekannt sind.

Die allgemeine Vorgehensweise bei der Berechnung der ersten beiden Momente des Maximums zweier unabhängiger stetiger Zufallsvariablen X und Y auf $\Omega \subseteq \mathbb{R}$ ist durch die Definition dieser Momente gegeben :

$$\begin{aligned} E[\max(X,Y)] &= \int_{\Omega} x f_{\max}(x)dx, \\ E[(\max(X,Y))^2] &= \int_{\Omega} x^2 f_{\max}(x)dx, \end{aligned}$$

wobei $f_{\max}(x)$ wieder die Dichte der Zufallsvariablen $\max(X,Y)$ bezeichnet, die gemäß (13) berechnet werden kann.

Es werden nun zunächst die Formeln zur exakten Berechnung der ersten beiden Momente angegeben. Die Herleitung dieser Formeln kann mit den oben angegebenen Definitionen dieser Momente und Gleichung (13) leicht nachvollzogen werden und soll daher an dieser Stelle nicht näher ausgeführt werden.

Formeln für die Varianzen werden jeweils nicht explizit angegeben, da sie zum Teil sehr umfangreich und unübersichtlich sind; die Varianz ergibt sich jedoch ganz allgemein direkt aus der

Differenz des zweiten Momentes und dem quadrierten Erwartungswert, also insbesondere gilt

$$var[\max(X,Y)] = E[(\max(X,Y))^2] - (E[\max(X,Y)])^2.$$

Im folgenden bezeichne die Zufallsvariable M jeweils abkürzend das Maximum der Zufallsvariablen X und Y, also

$$M := \max(X,Y).$$

Maximum zweier exponential verteilter Zufallsvariablen

Für $X \sim Exp(\lambda)$, $Y \sim Exp(\mu)$ gilt

$$\begin{aligned} E[M] &= \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda+\mu}, \\ E[M^2] &= 2\left(\frac{1}{\lambda^2} + \frac{1}{\mu^2} - \frac{1}{(\lambda+\mu)^2}\right). \end{aligned}$$

Maximum einer exponential und einer hyperexponential verteilten Zufallvariablen

Für $X \sim Exp(\lambda)$, $Y \sim H_2(\mu_1;\mu_2;q)$ gilt

$$\begin{aligned} E[M] &= \frac{1}{\lambda} + \frac{q}{\mu_1} + \frac{1-q}{\mu_2} - \frac{q}{\lambda+\mu_1} - \frac{1-q}{\lambda+\mu_2}, \\ E[M^2] &= 2\left(\frac{1}{\lambda^2} + \frac{q}{\mu_1^2} + \frac{1-q}{\mu_2^2} - \frac{q}{(\lambda+\mu_1)^2} - \frac{(1-q)}{(\lambda+\mu_2)^2}\right). \end{aligned}$$

Maximum einer exponential und einer Erlang-2 verteilten Zufallsvariablen

Für $X \sim Exp(\lambda)$, $Y \sim E_2(\mu)$ gilt

$$\begin{aligned} E[M] &= \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda+\mu} - \frac{2\mu}{(\lambda+2\mu)^2}, \\ E[M^2] &= \frac{2}{\lambda^2} + \frac{3}{2\mu^2} - \frac{2\lambda}{(\lambda+2\mu)^2} - \frac{12\mu}{(\lambda+2\mu)^3}. \end{aligned}$$

Maximum zweier Erlang-2 verteilter Zufallsvariablen

Für $X \sim E_2(\lambda)$, $Y \sim E_2(\mu)$ gilt

$$\begin{aligned} E[M] &= \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda+\mu} - \frac{\lambda\mu}{(\lambda+\mu)^3}, \\ E[M^2] &= \frac{3}{2}\left(\frac{1}{\lambda^2} + \frac{1}{\mu^2} - \frac{1}{(\lambda+\mu)^2} - \frac{2\lambda\mu}{(\lambda+\mu)^4}\right). \end{aligned}$$

Maximum einer Erlang-2 und einer hyperexponential verteilten Zufallsvariablen

Für $X \sim E_2(\lambda)$, $Y \sim H_2(\mu_1;\mu_2;q)$ gilt

$$E[M] = \frac{1}{\lambda} + \frac{q}{\mu_1} + \frac{1-q}{\mu_2} - \frac{q}{2\lambda+\mu_1} - \frac{1-q}{2\lambda+\mu_2} - \frac{2\lambda q}{(2\lambda+\mu_1)^2} - 2\lambda(1-q)(2\lambda+\mu_2)^2$$

$$E[M^2] = \frac{3}{2\lambda^2} + 2\left(\frac{q}{\mu_1^2} + \frac{1-q}{\mu_2^2} - \frac{q}{(2\lambda+\mu_1)^2} - \frac{1-q}{(2\lambda+\mu_2)^2}\right) - \frac{8\lambda q}{(2\lambda+\mu_1)^3} - \frac{8\lambda(1-q)}{(2\lambda+\mu_2)^3}.$$

Maximum zweier hyperexponential verteilter Zufallsvariablen

Für $X \sim H_2(\lambda_1;\lambda_2;p)$, $Y \sim H_2(\mu_1;\mu_2;q)$ gilt

$$E[M] = \frac{p}{\lambda_1} + \frac{1-p}{\lambda_2} + \frac{q}{\mu_1} + \frac{1-q}{\mu_2} - \frac{pq}{\lambda_1+\mu_1} - \frac{p(1-q)}{\lambda_1+\mu_2} - \frac{(1-p)q}{\lambda_2+\mu_1} - \frac{(1-p)(1-q)}{\lambda_2+\mu_2}.$$

$$\begin{aligned} E[M^2] &= 2\left(\frac{p}{\lambda_1^2} + \frac{1-p}{\lambda_2^2} + \frac{q}{\mu_1^2} + \frac{1-q}{\mu_2^2}\right) \\ &\quad -2\left(\frac{pq}{(\lambda_1+\mu 1)^2} + \frac{(1-p)q}{(\lambda_2+\mu_1)^2} + \frac{p(1-q)}{(\lambda_1+\mu_2)^2} + \frac{(1-p)(1-q)}{(\lambda_2+\mu_2)^2}\right). \end{aligned}$$

Zur Erprobung des Operators $MAXZV_{101}$ bzw. seiner Implementierung wurden insgesamt 384 Maxima berechnet. Dabei wurden die Parameter der Zufallsvariablen X und Y für jede der oben aufgeführten Kombinationen der Verteilungen so gewählt, daß die Erwartungswerte $E[X]$ bzw. $E[Y]$ jeweils alle Werte aus $\{0.2,\ 0.5,\ 1.0,\ 1.5,\ 2.0,\ 3.0,\ 4.0,\ 8.0\}$ annahmen. Für jede Kombination der Verteilungen wurden also 64 Maxima berechnet.

Es wurden die 101-Parameter $N := 0$ und $d := 0.05$ festgelegt und der Parameter L jeweils so gewählt, daß $Ld = E[X] + E[Y]$ galt, also $L := \frac{1}{d}(E[X]+E[Y]) = 20(E[X]+E[Y])$.

Die folgende Tabelle gibt eine Übersicht über die größten beobachteten relativen Fehler, woraus die Qualität der Implementierung von $MAXZV_{101}$ ersichtlich wird.

Verteilung von X	Verteilung von Y	relativer Fehler in %		
		$E[M]$	$E[M^2]$	$var[M]$
Exp	Exp	0.94	4.42	9.00
Exp	H_2	0.42	1.31	2.22
Exp	E_2	0.64	1.18	4.11
E_2	E_2	0.08	1.35	5.76
E_2	H_2	0.39	1.98	4.53
H_2	H_2	0.21	5.78	13.18

Tabelle 1: Größte beobachtete relative Fehler in %

2.3 (max,+)-lineare Systeme

Vielfach finden bei der Analyse allgemeiner stochastischer Ereignisgraphen offene (max,+)-lineare Systeme Anwendung. Diesen Systemen liegt die sogenannte (max,+)-Algebra zugrunde, deren Operatoren in weitgehender Analogie zur Maximumbildung bzw. Addition auf den reellen Zahlen definiert sind ([BCOQ92]).

Beispiele für die Anwendung solcher stochastischer Ereignisgraphen sind unter anderem zahlreiche Typen von Netzwerken, wie stochastische Petrinetze, Fork-Join-Netze oder, in dieser Arbeit bedeutend, Tandem-Warteschlangennetze in verschiedenen Variationen. Dabei entstehen durch die Modellierung mittels (max,+)-linearer Systeme keinerlei Restriktionen bezüglich der Verteilungen der Zufallsvariablen, die Bedien- und Zwischenankunftszeiten repräsentieren.

Es erfolgt nun zunächst eine formale Beschreibung von (max,+)-linearen Systemen und der zugrundeliegenden Algebra und dann eine äquivalente Darstellung zur iterativen Berechnung stationärer Größen.

Definition 2 (($\oplus, \otimes$)-Algebra) *Eine algebraische Struktur $(\mathcal{K}, \oplus, \otimes)$ heißt $(\oplus, \otimes)$-Algebra, falls gilt:*

- *$\oplus$ ist assoziativ, kommutativ und besitzt ein neutrales Element ε,*
- *$(\mathcal{K}, \otimes)$ ist eine Gruppe,*
- *$\otimes$ ist distributiv bezüglich $\oplus$,*
- *Das $\otimes$-neutrale Element e erfüllt $\varepsilon \otimes e = e \otimes \varepsilon = \varepsilon$,*
- *$\otimes$ hat Priorität vor $\oplus$.*

Definition 3 ((max,+)-Algebra) *Die algebraische Struktur $(\overline{\mathbb{R}}, max, +)$, wobei max die Maximumbildung auf $\overline{\mathbb{R}}$ mit neutralem Element $-\infty$ und $+$ die Addition auf $\overline{\mathbb{R}}$ mit neutralem Element 0 bezeichnet, ist eine $(\oplus, \otimes)$-Algebra und heißt die (max,+)-Algebra auf $\overline{\mathbb{R}}$ oder kurz (max,+)-Algebra.*

Bemerkung: Da $a \leq b$ für $\max(a, b) = b$ gilt, wird die (max,+)-Algebra auch *teilweise geordnete Gruppe* genannt.

Definition 4 (Reihenkomposition) *Sei A reelle $(p \times q)$-Matrix und B reelle $(q \times r)$-Matrix. Dann ist die reelle $(p \times r)$-Matrix $A \otimes B$ definiert durch*

$$(A \otimes B)_{i,j} := \bigoplus_{k=1}^{q} A_{i,k} \otimes B_{k,j} = \max_{k=1}^{q}\{A_{i,k} + B_{k,j}\}, \qquad i = 1, \ldots, p;\ j = 1, \ldots, r.$$

Definition 5 (Parallelkomposition) *Seien A und B reelle $(p \times q)$-Matrizen. Dann ist die reelle $(p \times q)$-Matrix $A \oplus B$ definiert durch*

$$(A \oplus B)_{i,j} := A_{i,j} \oplus B_{i,j} = \max(A_{i,j};\ B_{i,j}), \qquad i = 1, \ldots, p;\ j = 1, \ldots, q.$$

Bemerkung: Reihen- und Parallelkomposition von Matrizen mit den Operatoren $\otimes$ und $\oplus$ sind also analog zur allgemein geläufigen reellen Matrixmultiplikation bzw. Matrixaddition definiert.

Definition 6 ((max,+)-lineares System) *Sei $\{T_n\}$ eine isotone Folge reeller Zufallsvariablen, $\{A_n\}$ eine Folge stationärer reeller $(\alpha \times \alpha)$-Zufallsmatrizen und $\{B_n\}$ eine Folge stationärer reeller Zufallsvektoren der Dimension α. Eine Folge $\{X_n\}$ reeller α-dimensionaler Zufallsvektoren heißt offenes (max,+)-lineares stochastisches System oder kurz (max,+)-lineares System, falls die Rekursion*

$$X_{n+1} = A_n \otimes X_n \oplus B_{n+1} \otimes T_{n+1} \tag{14}$$

erfüllt ist.

Schreibt man die Rekursion in Komponentenform und benutzt dann die auf $\overline{\mathbb{R}}$ geläufigeren Operatoren max und $+$, so ergibt sich

$$\begin{aligned}
\begin{pmatrix} X_{n+1,1} \\ \vdots \\ X_{n+1,\alpha} \end{pmatrix} &= \begin{pmatrix} (A_n)_{11} & (A_n)_{12} & \cdots & (A_n)_{1\alpha} \\ \vdots & \vdots & & \vdots \\ (A_n)_{\alpha 1} & (A_n)_{\alpha 2} & \cdots & (A_n)_{\alpha\alpha} \end{pmatrix} \otimes \begin{pmatrix} X_{n,1} \\ \vdots \\ X_{n,\alpha} \end{pmatrix} \oplus \begin{pmatrix} B_{n+1,1} \\ \vdots \\ B_{n+1,\alpha} \end{pmatrix} \otimes T_{n+1} \\
&= \begin{pmatrix} \bigoplus_{k=1}^{\alpha} ((A_n)_{1k} \otimes X_{n,k}) \\ \vdots \\ \bigoplus_{k=1}^{\alpha} ((A_n)_{\alpha k} \otimes X_{n,k}) \end{pmatrix} \oplus \begin{pmatrix} B_{n+1,1} \otimes T_{n+1} \\ \vdots \\ B_{n+1,\alpha} \otimes T_{n+1} \end{pmatrix} \\
&= \begin{pmatrix} \max\{(A_n)_{11} + X_{n,1}, \ldots, (A_n)_{1\alpha} + X_{n,\alpha}\} \\ \vdots \\ \max\{(A_n)_{1\alpha} + X_{n,1}, \ldots, (A_n)_{\alpha\alpha} + X_{n,\alpha}\} \end{pmatrix} \oplus \begin{pmatrix} B_{n+1,1} + T_{n+1} \\ \vdots \\ B_{n+1,\alpha} + T_{n+1} \end{pmatrix} \\
&= \begin{pmatrix} \max(\max\{(A_n)_{11} + X_{n,1}, \ldots, (A_n)_{1\alpha} + X_{n,\alpha}\}, B_{n+1,1} + T_{n+1}) \\ \vdots \\ \max(\max\{(A_n)_{\alpha 1} + X_{n,1}, \ldots, (A_n)_{\alpha\alpha} + X_{n,\alpha}\}, B_{n+1,\alpha} + T_{n+1}) \end{pmatrix}. \qquad (15)
\end{aligned}$$

In allen Modellen, die mit Hilfe von (max,+)-linearen Systemen untersucht werden, repräsentiert die Folge $\{T_n\}$ die Ankunftszeitpunkte der einzelnen Kunden im System, das heißt, T_n ist der Zeitpunkt, zu dem der n-te Kunde in das System eintritt. Die Koordinaten $X_{n,i}$ eines Zufallsvektors der Folge $\{X_n\}$ repräsentieren ebenfalls Zeitpunkte, wie etwa den Beginn der n-ten Bedienung, die die i-te Bedienstation des Netzwerks durchführt.

Hierbei seien die Bedienstationen des Netzwerks numeriert, wobei die Numerierung für allgemeine Netzwerke, im Gegensatz zu Tandemsystemen im speziellen, keine Bedeutung für die Reihenfolge der Bedienung der Kunden hat. Durch die Matrizen A_n und die Vektoren B_n sind, je nach analysiertem Modell unterschiedlich, Bedienzeiten der Kunden repräsentiert. Eine genaue Spezifikation für Tandemsysteme wird in Kapitel 4 gegeben.

Somit lassen sich also aus gegebenen Ankunfts- und Bedienzeiten mit einem Startvektor X_0 die Bedienstartzeiten der einzelnen Bediener iterieren. Nun ist jedoch offensichtlich, daß die Folge der Bedienstartzeiten divergiert und keine wesentlichen Merkmale des Systems direkt aus ihr hervorgehen. Weitaus nützlicher sind daher die Zeitdifferenzen zwischen Ankunfts- und Bedienstartzeiten, die nun bestimmt werden.

Es sei also $\{W_n\}$ mit $W_{n,i} := X_{n,i} - T_n$ für $i \in \{1, \ldots, \alpha\}, n \in \mathbb{N}$ die Folge der Verweilzeiten bis zum Bedienstart in den Bedienern, das heißt, $W_{n,i}$ ist die Zeit zwischen Ankunft des n-ten Kunden und dem Start der n-ten Bedienung im i-ten Bediener. Für Tandemsysteme entspricht die n-te Bedienung in einem Bediener gerade der Bedienung des n-ten Kunden in diesem Bediener. Im allgemeinen Fall läßt sich jedoch nicht angeben, welcher Kunde die n-te Bedienung eines bestimmten Bedieners erhält.

Es wird nun (14) so modifiziert, daß eine Rekursion für diese beschriebenen Verweilzeiten resultiert.

