

# Warteschlangen

Christian Dombacher  
Nikolaus Lenaugasse 8  
A-2232 Deutsch-Wagram

07.04.1994

## Hinweise und Fehlerberichtigung

Diese Arbeit wurde 1994 auf einem Apple MacIntosh IIfx und einem Laser-Writer II NTX unter Verwendung von Freehand 3.0, Mathtype 3.1 und Page-maker 5.0 erstellt und gedruckt. Leider sind sämtliche Layoutsysteme damals wie heute erheblich von den verwendeten Zeichensätzen abhängig. Trotzdem ist es uns gelungen, die Arbeit weitestgehend wieder herzustellen, obwohl einige Zeichensätze nicht mehr zur Verfügung standen. Im Zuge dessen wurden einige Details in den Postscriptdateien per Hand ersetzt, um die Arbeit lesbarer zu machen.

Im Gegensatz zu meiner späteren Arbeit "Stationary Queueing Models" geht die vorliegende Arbeit verstärkt auf die Grundlagen der Warteschlangentheorie ein. Sie kann somit als Einstieg in meine späteren Arbeiten gesehen werden. Ein weiterer Unterschied besteht in der Wahl der Sprache, die gesamte Arbeit ist in Deutsch gehalten.

Wie bei jedem größeren Unterfangen haben sich auch hier kleinere Fehler eingeschlichen. Diese werden im Folgenden benannt und die entsprechenden Berichtigungen angegeben:

- Seite 11: Gleichung [1.12] sollte richtig  $X = \sum_{n=1}^{\infty} \mu_n p_n$  lauten.
- Seite 28: In Gleichung [2.32] sollte der obere Term für

$$p_0 = \left( 1 + \sum_{n=1}^m \frac{1}{n!} \rho^n + \sum_{n=m+1}^{\infty} \frac{1}{m! m^{n-m}} \rho^n \right)^{-1}$$

lauten, es fehlt die Inversion.

- Seite 30: Hier ist Formel [2.38] unvollständig, es gilt

$$\begin{aligned}\Pr \{\text{Stau}\} &= \Pr \{m \text{ oder mehr Kunden im System}\} \\ &= \sum_{n=m}^{\infty} p_n = \sum_{n=m}^{\infty} \frac{1}{m!m^{n-m}} \rho^n p_0 = \frac{1}{m!} \rho^m \frac{1}{1 - \rho/m} p_0\end{aligned}$$

- Seite 33 Absatz 2.2.4: Statt Anschnitt lies Abschnitt.
- Seite 41 Absatz 3.3.1: der Mittelwert der Bedienzeitverteilung ist  $\frac{1}{\mu}$  ( $\lambda$  beschreibt die Ankunftsrate!).
- Seite 59: In der zweiten Formel von unten ist der Konvergenzpunkt des Limes falsch angegeben, der richtige Ausdruck lautet  $\lim_{n \rightarrow \infty} W_q^{(n)}(t) = W_q(t)$ .
- Seiten 59 und 60: Hier tritt des öfteren (siehe Formeln [3.59], [3.60], [3.61] und [3.62]) der Ausdruck  $U(x)$  auf, dieser ist durch  $u(x)$  zu ersetzen.
- Seite 76: Hier wurde das Elementsymbol in Formel [B.5] nicht richtig gesetzt, stattdessen lies  $\forall x \in \mathbb{R}$ .

# Inhaltsverzeichnis

|   |    |
|---|----|
| 1. Einführung .....   | 1  |
| 1.1 Geschichte der Warteschlangentheorie .....  | 1  |
| 1.2 Charakterisierung der Warteschlangenprozesse .....  | 1  |
| 1.2.1 Ankunftsprozeß .....  | 2  |
| 1.2.2 Bedienprozeß .....  | 5  |
| 1.2.3 Warteschlangendisziplin .....   | 7  |
| 1.2.4 Systemkapazität .....   | 7  |
| 1.2.5 Anzahl der Bediener .....   | 7  |
| 1.2.6 Zusammensetzung der Bedienstufen .....  | 8  |
| 1.2.7 Notation .....  | 8  |
| 1.3 Leistungsgrößen von Warteschlangenprozessen .....   | 10 |
| 1.4 Deterministische Wartesysteme .....   | 13 |
| 2. Geburten/Todesmodelle .....  | 17 |
| 2.1 M/M/1-System .....  | 17 |
| 2.1.1 Entwicklung des M/M/1-Modells .....   | 17 |
| 2.1.2 Leistungsgrößen der M/M/1-Modells .....   | 21 |
| 2.1.3 Warteraumbeschränkung .....   | 23 |
| 2.1.4 Andere Warteschlangendisziplinen .....  | 27 |
| 2.2 Lastabhängige Systeme .....   | 27 |
| 2.2.1 Das M/M/m-System .....  | 28 |
| 2.2.2 Das M/M/m/K-System .....  | 31 |
| 2.2.3 Ein endloser Bediener (Infinite Server) .....   | 32 |
| 2.2.4 Ungeduldige Kunden .....  | 33 |
| 2.2.5 Das Zentralrechnermodell (Central Server CPU Model) .....   | 36 |
| 2.2.6 Ein lastabhängiges System mit zwei Betriebsgeschwindigkeiten .....  | 38 |
| 2.2.7 Weiterführende Literatur .....  | 40 |
| 3. Fortgeschrittene Modelle .....   | 41 |
| 3.1 Das M/G/1-Modell .....  | 41 |
| 3.1.1 Zustandswahrscheinlichkeiten des M/G/1-Modells .....  | 41 |
| 3.1.2 Leistungsgrößen des M/G/1-Modells .....   | 44 |
| 3.1.3 Weitere Aspekte des M/G/1-Modells .....   | 48 |
| 3.1.4 Spezialisierung auf das M/M/1-Modell .....  | 49 |
| 3.1.5 Spezialisierung auf das M/D/1-Modell .....  | 50 |
| 3.2 Das G/M/1-Modell .....  | 52 |
| 3.2.1 Spezialisierung auf das M/M/1-Modell .....  | 55 |
| 3.2.2 Spezialisierung auf das D/M/1-Modell .....  | 56 |
| 3.3 Das G/G/1 Modell .....  | 59 |
| Anhang  |    |
| A. Differentialgleichungen .....  | 62 |
| A.1 Klassifikation .....  | 62 |
| A.2 Lösungen .....  | 62 |
| A.3 Methoden zur Lösung von Differentialgleichungen .....   | 63 |
| A.3.1 Trennung der Variablen .....  | 63 |
| A.3.2 Die Methode des integrierenden Faktors .....  | 63 |
| A.3.3 Variation der Konstanten für lineare Differentialgleichungen der 1. Ordnung .....                                 | 64 |
| A.3.4 Die charakteristische Gleichung für linear homogene Differentialgleichungen<br>mit konstanten Koeffizienten ..... | 65 |
| A.3.5 Der unbestimmte Ansatz für spezielle Störfunktionen .....   | 67 |
| A.3.6 Variation der Konstanten .....  | 69 |
| A.3.7 Systeme von Differentialgleichungen .....   | 70 |
| A.3.8 Zurückführen auf Differentialgleichungen mit niederer Ordnung .....   | 71 |
| A.3.9 Die Verwendung von Operatoren .....   | 72 |

|   |     |
|---|-----|
| B. Differenzgleichungen .....   | 74  |
| B.1 Klassifikation .....  | 74  |
| B.2 Typen von Differenzgleichungen und deren Lösungen .....               | 75  |
| B.2.1 Lineare Differenzgleichungen 1. Ordnung .....                       | 75  |
| B.2.2 Lineare Differenzgleichungen mit konstanten Koeffizienten .....     | 77  |
| B.2.3 Systeme von Differenzgleichungen mit konstanten Koeffizienten ..... | 79  |
| C. Transformationen und erzeugende Funktionen .....                       | 82  |
| C.1 Laplace - Transformation .....  | 82  |
| C.2 Laplace-Stieltjes-Transformation .....                                | 85  |
| C.3 Momenterzeugende Funktionen .....                                     | 86  |
| C.4 Z-Transformation (Erzeugende Funktion) .....                          | 87  |
| C.5 Transformationen wichtiger Verteilungen .....                         | 90  |
| D. Wahrscheinlichkeitstheoretische Grundlagen .....                       | 94  |
| D.1 Zufallsvariablen .....  | 94  |
| D.2 Elementare Größen .....   | 98  |
| D.3 Beziehungen zwischen mehrdimensionalen Zufallsvariablen .....         | 99  |
| D.3.1 Unabhängigkeit .....  | 99  |
| D.3.2 Bedingte Wahrscheinlichkeit .....                                   | 100 |
| D.3.3 Zentraler Grenzwertsatz .....                                       | 101 |
| D.4 Mehrdimensionale elementare Größen .....                              | 102 |
| D.5 Statistische Ungleichungen und Gesetze .....                          | 103 |
| E. Stochastische Prozesse und stochastische Ketten .....                  | 107 |
| F. Markovprozesse und Markovketten .....                                  | 109 |
| F.1 Markovketten mit diskretem Parameterraum .....                        | 109 |
| F.2 Markovketten mit kontinuierlichem Parameterraum .....                 | 113 |
| F.3 Verwendung der Matrizenrechnung .....                                 | 115 |
| G. Bibliographie .....  | 117 |
| H. Symbolverzeichnis .....  | 121 |

# 1. Einführung

Ob in Supermärkten, in öffentlichen Ämtern oder im Verkehrsstau - überall wird gewartet. Und warum warten wir? Wohl deswegen, weil zu bestimmten Zeiten des Tages mehr Leute bedient werden wollen, als bedient werden können.

Gleichzeitig mit den Warteschlangen, die sich vor den Schaltern bilden, tauchen Fragen auf wie z.B.

- Wie lange müssen die Kunden warten?
- Wie lange wird die Schlange?
- Wie kann ich meine Leute einsetzen, um die Wartezeiten zu verkürzen?

In der Warteschlangentheorie wird versucht, auf all diese und andere Fragen Antwort zu finden.

## 1.1 Geschichte der Warteschlangentheorie

Als Vater der Warteschlangentheorie kann der Däne A.K. Erlang angesehen werden, der sich als erster an die mathematische Behandlung von Telefongesprächen wagte. Den Startschuß setzte er mit seiner 1909 veröffentlichten Arbeit *The Theory of Probabilities and Telephone Conversations*.

In den 30'er Jahren wurde der nächste Meilenstein in der Warteschlangentheorie gelegt, die Pollaczek-Khintchine-Formel. Diese von Felix Pollaczek entwickelte Formel ermöglichte erstmals eine stark vereinfachte Berechnung der Kunden in einem Wartesystem mit allgemeinen Annahmen.

Im Jahr 1951 veröffentlichte D.G. Kendall seine Arbeit über das Konzept der eingebetteten Markovketten. Von ihm stammt auch die noch heute gültige Notation für Warteschlangensysteme.

Fast zur gleichen Zeit entwickelte Lindley eine Gleichung zur Berechnung der mittleren Wartezeit in einem System mit sehr allgemeinen Annahmen.

Bis zum Jahr 1957 lag das Hauptaugenmerk auf einzelnen Bedienstationen. Jackson beschäftigte sich als erster eingehend mit Warteschlangennetzen und fand heraus, daß man die Lösung eines offenen Warteschlangennetzes unter bestimmten Bedingungen aus den Lösungen der einzelnen Warteschlangenknoten zusammensetzen kann (Produktformlösung). Gordon und Newell zeigten, daß auch geschlossene Netzwerke unter den selben Annahmen eine solche Produktformlösung besitzen.

Mit dem Auftreten von Computern und Computernetzwerken erkannte man die Warteschlangentheorie als leistungsstarkes Design- und Analyseinstrument.

## 1.2 Charakterisierung der Warteschlangenprozesse

Zur Charakterisierung von Warteschlangenprozessen benötigt man folgende Angaben:

- Ankunftsprozeß
  - ...Wie groß ist der Zeitabstand zwischen den einzelnen Kunden?
  - ...Kommt immer nur ein Kunde oder kommen auch Gruppen an?
- Bedienprozeß
  - ...Wie lange dauert die Bedienung?
  - ...Ist die Dauer der Bedienung von der Anzahl der Kunden abhängig?
- Warteschlangendisziplin
  - ...Wird einer nach dem anderen abgefertigt oder wird jemand bevorzugt?

- Systemkapazität  
...Wie viele Leute passen in das Wartezimmer?
- Anzahl der Bediener  
...Wie viele Bedienmöglichkeiten stehen zur Verfügung?  
...Wenn ein weiterer Bediener hinzugefügt wird - wie ändert sich die Wartezeit?
- Zusammensetzung der Bedienstufen  
...Wie viele Bediener müssen passiert werden und wie lange dauert das?

Es sind einige Fragen angeführt, die in Zusammenhang mit den jeweiligen Punkten immer wieder auftreten. Die Komponenten eines Warteschlangensystems sind in Abb.1.1 dargestellt.

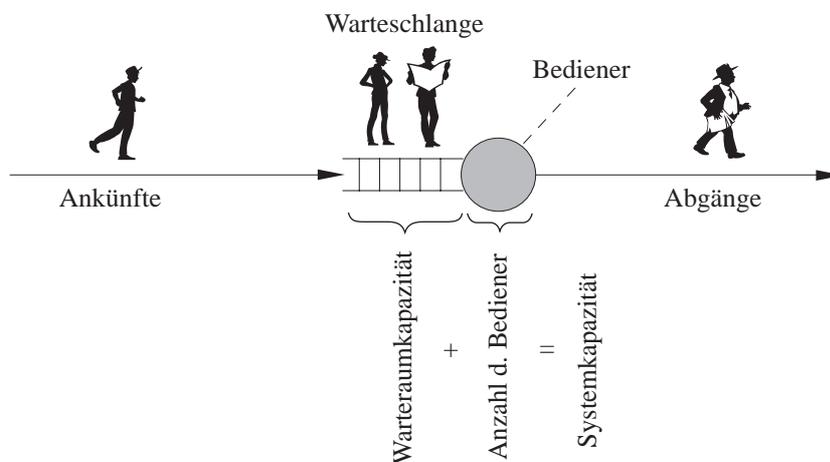


Abb 1.1 Schematische Darstellung eines Warteschlangenprozesses

### 1.2.1 Ankunftsprozeß

Die Beschreibung des Ankunftsprozesses kann auf zweierlei Arten erfolgen, zum einen durch Angabe der Verteilung der Ankünfte selbst, zum anderen durch Angabe der Verteilung der *Zwischenankunftszeiten*. Die *Zwischenankunftszeiten* sind die Zeiten zwischen den einzelnen Ankünften (siehe auch Abbildung 1.2)

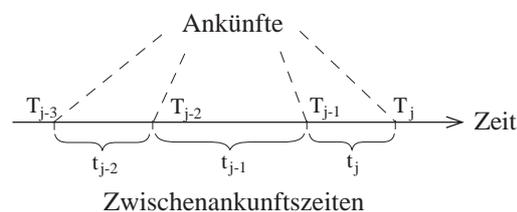


Abb 1.2. Zusammenhang Ankunftszeiten - Zwischenankunftszeiten

In der Warteschlangentheorie ist es üblich, die Verteilung der Zwischenankunftszeiten anzugeben. Die zugehörige Verteilung der Ankünfte läßt sich aus der Verteilung der Zwischenankunftszeiten ermitteln.

Bezeichnet  $t_j$  die Zeit zwischen zwei Ankünften zu den Zeitpunkten  $T_j$  und  $T_{j-1}$

$$t_j = T_j - T_{j-1} \quad [1.1]$$

und sind die  $(t_j)$  für alle  $j$  unabhängig und *exponentialverteilt* mit Parameter  $\lambda$

$$Pr\{t_j > t\} = e^{-\lambda t} \quad [1.2]$$

so ist die Anzahl der Ankünfte  $N_t$  im Zeitraum  $[0, t]$  *poissonverteilt*:

$$Pr\{N_t = j\} = \frac{(\lambda t)^j}{j!} e^{-\lambda t}$$

Die mathematische Herleitung ist in [GROS74] und [GROS85] in Kapitel 1.8 oder in [COOP72] in Kapitel 2.5 zu finden. Ausgehend von der kompletten Zufälligkeit der Ankünfte bietet [COX61] in Kapitel 1.3 einen anderen Ansatz.

Die Exponentialverteilung ist die wahrscheinlich wichtigste Verteilung in der Warteschlangentheorie, sie ist die einzige kontinuierliche Verteilung, die die *Markoveigenschaft* erfüllt. Diese besagt, daß die Vergangenheit keine Auswirkung auf die Zukunft hat. Im Falle der Verteilung der Ankünfte bedeutet dies, daß die seit der letzten Ankunft vergangene Zeit keinen Einfluß auf die Verteilung der Zeit bis zur nächsten Ankunft hat. Sind beispielsweise die Ankünfte von Zügen poissonverteilt und kommen im Mittel alle 30 Minuten Züge am Bahnhof an, so erwartet ein Reisender, der 25 Minuten nach der letzten Ankunft der Zuges am Bahnhof eintrifft, eine Wartezeit von 30 Minuten. Dies ist kein Widerspruch, da man von poissonverteilten Ankünften ausgeht. Im Falle deterministisch verteilter Ankünfte wäre die Wartezeit 5 Minuten. Diese Eigenschaft der *Gedächtnislosigkeit* erleichtert die mathematische Behandlung von Warteschlangen mit Poissonankünften, da die Zeit seit der letzten Ankunft im Modell nicht berücksichtigt werden muß.

Daher ist man auch versucht, einige mit der Exponentialverteilung verwandte Verteilungen auf diese zurückzuführen. Dazu gehören u.a. die *Erlangverteilung* und die *Hyperexponentialverteilung*. Bei der Bildung des Modells geht man davon aus, daß die Ankünfte gerade einen erlang- oder hyperexponentiell verteilten Bediener (siehe nächster Abschnitt) verlassen haben. Beide Verteilungen lassen sich durch Verknüpfung mehrerer Exponentialverteilungen erzeugen. Natürlich gibt es noch weitere Verteilungen, die sich auf diese Weise verwenden lassen. Dazu gehören die Hypoexponentialverteilung, die Coxverteilung und andere Erlangmischverteilungen. Der Vorteil all dieser Verteilungen besteht darin, daß sie leicht zur Approximation anderer, meist empirischer Verteilungsfunktionen herangezogen werden können. Ein solches Verfahren wird z.B. in [SCHM87] vorgestellt und in [SCHM89] erweitert. Mit der Parameterschätzung für Coxverteilungen beschäftigt sich [KRAM93].

Die wohl einfachste Form eines Ankunftsprozesses ist der Ankunftsprozeß mit *deterministisch* verteilten Zwischenankunftszeiten. Bis auf den Fall deterministisch verteilter Zwischenankunfts- und Bedienzeiten gestaltet sich

eine mathematische Behandlung schwieriger als bei Systemen mit poissonverteilten Ankünften. Dies kann darauf zurückgeführt werden, daß die deterministische Verteilung nicht gedächtnislos ist.

Bisher wurden nur kontinuierlich verteilte Zwischenankunftszeiten betrachtet, d.h. Ankünfte können zu jedem Zeitpunkt erfolgen. Besonders im Bereich der Computernetzwerke treten oft Architekturen auf, die eine fixe Paketlänge vorschreiben (z.B. ATM). Damit bildet die Zeit, ein solches Paket zu übertragen, einen Basisparameter des Modells. Die in anderen Warteschlangenmodellen kontinuierliche Zeit wird nun in *Slots* unterteilt und gibt daher Anlaß zur Verwendung von diskreten Verteilungen zur Beschreibung des Ankunftsprozesses. Das diskrete Analogon zur Exponentialverteilung ist die *geometrische Verteilung*. Bezeichnet  $p$  die Wahrscheinlichkeit einer Ankunft, so kann mit Hilfe der geometrischen Verteilung die Wahrscheinlichkeit für eine Ankunft genau im  $i$ -ten Slot berechnet werden:

$$Pr\{\text{Ankunft in Slot } i\} = (1 - p)^{i-1} p \quad [1.3]$$

Wie die Exponentialverteilung ist auch die geometrische Verteilung gedächtnislos. Bei geometrisch verteilten Zwischenankunftszeiten ist der Ankunftsprozeß ein *Bernoulliprozeß*.

$$\begin{aligned} Pr\{1 \text{ Ankunft}\} &= p \\ Pr\{\text{keine Ankunft}\} &= 1 - p \\ Pr\{\text{mehr als 1 Ankünfte}\} &= 0 \end{aligned} \quad [1.4]$$

Weiters von Interesse ist auch die *Binominalverteilung*, sie ermöglicht die Berechnung der Wahrscheinlichkeit von Kombinationen einer bestimmten Anzahl von freien und belegten Slots.

Eine Beschreibung von diskreten Ankunftsprozessen ist in [ROBE94] Kapitel 6.3 zu finden.

Doch nun zurück zur Poissonverteilung. Faßt man  $n$  verschiedene Poissonströme mit Zwischenankunftsverteilung  $1 - e^{-\lambda_i t}$  für  $i=1 \dots n$  zusammen, so erhält man wieder einen Poissonstrom, der nun die Zwischenankunftsverteilung  $1 - e^{-(\lambda_1 + \lambda_2 + \dots + \lambda_n)t}$  besitzt. Bei umgekehrtem Vorgang verhält es sich genauso. Wird ein Poissonstrom in  $n$  Ströme mit Wahrscheinlichkeit  $p_i$  für  $i=1 \dots n$  aufgespalten, so entstehen wieder  $n$  Poissonströme.

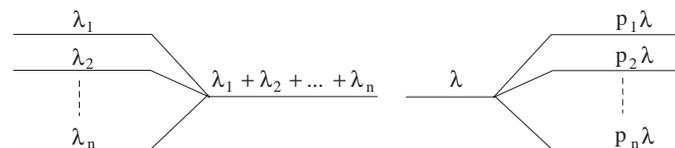


Abb 1.3. Verschmelzen und Aufspalten von Poissonströmen

Dabei bezeichnet  $\lambda$  die *Ankunftsrate*. Sie kann aus der mittleren Zwischenankunftszeit  $\bar{t}$  wie folgt ermittelt werden:

$$\lambda = \frac{1}{\bar{t}} \quad [1.5]$$

Eine weitere mögliche Interpretation von  $\lambda$  ist die mittlere Anzahl von Aufträgen pro Zeiteinheit.

Von technischer Seite aus kann man dem diskreten Bernoulliprozeß eine ähnliche Eigenschaft verleihen. Wer-

den  $n$  Bernoulliprozesse mit identischem Parameter  $p$  und gleich großen Slots verschmolzen (gemultiplext), so entsteht ein neuer Bernoulliprozeß. Die neue Slotgröße beträgt jetzt nur noch  $1/n$  der ursprünglichen Slotgröße. Die Übertragungsgeschwindigkeit in solchen Multiplexstrecken ist dann um vieles höher als bei der normalen Strecken. Nach dem Aufspalten entstehen dann wieder die ursprünglichen "langsamen" Bernoulliprozesse.

Auch bei allgemein verteilten Zwischenankunftszeiten sind noch Ergebnisse möglich. Weitaus angenehmer ist es jedoch, mit den zuvor genannten Approximationen zu arbeiten.

Neben der Verteilung der Zwischenankunftszeiten gibt es noch andere Kriterien, die den Ankunftsprozeß beeinflussen können. So gibt es verschiedene Formen von *Ungeduld* bei Kunden. Zum einen gibt es gehemmte Kunden, die bei zu langen Warteschlangen sich nicht in die Warteschlange eingliedern wollen und das Wartesystem wieder verlassen, zum anderen gibt es Kunden, die sich eine gewisse Zeit anstellen, dann die Gesamtwartezeit schätzen und das System gegebenenfalls wieder verlassen. Im ersten Fall wird eine "Hemmfunktion" eingeführt und im zweiten Fall wird die Abgangsrate (siehe nächster Abschnitt) erhöht. Neben den genannten gibt es natürlich auch noch andere Lösungsmöglichkeiten. Ein weiteres interessantes Phänomen sind die *Gruppenankünfte*. Bei den meisten Modellen wird der Einfachheit wegen von Einzelankünften ausgegangen.

### 1.2.2 Bedienprozeß

Der Bedienprozeß wird durch die *Verteilung der Bedienzeiten* charakterisiert. Jedoch wird auch hier nicht die mittlere Bedienzeit  $\bar{s}$ , sondern die *mittlere Bedienrate*  $\mu$  als Parameter für die meisten Verteilungen herangezogen. Der Zusammenhang ist wie folgt definiert:

$$\mu = \frac{1}{\bar{s}} \quad [1.6]$$

Die einfachsten Modelle lassen sich mit Hilfe der gedächtnislosen Exponentialverteilung erzeugen. Allerdings ist die Praxisnähe bei solchen Methoden stark in Frage gestellt. Darum ist man auch dazu übergegangen, verschiedene Modelle mit Hilfe von *hyperexponentiell* bzw. *erlangverteilten* Bedienern zu erstellen. Diese Bediener lassen sich durch parallel bzw. seriell verknüpfte exponentialverteilte Bedieneinheiten darstellen. Jeder dieser exponentialverteilten Bediener wird als Phase bezeichnet, man spricht also von einem hyperexponentiell bzw. erlangverteilten Bediener mit  $k$  Phasen. Dies ist in Abb 1.4 graphisch dargestellt.

Alle diese Verteilungen eignen sich zur Approximation von allgemeinen Verteilungen. Dabei kann die Exponentialverteilung zur Näherung von Verteilungen mit einem Variationskoeffizienten um 1 verwendet werden. Ist der Variationskoeffizient größer als 1, sollte die Hyperexponentialverteilung, ist er kleiner als 1 die Erlangverteilung herangezogen werden. Dabei ist die Bestimmung der einzelnen Parameter von großer Wichtigkeit. Die Wahl der richtigen Verteilung und die Schätzung der Parameter wird in [GROS85] Kapitel 6.6 behandelt.

Neben den oben genannten gibt es noch *Erlangverteilungen*, die sich mit Hilfe der Phasendarstellung erklären lassen. Dazu gehören die *Hypoexponential-* und die *Coxverteilung*. Erstere ist eine Verallgemeinerung der Erlangverteilung auf eine Verteilung mit  $k$  sequentiellen Phasen bei unterschiedlichen Bedienraten je Phase. Die Coxverteilung ist eine Verteilung mit  $k$  sequentiellen Phasen, wobei der Bediener nach jeder Phase verlassen werden kann. Eine Darstellung dieser Bedienverteilungen ist in [BOLC89] Kapitel 2.2.2 zu finden. Für Interessenten an einem Approximationsverfahren, welches Erlangverteilungen zur Näherung von empirischen Verteilungsfunktionen heranzieht ist [SCHM87] und [SCHM89] und [KRAM93] zu empfehlen.

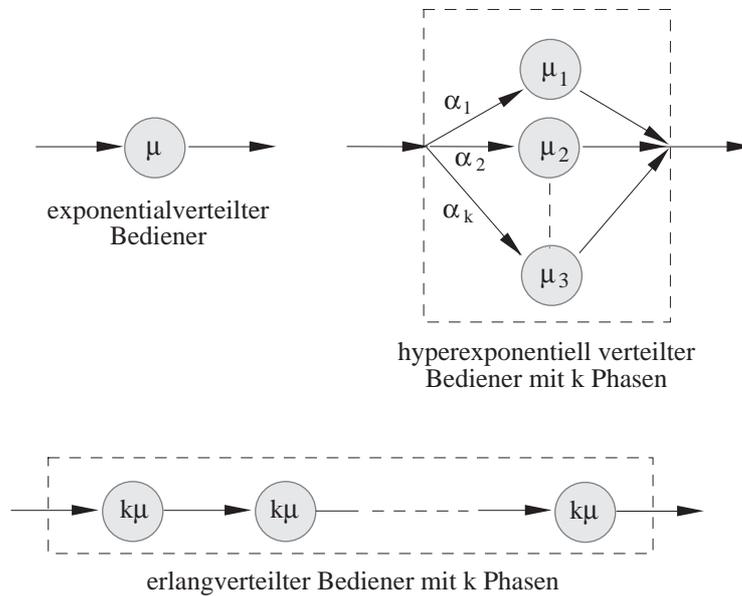


Abb 1.3 Verschiedene Bediener

Bisher wurden nur Systeme mit kontinuierlich verteilter Bedienzeit behandelt. Benötigt man jedoch ein Modell, in dem Ereignisse nur in bestimmten Zeitintervallen auftreten dürfen, so bietet sich die gedächtnislose *geometrische Verteilung* zur Beschreibung des Bedienverhaltens an. Dabei wird die Zeit in sogenannte *Slots* eingeteilt. (Näheres dazu: siehe voriger Abschnitt)

Weiters lassen sich Ereignisse für Modelle mit deterministischen bzw. allgemein verteilten Bedienzeiten ermitteln. Die Berechnung gestaltet sich meist kompliziert, da der Vorteil der Gedächtnislosigkeit nicht gegeben ist.

Eine tabellarische Darstellung der wichtigsten Verteilungen befindet sich im Anhang (Tab. D.1-D.3). Neben den Dichtefunktionen sind dort auch Erwartungswert und Varianz angeführt.