Sei $i \in \{1, \ldots, \alpha\}, n \in \mathbb{N}$, sowie $\tau_n := T_{n+1} - T_n$ die Zwischenankunftszeit zwischen dem n-ten und $(n+1)$-ten Kunden und sei $C(x)$ die $\alpha \times \alpha$-Matrix

$$C(x) := \begin{pmatrix} -x & & -\infty \\ & \ddots & \\ -\infty & & -x \end{pmatrix} \quad \text{für} \quad x \in \mathbb{R}.$$

Dann gilt

Satz 1 $X_{n+1} = A_n \otimes X_n \oplus B_{n+1} \otimes T_{n+1} \Leftrightarrow W_{n+1} = A_n \otimes W_n \otimes C(\tau_n) \oplus B_{n+1}$

Beweis

$$X_{n+1} = A_n \otimes X_n \oplus B_{n+1} \otimes T_{n+1}$$

$$\begin{pmatrix} X_{n+1,1} \\ \vdots \\ X_{n+1,\alpha} \end{pmatrix} = \begin{pmatrix} \max(\max\{(A_n)_{11} + X_{n,1}, \ldots, (A_n)_{1\alpha} + X_{n,\alpha}\}, B_{n+1,1} + T_{n+1}) \\ \vdots \\ \max(\max\{(A_n)_{\alpha 1} + X_{n,1}, \ldots, (A_n)_{\alpha\alpha} + X_{n,\alpha}\}, B_{n+1,\alpha} + T_{n+1}) \end{pmatrix}$$

$$\begin{pmatrix} X_{n+1,1} - T_{n+1} \\ \vdots \\ X_{n+1,\alpha} - T_{n+1} \end{pmatrix} = \begin{pmatrix} \max(\max\{(A_n)_{11} + X_{n,1}, \ldots, (A_n)_{1\alpha} + X_{n,\alpha}\} - T_{n+1}), B_{n+1,1} \\ \vdots \\ \max(\max\{(A_n)_{\alpha 1} + X_{n,1}, \ldots, (A_n)_{\alpha\alpha} + X_{n,\alpha}\} - T_{n+1}), B_{n+1,\alpha} \end{pmatrix}$$

Somit ergibt sich

$$\begin{pmatrix} W_{n+1,1} \\ \vdots \\ W_{n+1,\alpha} \end{pmatrix} = \begin{pmatrix} \max\{(A_n)_{11} + X_{n,1}, \ldots, (A_n)_{1\alpha} + X_{n,\alpha}\} - T_{n+1} \\ \vdots \\ \max\{(A_n)_{\alpha 1} + X_{n,1}, \ldots, (A_n)_{\alpha\alpha} + X_{n,\alpha}\} - T_{n+1} \end{pmatrix} \oplus B_{n+1}$$

$$= \begin{pmatrix} \max\{(A_n)_{11} + W_{n,1} + T_n, \ldots, (A_n)_{1\alpha} + W_{n,\alpha} + T_n\} - T_{n+1} \\ \vdots \\ \max\{(A_n)_{\alpha 1} + W_{n,1} + T_n, \ldots, (A_n)_{\alpha\alpha} + W_{n,\alpha} + T_n\} - T_{n+1} \end{pmatrix} \oplus B_{n+1}$$

$$= \begin{pmatrix} \max\{(A_n)_{11} + W_{n,1}, \ldots, (A_n)_{1\alpha} + W_{n,\alpha}\} + T_n - T_{n+1} \\ \vdots \\ \max\{(A_n)_{\alpha 1} + W_{n,1}, \ldots, (A_n)_{\alpha\alpha} + W_{n,\alpha}\} + T_n - T_{n+1} \end{pmatrix} \oplus B_{n+1}$$

$$= \begin{pmatrix} \max\{(A_n)_{11} + W_{n,1}, \ldots, (A_n)_{1\alpha} + W_{n,\alpha}\} - \tau_n \\ \vdots \\ \max\{(A_n)_{\alpha 1} + W_{n,1}, \ldots, (A_n)_{\alpha\alpha} + W_{n,\alpha}\} - \tau_n \end{pmatrix} \oplus B_{n+1}$$

$$= A_n \otimes W_n - \begin{pmatrix} \tau_n \\ \vdots \\ \tau_n \end{pmatrix} \oplus B_{n+1}.$$

Aus der Überlegung, daß $\{M_{i,k} + (C(x))_{k,j}\}$ für alle α-spaltige reelle Matrizen M und alle $x \in \mathbb{R}$ sein Maximum für $k = j$ annimmt, da alle Nichtdiagonalelemente von $C(x)$ den Wert $-\infty$ haben, folgt für beliebige α-spaltige reelle Matrizen M

$$(M \otimes C(\tau_n))_{i,j} = \max_{k=1}^{\alpha}\{M_{i,k} + (C(\tau_n))_{k,j}\} = M_{i,j} + (C(\tau_n))_{j,j} = M_{i,j} - \tau_n$$

und somit speziell für $M = A_n \otimes W_n$ aus obiger Gleichung

$$W_{n+1} = A_n \otimes W_n \otimes C(\tau_n) \oplus B_{n+1} \quad \square \tag{16}$$

Es ist jedoch neben dieser Rekursion keine geschlossene Formel zur direkten Berechnung eines stationären $W = (W_1, W_2, \ldots, W_\alpha)^T$ bekannt.

In [BS95] wird eine Potenzreihenentwicklung zur Approximation des Erwartungswertes von W für offene Systeme mit Poisson-Ankunftsprozeß auf der Basis von (max,+)-linearen Systemen vorgeschlagen.

Dabei wird ein probabilistischer Ausdruck für die Koeffizienten $c_k^i \in \mathbb{R}$ der Gleichung

$$E[W_i] = \sum_{k=0}^{m} c_k^i \lambda^k + O(\lambda^{m+1}), \quad m \in \mathbb{N}^+,\ i \in \{1, \ldots, \alpha\};$$

wobei λ die Ankunftsrate bezeichnet, hergeleitet, durch den die c_k^i als Erwartungswert eines Polynoms

$$p_{k+1} : \mathbb{R}^{k+1} \to \mathbb{R}$$

dargestellt werden können. Dieses Verfahren wird dann auf verschiedene Warteschlangennetze angewandt.

Im Vergleich zu vielen anderen Verfahren, die Erwartungswerte oder höhere Momente von Leistungsgrößen in Warteschlangennetzen durch Potenzreihen approximieren, hat dieses Verfahren den Vorteil, daß hiermit Netze mit beliebigen unabhängig identisch verteilten Bedienzeiten analysiert werden können, wohingegen die meisten der älteren Verfahren nur auf Netze mit ausschließlich unabhängig identisch exponential verteilten Bedienzeiten anwendbar sind.

Zudem ist das beschriebene Verfahren auf eine ganze Klasse von Systemen, die Klasse der (max,+)-linearen Systeme, anwendbar, die rein durch ihre strukturellen Eigenschaften definiert ist, und nicht etwa durch die Generatormatrix der zur Modellierung verwendeten Markovkette.

Dieser letzgenannte Vorteil der Analyse mittels (max,+)-linearer Systeme ermöglicht auch die bereits erwähnte leichte Modifizierbarkeit auch der später in der vorliegenden Arbeit speziell für Tandemsysteme entwickelten Verfahren auf andere Netze, die durch (max,+)-lineare Systeme beschrieben werden können. Es müssen lediglich jeweils die für die zu untersuchenden Netze geeigneten Spezifikationen der Komponenten eines (max,+)-linearen Systems angegeben und bei Verwendung der 101-Approximation die resultierende Rekursion bzw. Iteration so umgeformt werden, daß die Operatoren max und + nur auf unabhängige Zufallsvariablen angewandt werden.

Dabei besteht in dieser Umformung jeweils das Hauptproblem der Analyse, da einerseits bisher nicht garantiert werden kann, daß für alle durch (max,+)-lineare Systeme modellierbare Netze eine solche Umformung möglich ist und andererseits selbst unter der Voraussetzung, daß eine solche Darstellung mit den Operationen max und + ausschließlich auf unabhängigen Zufallsvariablen existiert, bisher kein festes Schema für eine entsprechende Umformung bekannt ist. Dieser Schritt der Analyse muß also jeweils 'von Hand' ausgeführt werden.

Dabei kommt erschwerend hinzu, daß für viele Warteschlangennetze Abhängigkeiten auftretender Größen wie Warte- oder Verweilzeiten einzelner Kunden nicht geklärt sind, sodaß man im Zweifelsfall jeweils die Abhängigkeit solcher Größen annehmen muß, um falsche Unabhängigkeitsannahmen zu vermeiden.

Ein großer Vorteil der Benutzung approximierender 101-Dichten im Vergleich zum Verfahren aus [BS95] besteht jedoch darin, daß der Ankunftsprozeß ein beliebiger Erneuerungsprozeß sein darf und nicht speziell ein Poisson-Prozeß sein muß.

3 Tandemsysteme

In diesem Kapitel erfolgt eine präzise Spezifikation der in der Einleitung bereits informell beschriebenen Tandemsysteme, eine Zusammenstellung von bisherigen Ergebnissen bei der Analyse solcher Tandemsysteme und deren Leistungsgrößen, sowie die Beschreibung zweier Näherungsverfahren und die Entwicklung einer ersten iterativen Approximation unter Verwendung von 101-Dichten.

3.1 Definitionen und Notationen

Definition 7 (Tandemsystem) *Ein Tandemsystem ist ein offenes Warteschlangennetzwerk, das die folgenden Bedingungen erfüllt :*

- *Jeder Knoten (=Bedienstation) besteht aus einem Bediener und einer FIFO-Warteschlange unendlicher Kapazität, die Bedienstrategie ist also FCFS.*
- *Die Bedienstationen sind von 1 bis* $\alpha \in \mathbb{N}^+$ *durchnumeriert.*
- *Ankünfte finden nur an Bedienstation 1 statt.*
- *Kunden, die in Bedienstation* $i \in \{1, \ldots, \alpha - 1\}$ *fertig bedient worden sind, gehen über in Bedienstation i+1.*
- *Kunden, die in Bedienstation* α *fertig bedient worden sind, verlassen das System.*

Die einzelnen Bediener eines Tandemsystems werden entsprechend der Reihenfolge, in der sie durchlaufen werden, mit $b_1, \ldots, b_\alpha$ und die Kunden im System entsprechend der Reihenfolge ihrer Ankunft mit $K_1, K_2, \ldots$ bezeichnet.

Da ein Kunde niemals einen vor ihm ins System eingetretenen Kunden überholen kann, ist die Austrittsreihenfolge aus dem System gleich der Eintrittsreihenfolge in das System, wobei jedoch die Zwischenabgangszeiten im allgemeinen verschieden von den Zwischenankunftszeiten sind.

Die Zwischenankunfts- und Bedienzeiten seien durch unabhängig identisch verteilte Zufallsvariablen modelliert. Sind diese Zwischenankunfts- oder Bedienzeiten näher spezifiziert, so wird eine Notation der Form $A \rightarrow /B_1/1 \rightarrow /B_2/1 \rightarrow \cdots \rightarrow /B_\alpha/1$ verwendet. Dabei steht A für die Verteilung der Zwischenankunftszeiten an der ersten Station und B_i, $i = 1, \ldots, \alpha$ für die Verteilung der Bedienzeit am Bediener der i-ten Station. Besitzen k hintereinander angeordnete Bedienstationen gleiche Bedienzeitverteilung B, so wird für diese Bedienstationen die abkürzende Schreibweise $(\rightarrow /B/c)^k$ benutzt.

In dieser Notation werden also insbesondere Einbediener-Tandemsysteme mit allgemein verteilten unabhängigen Zwischenankunfts- und Bedienzeiten, die in dieser Arbeit untersucht werden,

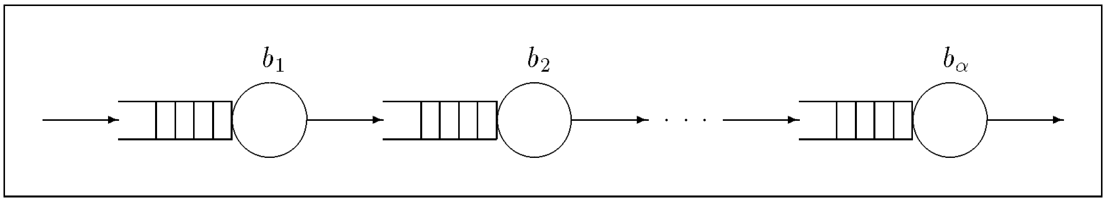

Abbildung 1: Tandemsystem mit α Bedienstationen

als $G(\to /G/1)^\alpha$-System bezeichnet. Entsprechend der erweiterten Kendall-Notation können mit der hier eingeführten Notation auch Varianten der oben definierten Tandemsysteme, wie etwa Systeme mit mehreren Bedienern je Bedienstation, beschränkter Warteschlangenkapazitäten oder von FCFS verschiedenen Bedienstrategien beschrieben werden. Solche Systeme sollen mit *modifiziertem Tandemsystem* bezeichnet werden.

Beispiel 6 *Man betrachte das System $E_2 \to /M/1 \to /U/3 \to /M/1 \to /H_2/2$. Hierbei handelt es sich um ein modifiziertes Tandemsystem mit vier Bedienstationen, wobei die Zwischenankunftszeiten an der ersten Station Erlang-2 verteilt sind. Die erste Bedienstation besteht aus einem Bediener, dessen Bedienzeitverteilung exponential ist, die zweite Bedienstation besteht aus drei Bedienern mit stetig gleichverteilten Bedienzeiten, die dritte Station aus einem Bediener mit exponential verteilten Bedienzeiten und die vierte Station schließlich besteht aus zwei Bedienern, deren Bedienzeiten hyperexponential verteilt sind.*
Da nicht explizit anders angegeben, arbeiten hier alle Warteschlangen nach dem FIFO-Prinzip und haben unendliche Kapazität.

Das nächste Beispiel bezieht auch die Möglichkeiten von FCFS verschiedener Bedienstrategien, endlicher Warteschlangenkapazitäten, sowie mehrerer identischer, aufeinanderfolgender Bedienstationen ein.

Beispiel 7 *Man betrachte das System $M \to /G/2/LCFS\ (\to /M/5)^3 \to /U/1/10/PRIO$. Dieses System besteht aus fünf Bedienstationen, wobei die Zwischenankünfte an der ersten Station exponential verteilt sind. Die erste Bedienstation besteht aus zwei Bedienern, deren Bedienzeit allgemein unabhängig identisch verteilt ist. Die Bedienstrategie bei unendlicher Warteschlangenkapazität ist LCFS, das heißt, der letzte Kunde, der diese Station betritt, wird zuerst bedient. Die drei folgenden Bedienstationen bestehen aus jeweils fünf Bedienern mit exponential verteilter Bedienzeit. Die Warteschlangenkapazität ist unendlich und die Bedienstrategie ist FCFS. Die fünfte Station schließlich besteht aus einem Bediener mit gleichverteilten Bedienzeiten und einer Warteschlange mit zehn Plätzen. Die Bedienstrategie erfolgt gemäß verschiedener Prioritäten der Kunden. Bei beschränkter Warteschlangenkapazität trete ein in der vorhergehenden Station vollständig bedienter Kunde erst dann aus dieser Station aus, wenn die endliche Warteschlange nicht voll ist. Er blockiert also den vorhergehenden Bediener.*

Zur formalen Beschreibung von Tandemsystemen mit $\alpha \in \mathbb{N}^+$ Bedienstationen werden die folgende Bezeichnungen verwendet, wobei $i \in \{1, \ldots, \alpha\}$ gelte:

$(k_1, \ldots, k_\alpha)$	ist der Zustand des Tandemsystems, wobei $k_i \in \mathbb{N}$ die Anzahl der Kunden in Bedienstation i angibt,
$P_t(k_1, \ldots, k_\alpha)$	ist die Wahrscheinlichkeit, daß sich das System zur Zeit t im Zustand $(k_1, \ldots, k_\alpha)$ befindet,
$P_{t,i}(k_i)$	ist die Wahrscheinlichkeit, daß sich zur Zeit t k_i Kunden in Bedienstation i befinden,
λ_i	ist die mittlere Ankunftsrate von Kunden an Station i,
μ_i	ist die mittlere Bedienrate von Kunden in Station i,
$\rho_i = \frac{\lambda_i}{\mu_i}$	ist die mittlere Auslastung der Station i.

Die Übergangswahrscheinlichkeiten der Kunden von Station i nach Station j $(i, j \in \{0, \ldots, \alpha\})$ ergeben sich für Tandemsysteme wie folgt :

$$p_{ij} = \begin{cases} 1, & \text{falls } j = i + 1, \\ 1, & \text{falls } j = 0 \wedge i = \alpha, \\ 0, & \text{sonst.} \end{cases}$$

Dabei gilt speziell

p_{0j}	=	die Wahrscheinlichkeit, daß ein Kunde an Station j in das System eintritt,
p_{i0}	=	die Wahrscheinlichkeit, daß ein Kunde nach Fertigstellung in Station i das System verläßt.

Somit erhält man für die mittleren Ankunftsraten und Knotenauslastungen

$$\lambda_i = \lambda \qquad \text{und} \qquad \rho_i = \frac{\lambda}{\mu_i} \qquad \text{für } i = 1, \ldots, \alpha.$$

Wie allgemein in offenen Warteschlangennetzen lautet die Stabilitätsbedingung

$$\rho_i < 1, \qquad i \in \{1, \ldots, \alpha\}.$$

Zur Beschreibung von Leistungsgrößen in Tandemsystemen werden die nachfolgenden Zufallsvariablen verwendet, wobei wiederum $i \in \{1, \ldots, \alpha\}$, sowie $n \in \mathbb{N}^+$ gelte :

$N_i(t)$	=	die Anzahl der Kunden in der Station i zur Zeit t,
$N(t)$	=	die Anzahl der Kunden im gesamten System zur Zeit t.
$W_{n,i}$	=	die Verweilzeit des Kunden K_n von seiner Ankunft im System bis zum Start seiner Bedienung in b_i,
$\tilde{Q}_{n,i}$	=	die Wartezeit des Kunden K_n in der Warteschlange der Station i,
$Q_{n,i}$	=	die Wartezeit des Kunden K_n in den Warteschlangen der Stationen $1, \ldots, i$,
$\tilde{V}_{n,i}$	=	die Verweilzeit des Kunden K_n in der Station i,
$V_{n,i}$	=	die Verweilzeit des Kunden K_n von seiner Ankunft im System bis zur Fertigstellung seiner Bedienung in b_i.

Speziell gilt also :

$V_{n,\alpha}$	=	die Gesamtverweilzeit des Kunden K_n im System,
$Q_{n,\alpha}$	=	Gesamtwartezeit des Kunden K_n im System.

Untersucht man stationäre Zustandswahrscheinlichkeiten und Leistungsgrößen, so entfallen die kunden- und zeitbeschreibenden Parameter n und t.

3.2 Resultate für spezielle Tandemsysteme

In diesem Abschnitt werden Resultate für zwei spezielle, das heißt in der Verteilung ihrer Zwischenankunfts- oder Bedienzeiten näher spezifizierte Tandemsysteme vorgestellt. Den größten Teil dieser Darstellung nimmt aufgrund seiner besonderen Bedeutung das System $\mathrm{M}(\rightarrow/\mathrm{M}/1)^\alpha$ ein. In der Literatur ist zudem das System $\mathrm{G}(\rightarrow/\mathrm{D}/1)^\alpha$ und dessen Spezialfall $\mathrm{M}(\rightarrow/\mathrm{D}/1)^\alpha$ ausführlich beschrieben.

Allgemeinere Systeme werden in der Literatur, auch aufgrund der wesentlich schwierigeren Analysierbarkeit, deutlich seltener behandelt, es gibt daher auch wenige konkrete Ergebnisse für solche Systeme.

So wird in [BIN88] für das System M/G/1→/M/1 die erzeugende Funktion der stationären gemeinsamen Warteschlangenlängen bestimmt. In [Yak73] wurden die Abhängigkeiten stationärer Zustandswahrscheinlichkeiten im System G/M/1→/M/1 untersucht; die dortigen Ergebnisse jedoch zum Teil in [KK92] widerlegt. Eine Übersicht zahlreicher weiterer Arbeiten über Tandemsysteme, die hier nicht zitiert werden, findet man in [DK85].