Zusätzlich zur Bedienzeitverteilung gibt es auch noch andere Kriterien, die das Bedienverhalten beeinflussen. So gibt es auch die Möglichkeit einer *zustandsabhängigen Bedienung*. Je nach Anzahl der Kunden im System wird der einzelne Kunde schneller oder langsamer bedient. Ein Schalterbeamter bemerkt, wie sich eine schier endlose Schlange aufbaut und beeilt sich entweder mit der Abfertigung, um bis Dienstschluß fertig zu sein oder er setzt die Arbeit entmutigt fort, da die Warteschlange ja doch nicht kleiner wird.

Weiters gibt es die Möglichkeit der *Gruppenbedienung*, d.h. ein Bediener fertigt mehrere Kunden auf einmal ab. Dabei sind zwei Fälle zu unterscheiden. Zum einen besteht die Möglichkeit, daß der Bediener auch weniger Kunden gemeinsam abfertigt, als ihm möglich ist, zum anderen wartet er, bis so viele Kunden eingetroffen sind, daß eine volle Gruppe bedient werden kann.

### 1.2.3 Warteschlangendisziplin

Die *Warteschlangendisziplin* gibt an, in welcher Art und Weise die einzelnen Kunden aus der Warteschlange zur Bedienung ausgewählt werden. Beispielhaft sind nun die wichtigsten Disziplinen genannt und beschrieben:

- FCFS (First-Come-First-Served) bzw. FIFO (First-In-First-Out)  
Die Kunden werden in der Reihenfolge ihrer Ankunft bedient.
- LCFS (Last-Come-First-Served) bzw. LIFO (Last-In-First-Out)  
Der zuletzt angekommene Kunde wird zuerst bedient. Diese Disziplin eignet sich zur Modellierung von Stapelproblemen.
- RSS (Random-Selection-For-Service) bzw. SIRO (Service-In-Random-Order)  
Die Auswahl des nächsten Kunden aus der Warteschlange erfolgt zufällig.
- PRI (Priority)  
Die einzelnen Kunden werden einer bestimmten Prioritätsklasse zugeordnet. Kunden einer höheren Prioritätsklasse werden Kunden einer niederen Prioritätsklasse vorgezogen. Bei Kunden mit gleicher Priorität erfolgt die Auswahl nach FCFS.  
Weiters werden folgende Prioritätsfunktionen unterschieden: Im *preemptiven* Fall darf ein Kunde mit höherer Priorität einen Kunden niederer Priorität aus der Bedieneinheit verdrängen, um selbst gleich abgefertigt zu werden. Der verdrängte Kunde wird nach bestimmten Gesetzmäßigkeiten wieder in die Warteschlange eingereiht. Im *nicht-preemptiven* Fall darf sich der Kunde mit der höchsten Priorität in der Warteschlange an erster Stelle einordnen. Besonders in letzter Zeit interessant geworden sind die *dynamischen* Prioritäten, d.h. die Prioritäten ändern sich in Abhängigkeit von der Zeit.
- RR (Round-Robin)  
Die Bedienung eines Kunden erfolgt innerhalb eines vorgegebenen Zeitintervalls. Überschreitet die Gesamtbedienzeit des Kunden die Länge dieses Intervalls, so wird dieser an letzter Stelle einer FSFS-Warteschlange eingereiht. Diese Strategie ist beispielsweise bei der Modellierung von synchronen Multiplexern interessant.

### 1.2.4 Systemkapazität

Bei vielen Modellen ist die Warteraumkapazität begrenzt. Wird die Systemkapazität (Warteraumkapazität + Kunden in Bedienung) erreicht, so werden neu ankommende Kunden abgewiesen.

Diese Situation beschreibt u.a. auch die einfachste Form eines gehemmten Kunden. Sobald das Limit erreicht ist, wird das System nicht betreten.

### 1.2.5 Anzahl der Bediener

Besitzt ein System mehrere *Bedieneinheiten*, so müssen zwei Fälle unterschieden werden. Zum einen verfügen die Bedienkanäle über eine gemeinsame Warteschlange, zum anderen verfügt jeder Kanal über eine eigene Warteschlange.

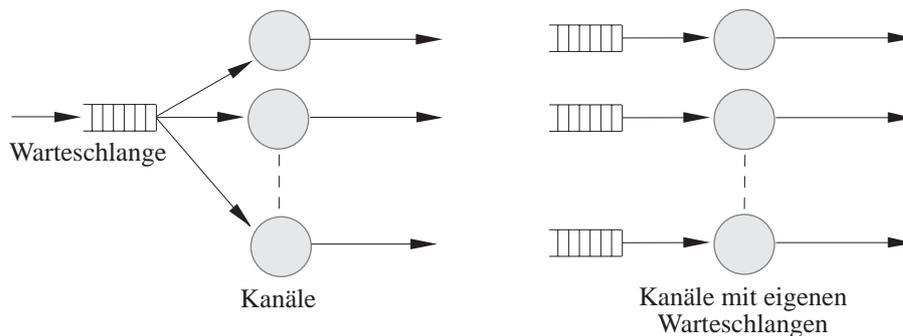


Abb 1.5 Unterscheidung von Mehrfachbedienern

Im zweiten Fall geht man allerdings von mehreren unabhängigen Einzelbedienern aus. Im ersten Fall kann ein äquivalenter lastabhängiger Einzelbediener erzeugt werden, der dem Mehrfachbediener entspricht. Abhängig von der Anzahl der Bediener  $m$  und der Anzahl der Kunden im System  $k$  ergibt sich folgende lastabhängige Bedienrate

$$\mu_k = \begin{cases} k\mu & \text{für } 1 \leq k < m \\ m\mu & \text{für } k \geq m \end{cases}$$

[1.7]

Dabei bezeichnet  $\mu$  die für jeden Bedienkanal gleiche mittlere Bedienrate.

### 1.2.6 Zusammensetzung der Bedienstufen

Besonders interessant im Zusammenhang mit Warteschlangennetzen ist die Zusammensetzung der Bedienstufen. Wie bereits erwähnt, lassen sich sequentiell angeordnete Bediener mit exponentiell verteilten Bedienzeiten zu einem erlang- bzw. hypoexponentiell verteilten Bediener zusammenfassen. Für eine Vielzahl von Modellen ist jedoch keine aggregierte Verteilung der einzelnen Bedienzeitverteilungen bekannt. Solche Warteschlangennetze können mit geeigneten Algorithmen analysiert werden.

### 1.2.7 Notation

Zur Beschreibung einfacher Warteschlangensysteme hat sich die folgende auf D.G. Kendall zurückgehende Kurzbeschreibung durchgesetzt.

Dabei bezeichnet  $A$  die Verteilung der Zwischenankunftszeiten und  $B$  die Verteilung der Bedienzeiten. Weiters drückt  $m$  die Anzahl der Bediener und  $K$  die maximale Anzahl der Kunden im System aus. Die Warteschlangendisziplin wird durch DIS beschrieben. In Tab. 1.1 sind alle Positionen und die gängigsten Kurzbezeichnungen zusammengefaßt.

| Position                                | Kapitelverweir | Symbol            | Beschreibung                            |
|---|----------------|-------------------|---|
| Verteilung der Zwischenankunftszeiten A | I.2.1          | M                 | Exponentialverteilung                   |
|   |                | D                 | Deterministische Verteilung             |
|   |                | $E_k$             | Erlangverteilung mit k Phasen           |
|   |                | $H_k$             | Hyperexponentialverteilung mit k Phasen |
|   |                | $C_k$             | Coxverteilung mit k Phasen              |
|   |                | GEO               | Geometrische Verteilung                 |
|   |                | G                 | Allgemeine Verteilung                   |
|   |                | GI                | Allgemeine unabhängige Verteilung       |
| Verteilung der Bedienzeiten B           | I.2.2          | M                 | Exponentialverteilung                   |
|   |                | D                 | Deterministische Verteilung             |
|   |                | $E_k$             | Erlangverteilung mit k Phasen           |
|   |                | $H_k$             | Hyperexponentialverteilung mit k Phasen |
|   |                | $C_k$             | Coxverteilung mit k Phasen              |
|   |                | GEO               | Geometrische Verteilung                 |
|   |                | G                 | Allgemeine Verteilung                   |
|   |                |                   |   |
| Anzahl der Bediener m                   | I.2.5          | 1, 2, 1, $\infty$ |   |
| Systemkapazität K                       | I.2.4          | 1, 2, 1, $\infty$ |   |
| Warteschlangendisziplin DJ              | I.2.3          | <u>FCFJ</u>       | <u>First-Come-First-Served</u>          |
|   |                | LCFJ              | Last-Come-First-Served                  |
|   |                | RJFJ              | Random-Selection-For-Service            |
|   |                | PRI               | Priority                                |
|   |                | RR                | Round-Robin                             |
|   |                | PJ                | Processor-Sharing                       |
|   |                | GD                | General Discipline                      |

Tab 1. 1 Warteschlangennotation

Dabei ist zu beachten, daß nur die ersten 3 Positionen angegeben werden müssen. Die anderen können, sofern sie von den in Tabelle 1.1 unterstrichenenen Standardwerten nicht abweichen, weggelassen werden.

### Beispiel 1.1

Bezeichnet M/D/2/100/PRI ein Warteschlangensystem, so können folgende Fakten abgelesen werden:

- Exponentialverteilte Zwischenankunftszeiten, d.h. poissonverteilte Ankünfte
- Deterministische Verteilung der Bedienzeiten
- 2 Bediener
- Systemkapazität 100, d.h. es gibt 98 Warteplätze
- Prioritätsvergabe als Strategie

Für ein GEO/GEO/1 System gilt:

- Geometrisch verteilte Zwischenankunftszeiten, d.h. der Ankunftsprozeß ist ein Bedienprozeß
- Geometrisch verteilte Bedienzeiten
- 1 Bediener
- Systemkapazität  $\infty$ , d.h. die Warteschlange puffert alle Ankünfte
- FCFS als Strategie

Ein M/M/1/ $\infty$  /LCFS-System kann auch so bezeichnet werden: M/M/1/LCFS

Natürlich können mit dieser Kurzschreibweise nicht alle Charakteristika beschrieben werden. Die zusätzlichen Details werden dann durch geeignete Symbole oder Ergänzungen beschrieben. Beispielsweise ist "M<sup>x</sup>/M/1" ein System mit Gruppenankünften. Für den zweiten Fall kann das "M/M/2-System mit heterogenen Servern" als Beispiel dienen. Es bezeichnet ein System, welches zwei Bedienkanäle mit unterschiedlichen Bedienraten besitzt.

### 1.3 Leistungsgrößen von Warteschlangenprozessen

Ziel bei der Modellierung von Warteschlangensystemen ist es, interessante Leistungsgrößen wie die Wartezeit oder die Anzahl der Kunden im System zu ermitteln. Dabei wird meist von einem *stationären System* ausgegangen. Dies bedeutet, daß das System schon eine unbestimmte Zeit in Betrieb ist und sich in einem Gleichgewichtszustand eingependelt hat. Dieses Einpendeln ist mit den gedämpften Schwingungen eines Regelsystems vergleichbar. Systeme, die sich nicht im Gleichgewicht befinden, sind nur schwer oder gar nicht zu analysieren.

In diesem Abschnitt werden einige wichtige Größen beschrieben und ihre Zusammenhänge erläutert.

- Die *Zustandswahrscheinlichkeit*  $p_n$  beschreibt die Wahrscheinlichkeit, daß sich ein stationäres Warteschlangensystem im Zustand  $n$  befindet:

$$p_n = Pr\{n \text{ Kunden im System}\} \quad [1.8]$$

Bei Systemen, die sich nicht im Gleichgewicht befinden, ist die zusätzliche Angabe eines Zeitparameters  $t$  notwendig.

- Die *Verkehrsintensität* beschreibt das Verhältnis von Ankunftsrate zu Bedienrate:

$$\rho = \lambda / \mu = \bar{s} / t \quad [1.9]$$

- Die *Auslastung*  $u$  gibt den Anteil der Gesamtzeit an, in der eine Bedieneinheit aktiv ist:

$$u = \lambda / m\mu = \rho / m \quad [1.10]$$

Besitzt das betrachtete Warteschlangensystem nur einen Bediener, so ist die Auslastung gleich der Verkehrsintensität. Beträgt die Auslastung 100%, so spricht man von einem *voll ausgelasteten Sy-*

stem. Liegt hingegen *Überlast* vor, d.h. die Auslastung ist größer als 100%, so handelt es sich um kein stationäres System. Die vereinfachten Formeln können infolgedessen nicht angewandt werden. Eine notwendige Bedingung für ein stabiles System mit unendlichem Warteraum lautet daher

$$u < 1 \quad [1.11]$$

Daraus ist ersichtlich, daß ein voll ausgelastetes System nicht stationär ist. Im Gegensatz zum Überlastfall, bei dem mehr Aufträge ankommen als abgefertigt werden können, ist die Erklärung nicht so intuitiv. Eine mögliche Interpretation wäre, daß es für den Bediener immer schwerer wird, die wachsende Schlange abzubauen, da die mittlere Bedienrate gleich der mittleren Ankunftsrate ist.

Bedingung [1.11] ist allerdings nicht immer ausreichend, um das untersuchte System als stabil zu bezeichnen. Systeme, bei denen die gemessenen, momentanen Auslastungen weit gestreut sind, neigen dazu, vom Normallastbereich in den Überlastbereich umzuschlagen. Die zu hohe momentane Auslastung führt dann zu einem Zusammenbruch des Systems. Diese Gefahr ist besonders bei Protokollen, die Kollisionen zulassen, gegeben (z.B. Ethernet). Es ist durchaus möglich, daß eine mittlere Auslastung von 90% zu einem Fehlverhalten führt. Werden allerdings Verfahren analysiert, die von vornherein keine Kollisionen zulassen (z.B. Token Ring), so besteht diese Gefahr nicht.

- Der *Durchsatz*  $X$  bezeichnet die mittlere Anzahl von Kunden, die pro Zeiteinheit bedient werden (*Abgangsrate*). Sofern ein System stabil ist, ist die Abgangsrate gleich der Ankunftsrate  $\lambda$ .

$$X = \sum_{n=1}^{\infty} n p_n \quad [1.12]$$

- Als mittlere *Antwort-* oder *Verweilzeit*  $W$  bezeichnet man die Zeit, die ein Kunde im Wartesystem verbringt.
- Die mittlere *Wartezeit*  $W_q$  gibt jene Zeit an, die sich ein Kunde bis zum Beginn seiner Bedienung in der Warteschlange aufhält. Es gilt folgender Zusammenhang:

$$W = W_q + \bar{s} = W_q + \frac{1}{\mu} \quad [1.13]$$

Die mittlere *Anzahl der Kunden*  $L$  im Wartesystem kann mit Hilfe der Definition des Erwartungswertes errechnet werden:

$$L = E(N) = \sum_{n=1}^{\infty} n p_n \quad [1.14]$$

Dabei bezeichnet  $N$  die Zufallsvariable "Kunden im System".

- Die mittlere *Warteschlangenlänge*  $L_q$  gibt die mittlere Anzahl der Kunden in der Warteschlange an.

- Das *Gesetz von Little* (1961):

Zwischen der mittleren Anzahl der Kunden im System und der mittleren Antwortzeit gibt es folgenden Zusammenhang:

$$L = \lambda W \quad [1.15]$$

Für die mittlere Warteschlangenlänge und die mittlere Wartezeit: gilt

$$L_q = \lambda W_q \quad [1.16]$$

Bemerkenswert ist, daß das Gesetz von Little für G/G/m-Systeme gültig ist. Auch Systemkapazität und Warteschlangendisziplin haben keinen Einfluß. Das Gesetz von Little kann, wie in folgender Zeichnung dargestellt, graphisch hergeleitet werden:

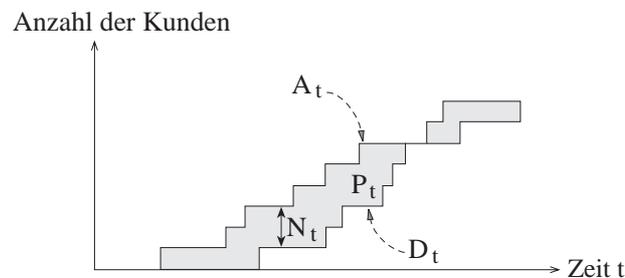


Abb 1.6 Ein Beweis für das Gesetz von Little

Innerhalb eines bestimmten Intervalls  $[0,t]$  bezeichnet  $A_t$  die Anzahl der Ankünfte und  $D_t$  die Anzahl der Abgänge. Die Differenz von  $A_t$  und  $D_t$  ergibt  $N_t$ , die Anzahl der Kunden im System zum Zeitpunkt  $t$ :

$$N_t = A_t - D_t$$

Bezeichnet  $\lambda_t$  die Ankunftsrate im Intervall  $[0,t]$ , so gilt

$$\lambda_t = \frac{A_t}{t}$$

Aus der Fläche  $R_t$  zwischen  $A_t$  und  $D_t$  kann die mittlere Anzahl der Kunden  $L_t$  zum Zeitpunkt  $t$  ermittelt werden

$$L_t = \frac{R_t}{t}$$

Dabei kann  $R_t$  als die Gesamtwartezeit im Intervall  $[0,t]$  interpretiert werden. Nun kann die mittlere Wartezeit jedes Kunden  $W_t$  leicht eruiert werden:

$$W_t = \frac{R_t}{A_t}$$

Werden die drei letzten Formeln zusammengefaßt, so ergibt sich

$$L_t = \frac{R_t}{t} = \frac{W_t A_t}{t} = W_t \lambda_t$$

Läßt man nun  $t \rightarrow \infty$  gehen, so kann daraus die bereits bekannte Gleichung [1.15] bestimmt werden. Mit Hilfe des Gesetzes von Little und Gleichung [1.13] lassen sich aus einem bekannten Wert die restlichen Leistungsgrößen berechnen (z.B. aus  $L$  folgen  $W$ ,  $W_q$  und schließlich  $L_q$ ).

Beweise für das Gesetz von Little sind fast überall in der Fachliteratur zu finden. Beispielsweise seien einige Bücher aufgezählt:

[GROS74], [GROS85], [KLEI74] und [LANG92].

Eine Zusammenstellung von Beweisen liefert [ROBE94] in Kapitel 2.3. Der Originaltext wurde von J.D.C. Little verfaßt und unter dem Namen *A Proof of the Queueing Formula  $L = \lambda W$  in Operations Research 9* im Jahr 1961 veröffentlicht. Eine gute Zusammenstellung der Leistungsgrößen für Warteschlangensysteme und -netze liefert [BOLC89] Kapitel 2.2.4.

## 1.4 Deterministische Wartesysteme

Die einfachsten Modelle in der Warteschlangentheorie sind die D/D/m-Systeme, da zu deren Beschreibung keine Verteilungsfunktionen nötig sind. Alle Systeme dieser Art lassen sich graphisch lösen. Der nun folgende Ansatz entspricht dem von [GROS85] Kapitel 1.7.

Dort wird von einem System mit einem Bediener und einer Kapazität  $K-1$  ausgegangen, es handelt sich infolgedessen um ein D/D/1/K-1 System. Ferner sind die Bedienzeiten größer als die Zwischenankunftszeiten, es gilt also  $\lambda > \mu$  bzw.  $\bar{s} > \bar{t}$ . Ein System dieser Art ohne Warteraumbeschränkung wäre überlastet. Bei voller Warteschlange werden Neuankünfte abgewiesen.

Bezeichnet  $N_t$  die Anzahl der Kunden zum Zeitpunkt  $t$ , so gilt für die Periode zwischen erster Ankunft und erster Abweisung

$$N_t = \{\text{Anzahl d. Ankünfte im Intervall } (0, t]\} - \{\text{Anzahl d. Abgänge im Intervall } (0, t]\}$$

$$= \frac{t}{\bar{t}} - \frac{t - \bar{t}}{\bar{s}} = \mu t - \left( \mu t - \frac{\mu}{\lambda} \right)$$

[1.17]

Zunächst wird ein deterministisches Wartesystem untersucht, bei dem die Bedienzeit ein Vielfaches der Ankunftszeit ist.

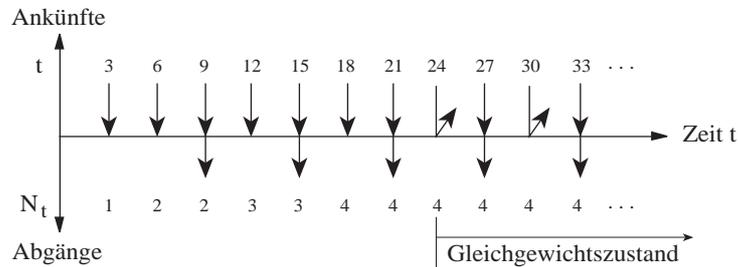


Abb 1.7 Ein D/D/1/4-System mit  $\bar{t} = 3, \bar{s} = 6$

Für ein D/D/1/4-System mit einer Zwischenankunftszeit von  $\bar{t} = 3$  und einer Bedienzeit von  $\bar{s} = 6$  ergibt sich aus Abbildung 1.7:

$$N_t = \begin{cases} 0 & (t < 3) \\ \frac{t}{3} - \left(\frac{t}{6} - \frac{1}{2}\right) & (3 < t < 24) \\ 4 & (t \geq 24) \end{cases}$$

Bevor die Wartezeit ermittelt werden kann, müssen einige neue Bezeichnungen eingeführt werden:

$n$  ... Laufnummer für die  $n$ -te erfolgreiche Ankunft

$S^{(n)}$  ... Servicezeit der  $n$ -ten Kunden,  $S^{(n)} = 6$  für  $n=1,2,\dots$

$T^{(n)}$  ... Zwischenankunftszeit zwischen dem  $n$ -ten und dem  $(n+1)$ -ten Kunden.

Im

Gleichgewichtszustand ist  $T^{(n)} = 6$ , sonst 3.

$W_q^{(n)}$  ... Wartezeit der  $n$ -ten Kunden

Im Gleichgewichtszustand gilt für  $W_q^{(n+1)}$

$$W_q^{(n+1)} = W_q^{(n)} + 6 - 6$$

sonst

$$W_q^{(n+1)} = W_q^{(n)} + 6 - 3$$

Da der erste Kunde nicht warten muß, gilt die Anfangsbedingung  $W_q^{(1)} = 0$ . Obige Ausdrücke können leicht erklärt werden. Die Wartezeit jedes neu angekommen Kunden erhöht sich um die Differenz von Bedien- und Zwischenankunftszeit. Zur Überprüfung kann Abbildung 1.7 herangezogen werden.

Allgemein läßt sich also formulieren:



$$N_t = \begin{cases} 0 & (0 \leq t < 3) \\ \frac{t}{3} - \frac{t-3}{7} & (3 \leq t < 15) \\ 3 & \left( \begin{array}{l} 21k + 15 \leq t < 21k + 17; \\ 21k + 18 \leq t < 21k + 31; \\ 21k + 33 \leq t < 21k + 36 \end{array} \right) \\ 2 & \left( \begin{array}{l} 21k + 17 \leq t < 21k + 18; \\ 21k + 37 \leq t < 21k + 33; \end{array} \right) \end{cases}$$

$$W_q^{(n)} = \begin{cases} 4(n-1) & (n < 5) \\ 13 & (n = 3k + 5) \\ 14 & (n = 3k + 6) \\ 12 & (n = 3k + 7) \end{cases}$$

mit  $k=0,1,2,\dots$

Systeme, die ihre Kunden schneller abfertigen, als sie ankommen, pendeln im Gleichgewichtszustand immer zwischen  $N_t = 1$  und  $N_t = 0$  oder bleiben im besten Fall bei  $N_t = 0$ . Es lassen sich Untersuchungen anstellen, wie schnell die Warteschlange bei einer gegebenen Anzahl von Kunden geleert wird. Dabei ist man am Zeitraum vor Erreichen des Gleichgewichtszustandes interessiert. Es wird also eine zeitabhängige, sogenannte *transiente* Analyse durchgeführt.

Wie bereits angesprochen, lassen sich deterministische Wartesysteme graphisch analysieren. Dies trifft auch bei erschwerten Annahmen zu, wie z.B. ungeduldige Kunden oder die Wahl einer anderen Warteschlangendisziplin als FCFS zu. Es können immer exakte Lösungen gefunden werden, da alle Einflußfaktoren genau bestimmt sind. Bei Systemen mit Verteilungsannahmen ist dies nicht der Fall.

## 2. Geburten/Todesmodelle

In diesem Kapitel werden einige Modelle vorgestellt, deren Leistungsgrößen mit Hilfe der Wahrscheinlichkeitsrechnung ermittelt werden. Den Anfang bildet das M/M/1-System, ein Modell, welches trotz restriktiver Annahmen eines der wichtigsten ist. Der Däne A.K. Erlang hatte herausgefunden, daß die Dauer eines Telefonsats annähernd exponentialverteilt und daß die Anzahl der Anrufer annähernd poissonverteilt ist.

### 2.1 M/M/1-System

Zur graphischen Beschreibung eines M/M/1-Systems kann das Zustandsübergangsdiagramm aus Abb. 2.1 herangezogen werden.

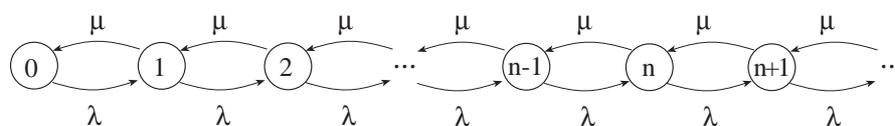


Abb 2. 1 Zustandsübergangsdiagramm eines M/M/1-Systems

Jeder Zustand stellt eine andere Anzahl von Kunden im System dar. So beschreibt Zustand 0 ein leeres System, Zustand 1 ein System mit einem Kunden usw. Die Zustandsübergänge bestehen aus der Ankunftsrate und der Abgangsrate. Mit Rate  $\lambda$  geht ein Zustand  $n$  in den Zustand  $n+1$  über und mit Rate  $\mu$  wechselt er wieder zurück. Zustandsübergänge gibt es nur zwischen benachbarten Zuständen, d.h. es gibt weder Gruppenankünfte noch Gruppenabfertigungen. Die Abgangsrate bleibt wie die Ankunftsrate konstant, es gibt also keine lastabhängiges Verhalten. Aus der Bezeichnung M/M/1 geht hervor, daß die Ankünfte poisson- und die Bedienzeiten exponentialverteilt sind. Die Warteschlange ist unendlich und damit auch die Anzahl der Zustände.

#### 2.1.1 Entwicklung des M/M/1-Modells

Zunächst denkt man sich die Zeitachse in viele kleine Intervalle der Länge  $\Delta t$  unterteilt. Nun besteht die Möglichkeit, daß ein Kunde im Intervall  $[t, t + \Delta t]$  ankommt oder auch nicht. Dabei ist ausgeschlossen, daß mehrere Kunden innerhalb dieses Intervalls eintreffen. Nun wird die Wahrscheinlichkeit, daß ein Kunde im Intervall  $[t, t + \Delta t]$  das System betritt proportional zu der Länge des Intervalls  $\Delta t$  gewählt. Den Proportionalitätsfaktor bildet die mittlere Ankunftsrate  $\lambda$ . Also gilt

$$\begin{aligned} Pr\{\text{genau 1 Ankomst in } [t, t + \Delta t]\} &= \lambda \Delta t \\ Pr\{\text{keine Ankomst in } [t, t + \Delta t]\} &= 1 - \lambda \Delta t \\ Pr\{\text{mehr als 1 Ankomst in } [t, t + \Delta t]\} &= 0 \end{aligned}$$

[2.1]

#### Bemerkung 2.1

Würde man mehrere Ankünfte innerhalb des Intervalls  $[t, t + \Delta t]$  erlauben, so müßte eine Funktion  $o(\Delta t)$  eingeführt werden. Da die Aussagekraft dieses Ausdrucks relativ klein ist und er im Zuge der nachfolgenden Grenzwertbildung verschwinden würde, wurde auf dessen Einführung in [2.1] verzichtet. Eine Herleitung, in der die Funktion  $o(\Delta t)$  berücksichtigt wird, befindet sich in [GROS74] Kapitel 2.1.

Eine ähnliche Annahme kann für den Bedienprozeß getroffen werden:

$$\begin{aligned} Pr\{\text{genau 1 Abgang in } [t, t + \Delta t]\} &= \mu\Delta t \\ Pr\{\text{kein Abgang in } [t, t + \Delta t]\} &= 1 - \mu\Delta t \\ Pr\{\text{mehr als 1 Abgang in } [t, t + \Delta t]\} &= 0 \end{aligned}$$

[2.2]

Bezeichnet  $p_n(t + \Delta t)$  die Wahrscheinlichkeit, daß sich  $n$  Kunden zum Zeitpunkt  $(t + \Delta t)$  im System befinden, so läßt sich diese mit Hilfe der in der Vorperiode  $t$  bestimmten Wahrscheinlichkeiten wie folgt ermitteln:

$$\begin{aligned} p_n(t + \Delta t) &= p_n(t) Pr\{\text{kein Abgang}\} Pr\{\text{keine Ankunft}\} \\ &\quad + p_{n-1}(t) Pr\{\text{genau 1 Ankunft in } [t, t + \Delta t]\} \\ &\quad + p_{n+1}(t) Pr\{\text{genau 1 Abgang in } [t, t + \Delta t]\} \end{aligned}$$

[2.3]

Lediglich für den Zustand 0 gilt etwas anderes:

$$\begin{aligned} p_0(t + \Delta t) &= p_0(t) Pr\{\text{keine Ankunft in } [t, t + \Delta t]\} \\ &\quad + p_1(t) Pr\{\text{genau 1 Abgang in } [t, t + \Delta t]\} \end{aligned}$$

[2.4]

Obige Gleichungen lassen sich so ausdrücken

$$p_n(t + \Delta t) = p_n(t)(1 - \lambda\Delta t)(1 - \mu\Delta t) + p_{n-1}(t)(\lambda\Delta t) + p_{n+1}(t)(\mu\Delta t)$$

[2.5]

Werden diese Ausdrücke umgeformt und läßt man  $\Delta t \rightarrow 0$  gehen, so erhält man

$$\begin{aligned} \frac{d p_n(t)}{d t} &= -(\lambda + \mu)p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t) \\ \frac{d p_0(t)}{d t} &= -\lambda p_0(t) + \mu p_1(t) \end{aligned}$$

[2.6]

Nun ist der Zeitabschnitt  $\Delta t$  so klein, daß nie mehr als 1 Ankunft bzw. 1 Abgang pro Zeitabschnitt erfolgen kann, außer es handelt sich um Gruppenankünfte. Diese wurden jedoch ausgeschlossen.