3.2.1 Das System $\mathbf{G}(\rightarrow/\mathbf{D}/1)^\alpha$

Gegeben sei ein $\mathrm{G}(\rightarrow/\mathrm{D}/1)^\alpha$-Tandemsystem. Die konstanten Bedienzeiten an den Bedienern $b_1, \ldots, b_\alpha$ seien mit $\sigma_1, \ldots, \sigma_\alpha$ bezeichnet. Alle wesentlichen Ergebnisse für dieses System stammen aus der Arbeit von Friedman ([Fri65]) oder lassen sich direkt daraus ableiten.

Für $\mathrm{G}(\rightarrow/\mathrm{D}/1)^\alpha$-Systeme ist die Verteilung des Abgangsprozesses unabhängig von der Anordnung der Bedienstationen, das heißt, ändert man die Reihenfolge der Bedienstationen, so bleibt der Abgangsprozeß unverändert.

Weiterhin zeigte Friedman, daß die Warteschlange der Station $i+1$ genau dann immer leer ist, wenn die Bedienzeit an der Station $i+1$ kleiner ist als die Bedienzeit an der Station i. Man sagt dann, Station i *dominiert* Station $i+1$. Diese Eigenschaft der Dominanz ist transitiv, für beliebige $i, j \in \{1, \ldots, \alpha\}$ gilt also

$$L_{Q_j} = 0 \Longleftrightarrow \sigma_j < \sigma_i \wedge i < j,$$

wobei L_{Q_j} für $j \in \{1, \ldots, \alpha\}$ die Anzahl der Kunden in der Warteschlange der Station j bezeichnet.

Will man nur das gesamte System analysieren, ohne die einzelnen Stationen näher zu untersuchen, so läßt sich nach Friedman ein $\mathrm{G}(\rightarrow/\mathrm{D}/1)^\alpha$-System zu einem System $\mathrm{G/D/1}\rightarrow/\mathrm{D}/\infty$ reduzieren, indem man die Stationen zunächst so anordnet, daß $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_\alpha$ gilt und dann die Stationen $2, \ldots, \alpha$ durch eine Station mit unendlich vielen Bedienern - also ohne Wartezeiten - und Bedienzeit $\sigma_2 + \sigma_3 + \cdots + \sigma_\alpha$ ersetzt. Die stationäre Gesamtwartezeit Q_α in einem $\mathrm{G}(\rightarrow/\mathrm{D}/1)^\alpha$-System ist also gleich der stationären Wartezeit in einem elementaren G/D/1-System mit Bedienzeit $\sigma = \max_i\{\sigma_i\}$ und die stationäre Gesamtverweilzeit V_α ergibt sich aus der Summe dieser Wartezeit und der Bedienzeiten $\sigma_1, \ldots, \sigma_\alpha$.

Weiter läßt sich bei der Analyse von $\mathrm{G}(\rightarrow/\mathrm{D}/1)^\alpha$-Systemen ohne Beschränkung der Allgemeinheit annehmen, daß $\sigma_1 \leq \sigma_2 \leq, \cdots, \leq \sigma_\alpha$, da sich für $\sigma_i \geq \sigma_{i+1}$ die Stationen i und $i+1$ zu einer Bedienstation mit einem Bediener und Bedienzeit $\sigma_i + \sigma_{i+1}$ zusammenfassen lassen.

Speziell für $M(\rightarrow/D/1)^\alpha$ mit Poisson-Ankunftsprozeß der Rate λ erhält man dann

$$\begin{aligned}
E[\tilde{Q}_i] &= \frac{\lambda}{2}\left(\frac{\sigma_i^2}{1-\lambda\sigma_i} - \frac{\sigma_{i-1}^2}{1-\lambda\sigma_{i-1}}\right), \\
E[\tilde{V}_i] &= E[\tilde{Q}_i] + \sigma_i, \\
E[Q_i] &= \frac{\lambda\sigma_i^2}{2(1-\lambda\sigma_i)}, \\
E[V_i] &= E[Q_i] + \sum_{j=1}^{i} \sigma_j.
\end{aligned}$$

3.2.2 Das System $M(\rightarrow/M/1)^\alpha$

Gegeben sei ein $M(\rightarrow/M/1)^\alpha$-Tandemsystem. Der Poisson-Ankunftsprozeß habe die Rate λ, die Parameter der unabhängigen exponentialverteilten Bedienzeiten an den Bedienern $b_1, \ldots, b_\alpha$ seien $\mu_1, \ldots, \mu_\alpha$. Es stellt sich heraus, daß diese Art von Tandemsystemen zur Klasse der *Produktformnetze* gehört, ihre stationären Zustandswahrscheinlichkeiten $P(k_1, \ldots, k_\alpha)$ sind also durch das Produkt der Randwahrscheinlichkeiten gegeben :

$$P(k_1, \ldots, k_\alpha) = \prod_{i=1}^{\alpha} P_i(k_i)$$

mit den Randwahrscheinlichkeiten

$$P_i(k_i) = P[N_i = k_i] = \left(\frac{\lambda}{\mu_i}\right)^{k_i}\left(1 - \frac{\lambda}{\mu_i}\right) = \rho_i^{k_i}(1-\rho_i).$$

Diese Produktformlösung folgt einerseits leicht aus dem Jackson-Theorem ([Jac63]), wurde andererseits aber bereits in [Jac56] in Analogie zur Analyse elementarer M/M/1-Systeme hergeleitet.

Dabei werden zunächst Differentialgleichungen für die transienten Zustandswahrscheinlichkeiten aufgestellt und dann, da von konstanten stationären Zustandswahrscheinlichkeiten ausgegangen werden kann, die Ableitungen

$$\frac{dP_t(k_1, \ldots, k_\alpha)}{dt} = 0$$

gesetzt, woraus schließlich obige Produktformlösung resultiert.

Man erhält also für die Warte- und Verweilzeiten

$$\begin{aligned}
E[\tilde{Q}_i] &= \frac{\rho_i}{\mu_i - \lambda}, \\
E[\tilde{V}_i] &= \frac{1}{\mu_i - \lambda}.
\end{aligned}$$

Die Erwartungswerte der Gesamtwarte- und Verweilzeiten $E[Q]$ und $E[V]$ erhält man durch Summation der $E[\tilde{Q}_i]$ bzw. $E[\tilde{V}_i]$.

Im folgenden werden nun noch einige weitere Ergebnisse für $M(\rightarrow/M/1)^\alpha$-Systeme angegeben.

Reich ([Rei57]) hat gezeigt, daß in Tandemsystemen mit Bedienstrategie FCFS und exponential verteilten Zwischenankunfts- und Bedienzeiten, sowie mit Ausnahme der ersten und letzten Station nur genau einem Bediener pro Station, also insbesondere auch für M(→/M/1)$^\alpha$-Systeme, die stationären Verweilzeiten $\tilde{V}_i$ in den einzelnen Stationen unabhängig sind. Dabei erweist sich als wesentlich, daß in den Stationen $2, \ldots, \alpha - 1$ nur ein Bediener vorhanden ist.

Für die Anzahl L_i der Kunden in den einzelnen Stationen zu den Ankunftszeitpunkten zeigte Reich für M(→/M/1)^2 im stationären Fall

$$P[L_1 = 0, L_2 = 0] = P[L_1 = 0]P[L_2 = 0]\frac{\mu_1 + \mu_2}{\mu_1 + \mu_2 - \lambda},$$

woraus die Abhängigkeit von L_1, L_2 folgt.

Unter der Voraussetzung, daß die Wartezeiten $\tilde{Q}_i$ in den einzelnen Stationen eines $M(\rightarrow /M/1)^\alpha$-Systems unabhängig sind, erhielt Nelson ([Nel58]) für die stationäre Gesamtwartezeit Q_α die kummulative Verteilungsfunktion

$$P[Q_\alpha \leq t] = P\left[\sum_{i=1}^{\alpha} \tilde{Q}_i \leq t\right] = 1 - \sum_{i=1}^{\alpha} A_{i\alpha} e^{-\mu_i(1-\rho_i)}$$

mit

$$A_{i\alpha} = P[\tilde{Q}_i > 0] \prod_{j=1, j\neq i}^{\alpha} \frac{1 - P[\tilde{Q}_j > 0]\mu_i(1-\rho_i)}{\mu_i(1-\rho_i) - \mu_j(1-\rho_j)}.$$

und

$$P[\tilde{Q}_i > 0] = (\mu_i - \lambda)e^{-(\mu_i - \lambda)q} dq.$$

Diese Unabhängigkeitsannahme wurde jedoch in [Bur64] für $\alpha = 2$ widerlegt, sodaß Nelsons Ergebnisse lediglich als Näherung angesehen werden können. An dieser Stelle soll darauf hingewiesen werden, daß solche oder ähnliche Unabhängigkeitsannahmen in der Literatur häufig ohne Beweis gemacht werden und oftmals falsch sind.

Es erscheint zunächst bemerkenswert, daß die stationären Wartezeiten $\tilde{Q}_i$ abhängig sind, während die stationären Verweilzeiten $\tilde{V}_i$ unabhängig sind. Dies folgt jedoch leicht aus der Wartezeitanalyse bei elementaren M/M/1-Systemen mit der dort nötigen Fallunterscheidung, ob das System zum Ankunftszeitpunkt eines neuen Kunden leer ist oder nicht ($L_1 = 0$ bzw. $L_1 > 0$), sowie aus der Abhängigkeit von L_1, L_2 und der Identität $P[\tilde{Q}_1 = 0, \tilde{Q}_2 = 0] = P[L_1 = 0, L_2 = 0]$.

Weber ([Web79]) hat gezeigt, daß die Verteilung des Abgangsprozesses in M(→/M/1)$^\alpha$-Systemen - wie in Tandemsystemen mit deterministischen Bedienzeiten - unabhängig von der Anordnung der Bedienstationen ist.

Dies folgt auch aus dem Resultat, daß im stationären Fall der Abgangsprozeß aus einem elementaren M/M/1-System ein Poissonprozeß mit gleicher Rate wie der Ankunftsprozeß ist ([Kel79]). Schon in [Mor55] wurde gezeigt, daß dieser Abgangsprozeß ein Poisson-Prozeß ist, jedoch blieb die Rate unklar.

Ebenso folgt aus [Kel79], daß die Verweilzeiten in den einzelnen Stationen jeweils exponentialverteilt mit Parameter $\mu_i - \lambda$ für $i \in \{1, \ldots, \alpha\}$ sind.

In [KK92] wurde aus den Eigenschaften eines Geburts-Todes-Prozesses, der das Verhalten der zweiten Station beschreibt, für $M(\rightarrow/M/1)^2$ die erzeugende Funktion der gemeinsamen Warteschlangenlänge zu den Ankunftszeitpunkten, sowie die Laplace-Transformierte der gemeinsamen stationären Wartezeitverteilung hergeleitet. Daraus folgt schließlich für den Erwartungswert der gemeinsamen Wartezeit

$$E[\tilde{Q}_1, \tilde{Q}_2] = \frac{\rho_1}{(1-\rho_1)(1-\rho_2)\mu_2}.$$

3.3 Approximationen für das System $G(\rightarrow/G/1)^\alpha$

In diesem Abschnitt werden zwei Approximationsverfahren für Leistungsmaße in Tandemsystemen vorgestellt, die später auch zum Vergleich mit den Analyseergebnissen der in dieser Arbeit entwickelten Algorithmen herangezogen werden. Desweiteren wird ein erstes iteratives Verfahren unter Verwendung von 101-Dichten entwickelt, daß auf der sogenannten *Erneuerungsannahme* basiert und die Bedienstationen eines Tandemsystems als separate G/G/1-Systeme behandelt, deren Ankunftsprozeß dem Abgangsprozeß der jeweils vorhergehenden Station entspricht.

Die QNET- bzw. ΠNET-Methode wurde für Warteschlangennetze recht allgemeinen Typs entworfen; die Analyse von Tandemsystemen ist daher nur eine von zahlreichen möglichen Anwendungen, wohingegen die MSA-Methode zunächst speziell zur Analyse von elementaren Einbediener- und Tandemsystemen konzipiert wurde, jedoch in Zukunft auch auf allgemeinere Netze erweitert werden soll.

Ältere Verfahren haben im Vergleich zu QNET/ΠNET und MSA in nahezu allen Fällen schlechter, oder zumindest nicht besser, abgeschnitten, sodaß hier nicht näher darauf eingegangen werden soll. Dabei wird häufig, wie auch im weiteren Verlauf dieser Arbeit, daß Verfahren QNA (Queueing Network Analyzer) aus [Whi83] zum Vergleich herangezogen.

3.3.1 Die Erneuerungsannahme

In vielen Warteschlangennetzen, insbesondere in Tandemsystemen, ist der Abgangsprozeß einer Bedienstation (eines Knotens) gleichzeitig der Ankunftsprozeß einer anderen Bedienstation. Ist dieser Abgangsprozeß ein Erneuerungsprozeß, dann kann die nachfolgende Bedienstation separat behandelt werden. Sind die Zwischenabgangszeiten unkorreliert, so kann die Approximation durch einen Erneuerungsprozeß eine Näherung für den Ankunftsprozeß der folgenden Station darstellen.

Man nimmt also an, daß der Abgangsprozeß einer Bedienstation ein Erneuerungsprozeß ist. Diese Annahme nennt man *Erneuerungsannahme*. Sie wurde erstmals in [Bur72] zur Diskussion gestellt. Es ist jedoch klar, daß der Abgangsprozeß eines G/G/1-Systems im allgemeinen kein Erneuerungsprozeß ist, sodaß hier also wieder eine falsche Unabhängigkeitsannahme gemacht wird. Eine notwendige, jedoch nicht hinreichende Bedingung dafür, daß der Abgangsprozeß ein Erneuerungsprozeß ist, ist seine Unkorreliertheit.

Daley ([Dal68]) gab als notwendige Bedingung für die Unkorreliertheit des Abgangsprozesses eines M/G/1-Systems an, daß die Bedienzeit exponential verteilt ist. Somit ist das einzige

M/G/1-System mit unkorreliertem Abgangsprozeß das M/M/1-System. In diesem Fall ist dieser Abgangsprozeß ein Poisson-Prozeß mit gleicher Rate wie der Ankunftsprozeß ([Kel79], vgl. 3.2.2), also ein Erneuerungsprozeß.

Für G/M/1-Systeme gab Daley eine notwendige und hinreichende Bedingung für die Unkorreliertheit des Abgangsprozesses an. Sei $\rho < 1$ die Auslastung des Systems, μ die Bedienrate und T^* die Laplace-Stieltjes-Transformierte der unabhängig identisch verteilten Zwischenankunftszeiten. Dann ist der Abgangsprozeß unkorreliert, genau dann, wenn

$$\rho = T^*(\mu(1-\rho)).$$

Weiterhin zeigte Daley jedoch, daß der Abgangsprozeß in G/M/1-Systemen unkorreliert sein kann, ohne daß er ein Erneuerungsprozeß ist.

Viele Approximationsverfahren für Tandemsysteme basieren auf der Erneuerungsannahme. Dies wird häufig damit gerechtfertigt, daß die Korrelation der Abgangsprozesse insbesondere bei geringer Auslastung oftmals nicht sehr signifikant ist. Es gibt jedoch Beispiele (siehe [LMT93]), die belegen, daß in einigen Fällen die Korrelation das Verhalten des jeweiligen Systems stark beeinflußt.

So weisen zum Beispiel die ansonsten sehr gute Approximationen liefernden Verfahren QNA ([Whi83]) und MSA ([GH96], vgl. 3.3.3) bei stark korrelierten Abgangsprozessen erhebliche Fehler bei der Berechnung des Erwartungswertes der stationären Wartezeiten in den einzelnen Stationen auf.

Ein auf der Erneuerungsannahme basierender Algorithmus unter Verwendung von 101-Dichten wird im letzten Abschnitt dieses Kapitels formuliert und später zum Vergleich herangezogen.

3.3.2 Die QNET/ΠNET-Methode

Die QNET-Methode ([HN90]) ist eine analytische Methode zur Berechnung stationärer Leistungsgrößen in offenen Warteschlangennetzen, deren α Knoten aus je einem Bediener mit FIFO-Warteschlange unendlicher Kapazität bestehen. Ankünfte sind allgemein verteilt und können an verschiedenen Knoten stattfinden. Verschiedene Kundenklassen sind erlaubt, die Reihenfolge, in der ein Kunde verschiedene Knoten durchläuft, ist jedoch festgelegt. Abgänge dagegen können nach Durchlauf jedes Knotens stattfinden. Mehrere Besuche eines Kunden am gleichen Knoten sind möglich, die Bedienzeiten sind allgemein verteilt und dürfen von der Vorgeschichte des Kunden abhängen. Man sieht leicht, daß Tandemsysteme alle genannten Voraussetzungen erfüllen und somit mit der QNET-Methode analysiert werden können.

Die Idee bei QNET besteht darin, das gegebene Originalnetzwerk durch ein *approximierendes Brownsches System* zu ersetzen und dann dessen stationäre Verteilung zu berechnen.

Konkreter gesagt wird eine α-dimensionale Brownsche Bewegung $\{W(t)\}$ für die transienten Wartezeiten der Kunden mit $W(t) = (W_1(t), \ldots, W_\alpha(t))$ durch gewisse geforderte Eigenschaften definiert, wobei $W_i(t)$ die Wartezeit eines Kunden, der zur Zeit t in die Station i eintritt, in der Warteschlange dieser Station beschreibt.

Die Berechnung einer stationären Verteilung W führt dann auf eine partielle Differentialgleichung, die numerisch gelöst wird. In dieser numerischen Lösung besteht das Hauptproblem und der größte Nachteil von QNET, denn eine Implementierung gibt es bisher nur für Netzwerke mit höchstens zwei Knoten. Gleichzeitig ist offensichtlich, daß für $\alpha > 2$ die Komplexität der Berechnung rapide wächst.

In [HN90] wurde deshalb für Netze mit mehr als zwei Knoten das auf parametrischer Dekomposition basierende Verfahren ΠNET vorgeschlagen, welches eine Produktformapproximation verwendet, um die stationäre Verteilung W der Brownschen Bewegung $\{W(t)\}$ zu berechnen, wobei die (im allgemeinen nicht bestehende, vgl. 3.2.2) Unabhängigkeit der Wartezeiten W_i angenommen wird. Die Erwartungswerte der W_i werden dann unter Verwendung der ersten und zweiten Momente der Zwischenankunfts- und Bedienzeiten approximiert.

Im Vergleich zu aus Simulation gewonnenen Werten liefert ΠNET deutlich schlechtere Werte als QNET, jedoch ähnlich gute Werte wie etwa QNA, sodaß für diesen Analyseansatz über approximierende Brownsche Systeme die Entwicklung effizienterer Verfahren zur Berechnung stationärer Verteilungen von Brownschen Bewegungen ein wichtiger Aspekt für die Zukunft darstellt. Um eine höhere Genauigkeit zu erzielen, könnte hier zudem bei Anwendung von ΠNET die Berücksichtigung höherer Momente, ähnlich wie bei der im nächsten Abschnitt beschriebenen Methode, ein erfolgversprechender Ansatz sein.

3.3.3 MSA - Eine Approximation höherer Ordnung

In [GH96] ist ein Verfahren zur Analyse von Leistungsgrößen in G/G/1-Systemen, und daraus resultierend in Tandemsystemen, durch Approximation höherer Ordnung unter Verwendung von MacLaurin'schen-Reihen angegeben (MSA=MacLaurin Series Approximation). Dabei müssen alle Momente der Bedienzeit endlich sein. Außerdem muß die Dichte der Zwischenankunftszeiten analytisch (holomorph) auf $\mathbb{R}_+$ sein.

Die erste Bedingung wird von allen in der Warteschlangentheorie gängigen Verteilungen erfüllt, während die zweite Bedingung für die Dichten der Verteilung der Zwischenankunftszeiten alle Dichten ausschließt, deren Träger nicht $\mathbb{R}_+$ ist, wie etwa als wohl geläufigste solche Verteilung die stetige Gleichverteilung. Insbesondere die in der Warteschlangentheorie besonders bedeutsamen Phasentypverteilungen erfüllen beide Bedingungen.

Mit diesem Verfahren können insbesondere die Momente der Zwischenabgangszeiten approximiert werden und somit unter der Erneuerungsannahme zur Näherung für den Ankunftsprozeß einer weiteren Bedienstation benutzt werden. So läßt sich das Verfahren auf Tandemsysteme erweitern, indem deren Bedienstationen isoliert als elementare G/G/1-Systeme betrachtet werden, deren Ankunftsprozeß ein Erneuerungsprozeß ist.

Das Verfahren approximiert für $k \in \mathbb{N}$ die ersten k Momente der Warte-, Verweil- und Zwischenabgangszeiten eines G/G/1-Systems durch eine MacLaurinsche Reihe, deren Koeffizienten von den ersten k Momenten der Bedien- und Zwischenankunftszeiten, sowie den ersten k Ableitungen der Dichte der Zwischenankunftszeiten abhängen. Diese Koeffizienten können durch rekursive Gleichungen berechnet werden. Da es sich dabei aber um Koeffizienten einer unendlichen Reihe handelt, gibt es unendlich viele dieser Koeffizienten, sodaß eine Näherung für die MacLaurin-Reihen berechnet werden muß. Dies erfolgt durch *rationale Approximation*, das heißt, die MacLaurin-Reihen werden durch einen Quotienten endlicher Summen approximiert. Die Anzahl der Summanden gibt dabei eine Ordnung oder Güte der Approximation an, also je mehr Summanden verwendet werden, desto besser ist die Näherung.

Die Anwendung dieses Verfahrens auf Tandemsysteme basiert auf der Erneuerungsannahme. Die näherungsweise berechneten Momente der Zwischenabgangszeiten werden bei diesem Verfahren mit denen einer Erlang-Verteilung identifiziert, das heißt, als Verteilung des Ankunftsprozeß der zweiten Bedienstation wird eine Erlang-Verteilung angenommen, deren ersten k Momente mit den approximierten ersten k Momenten des Abgangsprozesses der ersten Bedienstation identisch sind. Dann werden Warte- und Verweilzeitmomente der zweiten Station durch MacLaurin-Reihen approximiert. Analog erfolgt die Berechnung der Momente für weitere Bedienstationen.

Ein weiterer Nachteil dieses Verfahrens neben der Erneuerungsannahme besteht in der Berechnung der Erlang-Verteilung, die als Ankunftsprozeß benutzt wird. Um aus den angenäherten Momenten des Abgangsprozesses eine Erlang-Verteilung mit identischen Momenten zu berechnen, müssen auch bei diesem Verfahren nichtlineare Gleichungen gelöst werden.

Im Vergleich zu aus Simulation gewonnenen Werten liefert das Verfahren unter Verwendung der ersten vier bis sechs Momente aller genannten Leistungsmaße in fast allen Fällen bessere Werte als QNA. Schlechtere Werte als QNA liefert MSA lediglich für Systeme, bei denen die Abgangsprozesse eine große Korrelation aufweisen, also Systeme, für die die Erneuerungsannahme nicht annähernd erfüllt ist. Die durch die Erneuerungsannahme auftretenden Fehler fallen also offensichtlich bei MSA mit wachsender Korrelation der Zwischenabgangszeiten stärker ins Gewicht als bei QNA, obwohl auch QNA auf der Erneuerungsannahme basiert.

In Zukunft wird das Hauptaugenmerk auf effizienten Verfahren zur Berechnung der beschriebenen Erlang-Verteilungen, sowie in der Erweiterung des Verfahrens auf allgemeinere Netzwerke liegen.

3.3.4 Eine 101-Approximation unter Erneuerungsannahme

Ein erster Ansatz zur Analyse von Tandemsystemen mit Hilfe von 101-Dichten besteht darin, ausgehend von der Erneuerungsannahme, die einzelnen Bedienstationen eines Tandemsystems separat zu analysieren, wobei jede Bedienstation als G/G/1-System interpretiert wird. Die grobe Struktur eines entsprechenden Algorithmus zur Berechnung der Warte- und Verweilzeiten in Tandemsystemen läßt sich dann folgendermaßen beschreiben :

1) $T :=$ stationäre Zwischenankunftszeit der Kunden im System; $i := 1$;

2) $S_i :=$ stationäre Bedienzeit an der i-ten Station;

3) Berechne aus S_i und T stationäre Warte-, Verweil- und Zwischenabgangszeiten $\tilde{Q}_i,\ \tilde{V}_i,\ D_i$;

4) Falls $i < \alpha$, setze $T := D_i$; $i := i + 1$; und gehe zu 2;

5) Berechne für $i = 1, \ldots, \alpha$ stationäre Verweilzeiten $V_i := \tilde{V}_1 + \tilde{V}_2 + \cdots + \tilde{V}_i$;

Die Berechnung der stationären Warte- und Zwischenabgangszeiten aus dem 3. Schritt muß nun konkretisiert werden. Für ein G/G/1-System seien T und S die stationäre Zwischenankunfts- bzw. Bedienzeit, W_n die Wartezeit des n-ten Kunden, sowie D_n die Zeit zwischen den Abgängen des n-ten und des $(n+1)$-ten Kunden.

Die stationären Wartezeiten lassen sich mit Hilfe der Lindley-Gleichung (vgl.[Kle75]) berechnen. Es gilt

$$\begin{aligned} W_1 &= 0, \\ W_{n+1} &= \max(W_n + S - T,\ 0),\ n > 0. \end{aligned}$$

Somit lassen sich iterativ die Wartezeiten der Kunden berechnen. Treten stationäre Verhältnisse ein, so konvergiert die Folge $\{W_n\}$ gegen die stationäre Wartezeit W. Man kann also sukzessive die Wartezeiten der einzelnen Kunden berechnen bis die Abweichungen bestimmter Momente zweier aufeinanderfolgender W_n eine vorgegebene Schranke ϵ unterschreiten. Die Verweilzeiten ergeben sich jeweils aus der Summe der Warte- und der Bedienzeit eines Kunden.

Die Berechnung der stationären Zwischenabgangszeiten folgt aus der Überlegung, daß sich die Abgangszeitpunkte der Kunden aus der Summe ihrer Ankunftszeitpunkte, ihrer Wartezeit und ihrer Bedienzeit ergeben. Sei t_n der Ankunftszeitpunkt des n-ten Kunden, sowie d_n der Abgangszeitpunkt des n-ten Kunden. Dann gilt

$$d_n = t_n + S + W_n.$$

Damit folgt

$$\begin{aligned} D_n &= d_{n+1} - d_n \\ &= T + W_{n+1} - W_n \\ &= T + \max(W_n + S - T,\ 0) - W_n \\ &= T + \max(V_n - T,\ 0) - V_n + S \\ &= S + \max(T - V_n,\ 0). \end{aligned}$$

Treten stationäre Verhätnisse ein, so konvergiert die Folge $\{D_n\}$ gegen die stationäre Zwischenabgangszeit D. Mit den entsprechenden Bezeichnungen für Tandemsysteme ergibt sich für den 3. Schritt des Algorithmus folgende Konkretisierung :

3a) $\tilde{Q}_{1,i} := 0;\ n := 1;$

3b) $n := n+1;\ \tilde{Q}_{n,i} := \max(\tilde{Q}_{n-1,i} + S_i - T,\ 0);\ \tilde{V}_{n,i} := \tilde{Q}_{n,i} + S_i;\ D_{n,i} := S_i + \max(T - \tilde{V}_{n,i},\ 0);$

3c) Berechne $\Delta :=$'Abweichung aufeinanderfolgender Iterierter'; Falls $\Delta > \epsilon$ gehe zu 3b;

3d) $\tilde{Q}_i := \tilde{Q}_{n,i};\ D_i := D_{n,i};$

Es werden also sukzessive die Wartezeiten aufeinanderfolgender Kunden bis zum Eintreten stationärer Verhältnisse berechnet, also bis sich die Zufallsvariablen, die diese Wartezeiten repräsentieren um nicht mehr als eine vorgegebene maximale Abweichung $\epsilon \in \mathbb{R}^+$ unterscheiden. Die Differenz von Zufallsvariablen ist wiederum eine Zufallsvariable, läßt sich also nicht direkt mit diesem ϵ vergleichen, sodaß die Abweichung Δ als eine Abweichung von Momenten der Zufallsvariablen $\tilde{Q}_{n-1,i}$ und $\tilde{Q}_{n,i}$, $\tilde{V}_{n-1,i}$ und $\tilde{V}_{n,i}$ bzw. $D_{n-1,i}$ und $D_{n,i}$ definiert werden muß. Stationäre Verhältnisse werden damit also so definiert, daß die Summe der Abweichungen korrespondierender Momente zweier aufeinanderfolgender Kunden weniger als ϵ beträgt, also

$$\Delta := \sum_{j=0}^{k} \left(\left| E[\tilde{Q}^j_{n,i}] - E[\tilde{Q}^j_{n-1,i}] \right| + \left| E[\tilde{V}^j_{n,i}] - E[\tilde{V}^j_{n-1,i}] \right| + \left| E[D^j_{n,i}] - E[D^j_{n-1,i}] \right| \right),\ k \in \mathbb{N}.$$

Der letzte Schritt bei der Entwicklung eines Algorithmus zur 101-Approximation der Wartezeiten in Tandemsystemen unter der Erneuerungsannahme ist nun die Darstellung aller vorkommenden Zufallsvariablen durch ihre 101-Dichten. Mit den in 2.2 beschriebenen Operatoren ergibt sich schließlich der folgende Algorithmus, der eine Erweiterung des Algorithmus aus [Str89] zur

Berechnung der Wartezeiten in G/G/1-Systemen darstellt. Dabei bezeichne f_X die Dichte einer Zufallsvariablen X und γ_X die 101-Dichte von X.

Algorithmus 101-Approximation unter Erneuerungsannahme
$\gamma_T := \Gamma_{101} f_T$;
$\gamma_{-T} := \Gamma_{101} f_{-T}$;
$i := 1$; {Nummer der Bedienstation }
REPEAT
  $\gamma_S := \Gamma_{101} f_{S_i}$; {101-Dichte der Bedienzeit an Station i }
  $\gamma_C := \gamma_S \otimes_{101} \gamma_{-T}$;
  $n := 2$; {Nummer des Kunden }
  $\gamma_{\hat{Q}_i} := \gamma_C$; {Pseudowartezeit des 2. Kunden in Station i }
  REPEAT
    $(\gamma_{\tilde{Q}_i},\ p_0) := MAX_{101} \gamma_{\hat{Q}_i}$; {Wartezeit des n-ten Kunden in Station i }
    $\gamma_{\tilde{V}_i} := \Gamma_{101}(p_0 \gamma_S + \gamma_{\tilde{Q}_i} \otimes_{101} \gamma_S)$; {101-Dichte der Verweilzeit in Station i }
    $\gamma_{-\tilde{V}_i} := -\gamma_{\tilde{V}_i}$;
    $\gamma_{-\tilde{Q}_i} := -\gamma_{\tilde{Q}_i}$;
    $\gamma_{H1} := \gamma_T \otimes_{101} \gamma_{-\tilde{V}_i}$;
    $(\gamma_{H2},\ p_0) := MAX_{101} \gamma_{H1}$;
    $\gamma_{D_i} := \Gamma_{101}(p_0 \gamma_S + \gamma_{H2} \otimes_{101} \gamma_S)$; {Zeit zwischen Abgängen der Kunden n und $n+1$ }
    $n := n + 1$;
    $\gamma_{\hat{Q}_i} := \Gamma_{101}(p_0 \gamma_C + \gamma_{\tilde{Q}_i} \otimes_{101} \gamma_C)$; {Pseudowartezeit des n-ten Kunden in Station i }
  UNTIL $\Delta < \epsilon$;
  $i := i + 1$;
UNTIL $i > \alpha$;
$\gamma_{V_1} := \gamma_{\tilde{V}_1}$;
FOR $i := 2$ TO α **DO** $\gamma_{V_i} := \gamma_{V_{i-1}} \otimes_{101} \gamma_{\tilde{V}_i}$.

Bei der Beschreibung des Algorithmus wird deutlich, daß auch das transiente Verhalten untersucht werden kann, denn der Zähler 'n' gibt gerade die Nummer des aktuellen Kunden an. Somit können für jeden einzelnen Kunden beliebige Momente der Warte-, Verweil- und Abgangszeiten bestimmt werden. Außerdem gibt der Wert von n nach Durchlauf der inneren REPEAT-Schleife für jede Station an, nach wievielen Kunden stationäre Verhältnisse eingetreten sind. Soll nur das stationäre Verhalten des Systems oder der einzelnen Stationen untersucht werden, so kann also auf diesen Zähler n verzichtet werden.

Die Laufzeitanalyse für diesen und alle weiteren Algorithmen, die in dieser Arbeit entwickelt werden, beschränkt sich auf die Analyse der Verweilzeitberechnung. Dies stellt keine Einschränkung dar, da sich jeweils alle anderen betrachteten Größen daraus ergeben.

Als Laufzeit oder Aufwand wird dabei jeweils die Anzahl von Operationen auf Zufallsvariablen betrachtet. Um diese Anzahl zu bestimmen, muß man zunächst untersuchen, wieviele Operationen pro Iterationsschritt ausgeführt werden.

In der inneren REPEAT-Schleife, die die Berechnung von $\tilde{V}_{n+1}$ aus $\tilde{V}_n$ realisiert, werden unabhängig von i und n konstant viele Operationen ausgeführt. Die äußere REPEAT-Schleife enthält zusätzlich nur konstant viele Initialisierungen und wird für jede Station einmal durchlaufen. Somit ergibt sich ein Aufwand von $O(\alpha)$ Operationen zur Berechnung aller Verweilzeiten $\tilde{V}_{n+1,i},\ i = 1, \ldots, \alpha$. Die Verweilzeiten $V_{n+1,i}$ werden hier zwar nicht berechnet, ergeben sich jedoch analog zur Berechnung der entsprechenden stationären Verweilzeiten aus $\alpha - 1$ Faltungen der $\tilde{V}_{n+1,i}$. Es ergibt sich also als Aufwand für einen Iterationsschritt $V_n \rightarrow V_{n+1}$ die Anzahl von $O(\alpha)$ Operationen. Damit folgt insgesamt für die Berechnung von V_n ein Aufwand von $O(\alpha n)$.

Die bei der Erprobung des Algorithmus berechneten Näherungen werden in Kapitel 5 im Rahmen des Vergleichs der in dieser Arbeit entwickelten Analysealgorithmen für Tandemsysteme mit den Verfahren QNET/ΠNET und MSA, sowie exakten und aus Simulation gewonnenen Werten vorgestellt und kommentiert.

4 Iterative (max,+)-Analyse von Tandemsystemen

In diesem Kapitel erfolgt zunächst die Darstellung von Tandemsystemen als (max,+)-lineare Systeme, das heißt, die Komponenten eines (max,+)-linearen Systems werden für die Anwendung auf Tandemsysteme konkretisiert. Daraus resultiert eine spezielle Rekursion für Tandemsysteme, die eine Iterationsvorschrift zur sukzessiven Berechnung der Wartezeiten aufeinanderfolgender Kunden darstellt.

Aus dieser Iterationsvorschrift wird zunächst direkt ein Algorithmus zur 101-Approximation von Leistungsgrößen unter der Annahme der Unabhängigkeit der auftretenden Zufallsvariablen hergeleitet.

Dann wird aus der speziellen Rekursion für Tandemsysteme eine Iterationvorschrift zur Berechnung von Leistungsmaßen hergeleitet, die lediglich die Operationen Maximumbildung und Addition auf erwiesenermaßen unabhängigen stetigen Zufallsvariablen benötigt. Somit können approximierende 101-Dichten und die darauf definierten Operatoren $MAXZV_{101}$ und $\otimes_{101}$, wobei die Faltung unabhängiger Zufallsvariablen gerade der Addition dieser Zufallsvariablen entspricht, ohne möglicherweise falsche Zusatzannahmen eingesetzt werden. Daher wird schließlich ein weiterer Algorithmus zur 101-Approximation von Leistungsgrößen in G(→/G/1)$^{\alpha}$-Systemen entwickelt.

Weiterhin wird die allgemeine Konvergenz der transienten Größen gegen stationäre Größen untersucht, um die Konvergenz der Iterationsalgorithmen zu gewährleisten.

4.1 Tandemsysteme als spezielle (max,+)-lineare Systeme

Die in 2.3 eingeführte Gleichung (16) soll also nun für Tandemsysteme konkretisiert werden. Dazu ist es zunächst erforderlich, die Komponenten eines (max,+)-linearen Systems, also die Folgen $\{T_n\}, \{A_n\}, \{B_n\}$ und $\{X_n\}$, für Tandemsysteme zu spezifizieren. Es bezeichne $\sigma_{n,i}$ für $i \in \{1, \ldots, \alpha\}, n \in \mathbb{N}^+$ die Bedienzeit des Kunden K_n im Bediener b_i.