### Bemerkung 2.2

Das Gleichungssystem [2.6] kann auch mit Hilfe der Theorie der Markovketten (siehe Anhang) bestimmt werden. Den Kern bildet dabei folgendes Gleichungssystem

$$\mathbf{p}'(t) = \mathbf{p}(t)\mathbf{Q}$$

wobei  $\mathbf{p}'(t)$  den Vektor der einzelnen Zustandswahrscheinlichkeiten und  $\mathbf{Q}$  die Zustandsübergangsmatrix darstellt. Ein solcher Ansatz für Geburten/Todesmodelle wird in [BOLC89] Beispiel 3.1

vorgeführt. Ein Geburten/Todesmodell läßt im Gegensatz zum M/M/1-Modell zustandsabhängige Ankunfts- und Bedienraten zu. Das M/M/1-System ist also ein spezieller Geburten/Todesprozeß.

Das Gleichungssystem [2.6] ist auch Ausgangspunkt für eine *transiente* (=zeitabhängige) Analyse des M/M/1-Systems. Wir interessieren uns jedoch mehr für die *stationären* Eigenschaften. Unter der Annahme, das System habe sich eingependelt, läßt sich das Gleichungssystem so schreiben:

$$\begin{aligned} 0 &= -(\lambda + \mu)p_n + \lambda p_{n-1} + \mu p_{n+1} \\ 0 &= -\lambda p_0 + \mu p_1 \end{aligned} \quad [2.7]$$

Es entfällt also der Zeitindex und die Änderung mit der Zeit  $\frac{dp_n(t)}{dt} = 0$ .

### Bemerkung 2.3

Wenn man von vornherein annimmt, daß sich das betrachtete System im Gleichgewicht befindet, so kann das Gleichungssystem [2.7] direkt aus dem Zustandsübergangdiagramm in Abb. 2.1 abgelesen werden. Dabei werden die gesamten Zugänge eines Zustandes den gesamten Abgängen des selben Zustandes gleichgesetzt.

Mit Hilfe der Regel IN=OUT ergibt sich:

$$\begin{aligned} (\lambda + \mu)p_n &= \lambda p_{n-1} + \mu p_{n+1} \\ \lambda p_0 &= \mu p_1 \end{aligned} \quad [2.7a]$$

Dieser Ansatz wird in der Literatur oft unter den Begriffen *Stochastische Balance* und *Stochastisches Gleichgewicht* geführt. Einer der ersten, die diesen Ansatz verwendeten, war R.B. Cooper. Er beschrieb das stochastische Gleichgewicht in Kapitel 2.3 seines Buches [COOP72]. Heute greifen viele Autoren wie [GROS85] und [ROBE94] auf diese Technik zurück.

Wird Gleichungssystem [2.7] wie folgt umgeformt,

$$\begin{aligned} p_{n+1} &= \frac{\lambda + \mu}{\mu} p_n - \frac{\lambda}{\mu} p_{n-1} \\ p_1 &= \frac{\lambda}{\mu} p_0 \end{aligned} \quad [2.7b]$$

so entsteht ein System von Differenzgleichungen, welches iterativ gelöst werden kann. Mit Hilfe der Definition der *Verkehrsintensität*  $\rho = \frac{\lambda}{\mu}$  ergibt sich für die Wahrscheinlichkeiten  $p_n$  für  $n=1,2,\dots$

$$\begin{aligned} p_1 &= \rho p_0 \\ p_2 &= \rho^2 p_0 \\ p_3 &= \rho^3 p_0 \end{aligned} \quad [2.8]$$

$$p_n = \rho^n p_0 \quad [2.9]$$

Diese Lösung läßt sich mit Hilfe der mathematischen Induktion beweisen. Dazu siehe [GROS85] Kapitel 2.2.2.

Schließlich muß nur noch  $p_0$  bestimmt werden. Da die Summe aller Wahrscheinlichkeiten gleich 1 sein muß, gibt es nur eine mögliche Lösung für  $p_0$ . Es gilt

$$1 = \sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \rho^n p_0$$

Daraus folgt unmittelbar

$$p_0 = \frac{1}{\sum_{n=0}^{\infty} \rho^n}$$

Unter Verwendung der Definition einer geometrischen Reihe

$$\sum_{n=0}^{\infty} \rho^n = \frac{1}{1-\rho} \quad \text{für } \rho < 1$$

ergibt sich

$$p_0 = 1 - \rho \quad [2.10]$$

Da die Verkehrsintensität  $\rho$  gleich der Auslastung  $u$  ist und diese für Systeme mit unendlicher Warteschlange kleiner 1 sein muß, stellt die Verwendung der Definition der geometrischen Reihe keine Beschränkung dar. Die stationäre Lösung des M/M/1-Systems lautet also

$$p_n = \rho^n (1 - \rho) \quad \text{für } \rho < 1 \quad [2.11]$$

Anstatt des iterativen Ansatzes kann auch eine wahrscheinlichkeitserzeugende Funktion zur Lösung herangezogen werden. Dazu siehe [GROS85] Kapitel 2.2.2.

Ähnliche Darstellungen sind in [ALLE78] Kapitel 5.2.1., [KLEI74] Kapitel 3.2, [ROBE94] Kapitel 2.2 und [STOC78] Kapitel III.3. Letzterer beschreibt auch den Zusammenhang der Funktion  $o(\Delta t)$  mit der Exponentialverteilung genauer. Einen anderen Weg schlägt [SAAT61] in Kapitel 4.2 ein. Er löst das Differentialgleichungssystem [2.6] mit Hilfe von Besselfunktionen und leitet dann erst die stationären Zustandswahrscheinlichkeiten ab.

### 2.1.2 Leistungsgrößen des M/M/1-Modells

Die mittlere Anzahl von Kunden im System kann wie folgt ermittelt werden.

$$\begin{aligned}
 L &= \mathbb{E}(N) = \sum_{n=0}^{\infty} np_n = (1-\rho) \sum_{n=0}^{\infty} n\rho^n \\
 &= (1-\rho)\rho \sum_{n=1}^{\infty} n\rho^{n-1} \\
 &= (1-\rho)\rho \frac{d\left(\sum_{n=0}^{\infty} \rho^n\right)}{d\rho} \\
 &= (1-\rho)\rho \frac{d(1/(1-\rho))}{d\rho} \\
 &= \frac{(1-\rho)\rho}{(1-\rho)^2} = \frac{\rho}{(1-\rho)} = \frac{\lambda}{\mu - \lambda}
 \end{aligned}$$

[2.12]

Dabei ist anzumerken, daß  $\sum_{n=1}^{\infty} n\rho^{n-1}$  die Ableitung der geometrischen Reihe  $\sum_{n=0}^{\infty} \rho^n$  ist.

Mit Hilfe des Gesetzes von Little kann aus der mittleren Anzahl von Kunden im System die mittlere Antwortzeit ermittelt werden.

$$W = \frac{1}{\lambda} \cdot L = \frac{1}{\lambda} \cdot \frac{\rho}{1-\rho} = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu - \lambda}$$

[2.13]

Da die mittlere Wartezeit die Differenz von mittlerer Antwortzeit und mittlerer Servicezeit ist, ergibt sich

$$W_q = W - \bar{s} = \frac{1}{\mu(1-\rho)} - \frac{1}{\mu} = \frac{\rho}{\mu(1-\rho)} = \frac{\lambda}{\mu(\mu - \lambda)}$$

[2.14]

Die mittlere Wartezeit kann auch ohne Zuhilfenahme des Gesetzes von Little ermittelt werden. Dazu wird die *Wartezeitverteilung*  $W_q(t)$  bestimmt

$$W_q(t) = \begin{cases} 1-\rho & (t=0) \\ 1-\rho e^{-\mu(1-\rho)t} & (t>0) \end{cases}$$

[2.15]

Durch Erwartungswertbildung kann daraus dann die mittlere Wartezeit ermittelt werden. In [GROS85] Kapitel 2.2.4 ist der Entwicklungsverlauf zu finden.

Die mittlere Warteschlangenlänge  $L_q$  kann wieder mit Hilfe des Gesetzes von Little aus  $W_q$  bestimmt werden

$$L_q = \lambda W_q = \lambda \frac{\lambda}{\mu(\mu - \lambda)} = \frac{\rho^2}{1-\rho}$$

[2.16]

Zwischen der mittleren Anzahl von Kunden  $L$  im System und der mittleren Warteschlangenlänge  $L_q$  kann folgender Zusammenhang hergestellt werden

$$\begin{aligned} W &= \frac{1}{\lambda} L \\ W_q + \frac{1}{\mu} &= \frac{1}{\lambda} L \\ \frac{1}{\lambda} L_q + \frac{1}{\mu} &= \frac{1}{\lambda} L \\ L_q + \frac{\lambda}{\mu} &= L \\ L_q + \rho &= L \end{aligned}$$

[2.17]

Da das Gesetz von Little keine Verteilungsannahmen voraussetzt, gilt Gesetz [2.17] für alle Wartesysteme mit einem Bediener. Einzige Bedingung ist, daß das Gesetz von Little für das betrachtete System gilt. Ist der Warteraum beschränkt, muß die Ankunftsrate modifiziert werden (siehe 2.1.3).

### Beispiel 2.1

An einen Datenbankserver wird im Mittel alle 500 ms eine Anfrage gestellt. Da die Anfragen von einer Vielzahl von Benutzern gestellt werden, geht man von einer Poissonverteilung aus. Durch Messungen wurde herausgefunden, daß die Bearbeitung durchschnittlich 450 ms dauert. Die Streuung der Bearbeitungszeit entspricht in etwa der mittleren Bearbeitungszeit, d.h. es kann von einer Exponentialverteilung ausgegangen werden.

Aus  $\bar{t} = 0,5$  und  $\bar{s} = 0,45$  ergibt sich

$$\begin{aligned} \lambda &= 2, & \mu &= 2,2 \\ n = \rho &= \frac{\lambda}{\mu} = 0,9 = 90\% \end{aligned}$$

Damit ist der Datenbankserver zu 90% ausgelastet. Die Zustandswahrscheinlichkeiten können mit Hilfe von Gleichung [2.10] ermittelt werden.

$$\begin{aligned} p_0 &= 0,1 \\ p_1 &= 0,09 \\ p_2 &= 0,081 \\ p_3 &= 0,0729 \end{aligned}$$

Die mittlere Anzahl von Anfragen an das System beträgt

$$L = \frac{0,9}{1 - 0,9} = 9 \text{ Kunden}$$

Daraus folgt

$$W = \frac{1}{\lambda} L = \frac{1}{2} \cdot 9 = 4,5 \text{ sec}$$

Die Benutzer waren frustriert und so entschied man sich, eine schnellere Festplatte einzubauen. In Folge sank die Bearbeitungszeit um 100 ms. Die neuen Werte betragen nun

$$\begin{aligned} \lambda &= 2, & \mu &= 2.857 \\ n = \rho &= 0,7 = 70\% \\ L &= \frac{0,7}{1 - 0,7} = 2,3 \text{ Kunden} \\ W &= \frac{1}{\lambda} L \approx 1,16 \text{ sec} \end{aligned}$$

Dieses Beispiel zeigt, wie empfindlich ein System auf nur kleine Änderungen reagiert. Bei einer Reduktion der Auslastung um 20% kann die Wartezeit auf fast  $\frac{1}{4}$  der ursprünglichen Wartezeit verringert werden.

### 2.1.3 Warteraumbeschränkung

Wird die Warteraumkapazität eines M/M/1-Systems auf (K-1) festgelegt, so entsteht das M/M/1/K-System. Inklusive Zustand 0 gibt es nur noch K+1 Zustand, siehe dazu auch Abb 2.2

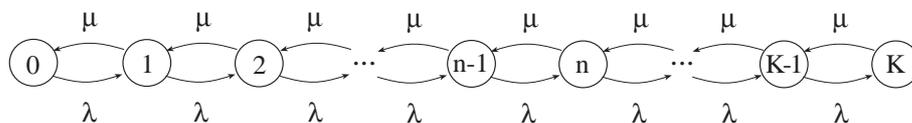


Abb 2.2 Zustandsübergangsdiagramm eines M/M/1/K-Modells

Da die Übergangsraten weiter konstant bleiben, gilt Gleichung [2.9]

$$p_n = \rho^n p_0$$

Lediglich  $p_0$  muß neu bestimmt werden

$$1 = \sum_{n=0}^K p_n = \sum_{n=0}^K \rho^n p_0$$

$$p_0 = \frac{1}{\sum_{n=0}^K \rho^n}$$

Dank der Definition der endlichen geometrischen Reihe erhalten wir

$$p_0 = \begin{cases} \frac{1-\rho}{1-\rho^{K+1}} & \text{für } \rho \neq 1 \\ \frac{1}{K+1} & \text{für } \rho = 1 \end{cases}$$

[2.18]

Eine Beschränkung der Auslastung und damit von  $\rho$  ist nicht nötig, da der begrenzte Warteraum ein Explodieren der Warteschlange verhindert. Mit Hilfe von Gleichung [2.9] und Gleichung [2.18] ergibt sich

$$p_n = \begin{cases} \frac{(1-\rho)\rho^n}{1-\rho^{K+1}} & \text{für } \rho \neq 1 \\ \frac{1}{K+1} & \text{für } \rho = 1 \end{cases}$$

[2.19]

Für die Ermittlung der mittleren Anzahl von Kunden im System wird genauso vorgegangen wie im Fall einer unbeschränkten Warteschlange

$$\begin{aligned} L &= \sum_{n=0}^K n p_n = p_0 \rho \sum_{n=0}^K n \rho^{n-1} \\ &= p_0 \rho \frac{d\left(\sum_{n=0}^K \rho^n\right)}{d\rho} = p_0 \rho \frac{d\left(\frac{1-\rho^{K+1}}{1-\rho}\right)}{d\rho} \\ &= p_0 \rho \frac{1 - (K+1)\rho^K + K\rho^{K+1}}{(1-\rho)^2} \\ L &= \frac{\rho[1 - (K+1)\rho^K + K\rho^{K+1}]}{(1-\rho^{K+1})(1-\rho)} && \text{für } \rho \neq 1 \\ L &= \frac{\sum_{n=0}^K n}{K+1} = \frac{K}{2} && \text{für } \rho = 1 \end{aligned}$$

[2.20]

Mit Hilfe des Gesetzes von Little erhält man wie gewohnt die mittlere Antwortzeit des Systems. Es ist lediglich die Ankunftsrate zu korrigieren

$$\lambda' = \lambda(1 - p_k) \quad [2.21]$$

Es gilt also

$$W = \frac{1}{\lambda'} L \quad [2.22]$$

Dabei bezeichnet  $\lambda'$  die *effektive Ankunftsrate*. Sie beschreibt die Ankünfte der Kunden, die das System auch betreten. Die mittlere Wartezeit erhält man wie gewohnt

$$W_q = W - \frac{1}{\mu} \quad [2.23]$$

Daraus folgt die mittlere Warteschlangenlänge

$$L_q = \lambda' W_q \quad [2.24]$$

Der mit Gleichung [2.17] vorgestellte Zusammenhang sieht auch etwas anders aus

$$\begin{aligned} L_q + \frac{\lambda'}{\mu} &= L \\ L_q + \frac{\lambda(1 - p_k)}{\mu} &= L \\ L_q + \rho \left( 1 - \frac{(1 - \rho)\rho^K}{1 - \rho^{K+1}} \right) &= L \\ L_q &= L - \rho \left( \frac{1 - \rho^K}{1 - \rho^{K+1}} \right) \end{aligned} \quad [2.25]$$

Beim Vergleich von Gleichung [2.17] und [2.10] bzw. [2.25] und [2.18] kann folgender allgemeiner Zusammenhang abgelesen werden

$$L_q = L - (1 - p_0) \quad [2.26]$$

Obiger Ausdruck gilt für alle Warteschlangensysteme mit einem Bediener. Ausgenommen sind Systeme mit Gruppenankünften und Gruppenabfertigungen.

Ähnliche Herleitungen sind in [GROS74] Kapitel 2.5 und [ROBE94] Kapitel 2.6 zu finden. Besonders interessant im Zusammenhang mit der Optimierung der Warteraumkapazität ist [MORS58]. Dort wird ein Modell vorgestellt, bei dem der Profit aus der Bedienung von Kunden dem Verlust von Kunden gegenüber gestellt wird. Auch für Systeme mit unbeschränkter Warteschlange wird dort ein Modell vorgestellt. Dieses wiegt die Bedienkosten gegen die mittlere Wartezeit ab.

### Beispiel 2.2

Betrachtet wird ein Teleshopping-System mit einem Bearbeiter. Im Mittel kommen 5 Anrufe pro Minute an und abgefertigt werden 6 Anrufe pro Minute. Sobald der Bearbeiter mit einer Bestellung beschäftigt ist, werden die Anrufer darauf hingewiesen und in eine Warteschlange eingereiht. Wenn mehr als 4 Kunden warten, werden Neuanrufe abgewiesen. Dieses Beispiel entspricht im Wesentlichen dem von [MINO93] Kapitel 5.1.2.

Mit Hilfe folgender Angaben

$$\lambda = 5 \frac{\text{Kunden}}{\text{min}}; \quad \mu = 6 \frac{\text{Kunden}}{\text{min}}; \quad K = 5 \text{ Kunden}$$

$$\rho = \frac{5}{6} = 0,8\bar{3}$$

läßt sich die mittlere Anzahl der Kunden im System errechnen (Gleichung [2.20]):

$$L = \frac{0,8\bar{3} [1 - 6 \cdot 0,8\bar{3}^5 + 5 \cdot 0,8\bar{3}^6]}{(1 - 0,8\bar{3}^6)(1 - 0,8\bar{3})} \approx 1,98 \text{ Kunden}$$

Die Wahrscheinlichkeit, daß ein Kunde abgewiesen wird, beträgt

$$p_5 = \frac{0,8\bar{3}^5 (1 - 0,8\bar{3})}{1 - (0,8\bar{3}^6)} \approx 0,1 = 10\%$$

Damit erhält man die effektive Ankunftsrate

$$\lambda' = 5(1 - 0,1) = 4,5$$

Mit dem Gesetz von Little ergibt das für die mittlere Antwortzeit

$$W = 1,98 \cdot \frac{1}{4,5} = 0,44 \text{ min}$$

Daraus folgt

$$W_q = 0,44 - \frac{1}{6} \approx 0,273 \text{ min}$$

$$L_q = 0,273 \cdot 4,5 \approx 1,23 \text{ Kunden}$$

In diesem System gehen im Mittel  $\lambda p_5$  Kunden pro Minute verloren, d.h. alle 2 Minuten wird ein Kunde abgewiesen.

### 2.1.4 Andere Warteschlangendisziplinen

Die Änderung der Warteschlangendisziplin hat auf  $p_n, L, L_q, W$  und  $W_q$  keine Auswirkungen. Für die Zustandswahrscheinlichkeiten ist die Reihenfolge der abgehenden Kunden unerheblich und bei den restlichen Größen handelt es sich um Mittelwerte. Es besteht aber sehr wohl ein Einfluß auf die Dichte- und Verteilungsfunktion von  $W$  bzw.  $W_q$ . Der Ausdruck [2.15] gilt also nur für Wartesysteme mit Disziplin FCFS. In [COOP72] Kapitel 6.3 werden die Disziplinen FCFS, LCFS und RSS für M/M/m-Systeme behandelt.

## 2.2 Lastabhängige Systeme

Sind Ankunfts- und/oder Bedienraten vom Systemzustand abhängig, spricht man von einem *lastabhängigen System*. Die Warteschlange ist unbeschränkt, Gruppenankünfte und -abfertigungen bleiben weiterhin ausgeschlossen. Damit ergibt sich folgender Zustandsgraph:

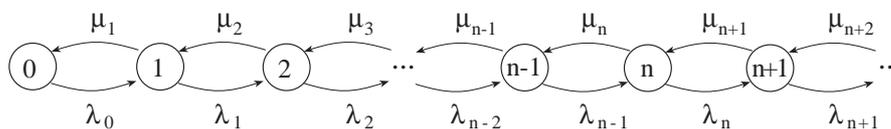


Abb 2.3 Zustandsübergangsdiagramm eines lastabhängigen Systems

Nun wird ein iterativer Ansatz gemacht

$$\begin{aligned}
 \lambda_0 p_0 &= \mu_1 p_1 & p_1 &= \frac{\lambda_0}{\mu_1} p_0 \\
 \lambda_1 p_1 &= \mu_2 p_2 & p_2 &= \frac{\lambda_1}{\mu_2} p_1 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} p_0 \\
 & \cdot & & \\
 \lambda_{n-1} p_{n-1} &= \mu_n p_n & p_n &= \frac{\lambda_{n-1}}{\mu_n} p_{n-1} = \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i} p_0
 \end{aligned}$$

[2.27]

Schließlich wird  $p_0$  wie gewohnt ermittelt

$$\begin{aligned}
 1 &= \sum_{n=0}^{\infty} p_n = p_0 + \sum_{n=1}^{\infty} p_n = p_0 + \sum_{n=1}^{\infty} \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i} p_0 \\
 p_0 &= \frac{1}{1 + \sum_{n=1}^{\infty} \prod_{i=1}^n \frac{\lambda_{i-1}}{\mu_i}}
 \end{aligned}$$

[2.28]

Obiges Modell wird auch oft als *Geburten/Todesmodell* bezeichnet. Werden Ankunfts- und Bedienraten konstant gewählt, so ergibt sich das M/M/1-Modell aus dem vorigen Abschnitt. Um Leistungsgrößen zu erhal-

ten, geht man von der mittleren Anzahl der Kunden im System aus

$$L = \sum_{n=1}^{\infty} n p_n \quad [2.29]$$

Wir definieren eine weitere Größe, den *Durchsatz*

$$X = \sum_{n=1}^{\infty} \mu_n p_n \quad [2.30]$$

und ermitteln mit Hilfe des Gesetzes von Little die mittlere Antwortzeit

$$W = L \cdot \frac{1}{X} = \frac{\sum_{n=1}^{\infty} n p_n}{\sum_{n=1}^{\infty} \mu_n p_n} \quad [2.31]$$

Werden Systeme mit beschränktem Warteraum untersucht, so muß lediglich die obere Grenze des Summationsoperators in den Gleichungen [2.28] bis [2.31] durch die Systemkapazität  $K$  ersetzt werden. Aus dem Geburten/Todesmodell lassen sich eine Reihe anderer Modelle ableiten. Einige davon werden nachfolgend behandelt.

### 2.2.1 Das M/M/m-System

Wie bereits in der Einführung (siehe [1.7]) angesprochen, läßt sich das System mit mehreren parallelen Bedienern als spezielles lastabhängiges System darstellen:

$$\lambda_n = \lambda \quad \text{für } n = 0, 1, 2, \dots$$

$$\mu_n = \begin{cases} n\mu & \text{für } 0 \leq n \leq m \\ m\mu & \text{für } n > m \end{cases}$$

Eingesetzt in Gleichung [2.27] und [2.28] ergibt sich

$$p_n = \begin{cases} \prod_{i=1}^n \frac{\lambda}{i\mu} p_0 = \frac{1}{n!} \rho^n p_0 & \text{für } 0 \leq n \leq m \\ p_m \prod_{i=m+1}^n \frac{\lambda}{m\mu} = \frac{1}{m! m^{n-m}} \rho^n p_0 & \text{für } n > m \end{cases} \quad [2.32]$$

$$p_0 = 1 + \sum_{n=1}^m \frac{1}{n!} \rho^n + \sum_{n=m+1}^{\infty} \frac{1}{m! m^{n-m}} \rho^n$$

$$= \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \sum_{n=m}^{\infty} \frac{1}{m! m^{n-m}} \rho^n \right)^{-1}$$

$$\begin{aligned}
&= \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m \sum_{n=m}^{\infty} \frac{1}{m^{n-m}} \rho^{n-m} \right)^{-1} \\
&= \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m \sum_{n=0}^{\infty} \frac{1}{m^n} \rho^n \right)^{-1} \\
&= \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m \left( \frac{1}{1 - \frac{\rho}{m}} \right) \right)^{-1}
\end{aligned}$$

[2.33]

Um eine Ermittlung der Leistungsgrößen zu vereinfachen, beginnt man mit  $L_q$ , da hier im Gegensatz zu  $L$  nur ein Teil der Formel [2.32] berücksichtigt werden muß

$$\begin{aligned}
L_q &= \sum_{n=m}^{\infty} (n-m) p_n = \sum_{n=m}^{\infty} (n-m) \frac{\rho^n}{m! m^{n-m}} p_0 \\
&= \frac{\rho^m p_0}{m!} \sum_{n=m}^{\infty} (n-m) \left( \frac{\rho}{m} \right)^{n-m} = \frac{\rho^m p_0}{m!} \sum_{i=1}^{\infty} i \left( \frac{\rho}{m} \right)^i \\
&= \frac{\rho^m p_0}{m!} \left( \frac{\rho}{m} \right) \sum_{i=1}^{\infty} i \left( \frac{\rho}{m} \right)^{i-1} = \frac{\rho^m p_0}{m!} \left( \frac{\rho}{m} \right) \frac{d}{d \left( \frac{\rho}{m} \right)} \left( \sum_{i=1}^{\infty} \left( \frac{\rho}{m} \right)^i \right) \\
&= \frac{\rho^m p_0}{m!} \left( \frac{\rho}{m} \right) p_0 \frac{d}{d \left( \frac{\rho}{m} \right)} \left( \frac{1}{1 - \frac{\rho}{m}} \right) = \frac{\frac{\rho^{m+1}}{m}}{m! \left( 1 - \frac{\rho}{m} \right)^2} p_0
\end{aligned}$$

[2.34]

Mit Hilfe des Gesetzes von Little und Gleichung [2.23] werden nun die restlichen Leistungsgrößen ermittelt

$$W_q = L_q \cdot \frac{1}{\lambda} = \frac{\frac{\rho^{m+1}}{m}}{\lambda m! \left( 1 - \frac{\rho}{m} \right)^2} p_0$$

[2.35]

$$W = W_q + \frac{1}{\mu} = \frac{\frac{\rho^{m+1}}{m}}{\lambda m! \left( 1 - \frac{\rho}{m} \right)^2} p_0 + \frac{1}{\mu}$$

[2.36]

$$L = \lambda W = \frac{\frac{\rho^{m+1}}{m}}{m! \left( 1 - \frac{\rho}{m} \right)^2} p_0 + \rho$$

[2.37]

Im folgenden Teil soll kurz eine Formel vorgestellt werden, die als *Erlang'sche Formel der 2. Art* (Erlang's C formula) bezeichnet wird. Sie gibt die Wahrscheinlichkeit an, daß Kunden warten müssen.

$$Pr\{\text{Stau}\} = Pr\{m \text{ oder mehr Kunden im System}\} \quad [2.38]$$

Auf die Substitution von  $p_0$  soll verzichtet werden, da dies nicht besonders zur Lesbarkeit beiträgt.