In allen mit (max,+)-linearen Systemen modellierbaren Warteschlangennetzen stellt, wie in 2.3 schon erwähnt, die stationäre Folge reeller Zufallsvariablen $\{T_n\}$ die Folge der Ankunftszeitpunkte der einzelnen Kunden im System und die stationäre Folge reeller Zufallsvektoren $\{X_n\}$ die Folge der Bedienstartzeiten der einzelnen Bedienstationen dar.

Für Tandemsysteme heißt das also

T_n = Ankunftszeitpunkt des Kunden K_n an der ersten Station des Systems,

$X_{n,i}$ = Zeitpunkt des Starts der Bedienung des Kunden K_n im Bediener b_i.

Damit folgt für $W_{n,i} := X_{n,i} - T_n$

$W_{n,i}$ = Verweilzeit des Kunden K_n von der Ankunft im System bis zum Bedienstart in b_i.

Dies stimmt also mit der in 3.1 gegebenen Definition von $W_{n,i}$ überein; die Folge $\{W_n\}$ ist also die Folge der Verweilzeiten des Kunden K_n bis zu seinen Bedienstarts in den einzelnen Bedienern.

Die stationäre Folge reeller Zufallsvektoren $\{B_n\}$ beschreibt die Mindestverweilzeit der Kunden im System bis zum Start ihrer Bedienungen in den einzelnen Bedienern, das heißt $B_{n,i}$ ist die Mindestverweilzeit des Kunden K_n von seiner Ankunft im System bis zum Start seiner Bedienung im Bediener b_i. Da ein Kunde, wenn er nie warten muß, gerade so lange im System verweilt, wie er bedient wird, folgt für diese Mindestverweilzeit

$$B_{n,i} = \sum_{k=1}^{i-1} \sigma_{n,k} \qquad \text{für } n \in \mathbb{N}^+, \ i,j = 1, \ldots, \alpha.$$

Die Koeffizienten der stationären Folge reeller Zufallsmatrizen $\{A_n\}$ geben Mindestverweilzeiten in Teilsystemen, an deren erster Station eine nichtleere Warteschlange vorgefunden wird, an. Genauer gesagt, gibt $(A_n)_{i,j}$ für $j \leq i$ die Mindestverweilzeit des Kunden K_{n+1} von seiner Ankunft an Station j bis zum Start seiner Bedienung in Station i, also durch den Bediener b_i an, unter der Voraussetzung, daß er in Station j warten muß. Damit muß K_{n+1} in Station j zusätzlich zu seiner eigenen Bedienzeit mindestens die Bedienzeit seines Vorgängers verweilen; in den folgenden Stationen bis zum Bedienstart in b_i muß er dann jeweils für mindestens die Dauer seiner eigenen Bedienung verweilen. Für $j > i$ kann ein Kunde nach Ankunft in Station j die Station i nie mehr durchlaufen, die Mindestverweilzeit wird dann auf den Wert $-\infty$ gesetzt, also ergibt sich insgesamt

$$(A_n)_{i,j} = \begin{cases} -\infty & \text{falls } i < j \\ \\ \sum_{k=j}^{i-1} \sigma_{n+1,k} + \sigma_{n,j} & \text{falls } i \geq j \end{cases} \qquad \text{für } n \in \mathbb{N}, \ i,j = 1, \ldots, \alpha.$$

(16) liefert dann in Komponentenschreibweise

$$\begin{pmatrix} W_{n+1,1} \\ \vdots \\ W_{n+1,i} \\ \vdots \\ W_{n+1,\alpha} \end{pmatrix} = \begin{pmatrix} \max(W_{n,1} + \sigma_{n,1} - \tau_n, 0) \\ \vdots \\ \max(\max_{j=1}^{i}\{W_{n,j} + \sum_{k=j}^{i-1} \sigma_{n+1,k} + \sigma_{n,j}\} - \tau_n, \sum_{k=1}^{i-1} \sigma_{n+1,k}) \\ \vdots \\ \max(\max_{j=1}^{\alpha}\{W_{n,j} + \sum_{k=j}^{\alpha-1} \sigma_{n+1,k} + \sigma_{n,j}\} - \tau_n, \sum_{k=1}^{\alpha-1} \sigma_{n+1,k}) \end{pmatrix}, \tag{17}$$

wobei $W_{n+1,i}$ die Zeit zwischen Ankunft des Kunden K_{n+1} im System und dem Start seiner Bedienung im Bediener b_i und τ_n die Zwischenankunftszeit zwischen den Kunden K_n und K_{n+1} bezeichnet.

Es ist bemerkenswert, daß die erste Komponente dieser Gleichung der Lindley-Gleichung zur Berechnung der Wartezeiten in elementaren G/G/1-Systemen entspricht, die z.B in [Kle75] näher beschrieben ist und bei der Herleitung des Algorithmus unter Erneuerungsannahme angewandt wird. Dies war zu erwarten, da für $\alpha = 1$ ein $G(\rightarrow/G/1)^{\alpha}$-System zu einem elementaren G/G/1-System degeneriert. Es unterstreicht die Komplexität der Wartezeitanalyse, daß selbst für diesen Spezialfall $\alpha = 1$ bis heute keine allgemeinen geschlossenen Formeln für die stationäre Verteilung der Warte- und Verweilzeit oder deren Momente bekannt sind, sondern lediglich für Varianten mit Einschränkungen bezüglich der Zwischenankunfts- und Bedienzeitverteilungen. Für allgemeine G/G/1-Systeme existieren lediglich Näherungsformeln, sowie obere und untere Schranken und einige Approximationsverfahren.

Gleichung (17) ist also eine Verallgemeinerung der Lindley-Gleichung auf höhere Dimensionen.

Um nun die Verweilzeit eines Kunden K_n in den Stationen $b_1, \ldots, b_i$ zu berechnen, muß zu $W_{n,i}$ lediglich die Bedienzeit von K_n in b_i addiert werden, also

$$V_{n,i} = W_{n,i} + \sigma_{n,i}, \quad \text{bzw. in Vektorschreibweise}$$

$$V_n = W_n + \begin{pmatrix} \sigma_{n,1} \\ \vdots \\ \sigma_{n,\alpha} \end{pmatrix}.$$

Insbesondere ist $V_{n,\alpha}$ die Verweilzeit des n-ten Kunden im System. Als Rekursionsgleichung für die Verweilzeiten ergibt sich

$$V_{n+1} = \begin{pmatrix} V_{n+1,1} \\ \vdots \\ V_{n+1,i} \\ \vdots \\ V_{n+1,\alpha} \end{pmatrix} = \begin{pmatrix} \max(V_{n,1} - \tau_n, 0) + \sigma_{n+1,1} \\ \vdots \\ \max(\max_{j=1}^{i}\{V_{n,j} + \sum_{k=j}^{i-1} \sigma_{n+1,k}\} - \tau_n, \sum_{k=1}^{i-1} \sigma_{n+1,k}) + \sigma_{n+1,i} \\ \vdots \\ \max(\max_{j=1}^{\alpha}\{V_{n,j} + \sum_{k=j}^{\alpha-1} \sigma_{n+1,k}\} - \tau_n, \sum_{k=1}^{\alpha-1} \sigma_{n+1,k}) + \sigma_{n+1,\alpha} \end{pmatrix}. \tag{18}$$

Da man bei der Berechnung stationärer Verhältnisse von stationären Zwischenankunfts- und Bedienzeiten ausgeht, kann man (18) nun dahingehend vereinfachen, daß die kundenabhängigen Größen durch entsprechende stationäre Größen ersetzt werden.

Es bezeichne T die Zufallsvariable, die wie die stationären Zwischenankunftszeiten verteilt ist und S_i, $i = 1, \ldots, \alpha$ die Zufallsvariable, die wie die stationäre Bedienzeit am Bediener b_i verteilt ist.

Dann erhält man aus (18)

$$V_{n+1} = \begin{pmatrix} V_{n+1,1} \\ \vdots \\ V_{n+1,i} \\ \vdots \\ V_{n+1,\alpha} \end{pmatrix} = \begin{pmatrix} \max(V_{n,1} - T, 0) + S_1 \\ \vdots \\ \max(\max_{j=1}^{i}\{V_{n,j} + \sum_{k=j}^{i-1} S_k\} - T, \sum_{k=1}^{i-1} S_k) + S_i \\ \vdots \\ \max(\max_{j=1}^{\alpha}\{V_{n,j} + \sum_{k=1}^{\alpha-1} S_k\} - T, \sum_{k=1}^{\alpha-1} S_k) + S_\alpha \end{pmatrix}. \tag{19}$$

4.1.1 Eine 101-Approximation unter Unabhängigkeitsannahme

Aus Gleichung (19) soll nun ein Algorithmus entworfen werden, der diese Iteration direkt umsetzt und die entsprechenden Zufallsvariablen durch Zufallsvariablen mit 101-Dichten approximiert. Dazu werde an dieser Stelle angenommen, daß die auftretenden Zufallsvariablen, auf die die Operationen Faltung und Maximumbildung angewandt werden, jeweils paarweise unabhängig sind. Dies ist eine notwendige Voraussetzung für die Anwendbarkeit der Operatoren $\otimes_{101}$ und $MAXZV_{101}$.

Im Vergleich zum Algorithmus unter Erneuerungsannahme aus 3.3.4 wird hier also nicht angenommen, daß die Ankunftsprozesse an den einzelnen Stationen Erneuerungsprozesse sind, sondern lediglich, daß die transienten Verweilzeiten paarweise unabhängig sind. Auch diese Annahme ist unbewiesen und möglicherweise falsch; es ist jedoch unklar, ob eventuelle Abhängigkeiten signifikanten Einfluß auf die Resultate haben.

Auch die mit dem Algorithmus zu dieser 101-Approximation unter Unabhängigkeitsannahme berechneten Näherungen werden in Kapitel 5 zum Vergleich herangezogen. Es soll nun wieder der Algorithmus schrittweise, zunächst informell, ohne Berücksichtigung der 101-Approximation, dann unter Einbeziehung von 101-Dichten und den darauf definierten Operatoren, beschrieben werden.

Um die mehrfach auftretenden Summen (Faltungen) von Bedienzeiten nicht für jeden Schritt neu berechnen zu müssen, ist es sinnvoll, diese in einem Feld zu speichern. Dabei gelte

$$S[i,j] := \sum_{k=i}^{j} S_k \qquad i,j \in \{1,\ldots,\alpha\},\ i \le j,$$

also speziell

$$S[i,i] = S_i, \qquad i \in \{1,\ldots,\alpha\}.$$

Die durch (19) beschriebene Iteration impliziert dann direkt den folgenden Algorithmus :

1) $T :=$ stationäre Zwischenankunftszeit der Kunden im System;

2) $S[i,i] :=$ stationäre Bedienzeit an der i-ten Station;

3) Berechne die Bedienzeitsummen $S[i,j] := \sum_{k=i}^{j} S_k = \sum_{k=1}^{j} S[k,k]$;

4) $n := 1$; Initialisiere Verweilzeiten des 1. Kunden $V_{1,i} := S[1,i]$;

5) $n := n+1$; Berechne gemäß Gleichung (19) V_n aus V_{n-1};

6) Berechne Δ:='Abweichung von V_n und V_{n-1}'; falls $\Delta > \epsilon$ gehe zu 5;

Analog zum Algorithmus unter Erneuerungsannahme wird dabei die Abweichung Δ als Summe der Beträge der Momentdifferenzen der Zufallsvariablen V_n und V_{n-1} definiert, also

$$\Delta := \sum_{j=0}^{k} \left| E[V_{n,i}^{j}] - E[V_{n-1,i}^{j}] \right|, \qquad k \in \mathbb{N}.$$

Die Berechnung von V_n aus V_{n-1} in Schritt 5 läßt sich folgendermaßen konkretisieren:

5a) $V_{n,1} := \max(V_{n-1,1} - T,\ 0) + S[1,1];\ i := 2;$

5b) $H1 := V_{n-1,1} + S[1,\ i-1];\ j := 2;$

5c) $H2 := V_{n-1,j} + S[j,\ i-1];$

5d) $H3 := \max(H1, H2);\ H1 := H2;$ falls $j \leq i$ $j := j+1$ und gehe zu 5c;

5e) $V_{n,i} := \max(H3 + (-T), S[1,\ i-1]) + S[i,i];$ falls $i \leq \alpha$ $i := i+1$ und gehe zu 5b;

Es kann nun also ein entsprechender Algorithmus mit Hilfe von 101-Dichten und deren Operatoren formuliert werden. Dabei bezeichne wieder f_X die Dichte einer Zufallsvariablen X und γ_X die 101-Dichte von X.

Algorithmus 101-Approximation unter Unabhängigkeitsannahme

$\gamma_{-T} := \Gamma_{101} f_{-T};$
FOR $i := 1$ TO α **DO** $\gamma_{S[i,i]} := \Gamma_{101} f_{S_i}$ **OD;**
FOR $i := 2$ TO $\alpha - 1$ **DO**
  FOR $j := i+1$ TO $\alpha - 1$ **DO**
    $\gamma_{S[i,j]} := \gamma_{S[i,j-1]} \otimes_{101} \gamma_{S[j,j]};$
  OD;
OD;
$n := 1;$
FOR $i := 1$ TO α DO $\gamma_{V_{neu,i}} := \gamma_{S[1,i]}$ **OD**; $\{V_{1,i} := \sum_{k=1}^{i} S_k\}$
REPEAT
  $n := n+1;$
  FOR $i := 1$ TO α DO $\gamma_{V_{alt,i}} := \gamma_{V_{neu,i}};$ **OD**;
  $\gamma_H := \gamma_{V_{alt,1}} \otimes_{101} \gamma_{-T};$
  $\gamma_H := MAX_{101}(p_0, \gamma_H);$
  $\gamma_{V_{neu,1}} := \gamma_H \otimes_{101} \gamma_{S[1,1]};$ $\{V_{n,1} := \max(V_{n-1,1} - T,\ 0) + S_1\}$
  FOR $i := 2$ TO α **DO**
    $\gamma_{H1} := \gamma_{V_{alt,1}} \otimes_{101} \gamma_{S[1,i-1]};$ $\{H1 := V_{n-1,1} + \sum_{k=1}^{i-1} S_k\}$
    FOR $j := 2$ TO $i-1$ **DO**
      $\gamma_{H2} := \gamma_{V_{alt,j}} \otimes_{101} \gamma_{S[j,i-1]};$ $\{H2 := V_{n-1,j} + \sum_{k=j}^{i-1} S_k\}$
      $\gamma_{H3} := MAXZV_{101}(\gamma_{H1}, \gamma_{H2});$
      $\gamma_{H1} := \gamma_{H3};$
    OD;
    $\gamma_{H1} := \gamma_{H3} \otimes_{101} \gamma_{-T};$
    $\gamma_{H2} := MAXZV_{101}(\gamma_{H1}, \gamma_{S[1,i-1]});$
    $\gamma_{V_{neu,i}} := \gamma_{H2} \otimes_{101} \gamma_{S[i,i]};$
  OD;
UNTIL $\Delta < \epsilon$.

Es ist klar, daß auch mit diesem Algorithmus die transienten Verweilzeiten untersucht werden können. Will man nur die stationären Verweilzeiten berechnen, so kann der Kundenzähler n entfallen. Es wird nun auch hier die durch die Anzahl von Operationen auf Zufallsvariablen bestimmte Laufzeit untersucht.

Der Initialisierungsteil vor der REPEAT-Schleife wird dominiert von der Berechnung der Bedienzeitsummen $S[i,j]$, woraus sich ein Aufwand von $O(\alpha^2)$ Operationen ergibt. Innerhalb der REPEAT-Schleife, wird der wesentliche Teil des Aufwands von den geschachtelten FOR-Schleifen bestimmt. Dabei wird die äußere dieser FOR-Schleifen $(\alpha-1)$-fach und die innere FOR-Schleife für jeden Wert $i = 2,\ldots,\alpha$ $(i-2)$-fach durchlaufen, woraus eine Anzahl von insgesamt $(\alpha-1)(1+2+\cdots+\alpha-2) = O(\alpha^3)$ Schleifendurchläufen folgt. Da jeweils konstant viele Operationen innerhalb aller Schleifen ausgeführt werden, ergibt sich für einen Durchlauf der REPEAT-Schleife, also der Berechnung von V_{n+1} aus V_n die Anzahl von $O(\alpha^3)$ Operationen und damit für die Berechnung von V_n der Aufwand $O(\alpha^3 n)$.

4.1.2 Eine 101-Approximation ohne Unabhängigkeitsannahmen

Bei den bisher hergeleiteten Algorithmen wurden jeweils Fehler aufgrund falscher Unabhängigkeitsannahmen in Kauf genommen. Bei der 101-Approximation unter Erneuerungsannahme geht man davon aus, daß die Abgangsprozesse aus den einzelnen Bedienstationen Erneuerungsprozesse sind und bei der 101-Approximation unter Unabhängigkeitsannahme wird die Unabhängigkeit aller transienter Verweilzeiten $V_{n,i},\ i \in \{1,\ldots,\alpha\},\ n \in \mathbb{N}$ vorausgesetzt.

Will man bei der Analyse von Tandemsystemen mit Hilfe von 101-Dichten solche Fehler ausschließen, muß man eine Darstellung der Verweilzeiten finden, so, daß die Operationen Faltung und Maximum ausschließlich auf beweisbar unabhängige Zufallsvariablen angewandt werden. Genauer gesagt müssen also die Verweilzeiten nur aus den Bedienzeiten $S_i,\ i \in \{1,\ldots,\alpha\}$ und den Zwischenabgangszeiten T berechenbar sein, da nur diese Größen nach Voraussetzung unabhängig sind.

Satz 2 (Berechnung transienter Verweilzeiten aus unabhängigen Größen) *Mit den bisher eingeführten Bezeichnungen, sowie $i \in \{1,\ldots,\alpha\},\ x_j \in \mathbb{N},\ j = 1,\ldots,i,\ n \in \mathbb{N}$ gilt*

$$V_{n,i} = \max_{1\le x_1\le\cdots\le x_i\le n}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach})$$

Beweis mittels vollständiger Induktion über n :

<u>$n = 1$</u>

$$\begin{aligned} V_{1,i} &= \sum_{j=1}^{i} S_j = \max(S_1+S_2+\cdots+S_i) \\ &= \max_{1\le x_1\le\cdots\le x_i\le 1}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach}) \end{aligned}$$

<u>$n \to n+1$</u>

$$V_{n+1,i} \overset{(19)}{=} \max(\max_{j=1}^{i}\{V_{n,j}+\sum_{k=j}^{i-1} S_k\} - T,\ \sum_{k=1}^{i-1} S_k) + S_i,$$

$$
\begin{aligned}
V_{n+1,i} &= \max(\max_{j=1}^{i}\{\max_{1\leq x_1\leq\cdots\leq x_i\leq n}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots\\
&\quad+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach})+\sum_{k=j}^{i-1}S_k\}-T,\ \sum_{k=1}^{i-1}S_k)+S_i\\
&= \max(\max_{j=1}^{i}\{\max_{1\leq x_1\leq\cdots\leq x_i\leq n}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots\\
&\quad+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach})+\sum_{k=j}^{i}S_k-T\},\ \sum_{k=1}^{i}S_k)\\
&\overset{1)}{=} \max(\max_{j=1}^{i}\{\max_{2\leq x_1\leq\cdots\leq x_i\leq n+1}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots\\
&\quad+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach})\},\ \sum_{k=1}^{i}S_k)\\
&\overset{2)}{=} \max(\max_{2\leq x_1\leq\cdots\leq x_i\leq n+1}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots\\
&\quad+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach}),\ \sum_{k=1}^{i}S_k)\\
&\overset{3)}{=} \max_{1\leq x_1\leq\cdots\leq x_i\leq n+1}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots\\
&\quad+\underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach}-\underbrace{T-\cdots-T}_{(x_i-1)-fach}).
\end{aligned}
$$

1) Nach Auswertung des inneren Maximums wird in jedem Fall die Zwischenankunftszeit T subtrahiert und mindestens die Bedienzeit S_i addiert. Daraus folgt die Inkrementierung der unteren Schranke für x_1, sowie der oberen Schranke für x_i.

2) Das betreffende Maximum hängt nicht mehr von j ab, da j innerhalb der Maximumbildung nicht mehr auftritt.

3) Die Bedienzeitsumme entspricht der gewünschten Form des Maximums für die spezielle Belegung $x_1 = x_2 = \cdots = x_i = 1$.

Die Gleichung aus dem obigen Satz ermöglicht also die direkte Berechnung transienter Verweilzeiten aus den unabhängigen stationären Bedien- und Zwischenankunftszeiten. Daraus ergibt sich schließlich die folgende Iteration zur sukzessiven Berechnung aufeinanderfolgender Verweilzeiten, wobei Faltungen und Maxima ausschließlich für unabhängige Zufallsvariablen berechnet werden.

$$
\begin{aligned}
V_{n+1,i} &= \max(V_{n,i}, \max_{1\le x_1\le\cdots\le x_i=n+1}(\underbrace{S_i+\cdots+S_i}_{x_1-fach}+\underbrace{S_{i-1}+\cdots+S_{i-1}}_{(x_2-x_1+1)-fach}+\cdots \\
&\quad + \underbrace{S_1+\cdots+S_1}_{(x_i-x_{i-1}+1)-fach} - \underbrace{T-\cdots-T}_{(x_i-1)-fach})) \qquad (20)\\
&= \max(V_{n,i}, \max_{1\le x_1\le\cdots\le x_i=n+1}(\sum_{k=1}^{x_1} S_i + \sum_{k=1}^{x_2-x_1+1} S_{i-1}+\cdots+\sum_{k=1}^{x_i-x_{i-1}+1} S_1 - \sum_{k=1}^{n} T)).
\end{aligned}
$$

Schrittweise soll nun wieder die Umsetzung der Iteration in einen Algorithmus angegeben und dessen Aufwand untersucht werden. Dabei fällt zunächst auf, daß ein erheblicher Teil der Berechnung von V_{n+1} aus V_n auf die Faltung identischer Zufallsvariablen entfällt. Es wäre ineffizient und würde zudem wohl die vorhandenen Speicherkapazitäten überschreiten, alle diese Summen von Zufallsvariablen zu speichern. Um jedoch nicht für jeden Schritt alle mehrfachen Faltungen identischer Zufallsvariablen neu berechnen zu müssen, werden für $k \in \mathbb{N}$ die 2^k-fachen Faltungen wie folgt vorab berechnet und gespeichert :

$$
\begin{aligned}
S_i[0] &:= S_i,\\
S_i[k] &:= S_i[k-1] + S_i[k-1],
\end{aligned}
$$

das heißt also

$$
S_i[k] = \underbrace{S_i+\cdots+S_i}_{2^k-fach}, \qquad i \in \{1,\ldots,\alpha\}.
$$

Als Aufwand zur Berechnung von $S_1[0],\ldots,S_1[k],\ldots,S_\alpha[0],\ldots,S_\alpha[k]$ ergibt sich dann der Aufwand von αk Faltungen, die vorab berechnet werden.

Um dann für beliebige $m \in \mathbb{N}$ die m-fachen Faltungen der jeweiligen Bedienzeiten zu berechnen, kann man jeweils folgendes Programmstück (Makro) ausführen :

Makro Berechne $S_i^{(m)}$;
$Hilf := m;\ S_i^{(m)} := 0;$
REPEAT
  $k := \lfloor \log_2(Hilf) \rfloor;\ \{2^k \le Hilf < 2^{k+1}\}$
  $S_i^{(m)} := S_i^{(m)} + S_i[k];$
  $Hilf := Hilf - k;$
UNTIL $Hilf = 0;$

Somit benötigt man für die m-fache Faltung bei vorab wie oben beschrieben gespeicherten $S_i[k]$ den Aufwand von $O(\log_2(m))$ Faltungen.

Beispiel 8 (2001-fache Faltung von X) *Dieses Beispiel soll für die 2001-fache Faltung einer Zufallsvariablen X mit $X[0] = X$, $X[k] = X[k-1] + X[k-1]$ den Ablauf des beschriebenen Makros verdeutlichen.*
1) Hilf=2001, k=10 $(2^k = 1024)$, $X^{(m)} = X[10]$
2) Hilf=977, k=9 $(2^k = 512)$, $X^{(m)} = X[10] + X[9]$
3) Hilf=465, k=8 $(2^k = 256)$, $X^{(m)} = X[10] + X[9] + X[8]$
4) Hilf=209, k=7 $(2^k = 128)$, $X^{(m)} = X[10] + X[9] + X[8] + X[7]$
5) Hilf=81, k=6 $(2^k = 64)$, $X^{(m)} = X[10] + X[9] + X[8] + X[7] + X[6]$
6) Hilf=17, k=4 $(2^k = 16)$, $X^{(m)} = X[10] + X[9] + X[8] + X[7] + X[6] + X[4]$
7) Hilf=1, k=0 $(2^k = 1)$, $X^{(m)} = X[10] + X[9] + X[8] + X[7] + X[6] + X[4] + X[0]$
8) Hilf=0.

Um nun den Aufwand eines Algorithmus ohne Unabhängigkeitsannahmen zu analysieren, genügt es, die obige Iteration näher zu untersuchen. Für jedes zur Maximumbildung herangezogene Element werden bei einem Iterationsschritt $V_{n,i} \rightarrow V_{n+1,i}$ jeweils $i+1$ verschiedene Zufallsvariablen höchstens $(n+1)$-fach gefaltet. Mit dem oben erläuterten Aufwand für mehrfache Faltungen ergibt sich also pro Element der Maximumbildung ein Aufwand von $(i+1)O(\log_2 n)$ Faltungen.

Die Anzahl von Elementen, die zur Maximumbildung herangezogen werden, entspricht der Anzahl der möglichen Belegungen von $1 \leq x_1 \leq \cdots \leq x_i = n+1$ bzw. der Anzahl möglicher Belegungen von $1 \leq x_1 \leq \cdots \leq x_{i-1} \leq n+1$. Dafür erhält man

$$\begin{aligned}\begin{pmatrix} n+i-1 \\ i-1 \end{pmatrix} &= \frac{(n+i-1)!}{(i-1)!n!} = \frac{(n+i-1)(n+i-2)\cdots(n+1)}{(i-1)!} \\ &= O(n^{i-1}).\end{aligned}$$

Damit folgt für den Aufwand für $V_{n,i} \rightarrow V_{n+1,i}$ $O(n^{i-1}\log n)$. Für den Gesamtaufwand für $V_n \rightarrow V_{n+1}$ erhält man die Anzahl von $O(n^0 \log n) + O(n \log n)O(n^2 \log n) + \cdots + O(n^{\alpha-1}\log n) = O(n^{\alpha-1}\log n)$ Operationen auf Zufallsvariablen. Zur Berechnung von V_n auf diese Art benötigt man also insgesamt $O(n^\alpha \log n)$ Operationen. Das Vermeiden fehlerhafter Annahmen wird also mit erheblichem Mehraufwand bezahlt.

Nun kann man zunächst wieder informell einen entsprechenden Algorithmus formulieren.

1) $T :=$ stationäre Zwischenankunftszeit der Kunden im System;

2) $S_i :=$ stationäre Bedienzeit an der i-ten Station $(i = 1, \ldots, \alpha)$;

3) Berechne für $K \in \mathbb{N}$ die Bedienzeitsummen $S_i[1], \ldots, S_i[K]$ $(i = 1, \ldots, \alpha)$;

4) $n := 1$; $T^{(1)} := T$; Initialisiere Verweilzeiten des 1. Kunden $V_{1,i} := S_1 + \cdots + S_i$ $(i = 1, \ldots, \alpha)$;

5) $n := n + 1$; $i := 1$;

6) Berechne $T^{(n)} := T^{(n-1)} + T$;

7) Berechne gemäß Gleichung (20) $V_{n,i}$ aus $V_{n-1,i}$;

8) Falls $i < \alpha$ setze $i := i + 1$ und gehe zu 7;

9) Falls $\Delta > \epsilon$ gehe zu 5.

Dabei sei Δ wie in 4.1.1 definiert und $k \in \mathbb{N}$ so gewählt, daß in jedem Fall weniger als 2^{K+1} Faltungen identischer Zufallsvariabler berechnet werden müssen. Ein solches K läßt sich vorab zwar nicht genau bestimmen, da die Anzahl der Faltungen identischer Zufallsvariablen von der Nummer n des jeweiligen Kunden abhängt, man kann jedoch problemlos ein hinreichend großes K finden. Die Konkretisierung von Schritt 7 läßt sich wie folgt formulieren.

7a) Setze $x_1 := x_2 := \cdots := x_{i-1} := 1;\ x_i := n;$

7b) $H1 := V_{n-1,i};$

7c) $H2 := S_i^{(x_1)} + S_{i-1}^{(x_2-x_1+1)} + \cdots + S_1^{(x_i-x_{i-1}+1)} - T^{(n)};$

7d) $H3 := \max(H1, H2);\ H1 := H3;$

7e) Finde nächste Belegung der x_j;

7f) Falls Belegung zulässig, gehe zu 7c;

7g) $V_{n,i} := H1;$

Dabei sollen die Belegungen der x_j in lexikographischer Ordnung abgearbeitet werden. Die neue Belegung wird dabei von dem folgenden Makro realisiert.

Makro Finde nächste Belegung;
$j := i - 1;$
WHILE $x_j = n$ **DO** $j := j - 1;$ **OD**;
IF $j = 0$ **THEN** $j := j + 1;$
$x_j := x_j + 1;$
FOR $k := j + 1$ **TO** $i - 1$ **DO** $x_k := x_j;$ **OD**;

Die letzte zulässige Belegung ist also $x_1 = \cdots = x_i = n$. Danach erhalten dann abschließend alle $x_j,\ j = 1, \ldots, i$ den Wert $n + 1$, also einen unzulässigen Wert. Dies wird dann als Zeichen für die Abarbeitung aller zulässigen Belegungen benutzt.

Es verbleibt die Formulierung des Algorithmus unter Verwendung von 101-Dichten.

Algorithmus 101-Approximation ohne Unabhängigkeitsannahmen
$\gamma_T := \Gamma_{101} f_T;$
FOR $i := 1$ **TO** α **DO**
  $\gamma_{S_i[0]} := \Gamma_{101} f_{S_i};$
  FOR $j := 1$ **TO** K **DO** $\gamma_{S_i[j]} := \gamma_{S_i[j-1]} \otimes_{101} \gamma_{S_i[j-1]};$ **OD**;
OD;
$n := 1;$
$\gamma_{V_{1,1}} := \gamma_{S_1[0]};$
FOR $i := 2$ **TO** α **DO** $\gamma_{V_{1,i}} := \gamma_{V_{1,i-1}} \otimes_{101} \gamma_{S_i[0]};$ **OD**;
$\gamma_{T^{(1)}} := \gamma_T;$
REPEAT
  $n := n + 1;$
  $\gamma_{T^{(n)}} := \gamma_{T^{(n-1)}} \otimes_{101} \gamma_T;\ \gamma_{-T^{(n)}} := -\gamma_{T^{(n)}};$
  FOR $i := 1$ **TO** α **DO**
   $x_i := n;$ **FOR** $j := 1$ **TO** $i - 1$ **DO** $x_j := 1;$ **OD**;
   $\gamma_{H1} := \gamma_{V_{n-1,i}};$
   REPEAT
    Berechne $\gamma_{S_i^{(x_i)}},\ \gamma_{S_{i-1}^{(x_2-x_1+1)}}, \ldots, \gamma_{S_1^{(x_i-x_{i-1}+1)}};$
    $\gamma_{H2} := (\cdots(\gamma_{S_i^{(x_i)}} \otimes_{101} \gamma_{S_{i-1}^{(x_2-x_1+1)}}) \otimes_{101} \cdots) \otimes_{101} \gamma_{S_1^{(x_i-x_{i-1}+1)}}) \otimes_{101} \gamma_{-T^{(n)}};$

$\gamma_{H3} := MAXZV_{101}(\gamma_{H1}, \gamma_{H2});$
$\gamma_{H1} := \gamma_{H3};$
Finde nächste Belegung;
UNTIL $x_1 = n + 1;$
$\gamma_{V_{n,i}} := \gamma_{H1};$
OD;
UNTIL $\Delta < \epsilon$.

Die Berechnung m-facher Faltungen der 101-Dichten von Bedienzeiten für beliebige $m \in \mathbb{N}$ erfolgt entsprechend dem Makro 'Berechne $S_i^{(m)}$', wobei jedoch die Zeile '$S_i^{(m)} := S_i^{(m)} + S_i[k]$' durch '$\gamma_{S_i^{(m)}} := \gamma_{S_i^{(m)}} \otimes_{101} \gamma_{S_i[k]}$' ersetzt wird.

4.2 Konvergenz transienter Größen gegen stationäre Größen

Die drei in dieser Arbeit entwickelten Algorithmen basieren alle auf dem Prinzip, stationäre Leistungsgrößen aus den transienten Leistungsgrößen, wie Warte- oder Verweilzeiten aufeinanderfolgender Kunden zu iterieren. Intuitiv scheint klar, daß jeweils die Folge der transienten Größen gegen stationäre Größen konvergiert, wenn die Stabilitätsbedingung $\rho_i < 1$ für alle $i \in \{1, \ldots, \alpha\}$ erfüllt ist. Dies läßt sich auch formal beweisen. Hierzu werden zunächst zwei wahrscheinlichkeitstheoretische Konvergenzbegriffe eingeführt.

Definition 8 (Schwache Konvergenz) *Sei $\{P_n\}$ eine Folge von Wahrscheinlichkeitsverteilungen auf $\mathbb{R}^\alpha$, $\alpha \in \mathbb{N}^+$. $\{P_n\}$ heißt schwach konvergent gegen eine Wahrscheinlichkeitsverteilung P auf $\mathbb{R}^\alpha$, falls für alle beschränkten stetigen Funktionen $f : \mathbb{R}^\alpha \to \mathbb{R}^\alpha$ gilt :*

$$\lim_{n\to\infty} \int f dP_n = \int f dP.$$

Man schreibt dann auch $P_n \xrightarrow{w} P$ (mit w für weak).

Definition 9 (Konvergenz in Verteilung) *Sei $\{X_n\}$ eine Folge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega_n, \mathcal{A}_n, P_n)$ mit Werten in $\mathbb{R}^\alpha$, $\alpha \in \mathbb{N}^+$. $\{X_n\}$ konvergiert in Verteilung gegen eine Zufallsvariable X auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$, falls die Folge $\{P_n\}$ schwach gegen P konvergiert, das heißt, $\{X_n\}$ konvergiert in Verteilung gegen X, falls $P_n \xrightarrow{w} P$.*

Beispiel 9 (Approximation der stetigen Gleichverteilung) *Sei $P := U(0,1)$ die stetige Gleichverteilung auf dem Intervall $[0,1]$ und $P_n := U(\{\frac{k}{n} : 1 \leq k \leq n\})$. Dann gilt für alle beschränkten stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$*

$$\int f dP_n = \sum_{k=1}^{n} \frac{1}{n} f(\frac{k}{n}) \quad \xrightarrow{n\to\infty} \quad \int_0^1 f(x)dx = \int f dP,$$

also $P_n \xrightarrow{w} P$.

Als weiteres Beispiel für die Anwendung der eingeführten Begriffe wird nun, ohne Beweis, der bekannte Satz von de Moivre-Laplace angeführt.

Beispiel 10 (Satz von de Moivre-Laplace) *Sei $\{X_n\}$ eine Folge von unabhängig identisch verteilten, quadratisch integrierbaren, reellen Zufallsvariablen mit Varianz $\sigma^2 \in \mathbb{R}^+$ und sei*

$$S_n := \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}$$

Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega_n, \mathcal{A}_n, P_n)$, $n \in \mathbb{N}$. Dann gilt:

(i) Die Folge von Verteilungen $\{P_n\}$ konvergiert schwach gegen die standardisierte Normalverteilung $N(0,1)$.

(ii) Die Folge von Zufallsvariablen $\{S_n\}$ konvergiert in Verteilung gegen eine $N(0,1)$-verteilte Zufallsvariable S.

Mit den eingeführten Begriffen kann man nun einen Konvergenzsatz formulieren. Gegeben sei ein $G(\rightarrow /G/1)^\alpha$-System. $\{V_n\}$ sei die Folge der Verweilzeiten bis zur Fertigstellung der Bedienungen in den einzelnen Stationen. Die Zufallsvariablen S_i, $i = 1, \ldots, \alpha$ seien wie die stationäre Bedienzeiten in den jeweiligen Bedienern verteilt und die Zufallsvariable T sei wie die stationären Zwischenankunftszeiten verteilt. Dann folgt

Satz 3 *Falls $E[S_i] < E[T]$ für alle $i \in \{1, \ldots, \alpha\}$ gilt, dann konvergiert $\{V_n\}$ in Verteilung unabhängig von den Anfangsbedingungen gegen eine eindeutig bestimmte stationäre Zufallsvariable V.*

Beweis folgt direkt aus [BB94], Seite 119.