### Beispiel 2.3

Betrachtet wird ein Netzwerk mit 4 Druckern. Diese sind über einen Printerserver miteinander verbunden. Im Mittel kommt alle 10 Minuten ein Druckauftrag an. Ein Ausdruck dauert durchschnittlich 20 Minuten, da es sich um sehr aufwendige, graphische Farbausdrucke handelt. Die Verteilungen sind annähernd so beschaffen, daß ein M/M/m-Modell zur Analyse herangezogen werden kann.

$$\text{Es gilt also } m = 4, \mu = 3 \text{ Jobs/h, } \lambda = 6 \text{ Jobs/h, } \rho = 2$$

Die Wahrscheinlichkeit, daß kein Drucker beschäftigt ist, beträgt

$$p_0 = \left( 1 + \frac{1}{1!}2 + \frac{1}{2!}2^2 + \frac{1}{3!}2^3 + \frac{1}{4!}2^4 \left( \frac{1}{1 - \frac{2}{4}} \right) \right)^{-1} \approx 0,13$$

Somit besteht eine Wahrscheinlichkeit von 87%, daß mindestens ein Drucker in Betrieb ist. Die mittlere Anzahl von Kunden im System und die mittlere Gesamtbearbeitungszeit betragen

$$L = \frac{2^5}{4! \frac{1}{2}} + 2 \approx 2,67 \text{ Jobs}$$

$$W = \frac{1}{6} \cdot 2,67 \approx 0,45 \text{h} = 27 \text{ Minuten}$$

Daraus lassen sich leicht die mittlere Wartezeit und die mittlere Länge der Warteschlange ablesen

$$W_q = 7 \text{ Minuten} \quad L_q \approx 0,67 \text{ Jobs}$$

### 2.2.2 Das M/M/m/K-System

Wie bereits angesprochen, lassen sich warteraumbeschränkte Systeme auch mit den Gleichungen für ein lastabhängiges System analysieren. Da aber nun eine endliche Zahl von Zuständen vorliegt, muß sich diese Grenze auch in den Formeln [2.28] bis [2.31] niederschlagen. Daraus geht aber auch hervor, daß für die Bestimmung der Zustandswahrscheinlichkeiten nur  $p_0$  neu berechnet werden muß. Gleichung [2.32] bleibt weiterhin gültig. Damit ergibt sich für  $p_0$

$$\begin{aligned}
 p_0 &= \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \sum_{n=m}^K \frac{1}{m! m^{n-m}} \rho^n \right)^{-1} \\
 &= \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m \sum_{n=0}^{\infty} \frac{1}{m^n} \rho^n \right)^{-1} \\
 &= \begin{cases} \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m \frac{1 - \left(\frac{\rho}{m}\right)^{K-m+1}}{1 - \frac{\rho}{m}} \right)^{-1} & \text{für } \frac{\rho}{m} \neq 1 \\ \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m (K - m + 1) \right)^{-1} & \text{für } \frac{\rho}{m} = 1 \end{cases}
 \end{aligned}$$

[2.39]

Völlig analog zu Gleichung [2.34] erhält man für  $L_q$ :

$$\begin{aligned}
 L_q &= \sum_{n=m}^K (n - m) p_n = \dots = \frac{\rho^m \left(\frac{\rho}{m}\right)}{m!} p_0 \frac{d}{d\left(\frac{\rho}{m}\right)} \left[ \frac{1 - \left(\frac{\rho}{m}\right)^{K-m+1}}{1 - \frac{\rho}{m}} \right] \\
 &= \frac{\rho^m \left(\frac{\rho}{m}\right)}{m! \left(1 - \frac{\rho}{m}\right)^2} p_0 \left[ 1 - \left(\frac{\rho}{m}\right)^{K-m+1} - \left(1 - \frac{\rho}{m}\right) (K - m + 1) \left(\frac{\rho}{m}\right)^{K-m} \right]
 \end{aligned}$$

[2.40]

Die restlichen Leistungsgrößen können mit Hilfe der Gleichungen [2.22] bis [2.24] ermittelt werden:

$$\begin{aligned}
 W_q &= \frac{1}{\lambda'} L_q & \text{mit } \lambda' &= \lambda(1 - p_K) \\
 W &= W_q + \frac{1}{\mu} \\
 L &= \lambda' W & \text{mit } \lambda' &= \lambda(1 - p_K)
 \end{aligned}$$

Ein Spezialfall des M/M/m/K-Modells ist das *Verlustsystem*. Dieses aus der Telephonie stammende System wird zur Analyse von Leitungsbündeln verwendet. Man geht davon aus, daß  $m$  Leitungen zwischen zwei Kno-

ten zur Verfügung stehen. Das heißt jedoch auch, daß die Kapazität  $K$  auf  $m$  Kunden (Anrufer) beschränkt ist. Es wird also  $K=m$  gewählt und in die Gleichungen [2.32] und [2.39] eingesetzt.

Es ergibt sich

$$p_n = \frac{1}{n!} \rho^n p_0$$

$$p_0 = \left( 1 + \sum_{n=1}^{m-1} \frac{1}{n!} \rho^n + \frac{1}{m!} \rho^m \right)^{-1} = \sum_{n=0}^m \frac{1}{n!} \rho^n$$

Durch Einsetzen erhält man die *Erlang'sche Formel der 1. Art*

$$p_n = \frac{\frac{\rho^n}{n!}}{\sum_{i=0}^m \frac{\rho^i}{i!}} \quad \text{für } 0 \leq n \leq m$$

[2.41]

Eine besonders interessante Größe ist die Wahrscheinlichkeit, daß das System voll ist.

$$p_K = p_m = \frac{\frac{\rho^m}{m!}}{\sum_{i=0}^m \frac{\rho^i}{i!}}$$

[2.42]

Dieser Ausdruck wird als *Erlang'sche Verlustformel* bezeichnet. Eine Besonderheit dieser Formel ist ihre Unabhängigkeit von der Bedienzeitverteilung. Gleichung [2.42] gilt also für  $M/G/m/m$ -Systeme genauso wie für  $M/M/m/m$ -Systeme. Der Beweis kann in [GROS85] Kapitel 5.2.2 nachgelesen werden. Er wird dort im Zuge der  $M/G/m$ -Systeme behandelt. Weitere derartige Modelle der Telephonie werden in [COOP72] Kapitel 3 dargestellt.

Beschreibungen von  $M/M/m/K$ -Systemen sind in der einschlägigen Literatur wie [GROS85] Kapitel 2.4 und [ALLE78] Kapitel 5.2.4 zu finden. Mit Optimierungsproblemen beschäftigt sich [MORS58] in Kapitel 4.

### 2.2.3 Ein endloser Bediener (Infinite Server)

Oft stellt sich die Frage nach dem Nutzen des  $M/M/\infty$ -Modells. Es erweist sich jedoch besonders in der Betrachtung von Warteschlangennetzen als brauchbar. So dient es einerseits als Verzögerungsknoten, andererseits der Modellierung von Terminals. Im täglichen Alltag wäre ein anderes Beispiel denkbar: Ein Besuch im Supermarkt kann als Netz, bestehend aus einem  $M/M/\infty$ -Modell und einem  $M/M/m/K$ -System, beschrieben werden. Ersteres modelliert das Zusammenstellen der Waren, zweiteres das Anstellen bei der Kassa. Dabei wird angenommen, daß  $m$  Kassen geöffnet sind und der Warteraum begrenzt ist.

Wir betrachten den endlosen Bediener als Spezialfall des Erlang'schen Verlustsystems M/M/m/m und lassen  $m$  gegen  $\infty$  gehen. Aus Gleichung [2.41] folgt

$$p_n = \frac{\frac{\rho^n}{n!}}{\sum_{i=0}^{\infty} \frac{\rho^i}{i!}}$$

Mit Hilfe der Reihendarstellung

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

folgern wir

$$p_n = \frac{\frac{\rho^n}{n!}}{e^\rho} = \frac{1}{n!} \rho^n e^{-\rho} \quad \text{für } n \geq 0$$

[2.43]

Natürlich gibt es noch weitere Ansätze für dieses Modell. So beschreibt [ROBE94] in Kapitel 2.7 die Herleitung mit Hilfe des Geburten/Todesmodells. Auch [GROS85] Kapitel 2.6 folgt diesem Ansatz. Die Leistungsgrößen können ermittelt werden, indem man Grenzwerte für die Leistungsgrößen eines M/M/m/m-Modells eruiert. Ergebnisse können aber auch durch einfaches Überlegen gefunden werden. Da sich keine Warteschlange bilden kann, d.h.  $L_q = 0$  und  $W_q = 0$ , ist die mittlere Antwortzeit gleich der mittleren Bedienzeit.

$$W = \bar{s} = \frac{1}{\mu}$$

[2.44]

Mit Hilfe des Gesetzes von Little folgt weiters

$$L = \lambda W = \frac{\lambda}{\mu} = \rho$$

[2.45]

### 2.2.4 Ungeduldige Kunden

In diesem Abschnitt sollen einige Arten von Ungeduld erläutert werden. Diese können unter Zuhilfenahme des lastabhängigen Bedieners modelliert werden. Zuerst werden *gehemmte Kunden* betrachtet. Diese sind mit steigender Warteschlangenlänge immer weniger bereit, das System zu betreten. Dazu wird die sogenannte *Hemmfunktion*  $b_n$ , eine monoton fallende Funktion eingeführt. Es gilt also

$$0 \leq \dots \leq b_{n+1} \leq b_n \leq \dots \leq b_1 \leq b_0 = 1$$

[2.46]

Für ein M/M/1-System mit gehemmten Kunden ergeben sich damit folgende Übergangsraten

$$\begin{aligned}\lambda_n &= \lambda b_n \\ \mu_n &= \mu\end{aligned}$$

Eingesetzt in die Gleichungen [2.27] und [2.28] folgt

$$p_n = \prod_{i=1}^n \frac{\lambda b_{i-1}}{\mu} p_0 = \rho^n p_0 \prod_{i=1}^n b_{i-1} \quad [2.47]$$

$$p_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \rho^n \prod_{i=1}^n b_{i-1}} \quad [2.48]$$

Dieses Modell ist eine Verallgemeinerung des M/M/1/K-Modells. Die Hemmfunktion lautet

$$b_n = \begin{cases} 1 & \text{für } 0 \leq n \leq K-1 \\ 0 & \text{für } n > K-1 \end{cases}$$

Weitere geeignete Hemmfunktionen sind

$$b_n = \frac{1}{n+1}$$

und

$$b_n = e^{-\frac{\alpha n}{\mu}}$$

Erstere wird in [KLEI75] Kapitel 3.3 vorgestellt. Es wird ein Modell untersucht, dessen Ankunftsprozeß durch gehemmte Kunden charakterisiert ist. Die zweite Hemmfunktion eignet sich besonders für Kunden, die vor Eintritt in das System die mittlere Gesamtwaitzeit mit  $\frac{n}{\mu} = n\bar{s}$  schätzen. Dafür benötigen sie allerdings eine Schätzung der mittleren Bedienzeit  $\bar{s}$ .

Eine weitere Form von Ungeduld ist der *Verzicht auf Bedienung*. Ein Kunde betritt das System, schätzt die mittlere Gesamtwaitzeit und verläßt gegebenenfalls das System. Dies hat eine modifizierte Abgangsrate zur Folge.

Die Verzichtsfunktion  $r(n)$  ist gegeben durch

$$r(n) = \begin{cases} 0 \\ \lim_{\Delta t \rightarrow 0} \left[ \frac{Pr \left\{ \begin{array}{l} \text{Der Kunde verzichtet auf Bedienung und verläßt das} \\ \text{System innerhalb } \Delta t, \text{ wenn } n \text{ Kunden im System sind} \end{array} \right\}}{\Delta t} \right] \end{cases} \quad [2.49]$$

Für ein M/M/1-System mit möglichem Verzicht auf Bedienung ergeben sich folgende Übergangsraten

$$\begin{aligned} \lambda_n &= \lambda \\ \mu_n &= \mu + r(n) \end{aligned}$$

Dies wird wieder in die Gleichungen [2.27] und [2.28] eingesetzt

$$p_n = \prod_{i=0}^{n-1} \frac{\lambda}{\mu + r(i)} p_0 = \lambda^n p_0 \prod_{i=1}^n \frac{1}{\mu + r(i)} \quad [2.50]$$

$$p_0 = \frac{1}{1 + \sum_{n=1}^{\infty} \lambda^n \prod_{i=1}^n \frac{1}{\mu + r(i)}} \quad [2.51]$$

Wiederum stellt

$$r(n) = e^{-\frac{\alpha n}{\mu}}$$

eine geeignete Verzichtsfunktion dar.

Natürlich lassen sich die beiden Arten von Ungeduld kombinieren, dazu siehe [GROS85] Kapitel 2.9.2. Die gezeigten Modelle können leicht auf mehr Bediener ausgedehnt werden, wobei die Bedienrate entsprechend zu modifizieren ist. Siehe dazu [RIO62] Kapitel 5.6. [MORS58] untersucht die Auswirkungen von gehemmten Kunden auf ein Optimierungsproblem in Kapitel 3. In [WAAL87] wird ein M/M/1/K-Modell mit ungeduldigen Kunden vorgestellt. Beim Eintritt in das System legt der Kunde die maximale Verweilzeit fest. Wird diese überschritten, so verläßt er das System ohne erfolgte Bedienung.

### 2.2.5 Das Zentralrechnermodell (Central Server CPU Model)

Das Zentralrechnermodell ist eigentlich ein kleines Warteschlangennetz, dennoch läßt es sich mit Hilfe eines lastabhängigen Systems modellieren.

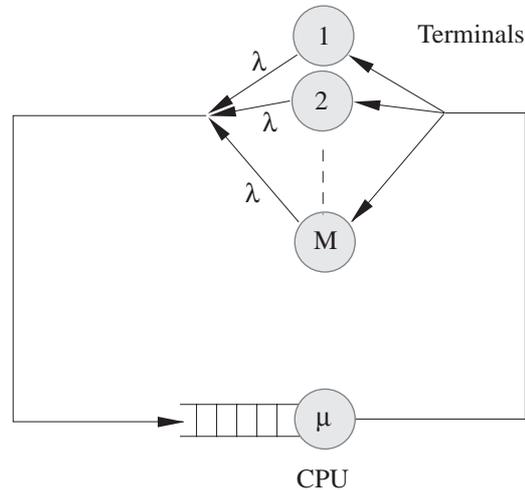


Abb 2.4 Schematische Darstellung eines Zentralrechnermodells

Man geht davon aus, daß jedes Terminal Jobs mit einer Rate  $\lambda$  generiert. Befinden sich nun  $n$  solcher Jobs im CPU-Knoten, so kann man davon ausgehen, daß noch  $M-n$  Terminals aktiv sind. Da es sich um ein geschlossenes Netz handelt, können Jobs das System weder betreten noch verlassen.

Damit ergeben sich folgende Übergangsraten

$$\lambda_n = \begin{cases} \lambda(M-n) & \text{für } 0 \leq n < M \\ 0 & \text{für } n \geq M \end{cases}$$

Daraus folgt

$$p_n = \prod_{i=1}^n \frac{\lambda(M-i+1)}{\mu} p_0 = \rho^n \frac{M!}{(M-n)!} p_0 \quad [2.52]$$

$$p_0 = \frac{1}{1 + \sum_{n=1}^M \rho^n \frac{M!}{(M-n)!}} \quad [2.53]$$

Damit ergibt sich für die mittlere Anzahl von Jobs im CPU-Knoten

$$L = \sum_{n=0}^M n p^n = \sum_{n=0}^M n \rho^n \frac{M!}{(M-n)!} p_0 = M! p_0 \sum_{n=0}^M \frac{n \rho^n}{(M-n)!} \quad [2.54]$$

Bevor das Gesetz von Little Anwendung findet, muß die effektive Ankunftsrate ermittelt werden. Wenn sich im CPU-Knoten  $n$  Jobs aufhalten, so beträgt die Ankunftsrate  $\lambda(M - n)$ . Da sich im Mittel  $L$  Jobs im CPU-Knoten aufhalten, ergibt sich die effektive Ankunftsrate zu

$$\lambda' = \lambda(M - L) \quad [2.55]$$

Nun können die restlichen Leistungsgrößen ermittelt werden

$$W = \frac{1}{\lambda'} L = \frac{L}{\lambda(M - L)} \quad [2.56]$$

$$W_q = W - \frac{1}{\mu} \quad [2.57]$$

$$L_q = \lambda' W_q = \lambda(M - L) W_q \quad [2.58]$$

Wird das Zentralrechnermodell auf  $m$  Bediener verallgemeinert, so entsteht das sogenannte *Machine Repairmen Modell*. Jede Ankunft entspricht dem Defekt einer von insgesamt  $M$  Maschinen. Sofort beginnt ein Servicetechniker mit der Reparatur. Insgesamt sind  $m$  Servicetechniker vorhanden. Die Dauer einer Reparatur beträgt im Mittel  $\bar{s} = \frac{1}{\mu}$ . Es ergeben sich folgende Übergangsraten.

$$\lambda_n = \begin{cases} \lambda(M - n) & \text{für } 0 \leq n < M \\ 0 & \text{für } n \geq M \end{cases}$$

$$\mu_n = \begin{cases} n\mu & \text{für } 0 \leq n < m \\ m\mu & \text{für } n \geq m \end{cases}$$

Daraus folgt

$$p_n = \begin{cases} \frac{M!}{n!(M - n)!} \rho^n p_0 & \text{für } 0 \leq n < m \\ \frac{M!}{(M - n)! m^{n-m} m!} \rho^n p_0 & \text{für } m \leq n \leq M \end{cases} \quad [2.59]$$

$$p_0 = \left[ 1 + \sum_{n=1}^{m-1} \frac{M!}{n!(M - n)!} \rho^n + \sum_{n=m}^M \frac{M!}{(M - n)! m^{n-m} m!} \rho^n \right]^{-1} \quad [2.60]$$

Bemerkenswert ist allerdings, daß die stationären Zustandswahrscheinlichkeiten des Machine Repairmen Modells unabhängig von der Verteilung der Ankünfte sind, es handelt sich somit um ein spezielles G/M/m-System. Lediglich die Zwischenankunftszeiten müssen im Mittel  $\bar{t} = 1/\lambda$  betragen.

Die mittlere Arbeitszeit von arbeitenden Servicetechnikern beträgt

$$L = \sum_{n=0}^M n p_n = p_0 \left[ \sum_{n=0}^{m-1} \frac{M!}{(n-1)!(M-n)!} \rho^n + \frac{M!}{m!} \sum_{n=m}^M \frac{n}{(M-n)! m^{n-m}} \rho^n \right] \quad [2.61]$$

Die restlichen Leistungsgrößen lassen sich dann mit Hilfe der Gleichungen [2.56] bis [2.58] ermitteln.

Dieses Modell kann noch weiter verallgemeinert werden, indem eine bestimmte Anzahl von Ersatzmaschinen, die beim Ausfall einer Maschine sofort deren Arbeit fortsetzen, mit einbezogen werden. Das so erweiterte Modell wird zusammen mit dem normalen Machine Repairmen Modell in [GROS85] Kapitel 2.7 behandelt. Eine sehr genaue Analyse in Kombination mit Warteraumbeschränkungen bietet [KLEI75] ab Kapitel 3.8 an. Eine Abhandlung über beide in diesem Abschnitt vorgestellten Modelle ist in [ALLE78] Kapitel 5.2.6 und 5.2.7 zu finden. Das Phänomen der Störung von Maschinen wird in [COX61] Kapitel IV einführend behandelt. Weitere Modellierungsansätze bieten [TRIV82] ab Kapitel 8.2.3 und [ROBE94] in Kapitel 2.10.

### 2.2.6 Ein lastabhängiges System mit zwei Betriebsgeschwindigkeiten

Das nun folgende Modell geht auf Karl Heinz F. Meyer zurück. Er behandelt es in seiner Arbeit [MEYE71] in Kapitel 2. Er untersucht ein System, das je nach Andrang zwei Arbeitsgeschwindigkeiten  $\bar{s}_1$  und  $\bar{s}_2$  kennt. Ausschlaggebend für den Wechsel zwischen den beiden Geschwindigkeiten ist eine bestimmte Anzahl von Kunden  $k$  im System. Wird  $k$  überschritten, so wechselt die Bedienrate auf  $\mu_2 = 1/\bar{s}_2$ . Fällt die Anzahl der Kunden unter  $k$ , so geht auch die Bedienrate auf  $\mu_1 = 1/\bar{s}_1$  zurück. Es gilt also

$$\lambda_n = \lambda$$

$$\lambda_n = \begin{cases} \mu_1 & \text{für } 1 \leq n \leq k \\ \mu_2 & \text{für } n > k \end{cases}$$

Damit ergeben sich folgende Zustandswahrscheinlichkeiten

$$p_n = \begin{cases} \rho_1^n p_0 & \text{für } 0 \leq n \leq k \\ \rho_1^k \rho_2^{n-k} p_0 & \text{für } n > k \end{cases} \quad [2.62]$$

$$\begin{aligned}
p_0 &= \left[ 1 + \sum_{n=1}^k \rho_1^n + \sum_{n=k+1}^{\infty} \rho_1^k \rho_2^{n-k} \right]^{-1} \\
&= \left[ \sum_{n=0}^k \rho_1^n + \rho_1^k \sum_{n=k+1}^{\infty} \rho_2^{n-k} \right]^{-1} \\
&= \left[ \frac{1 - \rho_1^{k+1}}{1 - \rho_1} + \rho_1^k \frac{\rho_2}{1 - \rho_2} \right]^{-1} \\
&= \frac{(1 - \rho_1)(1 - \rho_2)}{1 - \rho_2 - \rho_1^k(\rho_1 - \rho_2)}
\end{aligned}$$

[2.63]

mit  $\rho_1 = \lambda/\mu_1$  und  $\rho_2 = \lambda/\mu_2$ .  $p_0$  gibt die Wahrscheinlichkeit an, daß sich kein Kunde im System befindet. Meyer berechnet zudem noch die Wahrscheinlichkeiten  $p^{(i)}$  für die einzelnen Geschwindigkeitsstufen  $i=1,2$ . Diese kann man auch als den Zeitanteil beschreiben, den der Bediener benötigt, um die Kunden mit Rate  $\mu_i$  abzufertigen.

$$p^{(1)} = \sum_{n=1}^k p_n = \rho_1 \frac{(1 - \rho_1)^k (1 - \rho_2)}{1 - \rho_2 - \rho_1^k(\rho_1 - \rho_2)}$$

[2.64]

$$p^{(2)} = \sum_{n=k+1}^{\infty} p_n = \rho_2 \frac{\rho_1^k (1 - \rho_1)}{1 - \rho_2 - \rho_1^k(\rho_1 - \rho_2)}$$

[2.65]

Da alle Wahrscheinlichkeiten in Summe 1 ergeben müssen, gilt

$$p_0 + p^{(1)} + p^{(2)} = 1$$

Die mittlere Anzahl von Kunden im System errechnet sich wie folgt

$$\begin{aligned}
L &= \sum_{n=1}^k n p_n + \sum_{n=k+1}^{\infty} n p_n = \dots \\
&= \frac{\rho_1}{1 - \rho_1} - \frac{\rho_1 - \rho_2}{(1 - \rho_1)\rho_2} \left[ k + \frac{1}{1 - \rho_2} \right] p^{(2)}
\end{aligned}$$

[2.66]

Die restlichen Werte können nach folgendem Schema berechnet werden

$$W = \frac{1}{\lambda} L$$

[2.67]

$$L_q = L - (1 - p_0) = L - p^{(1)} - p^{(2)} \quad [2.68]$$

$$W_q = \frac{1}{\lambda} L_q \quad [2.69]$$

Der Zusammenhang in Gleichung [2.68] wurde im Abschnitt 2.1.3, Gleichung [2.26] behandelt. Meyer erweiterte das Modell in Kapitel 4 auch auf mehrere parallele Bedienkanäle. In Kapitel 5.1 beschäftigt er sich dann mit der Wahl eines kostenminimalen Parameters  $k$ . Ein ähnliches Modell behandelt auch [GROS85] Kapitel 2.8.

### 2.2.7 Weiterführende Literatur

Als Basisliteratur für die Theorie der Geburten/Todesmodelle können [GROS85], [KLEI75] und [ALLE78] angesehen werden. Letzterer bietet auch ein ganzes Kapitel über den praktischen Einsatz von Warteschlangen in der Leistungsbewertung von Computersystemen an. Allerdings handelt es sich bei den meisten Modellen um keine Geburten/Todesmodelle. Außerdem präsentiert [ALLE78] im Anhang eine übersichtliche Zusammenstellung aller behandelten Modelle. [KLEI75] und [GROS85] bieten hingegen eine sehr detaillierte und einprägsame Einführung in eine Vielzahl von Warteschlangensystemen.

Ohne viel mathematische Umschweife bietet [COX61] eine gut lesbare Einführung in die Theorie der Warteschlangen. Die aus der Telephonie stammenden Arbeiten [COOP72] und [RIO62] bringen eine Reihe von interessanten Ansätzen. Trotz der anderen Nomenklatur sind die dort vorgestellten Modelle denen der Warteschlangentheorie sehr ähnlich.

Einen anderen Weg schlagen [STOC78] und [TRIV82] ein. Sie behandeln einige Warteschlangenmodelle im Zuge der Analyse von Markovketten.

Mit optimalen Betriebsparametern beschäftigen sich [MORS58] und [MEYE71]. Letzterer stellt neben dem im Abschnitt 2.2.6 geschilderten Modell noch ein weiteres vor. Die beiden Geschwindigkeitsübergänge erfolgen jedoch nicht zum gleichen Zeitpunkt. Deshalb handelt es sich bei diesem System um kein Geburten/Todesmodell; es muß also ein anderer Lösungsweg gefunden werden. Eine Möglichkeit wäre, das in Bemerkung 2.3 angesprochene statistische Gleichgewicht vorauszusetzen und die stationären Gleichungen aus dem Zustandsübergangsdiagramm herzuleiten.

Eine sehr tiefgehende, mathematische Analyse der einzelnen Systeme offeriert [SAAT61]. Er leitet fast immer die stationäre Lösung aus der transienten ab.

Sehr viel neue und auch praxisorientierte Ansätze werden von [ROBE94] und [MINO93] vorgelegt. Eine sehr kurze Einführung mit Ausblick auf die Leistungsbewertung von Rechensystemen ist in [BOLC89] zu finden. Eine kurze, gut gegliederte Einführung dagegen beinhaltet [JAIN91].

### 3. Fortgeschrittene Modelle

In diesem Abschnitt werden Wartesysteme mit allgemeiner Ankunfts- und/oder Bedienzeitverteilung untersucht.

Ein geeignetes Hilfsmittel sind die *eingebetteten Markovketten*. Dazu betrachtet man die Zeitpunkte, zu denen ein Zustandsübergang erfolgt, da dort weder die Restbedienzeit, noch die verbleibende Zeit bis zur nächsten Ankunft eine Rolle spielt. Damit ist die Markoveigenschaft gegeben.

Einen ähnlichen Ansatz verwendete Lindley zur Ermittlung der Wartezeitverteilung für das G/G/1-System, der in Abschnitt 3.3 ansatzweise vorgestellt wird.

#### 3.1 Das M/G/1-Modell

Das M/G/1-Modell beschreibt ein System mit Poissonankünften, allgemeiner Bedienzeitverteilung  $B(\cdot)$  und einem Bediener. Das System wird unmittelbar vor Abgang eines Kunden betrachtet, d.h. die Restbedienzeit beträgt 0, der Kunde hat jedoch das System noch nicht verlassen.

##### 3.1.1 Zustandswahrscheinlichkeiten des M/G/1-Modells

Die mittlere Ankunftsrate ist wie beim M/M/1-Modell durch  $\lambda$  gegeben, die Bedienzeit besitzt die Verteilung  $B(\cdot)$  mit Mittelwert  $\frac{1}{\lambda}$ . Die Laplace-Transformierte der zugehörigen Dichte  $b(\cdot)$  lautet

$$\bar{b}(s) = \int_0^{\infty} b(x) e^{-sx} dx \quad [3.1]$$

Da die Ankünfte poissonverteilt sind, läßt sich die Wahrscheinlichkeit für  $a$  Ankünfte während einer Serviceperiode (= Zeit zwischen zwei Abgängen) mit der Länge  $t$  so beschreiben:

$$\begin{aligned} q_{a|t} &= Pr\{a \text{ Ankünfte während Serviceperiode mit Länge } t\} \\ &= \frac{e^{-\lambda t} (\lambda t)^a}{a!} \quad \text{für ein vorgegebenes } t \end{aligned}$$

Wird die Länge der Serviceperiode  $t$  als Zufallsvariable "Bedienzeit" angesehen, so läßt sich  $q_{a|t}$  über die Verteilung  $B(\cdot)$  ermitteln

$$\begin{aligned} q_a &= \int_0^{\infty} q_{a|t} b(t) dt \\ &= Pr\{a \text{ Ankünfte während einer Serviceperiode}\} \end{aligned} \quad [3.2]$$

Somit ergibt sich folgende Zustandsübergangsmatrix

$$\mathbf{Q} = \begin{pmatrix} q_0 & q_1 & q_2 & \cdot & \cdot & \cdot \\ q_0 & q_1 & q_2 & \cdot & \cdot & \cdot \\ 0 & q_0 & q_1 & \cdot & \cdot & \cdot \\ 0 & 0 & q_0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

[3.3]

Die Übergänge für einen Zustand  $n$  sind in Abbildung 3.1 graphisch dargestellt.