Obiger Satz sichert also die Existenz einer stationären Verweilzeitverteilung, falls die Erwartungswerte der einzelnen stationären Bedienzeiten jeweils kleiner als der Erwartungswert der stationären Zwischenankunftszeit sind.

Das heißt, wenn die Stabilitätsbedingung $\rho_i < 1$ erfüllt ist, konvergiert die Folge der transienten Verweilzeitvektoren in Verteilung gegen den eindeutig bestimmten stationären Verweilzeitvektor.

Aus der Konvergenz von $\{V_n\}$ gegen stationäres V folgt direkt die Konvergenz aller weiteren Warte- und Verweilzeiten, denn falls $\{V_n\}$ in Verteilung gegen stationäres V konvergiert, gilt

$$\begin{aligned}
V_{n,i} &= W_{n,i} + S_i, \; i \in \{1, \ldots, \alpha\} \Longrightarrow \{W_n\} \text{ konv. in Vert. gegen stationäres } W, \\
W_{n,i} &= Q_{n,i} + \sum_{j=1}^{i-1} S_j, \; i \in \{1, \ldots, \alpha\} \Longrightarrow \{Q_n\} \text{ konv. in Vert. gegen stationäres } Q, \\
Q_{n,i} &= \sum_{j=1}^{i} \tilde{Q}_{n,j}, \; i \in \{1, \ldots, \alpha\} \Longrightarrow \{\tilde{Q}_n\} \text{ konv. in Vert. gegen stationäres } \tilde{Q}, \\
\tilde{V}_{n,i} &= \tilde{Q}_{n,i} + S_i, \; i \in \{1, \ldots, \alpha\} \Longrightarrow \tilde{V}_n \text{ konv. in Vert. gegen stationäres } \tilde{V}.
\end{aligned}$$

Somit ist also die Konvergenz transienter Größen gegen stationäre Größen gesichert. Offen bleibt die Konvergenzgeschwindigkeit in Abhängigkeit von der Art der gewählten Iteration.

5 Vergleich der Algorithmen und Bewertung der Resultate

Um die drei vorgestellten 101-Approximationen und die mit deren Hilfe berechneten Resultate zu bewerten, wird zunächst ein Vergleich der möglichen Fehlerquellen und - in Tabelle 2 - der erforderlichen Anzahl von Operationen auf Zufallsvariablen angestellt.

Bei der 101-Approximation unter Erneuerungsannahme entstehen Fehler, falls die Abgangsprozesse aus den einzelnen Stationen keine Erneuerungsprozesse sind. Die einzigen Tandemsysteme, für die die Erneuerungsannahme erfüllt ist, sind $\mathrm{M}(\rightarrow/\mathrm{M}/1)^{\alpha}$-Systeme (vgl. 3.3.1).

Bei der 101-Approximation unter Unabhängigkeitsannahme entstehen Fehler, falls die transienten Verweilzeiten abhängig sind, da Faltung und Maximumbildung nur für unabhängige Zufallsvariablen definiert sind. Die Frage der Abhängigkeiten der transienten Verweilzeiten ist bis heute ungeklärt.

Bei der 101-Approximation ohne Unabhängigkeitsannahmen entstehen zwar keine Fehler aufgrund falscher Voraussetzungen, jedoch fallen hier die 101-Approximationsfehler stärker ins Gewicht, da wesentlich mehr Faltungen und Maxima berechnet werden müssen und sich die Fehler somit stärker fortpflanzen.

In allen drei 101-Approximationen treten Fehler aufgrund der nicht exakten Darstellung von Dichten als 101-Dichten, der Faltung, sowie der Maximumbildung von Zufallsvariablen mit 101-Dichten, oder einer ungünstigen Wahl der 101-Parameter N, L, d auf. Es ist klar, daß diese Fehler größer werden, je mehr Operationen auf Zufallsvariablen ausgeführt werden.

101-Approximation	$V_{n,i} \rightarrow V_{n+1,i}$	$V_n \rightarrow V_{n+1}$	V_n
Erneuerungsannahme	$O(1)$	$O(\alpha)$	$O(\alpha n)$
Unabhängigkeitsannahme	$O(\alpha^3)$	$O(\alpha^3)$	$O(\alpha^3 n)$
ohne Unabhängigkeitsannahmen	$O(n^{i-1} \log n)$	$O(n^{\alpha-1} \log n)$	$O(n^{\alpha} \log n)$

Tabelle 2: Anzahl der Operationen auf Zufallsvariablen

Es soll nun ein Vergleich zwischen den mittels der drei beschriebenen 101-Approximationen berechneten Werten und exakten bzw. aus anderen Verfahren resultierenden Ergebnissen angestellt werden. Exakte Werte kann man dabei lediglich für $\mathrm{M}(\rightarrow/\mathrm{M}/1)^{\alpha}$-Systeme bestimmen. Für alle anderen Arten von Tandemsystemen muß der Vergleich auf Beispiele beschränkt werden, anhand derer die Verfahren QNA, QNET/ΠNET und MSA erprobt wurden. Dabei wird die Güte der Approximationen jeweils an den in [HN90] bzw. [GH96] angegebenen Simulationsergebnissen gemessen. Diese Simulationen erfolgten jeweils für 50000 Kunden nach einer 'Aufwärmphase' von 10000 Kunden bei 40-facher Wiederholung und somit können die Mittelwerte dieser Simulationen als nahezu exakte Werte angesehen werden. Wenn also im weiteren Verlauf der Begriff Fehler verwendet wird, so ist jeweils die Abweichung vom Mittelwert der jeweiligen Simulation gemeint.

Die Abbruchbedingung bei den 101-Approximationen war jeweils eine relative Abweichung der untersuchten Momente zweier aufeinanderfolgender iterierter Zufallsvariablen von weniger als 10^{-5}.

Für $\alpha = 2$ wurden in [GH96] Systeme $E_2/E_2/1 \rightarrow /E_2/1$ und $E_2/E_2/1 \rightarrow /H_2/1$ untersucht. Dabei wurden jeweils die Auslastungen ρ_1 und ρ_2 mit Werten 0.2, 0.5 und 0.8 gewählt, also für beide Systemtypen jeweils neun verschiedene Systemspezifikationen (ρ_1, ρ_2) betrachtet. Für diese Beispiele sind in den Tabellen 3 und 4 die mittleren und maximalen Fehler zusammengestellt, die bei der Berechnung der ersten drei Momente der Wartezeiten $\tilde{Q}_2$ in der Warteschlange der zweiten Station beobachtet wurden. Dabei ist anzumerken, daß mittels QNA lediglich die ersten beiden Momente berechnet werden können.

Bei den 101-Approximationen hierzu wurden die Parameter $N := 0$, $d := 0.05$ festgelegt und der Parameter L so gewählt, daß $Ld = \max(E[T], E[S_1], E[S_2])$ galt, also $L := 20E[T]$.

Im folgenden seien die 101-Approximationen mit 101_{EA} (Erneuerungsannahme), 101_{UA} (Unabhängigkeitsannahme) und 101_{OA} (ohne Unabhängigkeitsannahmen) bezeichnet.

Maß	mittlerer Fehler					maximaler Fehler				
	QNA	MSA	101_{EA}	101_{UA}	101_{OA}	QNA	MSA	101_{EA}	101_{UA}	101_{OA}
$E[\tilde{Q}_2]$	6.57	2.77	5.45	5.16	4.89	10.35	6.13	10.76	12.52	14.98
$E[\tilde{Q}_2^2]$	22.18	6.29	10.89	12.03	10.96	37.50	16.97	24.32	36.60	40.76
$E[\tilde{Q}_2^3]$		10.44	15.71	19.47	21.34		28.92	42.76	68.14	83.92

Tabelle 3: Fehler in % für die Wartezeiten in der zweiten Station bei $E_2/E_2/1 \rightarrow /E_2/1$-Systemen

Maß	mittlerer Fehler					maximaler Fehler				
	QNA	MSA	101_{EA}	101_{UA}	101_{OA}	QNA	MSA	101_{EA}	101_{UA}	101_{OA}
$E[\tilde{Q}_2]$	12.34	2.44	2.56	10.74	4.61	33.33	3.77	3.92	25.13	13.61
$E[\tilde{Q}_2^2]$	46.22	4.62	3.85	28.91	8.18	82.76	8.09	5.58	60.56	30.88
$E[\tilde{Q}_2^3]$		8.06	5.74	42.10	17.56		17.55	9.76	106.14	59.03

Tabelle 4: Fehler in % für die Wartezeiten in der zweiten Station bei $E_2/E_2/1 \rightarrow /H_2/1$-Systemen

Für diese untersuchten Tandemsysteme mit 2 Bedienstationen zeigt sich, daß die drei 101-Approximationen im Mittel durchweg bessere Ergebnisse als QNA liefern. Bei 101_{EA} sind mit einer Ausnahme auch die maximalen Fehler kleiner als bei QNA und für $E_2/E_2/1 \rightarrow /H_2/1$-Systeme sowohl mittlere als auch maximale Fehler für das zweite und dritte Moment sogar kleiner als bei MSA.

Für 101_{UA} und 101_{OA} stellt sich heraus, daß die Approximationen nicht verläßlich sind, das heißt, sie weisen zum Teil erhebliche Fehler auf, obwohl sie im Mittel gute Ergebnisse liefern. Dazu ist jedoch anzumerken, daß bei beiden Verfahren große Fehler nur in sehr wenigen Fällen auftraten und somit sozusagen als 'Ausreißer' bezeichnet werden könnten. Die Ergebnisse konnten in diesen Fällen durch Veränderungen der 101-Parameter jeweils erheblich verbessert werden, sodaß 101_{OA} die besten Werte berechnete.

Die Wahl der 101-Parameter scheint also insbesondere für 101_{OA} einen wesentlich größeren Einfluß auf die Approximationsgüte zu haben als bei 101_{EA}. Es ist jedoch klar, daß bei der Analyse von Tandemsystemen, wie allgemein bei der Analyse von Warteschlangennetzen, die exakten Ergebnisse vorab natürlich nicht bekannt sind, sodaß man sich auf eine bestimmte Art der Parameterwahl festlegen muß. Diese Parameterwahl weiter zu verbessern, etwa ausgehend von den in 2.2.2 beschriebenen Ansätzen, muß künftig ein Hauptanliegen bei der Arbeit mit 101-Approximationen sein.

Um die Approximationen für Tandemsysteme mit mehr als zwei Bedienstationen zu testen, werden nun zwei Beispiele, sogenannte Fraker-Systeme, aus [HN90] betrachtet. Der Ankunftsprozeß an der ersten Station ist dabei jeweils ein Poisson-Prozeß mit Rate 1, die Bedienzeiten sind Erlang-k-verteilt mit $k \in \{1, 4, 8, \infty\}$ und die Auslastungen nehmen Werte $\rho_i \in \{0.3, 0.5, 0.7, 0.9\}$ an. Es stellt sich dabei das Problem, eine E_∞-Verteilung zu erzeugen. Das Analysewerkzeug VERENA erzeugt E_k-Verteilungen nur für $k \leq 60$, sodaß also bei den 101-Approximationen eine E_{60}-Verteilung als Näherung für eine E_∞-Verteilung angenommen werden muß. Tabelle 5 enthält die Spezifikationen für die beiden betrachteten Beispiele, wobei die Bedienzeit an der i-ten Station E_{k_i}-verteilt ist.

Tabelle 6 zeigt die Erwartungswerte der berechneten Gesamtwartezeiten Q_α für diese Systeme, sowie die Mittelwerte der entsprechenden Simulationen. In Klammern unter den Resultaten sind jeweils die relativen Abweichungen von diesen Mittelwerten angegeben.

System-		Nummern i der Bedienstationen							
Nummer		1	2	3	4	5	6	7	8
1	ρ_i	0.7	0.5	0.5	0.7	0.3	0.9	0.9	0.3
	k_i	8	1	∞	4	∞	1	8	4
2	ρ_i	0.9	0.7	0.3	0.9	0.3	0.5	0.5	0.6
	k_i	1	8	4	∞	8	4	∞	1

Tabelle 5: Systemspezifikation der 8-Station-Tandemsysteme

System-Nr.	Simulation	QNA	ΠNET	101_{EA}	101_{UA}	101_{OA}
1	13.62	11.76	12.86	11.77	6.03	8.92
		(13.7%)	(5.6%)	(13.6%)	(55.7%)	(34.5)%
2	10.53	12.31	9.91	12.83	6.86	6.50
		(16.9%)	(5.9%)	(20.8%)	(34.9%)	(38.3%)

Tabelle 6: Erwartungswerte der Gesamtwartezeiten für die 8-Station-Tandemsysteme

Die Beispiele zeigen, daß bei der Analyse von Tandemsystemen mit größerer Anzahl von Bedienstationen die mittels 101_{UA} und 101_{OA} berechneten Werte erhebliche Fehler aufweisen. Ein ähnlicher Effekt wurde bei der Analyse von M($\rightarrow$/M/1)$^\alpha$-Systemen mit $\alpha = 4, \ldots, 8$ deutlich. Dabei ist es im Gegensatz zu Tandemsystemen hierbei nur in wenigen Fällen gelungen, durch Veränderungen der 101-Parameter wesentliche Verbesserungen zu erreichen. Dies ist vermutlich darauf zurückzuführen, daß aufgrund der sehr großen Anzahl benötigter Operationen auf Zufallsvariablen mit 101-Dichten der Einfluß der 101-Parameter mit wachsender Anzahl von Stationen immer stärker zunimmt und somit jeweils nur sehr spezielle Parameterwahlen zu zufriedenstellenden Ergebnissen führen.

Desweiteren ist zur Analyse von M($\rightarrow$/M/1)$^\alpha$-Systemen lediglich zu sagen, daß 101_{EA} erwartungsgemäß gute Ergebnisse liefert, da für diese Systeme die Erneuerungsannahme erfüllt ist. Die folgende Tabelle gibt einen Überblick über die maximalen Fehler bei der Berechnung der Erwartungswerte der Wartezeiten $\tilde{Q}_i$ mittels 101_{EA}.

Station	1	2	3	4	5	6	7	8
max. Fehler in %	0.37	0.93	1.81	3.48	5.89	8.70	11.13	15.44

6 Zusammenfassung und Ausblick

Es soll nun kurz zusammengefaßt werden, welche neuen Ergebnisse in dieser Arbeit erzielt wurden, welche Verbesserungen in Zukunft in Betracht kommen und welche Fragen weiterhin offen bleiben.

Ein wesentlicher Fortschritt aller drei beschriebener 101-Approximationen gegenüber bisher bekannten Verfahren besteht darin, daß keinerlei Einschränkungen bezüglich der Anzahl der Bedienstationen und der Verteilungen von Zwischenankunfts- und Bedienzeiten bestehen. Zudem können von allen berechneten Leistungsgrößen, einschließlich transienter Größen, beliebige Momente berechnet werden, ohne daß ein zusätzliches Näherungsverfahren benötigt wird, da man für die approximierten Zufallsvariablen lediglich die Wahrscheinlichkeitsmassen mit der entsprechenden Potenz multiplizieren muß. Konkreter bedeutet dies, daß zur Momentberechnung keine nichtlinearen Gleichungen gelöst werden müssen.

Ein weiterer Vorteil, den die 101-Approximation ohne Unabhängigkeitsannahmen bietet, besteht darin, daß hierbei keine Fehler durch vorab gemachte falsche Voraussetzungen auftreten.

In der Zukunft können und sollten die 101-Approximationen, insbesondere durch detaillierte Untersuchung der Wahl der 101-Parameter, jedoch weiter verbessert werden. So ist beispielsweise weiterhin unklar, wie die Intervallbreite d bei der Zerlegung von $\mathbb{R}$ optimal gewählt werden kann. Ebenso kann die Wahl der Grenzen N und L weiter verfeinert werden und der Einfluß der 101-Parameter auf die Güte der Approximationen und die Konvergenzgeschwindigkeit der Iterationen näher analysiert werden. Dies stellt jedoch kein spezielles Problem der Analyse von Tandemsystemen dar, sondern ist noch ein generelles Problem bei der 101-Approximation.

Ein weiteres Ziel für die Zukunft ist die Erweiterung der beiden auf (max,+)-linearen Systemen basierenden 101-Approximationen auf allgemeinere Warteschlangennetze. Dabei ist zumindest bereits folgende grobe Vorgehensweise klar :

1. Darstellung des Netzes als (max,+)-lineares System,
   konkret: Bestimmung der Matrizen A_n und der Vektoren B_n;
2. Anpassung dieser Darstellung an Erfordernisse der 101-Approximation,
   konkret: Operationen max und + (Faltung) nur auf unabhängigen Zufallsvariablen.

Hierbei bleibt die Frage offen, ob es für jedes Netz eine solche Darstellung als (max,+)-lineares System gibt und ob für jedes (max,+)-lineare System eine äquivalente Darstellung existiert, die die Erfordernisse der 101-Approximation erfüllt. Für den Fall, daß man für ein Netz eine solche Darstellung garantieren kann, wäre es zudem wünschenswert, ein algorithmisches Schema zu finden, um die Darstellung mit abhängigen in eine Darstellung mit ausschließlich unabhängigen Zufallsvariablen zu transformieren.

A Das Modul RV2Plus

A.1 Definitionsmodul

```
DEFINITION MODULE RndVar2P;
(* Betriebssystem Linux, Compiler Mocka *)

FROM Math IMPORT imin, imax, exp, KonvertMExp;
FROM MathLib IMPORT real;
FROM SYSTEM IMPORT ADR;
FROM Reals IMPORT Real, RK, RealSize, RZugriff;
FROM RVBasics IMPORT ZufallsVar, Setze0Vert, VertVeraendert;
FROM RndVar2 IMPORT Moment0;
FROM RndVar3 IMPORT ErlangKVert;


PROCEDURE ErwartungStet(X:ZufallsVar):Real;
(* berechnet den Erwartungswert der stetigen Zufallsvariablen X *)


PROCEDURE Moment2Stet(X:ZufallsVar):Real;
(* berechnet das zweite Moment der stetigen Zufallsvariable X *)


PROCEDURE VarianzStet(X:ZufallsVar):Real;
(* berechnet die Varianz der stetigen Zufallsvariable X *)


PROCEDURE MomenteStet(X:ZufallsVar; VAR M:ARRAY OF Real; k:INTEGER);
(* berechnet die ersten k Momente der stetigen Zufallsvariable X *)


PROCEDURE MaxStetZV(X,Y:ZufallsVar; VAR M:ZufallsVar);
(* berechnet das Maximum M der stetigen Zufallsvariablen X und Y *)


PROCEDURE MinStetZV(X,Y:ZufallsVar; VAR M:ZufallsVar);
(* berechnet das Minimum M der stetigen Zufallsvariablen X und Y *)


END RndVar2P.
```

A.2 Implementationsmodul

```
IMPLEMENTATION MODULE RndVar2P;
(* Betriebssystem Linux, Compiler Mocka *)

FROM Math IMPORT imin, imax, exp, KonvertMExp;
FROM MathLib IMPORT real;
FROM SYSTEM IMPORT ADR;
FROM Reals IMPORT Real, RK, RealSize, RZugriff;
FROM RVBasics IMPORT ZufallsVar, Setze0Vert, VertVeraendert;
FROM RndVar2 IMPORT Moment0;
FROM RndVar3 IMPORT ErlangKVert;


PROCEDURE ErwartungStet(X:ZufallsVar):Real;
(* berechnet den Erwartungswert der stetigen Zufallsvariablen X *)

VAR n,l,d,EW,c1,c2,l1,l2,j:Real;
    Vert:RZugriff;
    i:INTEGER;
BEGIN
  d:=X^.delta;
  n:=real(X^.aStet);
  l:=real(X^.bStet+1);
  KonvertMExp(X^.Wlinks,X^.Mlinks,-n*d,-1.0,c1,l1);
  KonvertMExp(X^.Wrechts,X^.Mrechts,l*d,1.0,c2,l2);
  EW:=X^.Mlinks+X^.Mrechts;
  WITH X^ DO
    WITH Vert DO
      Adr:=VertStet+VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
      *RealSize;
      FOR i:=aStet TO bStet DO
        j:=real(i);
        EW:=EW+Elem^*delta*(j+0.5);
        Adr:=Adr+RealSize;
      END;
    END;
  END;
  RETURN EW;
END ErwartungStet;


PROCEDURE Moment2Stet(X:ZufallsVar):Real;
(* berechnet das zweite Moment der stetigen Zufallsvariable X *)

VAR n,l,d,l1,l2,c1,c2,m2,einDrittel,j:Real;
    Vert:RZugriff;
    i:INTEGER;
```

```
BEGIN
  d:=X^.delta;
  n:=real(X^.aStet);
  l:=real(X^.bStet+1);
  einDrittel:=RK(1.0/3.0);
  KonvertMExp(X^.Wlinks,X^.Mlinks,-n*d,-1.0,c1,l1);
  KonvertMExp(X^.Wrechts,X^.Mrechts,l*d,1.0,c2,l2);
  m2:=-c1*exp(l1*n*d)*(n*n*d*d/l1+2.0*n*d/l1/l1+2.0/l1/l1/l1);
  m2:=m2+c2*exp(-l2*l*d)*(l*l*d*d/l2+2.0*l*d/l2/l2+2.0/l2/l2/l2);
  WITH X^ DO
    WITH Vert DO
      Adr:=VertStet+VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
      *RealSize;
      FOR i:=aStet TO bStet DO
        j:=real(i);
        m2:=m2+Elem^*delta*delta*(j*j+j+einDrittel);
        Adr:=Adr+RealSize;
      END;
    END;
  END;
  RETURN m2;
END Moment2Stet;


PROCEDURE VarianzStet(X:ZufallsVar):Real;
(* berechnet die Varianz der stetigen Zufallsvariable X *)

BEGIN
  RETURN Moment2Stet(X)-ErwartungStet(X)*ErwartungStet(X);
END VarianzStet;


PROCEDURE MomenteStet(X:ZufallsVar; VAR M:ARRAY OF Real; k:INTEGER);
(* berechnet die ersten k Momente der stetigen Zufallsvariable X *)

CONST K = 30; (* maximale Anzahl berechenbarer Momente *)
VAR n,l,d,c1,c2,l1,l2,j,a,b,az,bz,dz,rz,mm:Real;
    m1,m2:ARRAY[1..K] OF Real;
    Vert:RZugriff;
    i,i0,i1,z,q:INTEGER;
BEGIN
  M[0]:=Moment0(X);
  d:=X^.delta;
  n:=RK(real(X^.aStet));
  l:=RK(real(X^.bStet+1));
  a:=n*d; b:=l*d;
  KonvertMExp(X^.Wlinks,X^.Mlinks,a,-1.0,c1,l1);
  KonvertMExp(X^.Wrechts,X^.Mrechts,b,1.0,c2,l2);
```

```
  m1[1]:=-c1*exp(l1*a)*(a*l1-1.0)/l1/l1;  (* 1.Moment auf linkem Intervall *)
  m2[1]:=c2*exp(-l2*b)*(l2*b+1.0)/l2/l2;  (* 1.Moment auf rechtem Intervall *)
  WITH X^ DO
    WITH Vert DO
      Adr:=VertStet+VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
      *RealSize;
      mm:=0.0;
      FOR i:= aStet TO bStet DO
        j:=real(i);
        mm:=mm+Elem^*delta*(j+0.5);
        Adr:=Adr+RealSize;
      END;
    END;
  END;
  M[1]:=m1[1]+m2[1]+mm;
  dz:=d; az:=a; bz:=b;
  FOR z:=2 TO k DO
    dz:=dz*d;  (* dz = d^z *)
    az:=az*a;  (* az = d^z = (n*d)^z *)
    bz:=bz*b;  (* bz = b^z = (l*d)^z *)
    rz:=real(z);
    m1[z]:=az*X^.Wlinks-rz/l1*m1[z-1];
    m2[z]:=bz*X^.Wrechts+rz/l2*m2[z-1];
    WITH X^ DO
      WITH Vert DO
        Adr:=VertStet+VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
        *RealSize;
        mm:=0.0;
        FOR i:=aStet TO bStet DO
          i0:=i; i1:=i+1;
          FOR q:=1 TO z DO
            i0:=i0*i;
            i1:=i1*(i+1);
          END;
          j:=real(i1-i0);
          rz:=real(z+1);
          mm:=mm+Elem^*dz*j/rz;
          Adr:=Adr+RealSize;
        END;
      END;
    END;
    M[z]:=m1[z]+m2[z]+mm;
  END;
END MomenteStet;


PROCEDURE MaxStetZV(X,Y:ZufallsVar; VAR M:ZufallsVar);
(* berechnet das Maximum M der stetigen Zufallsvariablen X und Y *)
```

```
VAR l1,l2,c1,c2,ll,f1,f2,n,l,d,x,y,k:Real;
    i,j:INTEGER;
    VertM,VertX,VertY:RZugriff;
BEGIN
  IF (X^.aStet=0) AND (X^.bStet=-1) THEN ErlangKVert(X,1,1.0/X^.Mrechts); END;
  IF (Y^.aStet=0) AND (Y^.bStet=-1) THEN ErlangKVert(Y,1,1.0/Y^.Mrechts); END;
  (* Bei exakten Expverteilungen intern mit Erlang-1-Verteilungen
     DME-0-Bereich bei exakter Expverteilung ist leer *)

  d:=X^.delta;
  n:=RK(real(X^.aStet));
  l:=RK(real(X^.bStet+1));
  Setze0Vert(M);

  (* Berechnung der DME-1-Parameter von X *)
  KonvertMExp(X^.Wlinks,X^.Mlinks,n*d,-1.0,c1,l1);
  KonvertMExp(X^.Wrechts,X^.Mrechts,l*d,1.0,c2,l2);
  ll:=l1+l2;

  (* Benutzten DME-0-Bereich von M festlegen *)
  M^.aStet:=X^.aStet; M^.bStet:=X^.bStet;
  IF Y^.aStet<M^.aStet THEN M^.aStet:=Y^.aStet; END;
  IF Y^.bStet>M^.bStet THEN M^.bStet:=Y^.bStet; END;

  (* W-Masse und 1.Moment von M auf linkem Intervall berechnen *)
  M^.Wlinks:=X^.Wlinks*Y^.Wlinks;
  M^.Mlinks:=(-n*d-1.0/ll)*M^.Wlinks;

  (* Berechnung der DME-0 *)
  f1:=X^.Wlinks;   (* f1,f2 im folgenden jeweils Summe aller W-Massen von *)
  f2:=Y^.Wlinks;   (* X,Y bis zum aktuell betrachteten inneren Intervall *)
  WITH M^ DO
    WITH VertM DO
    Adr:=VertStet+VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
    *RealSize;
      FOR i:=aStet TO bStet DO
         (* Finde Wert der DME-0 von X auf aktuellem Intervall *)
         WITH X^ DO
            WITH VertX DO
              Adr:=VertStet
              +VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
              *RealSize;
              FOR j:=1 TO i DO
                Adr:=Adr+RealSize;
              END;
              x:=Elem^;
            END;
```

```
          END;
          (* Finde Wert der DME-0 von Y auf aktuellem Intervall *)
          WITH Y^ DO
            WITH VertY DO
              Adr:=VertStet
              +VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
              *RealSize;
              FOR j:= 1 TO i DO
                Adr:=Adr+RealSize;
              END;
              y:=Elem^;
            END;
          END;
          Elem^:=x*f2+y*f1+x*y;  (* Berechne DME-0 von M aus DME-0 von *)
          f1:=f1+x; f2:=f2+y;    (* X und Y und den W-Massen f1, f2 *)
        Adr:=Adr+RealSize;
      END;
    END;
  END;

  (* Berechnung der DME-1-Parameter von X und Y auf rechtem Intervall *)
  KonvertMExp(X^.Wrechts,X^.Mrechts,l*d,1.0,c1,l1);
  KonvertMExp(Y^.Wrechts,Y^.Mrechts,l*d,1.0,c2,l2);
  ll:=l1+l2;

  (* W-Masse und 1.Moment von M auf rechtem Intervall berechnen *)
  M^.Wrechts:=X^.Wrechts+Y^.Wrechts-X^.Wrechts*Y^.Wrechts;
  M^.Mrechts:=c1*exp(-l1*l*d)/l1/l1*(l1*l*d+1.0);
  M^.Mrechts:=c2*exp(-l2*l*d)/l2/l2*(l2*l*d+1.0)+M^.Mrechts;
  M^.Mrechts:=c1*c2*exp(-ll*l*d)/l1/l2/ll*(ll*l*d+1.0)+M^.Mrechts;
  VertVeraendert(M);
END MaxStetZV;


PROCEDURE MinStetZV(X,Y:ZufallsVar; VAR M:ZufallsVar);
(* berechnet das Minimum M der stetigen Zufallsvariablen X und Y *)
(* analog zu MaxStetZV mit den entsprechenden Gleichungen f\"ur min *)

VAR l1,l2,c1,c2,ll,f1,f2,n,l,d,x,y,k:Real;
    i,j:INTEGER;
    VertM,VertX,VertY:RZugriff;
BEGIN
  IF (X^.aStet=0) AND (X^.bStet=-1) THEN ErlangKVert(X,1,1.0/X^.Mrechts); END;
  IF (Y^.aStet=0) AND (Y^.bStet=-1) THEN ErlangKVert(Y,1,1.0/Y^.Mrechts); END;
  (* Bei exakten Expverteilungen intern mit Erlang-1-Verteilungen
     DME-0-Bereich bei exakter Expverteilung ist leer *)

  d:=X^.delta;
```

```
n:=RK(real(X^.aStet));
l:=RK(real(X^.bStet+1));
KonvertMExp(X^.Wlinks,X^.Mlinks,n*d,-1.0,c1,l1);
KonvertMExp(X^.Wrechts,X^.Mrechts,l*d,1.0,c2,l2);
ll:=l1+l2;
Setze0Vert(M);
M^.aStet:=X^.aStet; M^.bStet:=X^.bStet;
IF Y^.aStet<M^.aStet THEN M^.aStet:=Y^.aStet; END;
IF Y^.bStet>M^.bStet THEN M^.bStet:=Y^.bStet; END;
M^.Wlinks:=X^.Wlinks*Y^.Wlinks-X^.Wlinks*Y^.Wlinks;
M^.Mlinks:=(-n*d-1.0/ll)*M^.Wlinks;
f1:=X^.Wlinks;
f2:=Y^.Wlinks;
WITH M^ DO
  WITH VertM DO
  Adr:=VertStet+VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
  *RealSize;
    FOR i:=aStet TO bStet DO
       WITH X^ DO
          WITH VertX DO
            Adr:=VertStet
            +VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
            *RealSize;
            FOR j:=1 TO i DO
              Adr:=Adr+RealSize;
            END;
            x:=Elem^;
          END;
        END;
        WITH Y^ DO
          WITH VertY DO
            Adr:=VertStet
            +VAL(CARDINAL,imax(imin(rStetDef-1,aStet-aStetDef),0))
            *RealSize;
            FOR j:= 1 TO i DO
              Adr:=Adr+RealSize;
            END;
            y:=Elem^;
          END;
        END;
        Elem^:=x*(1.0-f2)+y*(1.0-f1)+x*y;
        f1:=1.0-f1+x; f2:=1.0-f2+y;
      Adr:=Adr+RealSize;
    END;
  END;
END;
KonvertMExp(X^.Wrechts,X^.Mrechts,l*d,1.0,c1,l1);
KonvertMExp(Y^.Wrechts,Y^.Mrechts,l*d,1.0,c2,l2);
```

```
  ll:=l1+l2;
  M^.Wrechts:=X^.Wrechts*Y^.Wrechts;
  M^.Mrechts:=(l*d+1.0/ll)*M^.Wrechts;
  M^.Mrechts:=c2*exp(-l2*l*d)/l2/l2*(l2*l*d+1.0)+M^.Mrechts;
  M^.Mrechts:=c1*c2*exp(-ll*l*d)/l1/l2/ll*(ll*l*d+1.0)+M^.Mrechts;
  VertVeraendert(M);
END MinStetZV;


BEGIN
END RndVar2P.
```

Literatur

[BB94] F. Baccelli, P. Bremaud: *Elements of Queueing Theory: Palm-Martingale Calculus and Stochastic Recurrences.* Springer, Berlin, Heidelberg, New York, 1994.

[BCOQ92] F. Baccelli, G. Cohen, G.J. Olsder, J.P. Quadrat: *Synchronization and Linearity: An Algebra for Discrete Event Systems.* John Wiley & Sons, Chichester, 1992.

[BIN88] J.P.C. Blanc, R. Iasnogorodski, P. Nain: Analysis of the M/GI/1→./M/1 queueing model. *Queueing Systems 3*, 1988.

[BS95] F. Baccelli, V. Schmidt: Series expansion for Poisson driven (max,+)-linear systems. To Appear in Ann. Appl. Prob., 1995.

[Bur64] P.J. Burke: The dependence of delays in tandem queues. *Ann. Math. Statist. 35*, 1964.

[Bur72] P.J. Burke: Output processes and tandem queues. In J. Fox, Hrsg., *Proc. Symp. Computer-Communication Networks and Teletraffic.* John Wiley & Sons, 1972.

[Dal68] D.J. Daley: The correlation structure of the output process of some single-server queueing systems. *Ann. Math. Statist. 39*, 1968.

[DK85] R.L. Disney, D. König: Queueing networks: A survey of their random processes. *SIAM Review Vol. 27, No. 3*, 1985.

[Fri65] H.D. Friedman: Reduction methods for tandem queueing systems. *Operations Research 8*, 1965.

[GH96] M.K. Girish, J.Q. Hu: Higher order approximations for tandem queueing networks. *Queueing Systems 22*, 1996.

[HN90] J.M. Harrison, V. Nguyen: The QNET method for two-moment analysis of open queueing networks. *Queueing Systems 6*, 1990.

[Jac56] R.R.P. Jackson: Queueing processes with phase-type service. *J. Roy. Statist. Soc., B-18*, 1956.

[Jac63] J.R. Jackson: Jobshop-like queueing systems. *Management Science 10*, 1963.

[Kel79] F.P. Kelly: *Reversibility and Stochastic Networks.* John Wiley & Sons, 1979.

[KK92] F.I. Karpelevitch, A.Y. Kreinin: Joint distributions in Poissonian tandem queues. *Queueing Systems 12*, 1992.

[Kle75] L. Kleinrock: *Queueing Systems. Volume I: Theory.* John Wiley & Sons, 1975.

[KLR73] A.M. Kagan, Y.V. Linnik, C.R. Rao: *Characterization Problems in Mathematical Statistics.* John Wiley & Sons, 1973.

[LMT93] M. Livny, B. Melamed, A.K. Tsiolis: The impact of autocorrelation on queueing systems. *Management Science 39*, 1993.

[Mor55] P.M. Morse: Stochastic properties of waiting lines. *Operations Research 3*, 1955.

[NB93] R. Nitzgen, B. Bärk: *VERENA - Eine Programmbibliothek zum Rechnen mit Wahrscheinlichkeitsverteilungen, Zweite Version, Institut für Informatik, Universität Bonn*, 1993.

[Nel58] R. T. Nelson: Waiting-time distributions for application to a series of service centers. *Operations Research 6*, 1958.

[New79] G.F. Newell: *Approximate Behavior of Tandem Queues (Lecture Notes in Economics and Mathematical Systems Vol. 171.* Springer, Berlin, Heidelberg, New York, 1979.

[Rei57] E. Reich: Waiting times when queues are in tandem. *Ann. Math. Statist. 28*, 1957.

[Sie91] D. Siepmann: VERENA - Eine Programmbibliothek zum Rechnen mit Wahrscheinlichkeitsverteilungen und ihre Erprobung an Pollingsystemen. Diplomarbeit, Institut für Informatik, Universität Bonn, 1991.

[SJ80] J.E. Shore, R.W. Johnson: Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inf. Theory, IT-26, No. 1*, 1980.

[Str89] J.C. Strelen: Piecewise approximation of densities applying the principles of maximum entropy: Waiting times in G/G/1-Systems. In R. Puigjaner, D. Potier, Hrsg., *Modelling Techniques and Tools for Computer Performance Evaluation.* Plenum Press, 1989.

[Web79] R. Weber: The interchangeability of ./M/1 queues in series. *Jour. Appl. Prob. 16*, 1979.

[Whi83] W. Whitt: The Queueing Network Analyzer. *Bell. Sys. Tech. J. 62*, 1983.

[Yak73] Y.F. Yakushev: On two-stage service of recurrent input. *Proc. Acad. Sci. USSR, Engin. Cybern. 3*, 1973.