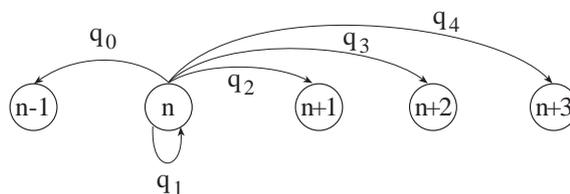


Abb 3.1 Zustandsübergänge eines M/G/1-Modells

Die stationären Zustandswahrscheinlichkeiten lassen sich mit Hilfe des Gleichungssystems (siehe Anhang F.3)

$$\mathbf{pQ} = \mathbf{p}$$

[3.4]

lösen. Daraus folgt

$$\begin{aligned} p_k &= p_0 q_k + \sum_{i=1}^{k+1} p_i q_{k-i+1} && \text{für } k = 0, 1, 2 \\ &= p_0 q_k + \sum_{i=0}^k p_{i+1} q_{k-i} \end{aligned}$$

[3.5]

Die Summe in Gleichung [3.5] erinnert an die Operation der Faltung,

$$h_i = \sum_{j=0}^{\infty} f_j q_{i-j}$$

die mit Hilfe der erzeugenden Funktionen auch so geschrieben werden kann

$$H(z) = F(z)G(z)$$

Wir definieren deshalb

$$P(z) = \sum_{i=0}^{\infty} p_i z^i \quad \text{mit } |z| \leq 1 \quad [3.6]$$

$$\begin{aligned} Q(z) &= \sum_{i=0}^{\infty} q_i z^i \\ &= \sum_{i=0}^{\infty} z^i \int_0^{\infty} \frac{e^{-\lambda t} (\lambda t)^i}{i!} b(t) dt \\ &= \int_0^{\infty} e^{-\lambda t + \lambda z t} b(t) dt \\ &= \bar{b}(\lambda(1-z)) \end{aligned} \quad [3.7]$$

Der in Gleichung [3.7] beschriebene Zusammenhang zwischen der erzeugenden Funktion der Übergangswahrscheinlichkeiten und der Laplace-Transformierten der Bedienzeitdichte wird später benötigt. Mit Hilfe der obigen Definition können wir Gleichung [3.5] nun so schreiben

$$P(z) = p_0 Q(z) + P_{k+1}(z) Q(z) \quad [3.8]$$

Lediglich der Ausdruck  $P_{k+1}(z)$  verhindert eine weitere Analyse. Man versucht deshalb diesen Ausdruck auf einen Term, der  $P(z)$  enthält zurückzuführen

$$\begin{aligned} P_{k+1}(z) &= \sum_{i=0}^{\infty} p_{i+1} z^i = z^{-1} \sum_{i=0}^{\infty} p_{i+1} z^{i+1} \\ &= -z^{-1} p_0 + z^{-1} \sum_{i=-1}^{\infty} p_{i+1} z^{i+1} \\ &= -z^{-1} p_0 + z^{-1} \sum_{i=0}^{\infty} p_i z^i \\ &= z^{-1} (P(z) - p_0) \end{aligned} \quad [3.9]$$

Eingesetzt in Gleichung [3.8] ergibt sich

$$\begin{aligned} P(z) &= p_0 Q(z) + (P(z) - p_0) z^{-1} Q(z) \\ P(z) (1 - z^{-1} Q(z)) &= p_0 (1 - z^{-1}) Q(z) \\ P(z) &= \frac{p_0 (1 - z^{-1}) Q(z)}{1 - z^{-1} Q(z)} \end{aligned}$$

Schließlich wird noch mit  $z/z$  multipliziert

$$P(z) = \frac{p_0 (1-z) Q(z)}{Q(z) - z} \quad [3.10]$$

Nun gilt es,  $p_0$  zu bestimmen. Es handelt sich bei  $P(z)$  und  $Q(z)$  um wahrscheinlichkeitserzeugende Funktionen, folglich gilt

$$P(1) = 1 \quad Q(1) = 1$$

Da der Erwartungswert für einen Zustandsübergang bei  $\frac{\lambda}{\mu} = \rho$  liegt, ergibt sich weiters

$$Q'(1) = \rho$$

Nun drückt man Gleichung [3.10] nach  $p_0$  aus und läßt  $z \rightarrow 1$  streben

$$p_0 = \lim_{z \rightarrow 1} \frac{P(z)[Q(z) - z]}{(1 - z)Q(z)}$$

Unter Anwendung der Regel von L'Hopital ergibt sich schließlich

$$p_0 = 1 - \rho \tag{3.11}$$

Mit Hilfe von Gleichung [3.7] und Gleichung [3.11] läßt sich Gleichung [3.10] wie folgt ermitteln

$$P(z) = \frac{(1 - \rho)(1 - z)\bar{b}(\lambda(1 - z))}{\bar{b}(\lambda(1 - z)) - z} \tag{3.12}$$

Sofern die Dichte der Bedienzeiten bekannt ist, kann deren Laplace-Transformierte in Gleichung [3.12] eingesetzt werden. Danach entwickelt man  $P(z)$  in eine Reihe und isoliert die Koeffizienten der  $z^i$  für alle  $i$ . Diese geben die Wahrscheinlichkeiten an, wonach sich unmittelbar vor einem Abgang  $i$  Kunden im System befinden. Es läßt sich zeigen, daß diese Wahrscheinlichkeiten nicht nur unmittelbar vor einem Abgang, sondern generell gelten. Auf den Beweis soll an dieser Stelle jedoch verzichtet werden, dazu siehe [GROS85] Kapitel 5.3.1 bzw. [ROBE94] Kapitel 2.12.3.

### 3.1.2 Leistungsgrößen des M/G/1-Modells

Obwohl sich mit Hilfe der eingebetteten Markovketten ein eleganter Lösungsweg zur Bestimmung der Zustandswahrscheinlichkeiten ergeben hat, eignet sich das Resultat nicht sehr zur Ermittlung der Leistungsgrößen. Deshalb wird ein Ansatz gewählt, bei dem die Zustandswahrscheinlichkeiten nicht unmittelbar vorliegen müssen. Den Ausgangspunkt bildet folgende Gleichung

$$N_{k+1} = \begin{cases} N_k - 1 + A_{k+1} & \text{für } N_k > 0 \\ A_{k+1} & \text{für } N_k = 0 \end{cases} \tag{3.13}$$

Dabei bezeichnet  $N_k$  die Anzahl der Kunden im System unmittelbar nach dem  $k$ -ten Abgang und  $A_{k+1}$  die Anzahl der Ankünfte zwischen dem  $k$ -ten und dem  $(k+1)$ -ten Abgang. Laut Gleichung [3.13] ergibt sich die

Anzahl der Kunden im System unmittelbar nach dem  $(k+1)$ -ten Abgang aus der Anzahl der Kunden nach dem  $k$ -ten Abgang minus dem abgegangenen Kunden plus den zwischen diesen Abgängen eingetroffenen Kunden. Unter Verwendung der Funktion

$$U(N_k) = \begin{cases} 1 & N_k > 0 \\ 0 & N_k = 0 \end{cases} \quad [3.14]$$

läßt sich Gleichung [3.13] wie folgt schreiben

$$N_{k+1} = N_k - U(N_k) + A_{k+1} \quad [3.15]$$

Nun bildet man den Erwartungswert

$$\mathbb{E}(N_{k+1}) = \mathbb{E}(N_k) - \mathbb{E}(U(N_k)) + \mathbb{E}(A_{k+1}) \quad [3.16]$$

Dabei bezeichnet  $\mathbb{E}(N_k)$  die mittlere Anzahl der Kunden im System unmittelbar nach dem  $k$ -ten Abgang. Wie schon zuvor bei den Zustandswahrscheinlichkeiten läßt sich auch hier zeigen, daß diese Einschränkung unnötig ist, es gilt also

$$\mathbb{E}(N_k) = L \quad \text{für alle } k$$

Daraus folgt

$$L = L - \mathbb{E}(U(N_k)) + \mathbb{E}(A_{k+1})$$

bzw.

$$\mathbb{E}(U(N_k)) = \mathbb{E}(A_{k+1}) \quad [3.17]$$

Weiters gilt

$$\mathbb{E}(U(N_k)) = \sum_{n=0}^{\infty} U(N_k) Pr\{N_k = n\} = \sum_{n=1}^{\infty} Pr\{N_k = n\}$$

Obiger Ausdruck gibt somit die Wahrscheinlichkeit an, daß der Bediener beschäftigt ist

$$\mathbb{E}(U(N_k)) = Pr\{\text{Bediener beschäftigt}\}$$

Die Wahrscheinlichkeit, daß der Bediener frei ist, leitet sich also aus Gleichung [3.11]

$$p_0 = 1 - \rho$$

Da die Summe dieser beiden Wahrscheinlichkeiten 1 ergeben muß, folgt unmittelbar

$$\mathbb{E}(U(N_k)) = \mathbb{E}[A_{k+1}] = 1 - p_0 = \rho \quad [3.18]$$

Um weitere Ergebnisse zu erzielen, muß Gleichung [3.15] quadriert werden.

$$N_{k+1}^2 = N_k^2 + U^2(N_k) + A_{k+1}^2 - 2N_k U(N_k) - 2A_{k+1} U(N_k) + 2A_{k+1}N_k$$

Nun wird erneut der Erwartungswert gebildet und das zweite Moment der Kunden im System gekürzt:

$$0 = \mathbb{E}(U^2(N_k)) + \mathbb{E}(A_{k+1}^2) - 2\mathbb{E}(N_k U(N_k)) - 2\mathbb{E}(A_{k+1} U(N_k)) + 2\mathbb{E}(A_{k+1}N_k) \quad [3.19]$$

Aus der Definition der Funktion  $U(N_k)$  [3.14] können folgende Zusammenhänge abgelesen werden

$$\begin{aligned} U^2(N_k) &= U(N_k) \\ N_k U(N_k) &= N_k \end{aligned}$$

Da die  $A_{k+1}$  von den  $N_k$  unabhängig sind, gilt zusätzlich

$$\mathbb{E}(A_{k+1} U(N_k)) = \mathbb{E}(A_{k+1})\mathbb{E}(U(N_k)) = \rho^2$$

Damit läßt sich Gleichung [3.19] wie folgt definieren

$$0 = \rho + \mathbb{E}(A_{k+1}^2) - 2L - 2\rho^2 + 2\rho L \quad [3.20]$$

Die Ankünfte sind poissonverteilt, daher hängt  $A_{k+1}$  nur von der Zeitspanne zwischen dem k-ten und dem (k+1)-ten Abgang (= Bedienzeit S) nicht aber von den Zeitpunkten k und (k+1) ab. Der Zeitindex kann also entfallen

$$\mathbb{E}(A_{k+1}^2) = \mathbb{E}(A^2)$$

Dieses zweite Moment der Ankünfte läßt sich mit Hilfe von Varianz und Erwartungswert ausdrücken

$$\mathbb{E}(A^2) = \text{Var}(A) + (\mathbb{E}(A))^2 = \text{Var}(A) + \rho^2 \quad [3.21]$$

Die Varianz kann wie folgt ermittelt werden (siehe auch [GROS85] Kapitel 5.1.1)

$$\text{Var}(A) = \mathbb{E}(\text{Var}(A|S)) + \text{Var}(\mathbb{E}(A|S)) \quad [3.22]$$

Für die Komponenten obiger Summe gilt

$$\begin{aligned}\mathbb{E}(\text{Var}(A|S)) &= \mathbb{E}(\lambda S) = \lambda \mathbb{E}(S) = \lambda \bar{s} = \rho \\ \text{Var}(\mathbb{E}(A|S)) &= \text{Var}(\lambda S) = \lambda^2 \text{Var}(S) = \lambda^2 \sigma_S^2\end{aligned}$$

Dabei bezeichnet  $\sigma_S^2$  die Varianz der Bedienzeit. Es ergibt sich

$$\text{Var}(A) = \rho + \lambda^2 \sigma_S^2$$

und weiters

$$\mathbb{E}(A^2) = \rho + \lambda^2 \sigma_S^2 + \rho^2$$

Dieser Ausdruck wird in Gleichung [3.20] eingesetzt

$$0 = 2\rho + \lambda^2 \sigma_S^2 + \rho^2 - 2L - 2\rho^2 + 2\rho L$$

Nach L ausgedrückt ergibt sich

$$L = \frac{2\rho - \rho^2 + \lambda^2 \sigma_S^2}{2(1 - \rho)} \quad [3.23]$$

bzw.

$$L = \rho + \frac{\rho^2 + \lambda^2 \sigma_S^2}{2(1 - \rho)}$$

Obige Ausdrücke stellen zwei äquivalente Formen der sogenannten Pollaczek-Khintchine-Mittelwertformel (P-K-Mittelwertformel) dar. Eine bemerkenswerte Eigenschaft dieser Formel ist, daß die Varianz der Bedienzeit ausreicht, um Ergebnisse für die mittlere Anzahl der Kunden im System zu erhalten.

Weiters läßt sich die Annahme, daß die mittlere Anzahl der Kunden im System mit steigender Varianz zunimmt, bestätigen.

Die weiteren Leistungsgrößen können mit Hilfe des Gesetzes von Little wie gewohnt errechnet werden.

$$W = \frac{1}{\lambda} L = \frac{1}{\mu} + \frac{\rho^2 + \lambda^2 \sigma_S^2}{2\lambda(1 - \rho)} \quad [3.24]$$

$$W_q = W - \frac{1}{\mu} = \frac{\rho^2 + \lambda^2 \sigma_S^2}{2\lambda(1 - \rho)} \quad [3.25]$$

$$L_q = \lambda W_q = \frac{\rho^2 + \lambda^2 \sigma_S^2}{2(1 - \rho)} \quad [3.26]$$

Natürlich lassen sich auch die Verteilungen für die mittlere Antwortzeit und die mittlere Wartezeit finden. Ihre transformierten Funktionen werden nun, ohne näher auf deren Ermittlung einzugehen, angegeben

$$\bar{W}(s) = \frac{(1 - \rho)s \bar{b}(s)}{\lambda \bar{b}(s) - \lambda + s} \quad [3.27]$$

$$\bar{W}_q(s) = \frac{(1 - \rho)s}{\lambda \bar{b}(s) - \lambda + s} \quad [3.28]$$

Auch diese beiden Formeln tragen den Namen Pollaczek-Khintchine-Formel.

Ähnliche und andere Ansätze für das M/G/1-Modell sind in der Literatur wie [COOP72] Kapitel 5.5, [STOC78] Kapitel III.7, [HAIG81] Kapitel 6.6 bis 6.14, [ALLE78] Kapitel 5.3.1, [KLEI75] Kapitel 5, [ROBE94] Kapitel 2.12 und [GROS85] Kapitel 5.1 zu finden.

Im Gegensatz zu den zuvor genannten Autoren leitet [SCHA73] Kapitel II.2 das M/G/1-Modell als Spezialfall vom G/G/1-Modell ab.

Transiente Charakteristiken werden von [PRAB65] u.a. in Kapitel 1.9, 2 und 3 behandelt. Einen Ansatz über die Verteilungen bietet [SAAT61] Kapitel 8.

### 3.1.3 Weitere Aspekte des M/G/1-Modells

Eine Möglichkeit, das Modell zu erweitern, besteht in der Einführung von *ungeduldigen Kunden*. Den Ausgangspunkt bildet Gleichung [3.2]. Sie wird wie folgt modifiziert

$$q_a = \int_0^\infty \frac{e^{-p_e \lambda t} (p_e \lambda t)^a}{a!} b(t) dt \quad [3.29]$$

Dabei bezeichnet  $p_e$  die Wahrscheinlichkeit, daß ein ankommender Kunde das System betritt. Durch die Einführung dieser Wahrscheinlichkeit wird aus dem ursprünglichen Poissonprozeß ein *gefilterter Poissonprozeß*.

Die Erweiterung auf mehr Bediener gestaltet sich nicht so einfach wie beim M/M/1-Modell. Eine Analyse des M/G/m-Modells mit Hilfe der eingebetteten Markovketten ist nicht mehr möglich, da die einzelnen Abgänge den Bedienern nicht mehr zugeordnet werden können. Damit hängt die Anzahl der Ankünfte nicht mehr nur von der Anzahl der Kunden im System nach einem Abgang ab. Eine Ausnahme bildet das M/G/m/m-System aus Kapitel 2.2.2. Weitere Ergebnisse für diesen Spezialfall sind in [GROS85] Kapitel 5.2 zu finden.

Durch Spezialisierung der Bedienzeitverteilung wird oft eine Analyse ermöglicht. Einige Beispiele dafür sind das M/M/m-, das M/D/m- und das M/E<sub>k</sub>/m-Modell. Das M/D/m-Modell wird von [PRAB65] Kapitel 1.7 und [GROS85] Kapitel 6.2 behandelt. Einige Ergebnisse für das E<sub>l</sub>/E<sub>k</sub>/m-Modell sind in [GROS85] Kapitel 3.3.3 zu finden. Wählt man  $l=1$ , so erhält man das M/E<sub>k</sub>/m-Modell.

Wird der Warteraum des M/G/1-Modells auf K Plätze beschränkt, so muß die Ankunftsrate in ähnlicher Weise wie beim M/M/1/K-Modell beschränkt werden, damit die P-K-Mittelwertformel ihre Gültigkeit behält

$$\lambda' = \lambda(1 - p_K)$$

Die Zustandswahrscheinlichkeiten werden wie beim M/G/1-Modell bestimmt, der einzige Unterschied besteht in der Zustandsübergangsmatrix [3.3]. Diese ist auf K-1 Plätze beschränkt. Es entsteht ein Gleichungssystem mit K Gleichungen und K Unbekannten, welches eindeutig gelöst werden kann.

Näheres zum M/G/1/K-Modell siehe [GROS85] Kapitel 5.1.8.

Einige Ergebnisse für das *M/G/1-Modell mit Gruppenankünften* sind in [GROS85] Kapitel 5.1.9 und [BHAT68] Kapitel 1 zu finden.

Das *M/G/1-Modell mit Prioritäten* wird von [WOLF89] Kapitel 8.5, [ALLE78] Kapitel 5.4 und [JAIS68] beschrieben.

### 3.1.4 Spezialisierung auf das M/M/1-Modell

Um die bisher erzielten Ergebnisse einigermaßen zu überprüfen, wird für die Bedienzeitverteilung eine Exponentialverteilung angenommen und das M/G/1-Modell so auf ein M/M/1-Modell reduziert. Die ermittelten Resultate können mit denen aus Kapitel 2.1 verglichen werden.

Die Laplace-Transformierte der Exponentialverteilung ist gegeben durch (siehe Tabelle (C.7, Spalte "Moment-erzeugende Funktionen" und setze  $s=-t$ ).

$$\bar{b}(s) = \frac{\mu}{\mu + s}$$

Eingesetzt in Gleichung [3.12] ergibt sich

$$P(z) = \frac{(1-\rho)(1-z) \frac{\mu}{\mu + \lambda(1-z)}}{\frac{\mu}{\mu + \lambda(1-z)} - z} = \frac{(1-\rho)(1-z)\mu}{\mu - z\mu - z\lambda(1-z)} = \frac{(1-\rho)\mu}{\mu - \lambda z} = \frac{1-\rho}{1-z\rho}$$

Diese Funktion entspricht der wahrscheinlichkeitserzeugenden Funktion eines M/M/1-Modells, was sich durch die Entwicklung in eine geometrische Reihe belegen läßt.

$$(1-\rho) \frac{1}{1-z\rho} = (1-\rho) \sum_{n=0}^{\infty} z^n \rho^n = \sum_{n=0}^{\infty} z^n p_n$$

Durch Isolation der  $z^n$  ergibt sich

$$p_n = \rho^n (1-\rho)$$

Dabei handelt es sich um die stationären Zustandswahrscheinlichkeiten des M/M/1-Modells.

Nun gilt es, die mittlere Anzahl von Kunden im System mit Hilfe der P-K Mittelwertformel zu eruiieren. Da die Bedienzeiten exponentialverteilt sind, liegt der Erwartungswert bei  $\frac{1}{\mu}$  und die Varianz bei  $\frac{1}{\mu^2}$  (siehe auch Tabelle D.3). Diese Werte werden nun in Gleichung [3.23] eingesetzt

$$L = \frac{2\rho - \rho^2 + \lambda^2 \frac{1}{\mu^2}}{2(1-\rho)} = \frac{\rho}{1-\rho}$$

Ein Vergleich mit Gleichung [2.12] zeigt, daß auch dieses Ergebnis korrekt ist.

### 3.1.5 Spezialisierung auf das M/D/1-Modell

Die Ergebnisse für das M/G/1-Modell lassen sich sehr gut auf das sonst schwer zu analysierende M/D/1-Modell anwenden. Diese Spezialisierung wird im folgenden gezeigt.

Die Dichtefunktion der deterministischen Verteilung ist gegeben durch

$$b(x) = \delta\left(x - \frac{1}{\mu}\right)$$

[3.30]

Die Stammfunktion von  $\delta(z)$  lautet  $\Delta(z)$  und besitzt folgendes Aussehen

$$\Delta(z) = \begin{cases} 0 & \text{für } z \leq 0 \\ 1 & \text{für } z > 0 \end{cases}$$

[3.31]

Die Laplace-Transformierte der Dichte lautet

$$\bar{b}(s) = e^{-\frac{s}{\mu}}$$

[3.32]

Diese in Gleichung [3.12] eingesetzt ergibt

$$P(z) = \frac{(1-\rho)(1-z) e^{-\frac{\lambda(1-z)}{\mu}}}{e^{-\frac{\lambda(1-z)}{\mu}} - z}$$

Nun wird mit  $\frac{e^{\rho(1-z)}}{e^{\rho(1-z)}}$  multipliziert

$$P(z) = \frac{(1-\rho)(1-z)}{1-z e^{\rho(1-z)}}$$

[3.33]

Obiger Ausdruck beschreibt die erzeugende Funktion für die Zustandswahrscheinlichkeiten eines M/D/1-Modells. Dieser wird nun in eine Reihe entwickelt

$$P(z) = (1 - \rho)(1 - z) \sum_{i=0}^{\infty} z^i e^{i\rho(1-z)}$$

Wird die Exponentialfunktion in eine Reihe entwickelt, der Ausdruck umgeformt und die Koeffizienten der  $z^i$  isoliert, so ergibt sich

$$\begin{aligned} p_0 &= 1 - \rho \\ p_1 &= (1 - \rho)(e^\rho - 1) \end{aligned}$$

Mit Hilfe dieser beiden Ausdrücke und noch weiteren Umformungen kann folgender Ausdruck errechnet werden

$$p_n = (1 - \rho) \sum_{i=0}^n \frac{(-\rho i)^{n-i} e^{\rho i}}{(n-i)!} - \sum_{i=0}^{n-1} \frac{(-\rho i)^{n-i-1} e^{\rho i}}{(n-i-1)!} \quad [3.34]$$

Weitaus einfacher gestaltet sich die Anwendung der P-K-Mittelwertformel. Da die Varianz der deterministischen Bedienzeitverteilungen  $\sigma_s^2 = 0$  ist, liefert Gleichung [3.23] folgendes Ergebnis

$$L = \rho + \frac{\rho^2}{2(1 - \rho)}$$

Nimmt man dann folgende Umformung vor,

$$L = \frac{2\rho(1 - \rho) + \rho^2}{2(1 - \rho)} = \frac{2\rho - \rho^2}{2(1 - \rho)} = \frac{\rho}{1 - \rho} - \frac{\rho^2}{2(1 - \rho)}$$

so läßt sich zeigen, daß die mittlere Anzahl der Kunden  $L_{M/D/1}$  eines M/D/1-Systems immer kleiner ist als die mittlere Anzahl der Kunden  $L_{M/M/1}$  in einem vergleichbaren M/M/1-System. Ist  $\rho = 0$ , so ist  $L_{M/M/1} = L_{M/D/1} = 0$ . Läßt man  $\rho \rightarrow 0$  gehen, so erweist sich, daß  $L_{M/D/1}$  und  $\frac{1}{2} L_{M/M/1}$  den selben Grenzwert anstreben.

Das M/D/1-Modell ist besonders für die Modellierung paketvermittelnder Knoten interessant. Will man allerdings ein gesamtes Netz analysieren, muß wieder auf Produktformlösungen zurückgegriffen werden, d.h. sämtliche Knoten werden als M/M/1-Modelle interpretiert. Die zuvor gemachten Grenzwertbetrachtungen haben gezeigt, daß die Leistungsgrößen des Netzes mit steigender Verkehrsintensität bzw. Auslastung je Knoten in zunehmendem Maße überschätzt werden.

### 3.2 Das G/M/1-Modell

Die hier vorliegende Analyse des G/M/1-Modells ähnelt in den Grundzügen der des M/G/1-Modells. Ein wesentlicher Unterschied liegt in den Regenerationszeitpunkten, das System wird unmittelbar vor den Ankünften von Kunden betrachtet. Bezeichnet  $a(t)$  die Verteilung der Zwischenankunftszeiten, so ergeben sich folgende Übergangswahrscheinlichkeiten

$$q_b = \Pr\{b \text{ Kunden bedient in einer Zwischenankunftsperiode}\}$$

$$= \int_0^{\infty} \frac{e^{-\mu t} (\mu t)^b}{b!} a(t) dt$$

[3.36]

und weiters die Übergangsmatrix

$$Q = \begin{pmatrix} 1 - q_0 & q_0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 1 - \sum_{k=0}^1 q_k & q_1 & q_0 & 0 & 0 & \cdot & \cdot & \cdot \\ 1 - \sum_{k=0}^2 q_k & q_2 & q_1 & q_0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$$

[3.37]

Die Übergänge für einen Zustand  $n$  sind in Abbildung 3.2 graphisch dargestellt:

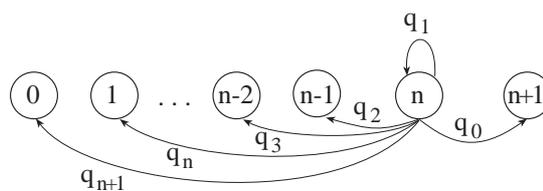


Abb 3.2 Zustandsübergänge einer G/M/1-Modells

Unter der Annahme, daß eine Gleichgewichtslösung existiert, läßt sich folgendes Gleichungssystem lösen

$$\mathbf{p}Q = \mathbf{p}$$

$$\mathbf{p}e = 1$$

[3.38]

Dabei bezeichnet  $\mathbf{e}$  einen Spaltenvektor, wobei alle Elemente gleich 1 sind. Daraus erklärt sich auch die Bedeutung des zweiten Ausdrucks, welcher der Zusatzbedingung entspricht

$$\sum_k p_k = 1$$

Die Lösung von [3.38] kann nach folgendem Schema ermittelt werden

$$p_k = \begin{cases} \sum_{i=0}^{\infty} p_{k+i-1} q_i & \text{für } k \geq 1 \\ \sum_{i=0}^{\infty} p_i \left(1 - \sum_{j=0}^i q_j\right) & \text{für } k = 0 \end{cases}$$

[3.39]

Im Gegensatz zum M/G/1-Modell entsprechen die Zustandswahrscheinlichkeiten unmittelbar vor Ankunft nicht den generellen Gleichgewichtswahrscheinlichkeiten.

Geht man davon aus, daß die Lösungen folgendes Aussehen haben

$$p_{k+1} = \gamma p_k$$

so ergibt sich

$$p_k = C \gamma^k$$

[3.40]

Diese Annahme läßt sich mit Hilfe der relativen Besuchsfrequenzen rechtfertigen, für eventuelle Detailfragen siehe [KLEI75] Kapitel 6.2. Setzt man nun [3.40] in den oberen Teil von [3.39] ein, so präsentiert sich die Formel wie folgt

$$\begin{aligned} p_k = C \gamma^k &= \sum_{i=0}^{\infty} p_{k+i-1} q_i \\ &= \sum_{i=k-1}^{\infty} p_i q_{i+1-k} \\ &= \sum_{i=k-1}^{\infty} C \gamma^i q_{i+1-k} \end{aligned}$$

Durch Elimination von C und  $\gamma^{k-1}$  erhält man

$$\gamma = \sum_{i=k-1}^{\infty} \gamma^{i+1-k} q_{i+1-k} = \sum_{i=0}^{\infty} \gamma^i q_i$$

Schließlich muß noch die Gleichung [3.36] eingesetzt werden

$$\gamma = \sum_{i=0}^{\infty} \gamma^i \int_0^{\infty} \frac{e^{-\mu t} (\mu t)^i}{i!} a(t) dt = \int_0^{\infty} e^{-\mu t(1-\gamma)} a(t) dt$$

[3.41]

Man kann  $\gamma$  auch mit Hilfe der Laplace-Transformierten der Ankunftsichte  $\bar{a}(s)$  ausdrücken

$$\gamma = \bar{a}(\mu(1 - \gamma)) \quad [3.42]$$

Nun muß nur noch die Konstante  $C$  bestimmt werden. Da die Summe aller Wahrscheinlichkeiten 1 ergeben muß, gilt

$$1 = \sum_{k=0}^{\infty} p_k = \sum_{k=0}^{\infty} C \gamma^k = C \sum_{k=0}^{\infty} \gamma^k = \frac{C}{1 - \gamma} \quad \text{für } 0 < \gamma < 1$$

Daraus folgt

$$C = 1 - \gamma \quad [3.43]$$

und weiters

$$p_k = (1 - \gamma) \gamma^k \quad \text{für } k \geq 0 \text{ und } 0 < \gamma < 1 \quad [3.44]$$

Dieses Ergebnis ähnelt stark dem des M/M/1-Modells, es tritt lediglich  $\gamma$  an die Stelle von  $\rho$ . Das ermöglicht die Verwendung der Formeln für die Leistungsgrößen des M/M/1-Modells. Dabei ist allerdings zu beachten, daß es sich um die Leistungsgrößen unmittelbar vor einer Ankunft handelt. Dennoch läßt sich zeigen, daß die Verteilungen für die Warte- und Antwortzeiten generell gelten. Siehe dazu [KLEI75] Kapitel 6.4.

Damit sind auch die Leistungsgrößen  $W$  und  $W_q$  nicht auf die Zeitpunkte unmittelbar vor den Ankünften beschränkt. Die Leistungsgrößen lauten

$$L^{(A)} = \frac{\gamma}{1 - \gamma} \quad [3.45]$$

$$L_q^{(A)} = \frac{\gamma^2}{1 - \gamma} \quad [3.46]$$

$$W = \frac{\gamma}{\lambda(1 - \gamma)} \quad [3.47]$$

$$W_q = \frac{\gamma}{\mu(1 - \gamma)} \quad [3.48]$$

Dabei bezeichnen  $L^{(A)}$  und  $L_q^{(A)}$  die mittlere Anzahl von Kunden im System bzw. die mittlere Warteschlangenlänge unmittelbar vor einer Ankunft. Bei genauerer Betrachtung von Gleichung [3.45] und [3.47] stellt sich heraus, daß das Gesetz von Little für  $L^{(A)}$  und  $W$  bzw. für  $L_q^{(A)}$  und  $W_q$  nicht gilt.

Dieser und auch andere Lösungsansätze sind in der einschlägigen Literatur, wie [ALLE78] Kapitel 5.3.2, [WOLF89] Kapitel 8.6, [HAIG81] Kapitel 6.15, [SAAT61] Kapitel 10, [PRAB65] Kapitel 2.9 und 3 und [COOP72] Kapitel 5.7 zu finden. Die Erweiterung auf mehrere Bediener schlägt sich besonders in den Formeln für  $\gamma$  und  $C$  nieder. Die weitere Analyse spaltet sich in zwei Teile: Entweder es sind  $m$  oder mehr Kunden im System und alle Bediener arbeiten, oder es stehen einer oder mehr Bediener frei. Im ersten Fall entspricht der Lösungsweg dem des  $G/M/1$ -Modells, es gilt allerdings  $k \geq m$  anstatt  $k \geq 0$  und die effektive Bedienrate beträgt  $\mu m$ . Der zweite Fall erfordert eine andere Behandlung, dazu siehe [KLEI75] Kapitel 6.5. Auch [GROS85] bietet in Kapitel 5.3.2 einige Ergebnisse für das  $G/M/m$ -Modell an. Ausgehend vom  $G/G/m$ -System untersucht [SCHA73] in Kapitel III.2 den Spezialfall des  $G/M/m$ -Systems.

Das  $G/M/1$ -Modell mit Gruppenabfertigung wird in [BHAT68] Kapitel 2 behandelt.

### 3.2.1 Spezialisierung auf das $M/M/1$ -Modell

Wie schon zuvor beim  $M/G/1$ -Modell wird in diesem Abschnitt versucht, die Ergebnisse des  $G/M/1$ -Modells im Falle eines  $M/M/1$ -Systems zu überprüfen.

Die momenterzeugende Funktion für exponentialverteilte Zwischenankunftszeiten kann aus Tabelle C.7 abgelesen werden, danach setzt man  $s=-t$  und erhält so die Laplace-Transformierte

$$\bar{a}(s) = \frac{\lambda}{s + \lambda}$$

Einsetzen in Gleichung [3.42] liefert

$$\gamma = \bar{a}(\mu(1 - \gamma)) = \frac{\lambda}{\mu - \mu\gamma + \lambda}$$

Durch Umformung erhält man folgende quadratische Gleichung,

$$\mu\gamma^2 - (\mu + \lambda)\gamma + \lambda = 0$$

die so zerlegt werden kann

$$(\gamma - 1)(\mu\gamma - \lambda) = 0$$

Daraus können nun die Ergebnisse abgelesen werden. Die erste Lösung  $\gamma = 1$  ist nicht maßgeblich, da sie die Bedingung für ein stabiles System  $0 < \gamma < 1$  verletzt. Die zweite Lösung lautet

$$\gamma = \frac{\lambda}{\mu} = \rho$$

Eingesetzt in Gleichung [3.44] ergibt sich

$$p_k = (1 - \rho)\rho^k$$

Ein Vergleich mit Gleichung [2.11] zeigt, daß es sich dabei um die Zustandswahrscheinlichkeiten des M/M/1-Modells handelt.

### 3.2.2. Spezialisierung auf das D/M/1-Modell

In diesem Abschnitt wird ein Modell mit deterministischen Zwischenankunftszeiten vorgestellt. Ergebnisse sind nur mit Hilfe numerischer Techniken möglich. Bezeichnet

$$\bar{a}(s) = e^{-\frac{s}{\lambda}} \quad [3.49]$$

die Laplace-Transformierte der deterministischen Verteilung (näheres dazu siehe Kapitel 3.1.5), so ergibt sich laut Gleichung [3.42]

$$\gamma = \bar{a}(\mu(1 - \gamma)) = e^{-\frac{\mu - \mu\gamma}{\lambda}} = e^{-\frac{(1 - \gamma)}{\rho}} \quad [3.50]$$

Dieses Gleichungssystem kann nur für bestimmte  $\rho$  numerisch ermittelt werden.

#### Beispiel 3.1

Gegenstand der Analyse ist ein Meßdatenerfassungssystem. Von einer Sensorstation werden Daten in bestimmten Zeitintervallen (alle 4 Sekunden) an einen Rechner gemeldet. Dort werden sie in eine Datenbank eingetragen. Die Dauer dieses Vorganges ist exponentialverteilt und beträgt im Mittel 2 Sekunden. Daraus ergibt sich eine Verkehrsintensität von 0,5. Mit Hilfe eines numerischen Verfahrens kann  $\gamma$  mit 0,2 angenähert werden. Somit erhält man folgende Näherungen für die Zustandswahrscheinlichkeiten.

$$\begin{aligned} p_0 &\approx (1 - \gamma) = 0,8 \\ p_1 &\approx (1 - \gamma)\gamma = 0,16 \\ p_2 &\approx (1 - \gamma)\gamma^2 = 0,032 \\ &\cdot \\ &\cdot \\ &\cdot \\ p_k &\approx 0,8 \cdot 0,2^k \end{aligned}$$

Wird die Laplace-Transformierte der deterministischen "Dichtefunktion" in eine Reihe entwickelt,

so kann eine Approximationsformel für das D/M/1-Modell abgeleitet werden:

$$e^{-\frac{s}{\lambda}} \approx 1 - \frac{s}{\lambda} + \frac{s^2}{2\lambda^2} \quad [3.51]$$

Einsetzen in Gleichung [3.42] liefert

$$\gamma = 1 - \frac{1 - \gamma}{\rho} + \frac{(\mu - \mu\gamma)^2}{2\lambda^2} = 1 - \frac{1 - \gamma}{\rho} + \frac{1 - 2\gamma + \gamma^2}{2\rho^2}$$

Es ergibt sich die quadratische Gleichung

$$\gamma^2 - \gamma(2 - 2\rho + \rho^2) + 2\rho^2 - 2\rho + 1 = 0$$

mit ihren Lösungen

$$\gamma_{1/2} = \frac{2 - 2\rho + \rho^2}{2} \pm \rho \frac{2 - \rho}{2}$$

Die erste Lösung  $\gamma_1 = 1$  verletzt die Stabilitätsbedingung  $0 < \gamma < 1$  und wird daher ausgeschlossen. Die zweite Lösung lautet

$$\gamma_2 = 1 - 2\rho + \rho^2 \quad [3.52]$$

und kann zur Berechnung der Zustandswahrscheinlichkeiten herangezogen werden.

$$p_0 = 1 - \gamma = 2\rho - \rho^2$$

$$p_k = (2\rho - \rho^2)(1 - 2\rho + \rho^2)^k$$

Mit Hilfe von Gleichung [3.52] lassen sich Approximationen für die Leistungsgrößen ermitteln

$$L^{(A)} = \frac{1 - 2\rho + \rho^2}{2\rho - \rho^2} \quad [3.53]$$

$$L_q^{(A)} = \frac{(1 - 2\rho + \rho^2)^2}{2\rho - \rho^2} \quad [3.54]$$

$$W = \frac{1 - 2\rho + \rho^2}{\lambda(2\rho - \rho^2)} \quad [3.55]$$

$$W_q = \frac{1 - 2\rho + \rho^2}{\mu(2\rho - \rho^2)} \quad [3.56]$$

Beispiel 3.2:

Werden die Zahlen aus Beispiel 3.1 eingesetzt, so ergeben sich folgende Näherungswerte für die Zustandswahrscheinlichkeiten

$$p_0 = 0,75$$

$$p_1 = 0,1875$$

$$p_2 = 0,046875$$

Der Fehler der Näherungslösungen kann reduziert werden, indem in der Reihenentwicklung der Exponentialfunktion [3.51] mehr Glieder berücksichtigt werden.

---

### 3.3 Das G/G/1-Modell

Obwohl dieses Modell sehr allgemein gehalten ist und damit eine Vielzahl anderer Modelle miteinschließt, lassen sich noch einige Ergebnisse erzielen. Die Analyse des G/G/1-Modells geht auf Lindley (1952) zurück und wird hier nur in den wichtigsten Punkten behandelt. Den Ausgangspunkt bildet Gleichung [1.18] aus der Einführung

$$W_q^{(n+1)} = \begin{cases} W_q^{(n)} + S^{(n)} - T^{(n)} & \text{für } W_q^{(n)} + S^{(n)} - T^{(n)} > 0 \\ 0 & \text{für } W_q^{(n)} + S^{(n)} - T^{(n)} \leq 0 \end{cases} \quad [3.57]$$

Dabei bezeichnet  $W_q^{(n)}$  die Wartezeit des n-ten Kunden. Die Bedienzeiten  $S^{(n)}$  sind nicht negative, unabhängige, identisch verteilte Zufallsgrößen mit Verteilungsfunktion  $B(\cdot)$  und Dichte  $b(\cdot)$ . Wie bisher wird die Laplace-Transformierte von  $b(\cdot)$  mit  $\bar{b}(s)$  bezeichnet. Die Zwischenankunftszeiten  $T^{(n)}$  müssen voneinander und von den Bedienzeiten unabhängig sein. Sie besitzen die Verteilung  $A(\cdot)$  und die Dichte  $a(\cdot)$ , die Laplace-Transformierte der Dichte ist durch  $\bar{a}(s)$  gegeben. Nun definiert man eine weitere Zufallsgröße

$$U^{(n)} = S^{(n)} - T^{(n)}$$

mit Verteilungsfunktion  $U(\cdot)$  und Dichte  $u(\cdot)$ . Die Laplace-Transformierte  $\bar{u}(s)$  läßt sich durch Faltung ermitteln

$$\bar{u}(s) = \bar{a}(-s)\bar{b}(s) \quad [3.58]$$

Ist die Wartezeitverteilung des n-ten Kunden  $W_q^{(n)}(t)$  bekannt, so kann auch die Wartezeitverteilung für den (n+1)-ten Kunden aus Gleichung [3.57] abgeleitet werden

$$W_q^{(n+1)}(t) = \begin{cases} \int_{-\infty}^t W_q^{(n)}(t-x)U(x)dx & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases} \quad [3.59]$$

Unter der Annahme eines stationären Gleichgewichtes,

$$\lim_{n \rightarrow \infty} W_q^{(n)}(t) = W_q(t)$$

das bei  $\rho < 1$  gegeben ist, muß dieses folgender Gleichung genügen

$$W_q(t) = \begin{cases} \int_{-\infty}^t W_q(t-x)U(x)dx & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases} \quad [3.60]$$

Diese Gleichung ist nach Lindley benannt und gehört zur Klasse der Wiener-Hopf Integralgleichungen. Allgemeiner Lösungsansätze als der nun folgende sind in [RIO62] Kapitel 4.5, [WOLF89] Kapitel 9.12 und [KLEI75]

Kapitel 8.2 zu finden. Das Integral auf der rechten Seite von Gleichung [3.60] besitzt einen Wert ungleich 0. Dieser wirkt sich für alle  $t \geq 0$  auf den Wert von  $W_q(t)$  aus. Man definiert deshalb

$$W_q^-(t) = \begin{cases} \int_{-\infty}^t W_q(t-x)U(x)dx & \text{für } t < 0 \\ 0 & \text{für } t \geq 0 \end{cases} \quad [3.61]$$

Diese Gleichung beschreibt den Wert des obigen Integrals für negative  $t$ . Gleichung [3.60] und [3.61] ergeben in Summe

$$W_q^-(t) + W_q(t) = \int_{-\infty}^t W_q(t-x)U(x)dx \quad [3.62]$$

Bezeichnen  $\overline{W}_q^-(s)$  und  $\overline{W}_q(s)$  die Laplace-Transformierten der Funktionen  $W_q^-(t)$  und  $W_q(t)$ , so läßt sich Gleichung [3.62] wie folgt angeben

$$\overline{W}_q^-(s) + \overline{W}_q(s) = \overline{W}_q(s)\overline{U}(s) = \overline{W}_q(s)\overline{a}(-s)\overline{b}(s) \quad [3.63]$$

Zwischen der Laplace-Transformierten  $\overline{w}_q(s)$  der noch unbekanntes Dichte der Wartezeit (= Laplace-Stieltjes-Transformierte der Wartezeitverteilung  $W_q(t)$ ) und der Laplace-Transformierten der Wartezeitverteilung  $\overline{W}_q(s)$  läßt sich folgender Zusammenhang (siehe Anhang C.2, Gleichung [C.4]) herstellen

$$\overline{w}_q(s) = s\overline{W}_q(s)$$

In Gleichung [3.63] eingesetzt ergibt sich

$$\overline{W}_q(s) + \frac{1}{s}\overline{w}_q(s) = \frac{1}{s}\overline{w}_q(s)\overline{a}(-s)\overline{b}(s)$$

Nach weiteren Umformungen folgt schließlich

$$\overline{w}_q(s) = \frac{s\overline{W}_q(s)}{\overline{a}(-s)\overline{b}(s) - 1} \quad [3.64]$$

Das Kernproblem der Anwendung dieser Formel für bestimmte Wartesysteme stellt die Bestimmung der Funktion  $\overline{W}_q^-(s)$  dar. Dies ist jedoch Aufgabe der Funktionentheorie und wird hier nicht weiter erläutert, besonders, da dies den Rahmen dieser Arbeit sprengen würde. Weitere Darstellungen des G/G/1-Modells sind in [GROS85] Kapitel 6.1, [COOP72] Kapitel 6.2, [WOLF89] Kapitel 9 und [KLEI75] Kapitel 8 zu finden. Eine Möglichkeit, die oben genannten komplexen mathematischen Betrachtungen zu umgehen, besteht in der Verwendung von Erlangmischverteilungen zur Approximation der Zwischenankunfts- und Bedienzeiten. Statt dessen können

auch Näherungen und Grenzwerte für bestimmte Leistungsgrößen ermittelt werden. Mit Hilfe des folgenden Ausdrucks können Grenzen für die Wartezeit eines G/G/1-Systems errechnet werden

$$\frac{\lambda^2 \sigma_S^2 + \rho(\rho - 2)}{2\lambda(1 - \rho)} \leq W_q \leq \frac{\lambda(\sigma_T^2 + \sigma_S^2)}{2(1 - \rho)} \quad [3.65]$$

Dabei bezeichnet  $\sigma_T^2$  die Varianz der Zwischenankunftszeiten und  $\sigma_S^2$  die Varianz der Bedienzeiten. Erhält man für die untere Grenze einen negativen Wert, so ist dieser durch 0 zu ersetzen. Die genaue Ermittlung der Grenzen [3.65] wird in [GROS85] Kapitel 7.1 beschrieben.

Mit Hilfe der *Allen-Cunneen Formel* kann ein Näherungswert für die mittlere Wartezeit eines G/G/m-Systems errechnet werden

$$W_q \approx \frac{C(m, \rho)}{\mu(m - \rho)} \left( \frac{\lambda^2 \sigma_T^2 + \mu^2 \sigma_S^2}{2} \right) \quad [3.66]$$

Dabei bezeichnet  $C(m, \rho)$  die *Erlang'sche Formel der 2. Art (Erlang Delay Formula)*. Sie gibt die Wahrscheinlichkeit an, daß alle Bediener eines vergleichbaren M/M/m-Systems belegt sind

$$C(m, \rho) = \sum_{n=m}^{\infty} p_n = \frac{\frac{\rho^m}{(m-1)!(m-\rho)}}{\sum_{n=0}^{m-1} \frac{\rho^n}{n!} + \frac{\rho^m}{(m-1)!(m-\rho)}} \quad [3.67]$$

Gleichung [3.66] liefert genaue Resultate für das M/M/m- und das M/G/1-System. Auch sonst werden gute Ergebnisse erzielt. Näheres zur Allen-Cunneen Formel ist in [ALLE78] Kapitel 5.5.1 zu finden.

## A. Differentialgleichungen

Eine Differentialgleichung besteht aus einer Funktion und ihren Ableitungen. Gesucht wird eine Funktion, die dieser Differentialgleichung genügt.

### A.1 Klassifikation

Eine Differentialgleichung heißt *gewöhnlich*, wenn die gesuchte Funktion nur von einer unabhängigen Variablen abhängt (Im Gegensatz zur *partiellen* Differentialgleichung, dort ist eine Abhängigkeit von weiteren Variablen gegeben). Eine gewöhnliche lineare Differentialgleichung hat beispielsweise folgendes Aussehen:

$$a_0(x) \frac{d^n y}{dx^n} + a_1(x) \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_{n-1}(x) \frac{dy}{dx} + a_n(x) y = f(x) \quad [\text{A.1}]$$

Eine Differentialgleichung hat die *Ordnung*  $n$ , wenn eine Ableitung  $n$ -ter Ordnung, aber keine Ableitung höherer Ordnung vorkommt.

Beispiel A.1 Eine lineare Differentialgleichung 2. Ordnung könnte so aussehen:

$$y'' + 3y' + 2y = 6e^x$$

Die Funktion  $f(x)$  aus Gleichung [A.1] bezeichnet man als *Störfunktion*. Falls keine Störfunktion existiert ( $f(x) = 0$ ), dann heißt die Differentialgleichung *homogen* (im Gegensatz zu *inhomogen* :  $f(x) \neq 0$ ). Weiters existieren auch noch nichtlineare Differentialgleichungen, auf die wir hier aber nicht näher eingehen wollen.

### A.2 Lösungen

Eine Differentialgleichung hat unendlich viele Lösungen, aus denen eine, durch Angabe einer Anfangsbedingung, z.B.  $y(x_0) = y_0$ , ausgewählt werden kann (*partikuläre Lösung*).

Beispiel A.2

Gegeben ist die Differentialgleichung  $y' = y$  mit den Lösungen  $y = C e^x$ .

Es existieren unendlich viele Lösungen, da  $C$  eine beliebig wählbare Konstante bezeichnet.

Aus der Anfangsbedingung  $y(0) = 1$  folgt  $1 = C e^0$ . Damit wird eine Kurve  $y = e^x$  aus der Lösungsschar ausgewählt.

Enthält eine Lösung eine oder mehrere Konstanten, je nach Lösungsverfahren und Ordnung, so heißt diese Lösung *allgemeine Lösung*. Die *Gesamtheit der Lösungen* ist allerdings nicht immer nur durch die allgemeine Lösung alleine gegeben, wie das folgende Beispiel zeigt.

Beispiel A.3

Gesucht ist die Lösung der Differentialgleichung  $y' = \sqrt[3]{y^2}$  mit Anfangsbedingung  $y(2) = 0$ . Neben der partikulären Lösung

$$y = \frac{1}{27}(x - 2)^3$$

gibt es noch eine weitere Lösung  $y = 0$ , die die Anfangsbedingung erfüllt.

Diese zusätzliche Lösung des obigen Beispiels heißt *singuläre* Lösung.

## A.3 Methoden zur Lösung von Differentialgleichungen

### A.3.1 Trennung der Variablen

Gegeben ist eine Funktion mit folgendem Aussehen

$$\frac{dy}{dx} = f(x)g(y) \quad [\text{A.2}]$$

Dann kann die Lösung wie folgt ermittelt werden

$$\int \frac{dy}{g(y)} = \int f(x) dx$$

Gleichung [A.2] wird, vereinfacht ausgedrückt, mit  $dx$  multipliziert, durch  $g(y)$  dividiert und anschließend mit einem Integral versehen.

#### Beispiel A.4

$$\begin{aligned} yy' &= 3x^2 + 2e^x \\ y \frac{dy}{dx} &= 3x^2 + 2e^x \\ \int y dy &= \int (3x^2 + 2e^x) dx \\ \frac{y^2}{2} &= x^3 + 2e^x + C \end{aligned}$$

### A.3.2 Die Methode des integrierenden Faktors

Eine Differentialgleichung der Form

$$\frac{dy}{dx} + a(x)y = f(x) \quad [\text{A.3}]$$

wird mit einer Funktion  $g(x)$  (dem integrierenden Faktor) multipliziert und in folgende Form gebracht:

$$\frac{d(gy)}{dx} = gf \quad [\text{A.4}]$$

Schließlich wird die Lösung mittels Trennung der Variablen ermittelt:

$$y = \frac{1}{g} \int g f dx + \frac{C}{g} \quad [\text{A.5}]$$

Die Anwendung der Produktregel auf Gleichung [A.4] liefert

$$\frac{dy}{dx} + \frac{y}{g} \frac{dg}{dx} = f$$

Gleichgesetzt mit Gleichung [A.3] ergibt das

$$\frac{1}{g} \frac{dg}{dx} = a(x)$$

Integriert und aufgelöst folgt

$$g = e^C e^{\int a(x) dx} \quad [A.6]$$

Da nur eine partikuläre Funktion gesucht wird, kann  $C=0$  gesetzt werden. Das Einsetzen von Gleichung [A.6] in Gleichung [A.5] ergibt die endgültige Lösung:

$$y = e^{-\int a(x) dx} \int e^{\int a(x) dx} f(x) dx + C e^{-\int a(x) dx}$$

### A.3.3 Variation der Konstanten für lineare Differentialgleichungen der 1. Ordnung

Betrachten wir noch einmal eine Gleichung der Form

$$\frac{dy}{dx} + a(x)y = f(x) \quad [A.3]$$

Die Lösungsgesamtheit dieser Differentialgleichung besteht aus der allgemeinen Lösung der zugehörigen homogenen Differentialgleichung  $y_h(x)$  und einer beliebigen partikulären Lösung  $y_p(x)$ . Also gilt

$$y(x) = y_h(x) + y_p(x)$$

Die Lösung der zugehörigen homogenen Differentialgleichung

$$\frac{dy}{dx} + a(x)y = 0$$

läßt sich wie gewohnt durch Trennung der Variablen finden. Ausgehend von der allgemeinen Lösung der zugehörigen homogenen Differentialgleichung führt der Ansatz

$$y_p(x) = C(x) e^{-\int a(x) dx}$$

letztendlich zu folgendem Ergebnis:

$$C(x) = \int f(x) e^{-\int a(x) dx}$$

Beispiel A.5

$$y' - \frac{1-x}{x}y = 4x^2$$

Lösung der zugehörigen homogenen Gleichung

$$\begin{aligned} \frac{dy}{dx} - \frac{1-x}{x}y &= 0 \\ \int \frac{dy}{y} &= \int \frac{1-x}{x} dx \\ \ln|y| &= \ln|x| - x + \ln|C| \quad \text{mit } C \neq 0 \\ y &= Cx e^{-x} \end{aligned}$$

Durch Variation der Konstanten gelangt man zu folgendem Ansatz

$$y_p(x) = C(x)x e^{-x}$$

Einsetzen in die Ausgangsgleichung ergibt

$$\begin{aligned} C'(x)x e^{-x} + C(x)(e^{-x} - x e^{-x}) - \frac{1-x}{x}C(x)e^{-x} &= 4x^2 \\ C'(x)x e^{-x} &= 4x^2 \\ C(x) &= 4 \int x e^x = 4(x e^x - e^x) + C_1 \\ \text{Mit } C_1 = 0 \text{ folgt} \\ y_p(x) &= 4(x^2 - x) \end{aligned}$$

Die Lösungsgesamtheit lautet daher

$$y(x) = Cx e^{-x} + 4(x^2 - x) \quad \text{mit } C \in \mathbb{R}$$

### A.3.4 Die charakteristische Gleichung für linear homogene Differentialgleichungen mit konstanten Koeffizienten

Werden die Koeffizienten von Gleichung [A.1] konstant gewählt und ist  $f(x) = 0$ , so entsteht

$$a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_{n-1} \frac{dy}{dx} + a_n y = 0$$

[A.7]

Dies ist eine linear homogene Gleichung mit konstanten Koeffizienten. Die Form von Gleichung [A.7] läßt darauf schließen, daß die Lösung  $e^{rx}$  enthält, da es sich dabei um die einzige Funktion handelt, deren Ableitung ein Vielfaches von sich selbst ist, das heißt

$$\frac{d^n e^{rx}}{dx^n} = r^n e^{rx}$$

Daraus ergibt sich

$$(a_0 r^n + a_1 r^{n-1} + \dots + a_{n-1} r + a_n) e^{rx} = 0 \quad [\text{A.8}]$$

Für den Fall, daß  $e^{rx} \neq 0$  ist, muß also die *charakteristische Gleichung*

$$a_0 r^n + a_1 r^{n-1} + \dots + a_{n-1} r + a_n = 0 \quad [\text{A.9}]$$

gelöst werden. Falls  $n$  unterschiedliche Wurzeln  $r_1, r_2, \dots, r_n$  existieren, kann Gleichung [A.8] wie folgt ausgedrückt werden

$$(r - r_1)(r - r_2) \dots (r - r_n) e^{rx} = 0$$

Die allgemeine Lösung ist dann durch

$$y = \sum_{i=1}^n C_i e^{r_i x} = C_1 e^{r_1 x} + C_2 e^{r_2 x} + \dots + C_n e^{r_n x}$$

gegeben. Bei Mehrfachwurzeln wird wie folgt vorgegangen: Angenommen, die Wurzel  $r_1$  kommt  $k$  mal vor, so ergibt sich

$$(r - r_1)^k (r - r_2) \dots (r - r_{n-k}) e^{rx} = 0$$

und für die allgemeine Lösung

$$\begin{aligned} y &= \sum_{i=1}^k C_i x^{k-1} e^{r_1 x} + \sum_{i=k+1}^n C_i e^{r_i x} \\ &= C_1 e^{r_1 x} + C_2 x e^{r_1 x} + \dots + C_k x^{k-1} e^{r_1 x} + C_{k+1} e^{r_2 x} + \dots + C_n e^{r_{n-k+1} x} \end{aligned}$$

Beispiel A.6

$$y''' - 4y'' + 5y' - 2y = 0$$

Die charakteristische Gleichung lautet

$$r^3 - 4r^2 + 5r - 2 = 0$$

Eine Lösung  $r_1 = 1$  (durch Einsetzen leicht nachzuprüfen). Die Gleichung reduziert sich mittels Polynomdivision durch  $(r-1)$  auf

$$r^2 - 3r + 2 = 0$$

Die beiden weiteren Lösungen erhält man wie folgt

$$r_{2,3} = \frac{3}{2} \pm \sqrt{\frac{9}{4} - 2} = \frac{3}{2} \pm \frac{1}{2}$$

$$r_2 = 2; \quad r_3 = 1$$

Die zerlegte charakteristische Gleichung lautet also

$$(r-2)(r-1)^2 = 0$$

Die allgemeine Lösung ist

$$y = C_1 e^{2x} + C_2 e^x + C_3 x e^x$$

**A.3.5 Der unbestimmte Ansatz für spezielle Störfunktionen**

Gegeben ist eine Gleichung der Form

$$a_0 \frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + \dots + a_{n-1} \frac{dy}{dx} + a_0 y = f(x)$$

[A.10]

wobei  $f(x)$  aus den Ausdrücken  $x^m$ ,  $\sin bx$ ,  $\cos bx$ ,  $e^{bx}$  oder Kombinationen dieser Ausdrücke besteht. Nun kann ein unbestimmter Ansatz mit anschließendem Koeffizientenvergleich durchgeführt werden. Doch vorher definieren wir die Familie einer Funktion  $f(x)$  und ihrer Ableitungen (ohne zusätzliche Konstanten).

| Störfunktion | Familie                                  |
|--------------|--|
| $x^m$        | $x^m, x^{m-1}, x^{m-2}, \dots, x_1, x_0$ |
| $e^{bx}$     | $e^{bx}$                                 |
| $\sin bx$    | $\sin bx, \cos bx$                       |
| $\cos bx$    | $\sin bx, \cos bx$                       |

Tab A.1 Partikuläre Versuchslösungen für Differentialgleichungen

Für jede Teilfunktion der Störfunktion wird die entsprechende Familie in den Ansatz aufgenommen. Doppelte Familien oder Teilmengen anderer Familien können weggelassen werden. Falls ein Element einer Familie Teil der allgemeinen Lösung ist, wird diese Familie solange mit  $x$  multipliziert, bis kein Element mehr in die allgemeine Lösung eingesetzt werden kann. Schließlich wird die so angesetzte partikuläre Lösung in die Differentialgleichung eingesetzt und ein Koeffizientenvergleich durchgeführt.

Beispiel A.7

$$y''' - y' = 2x + 1 - 4 \cos x + 2 e^x$$

Die allgemeine Lösung der zugehörigen homogenen Gleichung lautet

$$y = C_1 + C_2 e^x + C_3 e^{-x}$$

Folgende Familien sind in der Störfunktion enthalten

$$\{x, 1\}, \{1\}, \{\cos x, \sin x\}, \{e^x\}$$

Die Familie  $\{1\}$  ist Teilmenge der Familie  $\{x, 1\}$  und kann somit entfallen. Weiters sind die Familien  $\{x, 1\}$  und  $\{e^x\}$  Teil der allgemeinen Lösung und werden daher beide mit  $x$  multipliziert.

Folgende Familien werden nun für den Ansatz verwendet

$$\{x^2, x\}, \{\cos x, \sin x\}, \{x e^x\}$$

$$y_p = Ax^2 + Bx + C \cos x + D \sin x + Ex e^x$$

Einsetzen von  $y_p$  in die Ausgangsgleichung liefert

$$C \sin x - D \cos x + E(x e^x + 3 e^x)$$

$$- [2Ax + B - C \sin x + D \cos x + E(x e^x + e^x)]$$

$$= 2x + 1 - 4 \cos x + 2 e^x$$

Vereinfacht ergibt dies

$$-2Ax - B + 2C \sin x - 2D \cos x + 2E e^x = 2x + 1 - 4 \cos x + 2 e^x$$

Nach einem Koeffizientenvergleich erhält man

$$A = -1; B = -1; C = 0; D = 2; E = 1$$

Daraus folgt die partikuläre Lösung

$$y_p = -x^2 - x + 2 \sin x + x e^x$$

und die Lösungsgesamtheit

$$y = C_1 + C_2 e^x + C_3 e^{-x} - x^2 - x + 2 \sin x + x e^x$$

### A.3.6 Variation der Konstanten

Wird die Gleichung [A.1] durch  $a_0(x)$  dividiert, so erhält man

$$\frac{d^n y}{d x^n} + b_1(x) \frac{d^{n-1} y}{d x^{n-1}} + \dots + b_{n-1}(x) \frac{d y}{d x} + b_n(x) y = g(x)$$

[A.11]

Setzt sich die Lösung der zugehörigen homogenen Gleichung aus  $n$  Teillösungen  $y_1(x), \dots, y_n(x)$  zusammen,

$$y = C_1 y_1 + C_2 y_2 + \dots + C_n y_n$$

[A.12]

so kann die partikuläre Lösung wie folgt ausgedrückt werden

$$y_p = C_1(x) y_1(x) + C_2(x) y_2(x) + \dots + C_n(x) y_n(x)$$

[A.13]

Um Ergebnisse für die Funktion  $C_1(x), \dots, C_n(x)$  zu erhalten, muß folgendes Gleichungssystem gelöst werden:

$$\begin{aligned} C'_1 y_1 + C'_2 y_2 + \dots + C'_n y_n &= 0 \\ C_1 y'_1 + C_2 y'_2 + \dots + C_n y'_n &= 0 \\ C'_1 y_1^{(n-2)} + C'_2 y_2^{(n-2)} + \dots + C'_n y_n^{(n-2)} &= 0 \\ C'_1 y_1^{(n-1)} + C'_2 y_2^{(n-1)} + \dots + C'_n y_n^{(n-1)} &= g \end{aligned}$$

[A.14]

Schließlich werden die  $C'_i$  integriert, um die  $C_i$  zu erhalten.

Für eine linear inhomogene Differentialgleichung mit konstanten Koeffizienten

$$\frac{d^2y}{dx^2} + b_1 \frac{dy}{dx} + b_2 y = f(x) \quad [\text{A.15}]$$

machen wir wieder den gewohnten Ansatz

$$y_p(x) = C_1(x) y_1(x) + C_2(x) y_2(x) \quad [\text{A.16}]$$

mit der Zusatzbedingung

$$C_1'(x) y_1(x) + C_2'(x) y_2(x) = 0$$

Daraus folgt, ohne hier näher darauf einzugehen

$$C_1(x) = \int \frac{-y_2 f(x)}{y_1 y_2' - y_1' y_2} dx \quad \text{und} \quad C_2(x) = \int \frac{y_1 f(x)}{y_1 y_2' - y_1' y_2} dx \quad [\text{A.17}]$$

Aus diesen beiden Funktionen kann schließlich  $y_p(x)$  bestimmt werden.

### A.3.7 Systeme von Differentialgleichungen

Als ein System von Differentialgleichungen bezeichnet man zwei oder mehrere Gleichungen mit mehreren abhängigen und einer unabhängigen Variablen. Durch Umformen wird versucht, aus den gegebenen Gleichungen eine Gleichung zu ermitteln, die mit den bisher besprochenen Methoden gelöst werden kann.

#### Beispiel A.8

$$x y' = y$$

$$x z z' + x^2 + y^2 = 0$$

Löse 1. Gleichung

$$\frac{dy}{dx} = \frac{y}{x}$$

$$y = C_1 x$$

$$\int \frac{dy}{y} = \int \frac{dy}{y}$$

mit  $C_1 \in \mathbb{R}$

Setze in 2. Gleichung ein und löse diese

$$\begin{aligned}
 x z z' + x^2 + C_1^2 x^2 &= 0 \\
 z' &= \frac{-x - C_1 x}{z} & \int z dz &= \int (x + C_1^2 x) dx \\
 \frac{z^2}{2} &= -\frac{1}{2} x^2 - \frac{C_1^2}{2} x^2 + \frac{C_2}{2} & \text{mit } C_2 &\in \mathbb{R} \\
 z &= \pm \sqrt{C_2 - x^2(1 + C_1^2)}
 \end{aligned}$$

### A.3.8 Zurückführen auf Differentialgleichungen mit niederer Ordnung

Falls die Differentialgleichung die Form

$$y'' = f(x)$$

hat, so kann sie auf eine Differentialgleichung 1. Ordnung der Form

$$y' = F(x)$$

zurückgeführt werden.  $F(x)$  bezeichnet die Stammfunktion von  $f(x)$ .

Enthält eine Differentialgleichung 2. Ordnung nur die 1. und 2. Ableitung einer Funktion, so erhält man nach der Substitution  $z = y'$  eine Differentialgleichung 1. Ordnung mit der abhängigen Variable  $z$ .

Kommt in einer Differentialgleichung 2. Ordnung die unabhängige Variable nicht vor, so setzt man

$$y' = p \quad \text{und damit} \quad y'' = p \frac{dp}{dy}$$

und erhält dadurch eine Differentialgleichung 1. Ordnung in  $y, p, \frac{dp}{dy}$ , wobei  $y$  nun die unabhängige und  $p$  die abhängige Variable darstellt.

Ist eine Differentialgleichung 2. Ordnung der Form

$$\frac{d^2 y}{dx^2} + b_1 \frac{dy}{dx} + b_2 y = f(x)$$

[A.18]

gegeben und ist eine homogene Lösung  $y_1$  bekannt, so kann  $y$  durch

$$y = y_1 p$$

[A.19]

ausgedrückt werden. Allerdings muß  $p$  aus folgender Differentialgleichung ermittelt werden:

$$y_1 \frac{d^2 p}{dx^2} + \left( 2 \frac{dy_1}{dx} + b_1 y_1 \right) \frac{dp}{dx} = f(x)$$

Diese kann hingegen wieder auf eine Differentialgleichung 1. Ordnung zurückgeführt werden.

$$\begin{aligned} \frac{dp}{dx} &= q \\ y_1 \frac{dq}{dx} + \left( 2 \frac{dy_1}{dx} + b_1 y_1 \right) q &= f(x) \end{aligned}$$

[A.20]

Ist  $q$  ermittelt, so wird  $p$  durch Integration berechnet und in Gleichung [A.9] eingesetzt, um die Lösung für  $y$  zu erhalten.

### A.3.9 Die Verwendung von Operatoren

Bevor die Verwendung von Operatoren an einem Beispiel (siehe Beispiel A.7) erläutert werden kann, muß folgende Schreibweise vereinbart werden

$$y' = Dy \quad ; \quad y'' = D^2 y \quad ; \quad y''' = D^3 y \quad \text{etc.}$$

#### Beispiel A.9

$$y''' - y' = 2x + 1 - 4 \cos x + 2 e^x$$

In Operatorschreibweise:

$$\begin{aligned} D^3 y - Dy &= 2x + 1 - 4 \cos x + 2 e^x \\ D(D+1)(D-1)y &= 2x + 1 - 4 \cos x + 2 e^x \end{aligned}$$

Es wird eine neue Variable eingeführt

$$y_1 = (D+1)(D-1)y$$

Dies entspricht

$$Dy_1 = 2x + 1 - 4 \cos x + 2 e^x$$

daraus folgt:

$$\frac{dy_1}{dx} = 2x + 1 - 4 \cos x + 2e^x$$

$$y_1 = \int (2x + 1 - 4 \cos x + 2e^x) dx + C_1$$

$$y_1 = x^2 + x - 4 \sin x + 2e^x + C_1$$

Resubstituiert

$$(D+1)(D-1)y = x^2 + x - 4 \sin x + 2e^x + C_1$$

Es wird eine neue Variable eingeführt

$$y_2 = (D-1)y$$

daraus folgt:

$$(D+1)y_2 = x^2 + x - 4 \sin x + 2e^x + C_1$$

Dies entspricht

$$\frac{dy_2}{dx} + y_2 = x^2 + x - 4 \sin x + 2e^x + C_1$$

Diese Differentialgleichung 1. Ordnung kann z.B. mit der Methode des integrierenden Faktors gelöst werden:

$$y_2 = e^{-x} \int e^x (x^2 + x - 4 \sin x + 2e^x + C_1) dx + C_2 e^{-x}$$

$$y_2 = x^2 - x + 1 - 2 \sin x + 2 \cos x + e^x + C_1 + C_2 e^{-x}$$

Resubstituiert:

$$(D-1)y = x^2 - x + 1 - 2 \sin x + 2 \cos x + e^x + C_1 + C_2 e^{-x}$$

Wiederum erhalten wir eine Differentialgleichung 1. Ordnung, die mit der Methode des integrierenden Faktors gelöst werden kann

$$y_2 = e^x \int e^{-x} (x^2 - x + 1 - 2 \sin x + 2 \cos x + e^x + C_1 + C_2 e^{-x}) dx + C_3 e^x$$

$$y = -x^2 - x + 2 \sin x + x e^x - 2 - C_1 - \frac{C_2}{2} e^{-x} + C_3 e^x$$

Werden nun die Vorzeichen, Konstanten und Brüche in die Konstanten  $C_1$ ,  $C_2$  und  $C_3$  eingerechnet, so ergibt sich

$$y = -x^2 - x + 2 \sin x + x e^x + C_1 + C_2 e^{-x} + C_3 e^x$$

## B. Differenzgleichungen

Manche Prozesse laufen nicht kontinuierlich, sondern stufenförmig ab. Daher sind Differentialgleichungen kein geeignetes Mittel, um solche Prozesse zu beschreiben. Das diskrete Analogon zu den Differentialgleichungen sind die *Differenzgleichungen*, sie beschreiben, wie die Kennzahlen einer Prozeßstufe aus den Kennzahlen der vorangegangenen Prozeßstufen zu berechnen sind.

Wenn der Zustand eines Prozesses zum Zeitpunkt  $t$  mit  $y_t$  und der Zustand zum Zeitpunkt  $t+1$  mit  $y_{t+1}$ , usw. bezeichnet wird, so kann die *erste* (endliche) *Differenz* als

$$\Delta y = y_{t+1} - y_t$$

definiert werden. Die zweite Differenz ist dann durch

$$\Delta^2 y = \Delta(\Delta y) = y_{t+2} - y_{t+1} - (y_{t+1} - y_t) = y_{t+2} - y_t$$

und die  $n$ -te Differenz durch

$$\Delta^n y = \Delta(\Delta^{n-1} y) = y_{t+n} - y_t$$

gegeben. Die  $n$ -te Differenz entspricht also der  $n$ -ten Ableitung einer Differentialgleichung. Dank dieser Ähnlichkeit sind einige Theorien der Differentialgleichung übertragbar. (z. B. die Gregory-Newton Formel, sie entspricht der Taylor Formel für Differentialgleichungen)

### B.1 Klassifikation

Die Einteilung der Differenzgleichungen erfolgt völlig analog zu der der Differentialgleichungen. Zum einen gibt es *gewöhnliche*, zum anderen *partielle* Differenzgleichungen. In diesem Abschnitt werden allerdings nur einige wenige Arten von gewöhnlichen Differenzgleichungen behandelt. Eine lineare Differenzgleichung mit

konstanten Koeffizienten hat folgendes Aussehen:

$$y_{t+n} + a_1 y_{t+n-1} + \dots + a_{n-1} y_{t+1} + a_n y_t = f_t \quad [\text{B.1}]$$

$f_t$  bezeichnet die schon von den Differentialgleichungen bekannte *Störfunktion*. Ist diese gleich 0, so spricht man von einer homogenen, andernfalls von einer inhomogenen, d.h.  $f_t \neq 0$ , Differenzgleichung. Weiters ist abzusehen, daß Gleichung [B.1] *Ordnung*  $n$  besitzt.

## B.2 Typen von Differenzgleichungen und deren Lösungen

### B.2.1 Lineare Differenzgleichungen 1. Ordnung

Das nun folgende Beispiel soll illustrieren, daß eine Lösung für einfache Typen von Differenzgleichungen schrittweise, auch ohne Kenntnis der allgemeinen Lösungsformel, ermittelt werden kann.

#### Beispiel B.1

Jemand legt einen Betrag  $S_0$  auf ein Kapitalsparbuch zum Zinssatz  $p$ . Es stellt sich die Frage, wie hoch ist das Vermögen nach  $n$  Jahren unter der Voraussetzung, daß kein weiterer Betrag einbezahlt wird. Das Vermögen setzt sich lediglich aus dem einbezahlten Betrag und den Zinsen des Vorjahres zusammen.

$$S_{t+1} = S_t + S_t \cdot p$$

Daraus folgt

$$S_{t+1} = (1 + p)S_t$$

Ausgehend von der Einlage läßt sich das Vermögen nach  $t$  Jahren schrittweise für alle  $t=1 \dots n$  ermitteln:

$$S_1 = (1 + p)S_0$$

$$S_2 = (1 + p)S_1 = (1 + p)^2 S_0$$

$$S_n = (1 + p)^n S_0$$

Dabei gibt die letzte Zeile das Vermögen nach  $n$  Jahren an.

#### Beispiel B.2

Betrachten wir einen weiteren Fall, bei dem jemand einen Betrag  $S_0$  auf ein Kapitalsparbuch einbezahlt und am Ende eines jeden Jahres diesen noch um einen gleichbleibenden Betrag  $A$  (Annuität) erhöht. Daraus folgt

$$S_{t+1} = (1 + p)S_t + A$$

Wieder wird schrittweise vorgegangen:

$$S_1 = (1 + p)S_0 + A$$

$$S_2 = (1 + p)S_1 + A = (1 + p)^2 S_0 + (1 + p)A + A = (1 + p)^2 S_0 + A[(1 + p) + 1]$$

$$S_n = (1 + p)^n S_0 + A \sum_{i=0}^{n-1} (1 + p)^i = (1 + p)^n S_0 - A \left( \frac{1 - (1 + p)^n}{p} \right)$$

Wird am Ende des Jahres nur ein Betrag  $A$  eingelegt, d.h.  $S_0 = 0$ , so ergibt sich für das Vermögen  $S_n$  nach  $n$  Jahren

$$S_n = A \sum_{i=0}^{n-1} (1+p)^i = A \left( \frac{1 - (1+p)^n}{-p} \right)$$

Wird diese Einzahlung hingegen zu Beginn des Jahres getätigt, so ergibt sich

$$S_n = A \sum_{i=1}^n (1+p)^i = A(1+p) \left( \frac{1 - (1+p)^n}{-p} \right)$$

Betrachten wir nun die allgemeine Lösung einer Differenzgleichung erster Ordnung der Art

$$y_{t+1} - a y_t = 0 \quad [\text{B.2}]$$

Die Lösung läßt sich aus Beispiel B.1 ableiten und lautet

$$y_t = a^t y_0$$

Tritt anstelle des konstanten Faktors  $a$  eine Funktion  $a_t$ , so lautet die Differenzgleichung

$$y_{t+1} - a_t y_t = 0 \quad [\text{B.3}]$$

und deren Lösung

$$y_t = y_0 \prod_{i=0}^{t-1} a_i$$

Gleichung [B.3] bezeichnet eine lineare homogene Differenzgleichung 1. Ordnung und Gleichung [B.2] stellt den Spezialfall mit konstanten Koeffizienten dar.

Sehen wir uns jetzt die Lösung einer inhomogenen Differenzgleichung an

$$y_{t+1} - a_t y_t = f_t \quad [\text{B.4}]$$

Sie lautet

$$y_t = y_0 \prod_{i=0}^{t-1} a_i + \sum_{i=0}^{t-1} f_i \prod_{j=i+0}^{t-1} a_j \quad \text{mit} \prod_{i=x}^x a_i = 1 \quad \forall x \in \mathbb{R} \quad [\text{B.5}]$$

Falls konstante Koeffizienten vorliegen,

$$y_{t+1} - a y_t = f_t \quad [\text{B.6}]$$

so vereinfacht sich die Lösung zu

$$y_t = y_0 a^t + \sum_{i=0}^{t-1} a^{t-1-i} f_i \quad [\text{B.7}]$$

Liegt auch noch eine konstante Störfunktion  $f_t = f$  vor, so kann die Lösung nochmals vereinfacht werden (siehe auch Beispiel B.2)

$$y_t = \begin{cases} y_0 a^t + \frac{1-a^t}{1-a} f & \text{für } a \neq 1 \\ y_0 + t \cdot f & \text{für } a = 1 \end{cases} \quad [\text{B.8}]$$

Bei genauerer Betrachtung der Lösung [B.5], [B.7] und [B.8] können wir feststellen, daß sich die Lösung aus einer allgemeinen und einer partikulären Lösung zusammensetzt. Bei Differenzgleichungen höherer Ordnung wird eine getrennte Ermittlung dieser beiden Lösungen ähnlich der der Differentialgleichungen nötig.

### B.2.2 Lineare Differenzgleichungen mit konstanten Koeffizienten

Für Gleichung [B.1] läßt sich die Lösung der zugehörigen homogenen Differenzgleichung

$$y_{t+n} + a_1 y_{t+n-1} + \dots + a_{t-1} y_{t+1} + a_n y_t = 0 \quad [\text{B.9}]$$

z. B. mit Hilfe von Operatoren oder durch die charakteristische Gleichung

$$r^n + a_1 r^{n-1} + \dots + a_{n-1} r + a_n = 0 \quad [\text{B.10}]$$

finden. Die allgemeine Lösung setzt sich aus den Wurzeln der charakteristischen Gleichung zusammen:

$$y_t = C_1 r_1^t + C_2 r_2^t + \dots + C_n r_n^t \quad [\text{B.11}]$$

Bei mehrfach vorhandenen Wurzeln wird ähnlich verfahren wie bei Differentialgleichungen. Sind  $k$  gleiche Wurzeln vorhanden, so werden sie wie folgt behandelt:

$$y_t = C_1 r_1^t + C_2 t r_1^t + C_3 t(t-1) r_1^t + \dots + C_k t(t-1)\dots(t-k+1) r_1^t \\ + C_{k+1} r_2^t + C_{k+2} r_3^t + \dots + C_n r_{n-k+1}^t$$

Bei Verwendung von Operatoren wird der Operator  $D$  wie folgt definiert

$$Dy_t = y_{t+1} \\ D^2 y_t = y_{t+2}$$

Daraus ergibt sich folgender Zusammenhang zwischen  $D$  und  $\Delta$

$$D^n = (\Delta + 1)^n$$

Gleichung [B.9] in Operatorschreibweise lautet

$$(D^n + a_1 D^{n-1} + \dots + a_{n-1} D + a_n) y_t = 0$$

[B.12]

### Beispiel B.3

$$y_{t+2} + 6y_{t+1} + 9y_t = 0$$

Die charakteristische Gleichung

$$\begin{aligned} r^2 + 6r + 9 &= 0 \\ (r + 3)(r + 3) &= 0 \end{aligned}$$

hat zwei Wurzeln  $-3$  und  $-3$ . Daraus folgt die allgemeine Lösung

$$(r + 3)(r + 3) = 0$$

Die partikuläre Lösung einer inhomogenen Gleichung der Form [B.1] lässt sich mit den von den Differentialgleichungen her bekannten *unbestimmten Ansatz* finden. Für welche Störfunktion welche Ansätze gemacht werden, zeigt Tabelle B.1

| Störfunktion $f_t$   | Versuchslösung   |
|----------------------|--|
| 1                    | A  |
| $t^i$                | $A_0 + A_1 t + \dots + A_i t^i$ für $i \in \mathbb{N}$ |
| $s^t$                | $A s^t$ für $s \in \mathbb{R}$                         |
| $\sin(st), \cos(st)$ | $A \sin(st) + B \cos(st)$                              |

Tab B.1 Partikuläre Versuchslösungen für Differenzgleichungen

Wiederum wird für jeden Teil der Störfunktion die entsprechende Versuchslösung in den Ansatz aufgenommen. Doppelte Versuchslösungen und Teilmengen anderer Versuchslösungen können weggelassen werden. Falls ein Element einer Versuchslösung Teil der allgemeinen Lösung ist, wird diese Versuchslösung so lange mit  $t$  multipliziert, bis dies nicht mehr der Fall ist. Die erhaltene partikuläre Lösung wird nun in die Differenzgleichung eingesetzt, um die Koeffizienten bestimmen zu können.

Beispiel B.4

$$y_{t+2} + 6y_{t+1} + 9y_t = 16t^2$$

Aus Beispiel B.3 kennen wir die allgemeine Lösung

$$y_t = C_1(-3)^t + C_2t(-3)^t$$

Wir setzen die partikuläre Lösung wie folgt an (siehe Tabelle B.1)

$$y_{t,p} = A_0t^2 + A_1t + A_2$$

Eingesetzt in die Ausgangsgleichung ergibt dies

$$\begin{aligned} A_0(t+2)^2 + A_1(t+2) + A_2 + 6[A_0(t+1)^2 + A_1(t+1) + A_2] \\ + 9[A_0t^2 + A_1t + A_2] &= 16t^2 \\ 16A_0t^2 + (16A_0 + 16A_1)t + 10A_0 + 8A_1 + 16A_2 &= 16t^2 \end{aligned}$$

Daraus folgt

$$\begin{aligned} A_0 &= 1; \quad A_1 = -1; \quad A_2 = -\frac{1}{8} \\ y_{t,p} &= t^2 - t - \frac{1}{8} \end{aligned}$$

Somit ergibt sich

$$y_t = C_1(-3)^t + C_2(-3)^t + t^2 - t - \frac{1}{8}$$

**B.2.3 Systeme von Differenzgleichungen mit konstanten Koeffizienten**

Wir werden nun die Operatorschreibweise verwenden, um eine Lösung für ein Differenzgleichungssystem herbeizuführen. Folgendes Beispiel soll die Vorgehensweise verdeutlichen:

Beispiel B.5

Gegeben ist ein System von Differenzgleichungen

$$\begin{aligned} y_{t+1} - 3y_t + z_{t+1} - 3z_t &= 2 \\ 2y_{t+1} - 5y_t + 3z_{t+1} - 3z_t &= 6 \cdot 4^t \end{aligned}$$

In Operatorschreibweise

$$\begin{aligned}(D-3)y_t + (D-3)z_t &= 2 \\ (2D-5)y_t + 3(D-3)z_t &= 6 \cdot 4^t\end{aligned}$$

Die charakteristische Gleichung der zugehörigen homogenen Differenzgleichung lautet:

$$\begin{vmatrix} (D-3) & (D-3) \\ (2D-5) & (3D-3) \end{vmatrix} = D^2 - D - 6 = 0$$

Es ergeben sich die Wurzeln  $D_1 = 3$  und  $D_2 = -2$  und somit die allgemeinen Lösungen

$$\begin{aligned}y_t &= C_1(3)^t + C_2(-2)^t \\ z_t &= C_3(3)^t + C_4(-2)^t\end{aligned}$$

Nun werden folgende Ansätze für die partikuläre Lösung gemacht

$$\begin{aligned}y_{t,p} &= A_0 + A_1 \cdot 4^t \\ z_{t,p} &= A_2 + A_3 \cdot 4^t\end{aligned}$$

Eingesetzt in die Ausgangsgleichungen und vereinfacht ergibt sich

$$\begin{aligned}-2A_0 - 2A_2 + (A_1 + A_3)4^t &= 2 \\ -3A_0 + (3A_1 + 9A_3) &= 6 \cdot 4^t\end{aligned}$$

Es ergeben sich folgende Koeffizienten

$$A_0 = 0; \quad A_1 = -1; \quad A_2 = -1; \quad A_3 = 1$$

und somit die Lösungen

$$\begin{aligned}y_t &= C_1(3)^t + C_2(-2)^t - 4^t \\ z_t &= C_3(3)^t + C_4(-2)^t - 1 + 4^t\end{aligned}$$

Werden die allgemeinen Lösungen in das Ausgangsgleichungssystem eingesetzt, so ergeben sich folgende Zusammenhänge

$$C_3 = -\frac{1}{6}C_1 \quad \text{und} \quad C_4 = -C_2$$

Damit erhält man eine Lösungsgesamtheit von

$$y_t = C_1(3)^t + C_2(-2)^t - 4^t$$
$$z_t = -\frac{1}{6}C_1(3)^t - C_2(-2)^t - 1 + 4^t$$

#### Literaturhinweise

Eine kurze Zusammenfassung bieten [GROS74] und [GROS85] in Anhang 2.

Eine detaillierte Darstellung, die u.a. auch die Stabilität der Lösungen untersucht, ist in [DOR288] in Kapitel V zu finden.

---

## C. Transformationen und erzeugende Funktionen

Transformationen haben ihren Ursprung in der Elektrotechnik und der Physik. Entweder sind sie Teil des untersuchten physikalischen Systems oder sie werden bei der analytischen Betrachtung herangezogen. Im folgenden sollen die wichtigsten Transformationen für die Warteschlangentheorie erläutert werden.

### C.1 Laplace - Transformation

Die *Laplace-Transformation*  $\bar{f}(t)$  ist definiert als

$$\bar{f}(s) = \int_0^{\infty} e^{-st} f(t) dt \quad \text{mit} \quad \text{Real}(s) > s_0 \quad [\text{C.1}]$$

Das Integral konvergiert für jedes komplexe  $s$ , dessen Realanteil größer als ein bestimmter Wert  $s_0$  ist. Die wichtigsten *Laplace-Transformierten* sind in Tabelle C.1 zu finden. Bevor anhand eines konkreten Beispiels gezeigt wird, wie eine Differentialgleichung mittels *Laplace-Transformation* gelöst werden kann, werden noch einige wichtige Eigenschaften behandelt.

#### Satz C.1

Gegeben sei eine Funktion

$$g(t) = \sum_{i=1}^n a_i f_i(t) \quad \text{mit} \quad a_i \text{ beliebig konstant,}$$

so ist deren Laplace-Transformierte  $\bar{g}(t)$  gegeben durch

$$\bar{g}(s) = \sum_{i=1}^n a_i \bar{f}_i(s)$$

Die Laplace-Transformierte einer Summe von Funktionen besteht also aus der Summe der Laplace-Transformierten dieser Funktion.

#### Satz C.2

Gegeben sind zwei unabhängige kontinuierliche Zufallsvariablen  $X_1$  und  $X_2$  mit ihren Dichtefunktionen  $f_1(\cdot)$  und  $f_2(\cdot)$ . Es gilt

$$X_3 = X_1 + X_2 \quad f_3(t) = \int_{-\infty}^{\infty} f_1(t-k) f_2(k) dk \quad \bar{f}_3(s) = \bar{f}_1(s) \bar{f}_2(s)$$

Dadurch ist es uns möglich geworden, den komplizierten Vorgang der Faltung durch eine Multiplikation zu ersetzen.

Satz C.3

Wenn eine kontinuierliche Zufallsvariable mit Dichtefunktion  $f(\cdot)$  und deren Laplace-Transformierten  $\bar{f}(s)$  gegeben ist, sind die Momente von X gegeben durch

$$\mathbb{E}(X^k) = (-1)^k \frac{d^k \bar{f}(s)}{ds^k} \Big|_{s=0} \quad \text{mit } k = 1, 2, 3, \dots \text{ ①}$$

Weitere wichtige Eigenschaften von Laplace-Transformationen können aus Tabelle C.2 abgelesen werden. Doch nun zu dem bereits angekündigten Beispiel:

Beispiel C.1

Gegeben sei eine Differentialgleichung

$$y' + a y = e^{-bt}$$

mit der Anfangsbedingung  $y(0) = 0$ . Der genaue Lösungsweg ist in Abbildung 1 dargestellt.

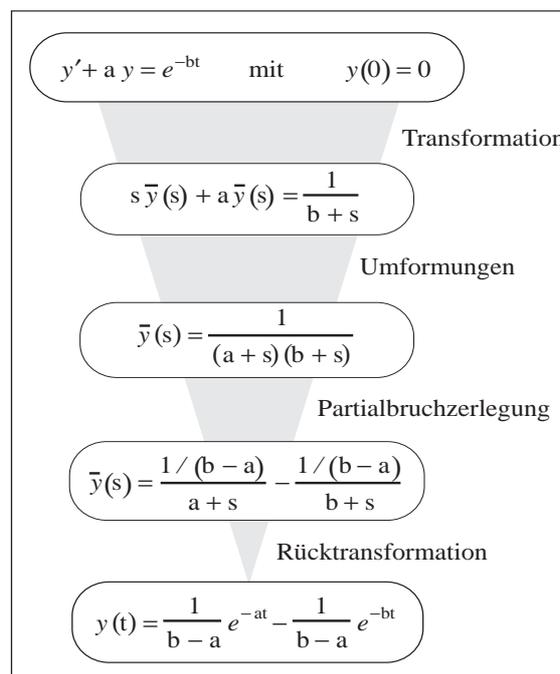


Abb C.1 Lösung einer Differentialgleichung mittels Laplace-Transformation

① bedeutet: anstelle  $s=0$ , d.h. nachdem die Ableitung gebildet wurde, muß  $s=0$  gesetzt werden.

Die Partialbruchzerlegung erfolgt dabei nach folgender Formel:

$$\bar{f}(s) = \sum_{i=1}^k \frac{C_i}{s + a_i}$$

Für das Beispiel bedeutet das

$$\begin{aligned} \frac{1}{(a+s)(b+s)} &= \frac{C_1}{s+a} + \frac{C_2}{s+b} \\ &= \frac{(C_1b + C_2a) + (C_1 + C_2)s}{(a+s)(b+s)} \end{aligned}$$

Es folgen

$$C_1b + C_2a = 1 \quad \text{und} \quad C_1 + C_2 = 0$$

Durch Lösen dieses Gleichungssystems ergeben sich

$$C_1 = \frac{1}{b-a} \quad \text{und} \quad C_2 = -\frac{1}{b-a}$$

| $f(t), t \geq 0$ | $\bar{f}(s)$  |
|------------------|---|
| $f(t)$           | $f(s) = \int_0^{\infty} f(t) e^{-st} dt$                      |
| $c$              | $\frac{c}{s}$   |
| $t$              | $\frac{1}{s^2}$   |
| $t^a$            | $\frac{\Gamma(a+1)}{s^{a+1}} \quad \text{mit } a > -1$        |
| $t^a$            | $\frac{a!}{s^{a+1}} \quad \text{mit } a = 0, 1, 2, \dots$     |
| $t^{-1/2}$       | $\frac{\sqrt{\pi}}{\sqrt{s}}$                                 |
| $e^{-at}$        | $\frac{1}{s+a}$   |
| $t e^{-at}$      | $\frac{1}{(s+a)^2}$   |
| $t^b e^{-at}$    | $\frac{b!}{(s+a)^{b+1}} \quad \text{mit } b = 0, 1, 2, \dots$ |

Tab C.1 Die wichtigsten Laplace-Transformierten

|                           |  |           |
|---------------------------|--|-----------|
| $a f(t) + b g(t)$         | $a \bar{f}(s) + b \bar{g}(s)$                        |           |
| $f'(t)$                   | $s \bar{f}(s) - f(0)$                                | Ableitung |
| $f''(t)$                  | $s^2 \bar{f}(s) - s f(0) - f'(0)$                    |           |
| $f^{(n)}(t)$              | $s^n \bar{f}(s) - \sum_{i=1}^n s^{n-i} f^{(i-1)}(0)$ |           |
| $\int_0^t f(k) dk$        | $\frac{1}{s} \bar{f}(s)$                             | Integral  |
| $\int_0^t f(t-k) g(k) dk$ | $\bar{f}(s) \bar{g}(s)$                              | Faltung   |
| $e^{at} f(t)$             | $\bar{f}(s-a)$                                       |           |
| $t f(t)$                  | $-\bar{f}'(s)$                                       |           |
| $t^n f(t)$                | $(-1)^n \bar{f}^{(n)}(s)$                            |           |
| $\frac{1}{t} f(t)$        | $\int_s^\infty \bar{f}(k) dk$                        |           |

Tab C.2 Eigenschaften von Laplace-Transformationen

## C.2 Laplace-Stieltjes-Transformation

Bildet man die Ableitung der Laplace-Transformation [C.1], so erhält man die *Laplace-Stieltjes-Transformation*

$$f^*(s) = \int_0^\infty e^{-st} df(t) \quad [\text{C.2}]$$

Die Verwandtschaft der beiden Transformationen kann also folgendermaßen beschrieben werden

$$f^*(s) = \bar{f}'(s) = s \bar{f}(s) - f(s) \quad [\text{C.3}]$$

Daher können zur Transformation die Tabellen für Laplace-Transformationen herangezogen werden.

Soll nun die Laplace-Stieltjes-Transformierte der Verteilungsfunktion  $F(\cdot)$  einer kontinuierlichen Zufallsvariable gebildet werden, so gelten folgende Bedingungen:

$$F(0) = 0 \quad \text{und} \quad F(\infty) = 1$$

Daraus ergibt sich

$$F^*(s) = s \bar{F}(s) \quad [\text{C.4}]$$

Beispiel C.2

Am Beispiel der Exponentialverteilung

$$F(t) = 1 - e^{-\Theta t}$$

ergibt sich laut Gleichung [C.4]

$$F^*(s) = s\bar{F}(s) = s\left(\frac{1}{s} - \frac{1}{s + \Theta}\right) = \frac{\Theta}{s + \Theta}$$

Die Laplace-Stieltjes-Transformierte einer Verteilungsfunktion  $F(\cdot)$  kann auch ermittelt werden, indem die Laplace-Transformierte ihrer Dichtefunktion  $f(\cdot)$  gebildet wird.

$$F^*(s) = \bar{f}(s)$$

[C.5]

Beispiel C.3

Für die Exponentialverteilung aus Beispiel C.2 mit Dichtefunktion

$$f(t) = \Theta e^{-\Theta t}$$

ergibt sich somit die selbe Lösung wie in Beispiel C.2

$$F^*(s) = \bar{f}(s) = \Theta\left(\frac{1}{s + \Theta}\right) = \frac{\Theta}{s + \Theta}$$

**C.3 Momenterzeugende Funktionen**

Wird Satz C.3 genauer betrachtet, so stellt sich die Frage, ob die Momente nicht noch etwas einfacher errechnet werden können. Die Antwort heißt *momenterzeugende Funktion*:

$$M(s) = \mathbb{E}(e^{sX}) = \begin{cases} \int_0^{\infty} e^{sx} f(x) dx & \text{für kont. X mit Dichte } f(\cdot) \\ \sum_{x_i} e^{\Theta x_i} p(x_i) & \text{für diskrete X} \end{cases}$$

[C.6]

Wird die Dichtefunktion  $f(\cdot)$  einer Zufallsvariablen herangezogen, so läßt sich die momenterzeugende Funktion aus der Laplace-Transformierten der Dichtefunktion ableiten:

$$M(s) = \bar{f}(-s) \quad \textcircled{2}$$

[C.7]

Nun lassen sich die Momente nach folgender Formel berechnen:

$$\mathbb{E}(X^k) = \left. \frac{d^k M(s)}{ds^k} \right|_{s=0} \quad [\text{C.8}]$$

#### Beispiel C.4

Für die Exponentialverteilung lautet die momenterzeugende Funktion

$$M(s) = \bar{f}(-s) = \frac{\Theta}{\Theta - s}$$

Der Erwartungswert (das 1. Moment) läßt sich wie folgt ermitteln:

$$\mathbb{E}(.) = \left. \frac{dM(s)}{ds} \right|_{s=0} = \left. \frac{\Theta}{(\Theta - s)^2} \right|_{s=0} = \frac{1}{\Theta}$$

### C.4 Z-Transformation (Erzeugende Funktion)

Die *Z-Transformation*  $F(z)$  einer diskreten Funktion  $f_t$  ist wie folgt definiert

$$F(z) = \sum_{t=0}^{\infty} f_t z^t = f_0 + f_1 z + f_2 z^2 + \dots \quad [\text{C.9}]$$

Die wichtigsten Z-Transformierten sind in Tabelle C.3 aufgelistet. Die Z-Transformation besitzt einige mathematisch günstige Eigenschaften, auf die hier jedoch nicht genauer eingegangen werden soll, da sie denen der Laplace-Transformation äußerst ähnlich sind. Eine Auflistung dieser Eigenschaften findet sich in Tabelle C.4. Werden diskrete Verteilungsfunktionen betrachtet, so lassen sich einige weitere Bedingungen ableiten. Es entsteht die *wahrscheinlichkeitserzeugende Funktion*

$$P(z) = \sum_{n=0}^{\infty} p_n z^n \quad [\text{C.10}]$$

mit den Nebenwirkungen

$$P(0) = p_0 \quad \text{und} \quad P(1) = \sum_{n=0}^{\infty} p_n = 1 \quad [\text{C.11}]$$

Eine weitere günstige Eigenschaft ergibt sich aus der Ableitung der wahrscheinlichkeitserzeugenden Funktion:

$$P'(1) = \sum_{n=0}^{\infty} n p_n = \mathbb{E}(N) \quad [\text{C.12}]$$

Auf ähnliche Weise lassen sich auch noch weitere Momente ermitteln.

Bisher wurde der Variablen  $z$  keine Bedeutung beigemessen. Dies ändert sich jedoch, wenn ein System untersucht wird, das durch Kundenankünfte und Kundenabgänge gekennzeichnet ist (z.B. Kaufhauskassen). Be-

zeichnet  $p_n(t)$  die Wahrscheinlichkeit, daß  $n$  Kunden zum Zeitpunkt  $t$  im System sind, und werden Kunden mit Wahrscheinlichkeit  $(1-z)$  markiert, d.h.

$$\begin{aligned} Pr\{\text{Kunde markiert}\} &= 1 - z \\ Pr\{\text{Kunde nicht markiert}\} &= z, \end{aligned}$$

dann bezeichnet  $p_n(t)z^n$  die Wahrscheinlichkeit, daß zum Zeitpunkt  $t$   $n$  nicht markierte Kunden im System sind. Weiters folgt die Wahrscheinlichkeit  $P_{z,t}$ , daß keine markierten Kunden im System sind:

$$P_{z,t} = \sum_{n=0}^{\infty} p_n(t)z^n \quad [\text{C.13}]$$

Diese Methode heißt deshalb auch das *Markieren von Kunden*. Wenden wir uns nun einem Beispiel zu, das die Verwendung von wahrscheinlichkeitserzeugenden Funktionen illustrieren soll:

#### Beispiel C.5

Gegeben ist eine Poisson-Verteilung mit

$$p_n = \frac{e^{-\lambda} \lambda^n}{n!}$$

Die erzeugende Funktion lautet

$$P(z) = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n z^n}{n!} = e^{-\lambda} e^{\lambda z} = e^{-\lambda(1-z)}$$

Dann läßt sich der Erwartungswert wie folgt berechnen

$$E(.) = P'(1) = \lambda e^{-\lambda(1-z)} \Big|_{z=1} = \lambda$$

Abschließend die tabellarische Aufstellung der wichtigsten Z-Transformierten und ihrer Eigenschaften.

$f_t$  mit  $t = 0, 1, 2, \dots$       $F(z)$

$f_t$       $F(z) = \sum_{t=0}^{\infty} f_t z^t$

$c$       $\frac{c}{1-z}$

$t$       $\frac{z}{(1-z)^2}$

$t^2$       $\frac{z(1+z)}{(1-z)^3}$

|   |                             |
|---|-----------------------------|
| $f_t$ mit $t = 0, 1, 2 \dots$           | $F(z)$                      |
| $ca^t$                                  | $\frac{c}{1-az}$            |
| $ta^t$                                  | $\frac{az}{(1-az)^2}$       |
| $t^2 a^t$                               | $\frac{az(1+az)}{(1-az)^3}$ |
| $(t+1)$                                 | $\frac{1}{(1-z)^2}$         |
| $(t+1)a^t$                              | $\frac{1}{(1-az)^2}$        |
| $\frac{1}{n!}(t+n)(t+n-1)\dots(t+1)a^t$ | $\frac{1}{(1-az)^{n+1}}$    |
| $\frac{1}{t!}$                          | $e^z$                       |

Tab C.3 Die wichtigsten Z-Transformierten

|                                  |   |                             |
|----------------------------------|---|-----------------------------|
| $a f_t + b g_t$                  | $a F(z) + b G(z)$                                   |                             |
| $a^t f_t$                        | $F(az)$   |                             |
| $f_{t+1}$                        | $\frac{1}{z}[F(z) + f_0]$                           |                             |
| $f_{t+k}$ mit $k > 0$            | $\frac{F(z)}{z^k} - \sum_{i=1}^k z^{i-k-1} f_{i-1}$ |                             |
| $f_{t-1}$                        | $z F(z)$  |                             |
| $f_{t-k}$ mit $k > 0$            | $z^k F(z)$  |                             |
| $t f_t$                          | $z \frac{dF(z)}{dz}$                                |                             |
| $t(t-1)(t-2)\dots(t-n+1) f_t$    | $z^n \frac{d^n F(z)}{dz^n}$                         |                             |
| $\sum_{k=0}^t f_{t-k} g_k$       | $F(z) G(z)$   | Faltung                     |
| $f_n - f_{n-1}$                  | $(1-z) F(z)$  |                             |
| $\sum_{k=0}^n f_k$               | $\frac{F(z)}{1-z}$                                  | Wahrscheinlichkeitsfunktion |
| $\sum_{k=0}^{\infty} f_k$        | $F(1)$  |                             |
| $\sum_{k=0}^{\infty} (-1)^k f_k$ | $F(-1)$   |                             |
| $f_0$                            | $F(0)$  | Anfangswert                 |

Tab C.4 Eigenschaften von Z-Transformationen

## C.5 Transformationen wichtiger Verteilungen

In den meisten Fällen müssen die Transformaten von Verteilungen gebildet werden (Siehe dazu auch Beispiel C.2 und C.4). Um unnötigen Rechenaufwand zu vermeiden, sind auf den folgenden Seiten die wichtigsten Verteilungen mit deren moment- und wahrscheinlichkeitserzeugenden Funktionen zusammengefaßt.

| Name      | Wahrscheinlichkeitsfunktion<br>$p(x)$   | Parameter                              | Moment –<br>erzeugende Funktion<br>$\mathbb{E}(e^{tx})$ | Wahrscheinlichkeits –<br>erzeugende Funktion<br>$\sum_{m=0}^{\infty} p(m)z^m$ |
|-----------|---|--|---|---|
| Bernoulli | $p(x) = \begin{cases} p & \text{für } x = 1 \\ 1 - p & \text{für } x = 0 \end{cases}$ | $0 \leq p \leq 1$                      | $pe^t + 1 - p$  | $1 - p + pz$  |
| Binom.    | $p(x) = \binom{n}{x} p^x (1-p)^{n-x}$<br>mit $x = 0, 1, \dots, n$                     | $n = 1, 2, \dots$<br>$0 \leq p \leq 1$ | $(pe^t + 1 - p)^n$                                      | $(pz + 1 - p)^n$  |
| Poisson   | $p(x) = \frac{e^{-\lambda} \lambda^x}{x!}$<br>mit $x = 0, 1, 2, \dots$                | $\lambda > 0$                          | $e^{\lambda(e^t - 1)}$                                  | $e^{-\lambda(1-z)}$   |
| Geom.     | $p(x) = p(1-p)^{x-1}$<br>mit $x = 1, 2, \dots$  | $0 \leq p \leq 1$                      | $\frac{p}{1 - p - (1-p)^2 e^t}$                         | $\frac{pz}{1 - (1-p)z}$   |
| Neg.Bin.  | $p(x) = \binom{k+x-1}{x} p^k (1-p)^x$<br>mit $x = 0, 1, 2, \dots$                     | $k > 0$<br>$0 \leq p \leq 1$           | $\left( \frac{p}{1 - (1-p)e^t} \right)^k$               | $\left( \frac{p}{1 - (1-p)z} \right)^k$                                       |

Tab C.5 Erzeugende Funktionen der wichtigsten diskreten Verteilungen

| Name             | Dichtefunktion   | Parameter                                | Moment –<br>erzeugende Funktion   |
|------------------|--|--|---|
|                  | $f(x)$   |  | $E(e^{tx})$   |
| Uniform          | $f(x) = \frac{1}{b-a}$<br>mit $a \leq x \leq b$  | $-\infty < a < b < \infty$               | $\frac{e^{tb} - e^{ta}}{t(b-a)}$  |
| Normal           | $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(x-\mu)}{\sigma}\right]^2}$<br>mit $-\infty < x < \infty$ | $-\infty < \mu < \infty$<br>$\sigma > 0$ | $e^{t\mu + \frac{1}{2}t^2\sigma^2}$   |
| Chi <sup>2</sup> | $f(x) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{(n/2)-1} e^{-x/2}$<br>mit $x > 0$  | $n = 1, 2, \dots$                        | $\left(\frac{1/2}{1/2-t}\right)^{n/2}$  |
| Beta             | $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (x-1)^{\beta-1}$<br>mit $0 < x < 1$        | $\alpha, \beta > 0$                      | $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)} \sum_{j=0}^{\infty} \frac{\Gamma(\alpha+j)t^j}{\Gamma(\alpha+\beta+j)\Gamma(j+1)}$ |

Tab C.6 Momenterzeugende Funktionen der wichtigsten kontinuierlichen Verteilungen

| Name        | Dichtefunktion   | Parameter   | Moment –<br>erzeugende Funktion                   |
|-------------|--|---|---|
|             | $f(x)$   |   | $E(e^{tx})$                                       |
| Exponential | $f(x) = \theta e^{-\theta x}$<br>mit $x > 0$   | $\theta > 0$  | $\frac{\theta}{\theta - t}$                       |
| Gamma       | $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$<br>mit $x > 0$                             | $\alpha, \beta > 0$   | $\left(\frac{1/\beta}{1/\beta - t}\right)^\alpha$ |
| Erlang – k  | $f(x) = \frac{(\theta k)^k}{(k-1)!} x^{k-1} e^{-k\theta x}$<br>mit $x > 0$   | $\theta > 0$<br>$k = 1, 2, \dots$                           | $\left(\frac{k\theta}{k\theta - t}\right)^k$      |
| Hyperexp.   | $f(x) = \sum_{i=1}^k \alpha_i \mu_i e^{-\mu_i x}$<br>mit $x > 0$   | $\mu_i > 0; \alpha_i \geq 0$<br>$\sum_{i=1}^k \alpha_i = 1$ | $\prod_{i=1}^k \frac{\alpha_i \mu_i}{\mu_i - t}$  |
| Hypoexp.    | $f(x) = \sum_{i=1}^k \alpha_i \mu_i e^{-\mu_i x}$<br>mit $\alpha_i = \prod_{j \neq i} \frac{\mu_j}{\mu_j - \mu_i}$ | $\mu_i > 0$   | $\prod_{i=1}^k \frac{\mu_i}{\mu_i - t}$           |

Tab C.7 Momenterzeugende Funktionen der wichtigsten kontinuierlichen Verteilungen

### Literaturhinweise

Eine kurze Zusammenfassung über die Laplace-Transformation bieten [GROS74], [GROS85] und [TRIV82] im Anhang, [BOLC89] in Abschnitt 2.1.2.2, [SCHA73] in Kapitel 1.1 und [ALLE78] in Kapitel 2.9. Die Laplace-Stieltjes-Transformation wird lediglich von [GROS74], [GROS85] und [ALLE 78] angeschnitten. Sehr ausführlich behandelt wird diese Materie von [KLEI75] in Appendix 1.

Informationen über die Z-Transformation sind ebenfalls bei [KLEI75] in Appendix 1, [GROS74] und [GROS85] im Anhang, [TRIV82], in Kapitel 2.7, [ALLE78] in Kapitel 2.9, [BOLC89] in Abschnitt 2.1.2.1 und [COOP72] in Kapitel 2-4 zu finden.

Das Markieren von Kunden ist bei [COOP72] in Kapitel 2-4 und bei [KLEI75] in Kapitel 7 (The Method of Collective Marks) zu finden.

Die momenterzeugende Funktion wird in [TRIV82] in Kapitel 4.4 und bei [GROS74], [GROS85] im Anhang behandelt.

Vollständige Tabellierungen sind in [GROS74], [GROS85] und in diversen Formelsammlungen wie [BART89] zu finden.

---

## D. Wahrscheinlichkeitstheoretische Grundlagen

Bei vielen Untersuchungen und Experimenten sind genaue Aussagen oft unmöglich. Die Wahrscheinlichkeitstheorie und die Statistik beschäftigen sich mit diesen Unsicherheiten. Wir werden in diesem Abschnitt einige grundsätzliche Sachverhalte und Beziehungen erläutern.

### D.1 Zufallsvariablen

Wird ein statistisches Experiment durchgeführt, so läßt sich der Versuchsausgang als *Ereignis*,<sup>①</sup> welches mit einer bestimmten Wahrscheinlichkeit eintritt, darstellen. Eines der einfachsten Experimente ist das *Bernoulli Experiment*. Es kennt nur zwei Ereignisse, die mit Wahrscheinlichkeiten  $p$  und  $(1-p)$  eintreten.

#### Beispiel D.1

Peter und Paul werfen eine Münze. Da anzunehmen ist, daß es sich um eine unverfälschte Münze handelt, ist die Wahrscheinlichkeit für das Ereignis "Kopf" gleich der für das Ereignis "Zahl", also 50%.

Die Summe aller möglichen Ereignisse ergibt den *Ereignisraum*. Bei einem Experiment muß immer ein Ereignis aus dem Ereignisraum eintreten, mit anderen Worten, die Wahrscheinlichkeit, daß ein Ereignis aus dem Ereignisraum eintritt, ist gleich 1. Beim Bernoulli Experiment besteht der Ereignisraum aus zwei Elementen  $\{0, 1\}$ .<sup>②</sup> Nun kann man ein statistisches Experiment durch eine Funktion, die einen bestimmten Wert aus dem Ereignisraum annimmt, beschreiben. Eine solche Funktion heißt *Zufallsvariable*. Für das Bernoulli Experiment gilt also

$$Pr\{X = 0\} = 1 - p \quad \text{und} \quad Pr\{X = 1\} = p$$

Eine Zufallsvariable kann auch mehrere Werte annehmen, so gibt es z.B. beim Werfen eines Würfels sechs mögliche Versuchsgänge. Kann eine Zufallsvariable nur bestimmte Werte annehmen, so spricht man von einer *diskreten Zufallsvariablen*. Im Gegensatz dazu stehen die *kontinuierlichen Zufallsvariablen*, sie können alle Werte innerhalb eines bestimmten Bereiches annehmen. Ein Beispiel dafür sind Wasserstandsmessungen.

Vorerst werden wir uns nur mit den diskreten Zufallsvariablen beschäftigen. Es ist üblich, die möglichen Ausgänge des Experimentes durch ganze Zahlen zu beschreiben. Die Darstellung mittels ganzer Zahlen ist aber keine Vorbedingung für ein diskretes Experiment.

Die Wahrscheinlichkeiten für jeden Versuchsausgang können nun in einer *Wahrscheinlichkeitsfunktion* zusammengefaßt werden.

$$p(x) = Pr\{X = x\} \tag{D.1}$$

Weiters muß gelten (siehe vorne)

$$\sum_x Pr\{X = x\} = 1 \tag{D.2}$$

<sup>①</sup>Zusammenfassen mehrerer Versuchsgänge ist möglich, wenn Untergliederung unwichtig ist.

<sup>②</sup> $\{0\}, \{1\}$  sind je nach Experiment als "gut"- "schlecht", "Kopf"- "Zahl", etc. auslegbar.

Wird eine Wahrscheinlichkeitsfunktion kumuliert, so erhält man die (diskrete) *Verteilungsfunktion*.

$$P(x) = \sum_{k=0}^x p(k) = \sum_{k=0}^x Pr\{X = k\} = Pr\{X \leq x\}$$

[D.3]

### Beispiel D.2

Beim Werfen eines Würfels gibt es 6 weitere Möglichkeiten, jede davon mit einer Wahrscheinlichkeit von  $\frac{1}{6}$ , d.h.

$$Pr\{X = k\} = \frac{1}{6} \quad \text{für } k = 1, 2, 3, 4, 5, 6$$

Die Zufallsvariable  $X$  steht für "Die Augenzahl lautet". Die Wahrscheinlichkeitsfunktion formuliert man daher als

$$p(k) = \frac{1}{6}$$

Für die Verteilungsfunktion ergibt sich

$$\begin{array}{lll} P(1) = \frac{1}{6}; & P(2) = \frac{2}{6}; & P(3) = \frac{3}{6}; \\ P(4) = \frac{4}{6}; & P(5) = \frac{5}{6}; & P(6) = 1; \end{array}$$

allgemein ausgedrückt

$$P(k) = \frac{k}{6} \quad \text{für } k = 1, 2, 3, 4, 5, 6$$

Beide Funktionen sind in Abbildung 1 graphisch dargestellt.

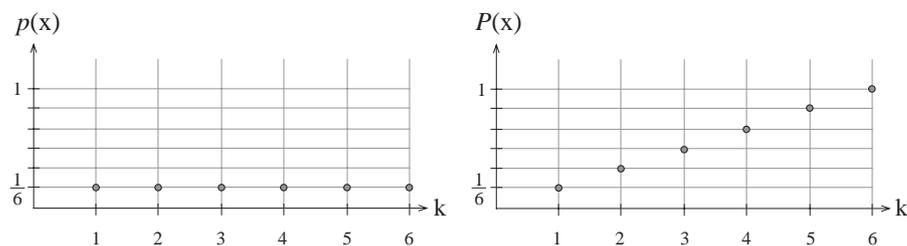


Abb D.1 Wahrscheinlichkeits- und Verteilungsfunktion eines Würfelexperimentes

Zu den einfachsten diskreten Verteilungen zählt die *Bernoulli Verteilung*, sie beschreibt ein Zufallsexperiment mit nur zwei Ausgängen

$$p(x) = \begin{cases} p & \text{für } x = 1 \\ 1 - p & \text{für } x = 0 \end{cases}$$

mit  $0 \leq p \leq 1$

Wird ein Bernoulli Experiment  $n$  mal durchgeführt und sind davon  $x$  Experimente ein Treffer, d.h.  $x$  Einsen und  $n-x$  Nullen, dann entsteht die *Binominalverteilung*

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

mit  $0 \leq p \leq 1$

Wird das Bernoulli Experiment hingegen nur so lange durchgeführt, bis zum ersten Mal ein Treffer vorliegt, d.h. bei  $x$  Versuchen eine Eins und  $x-1$  Nullen, so kann dieser Vorgang durch die *geometrische Verteilung* beschrieben werden.

$$p(x) = p(1-p)^{x-1}$$

mit  $0 \leq p \leq 1$

Weitere diskrete Verteilungen sind in Tabelle D.1 aufgelistet.

Wie bereits erwähnt, gibt es Zufallsvariablen, die alle Werte innerhalb eines bestimmten Intervalls annehmen können. Diese kontinuierlichen Zufallsvariablen werden durch ihre *Dichtefunktionen* bzw. durch ihre (kontinuierliche) *Verteilungsfunktion* beschrieben. Die Dichtefunktion  $f(x)$  entspricht der Wahrscheinlichkeitsfunktion für diskrete Zufallsvariablen. Für die kontinuierliche Verteilungsfunktion  $F(x)$  gilt wie bei der diskreten Verteilungsfunktion:

$$F(x) = Pr\{X \leq x\} \tag{D.4}$$

Da es sich bei der Verteilungsfunktion um eine monoton steigende Funktion (d.h. diese Funktion bleibt gleich oder nimmt zu) handelt, ergeben sich folgende Sachverhalte für  $x < y$ :

$$\begin{aligned} F(x) &\leq F(y) \\ P(x < X \leq y) &= F(y) - F(x) \end{aligned} \tag{D.5}$$

Wird die Verteilungsfunktion  $F(x)$  abgeleitet, so erhält man die Dichte  $f(x)$

$$f(x) = \frac{dF(x)}{dx} \tag{D.6}$$

mit folgenden Eigenschaften

$$f(x) \geq 0 \quad \text{für alle } x$$

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

$$Pr\{X = x\} = \int_x^x f(x) dx = 0$$

$$Pr\{X > x_1\} = \int_{x_1}^{\infty} f(x) dx$$

$$Pr\{x_1 \leq X \leq x_2\} = \int_{x_1}^{x_2} f(x) dx$$

[D.7]

Den Zusammenhang zwischen Dichte und Verteilung erläutert Abbildung D.2.

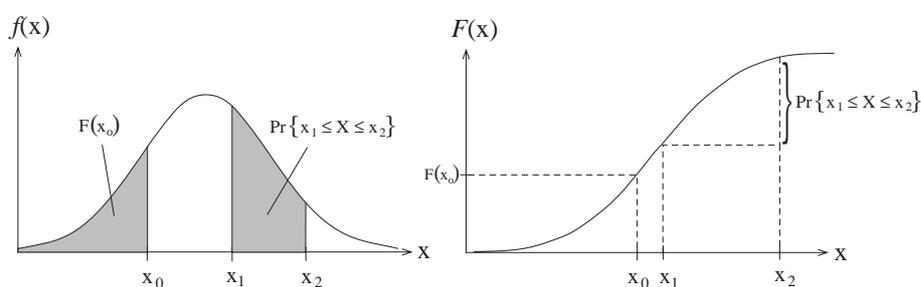


Abb D.2 Zusammenhang zwischen Dichte und Verteilung

Eine der bekanntesten und wichtigsten kontinuierlichen Verteilungen ist die *Normalverteilung*. Diese besitzt zwei Parameter  $\mu$  und  $\sigma^2$ . Die Wahrscheinlichkeit einer solchen mit  $(\mu, \sigma^2)$  normalverteilten Zufallsgröße  $z$  kann durch Normierung dieser

$$x = \frac{z - \mu}{\sigma}$$

aus der *Standard Normalverteilung*

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{s^2}{2}} ds$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

[D.8]

ermittelt werden. Die Kurven für Dichte und Verteilung der Standard-Normalverteilung sind in Abbildung D.3 dargestellt

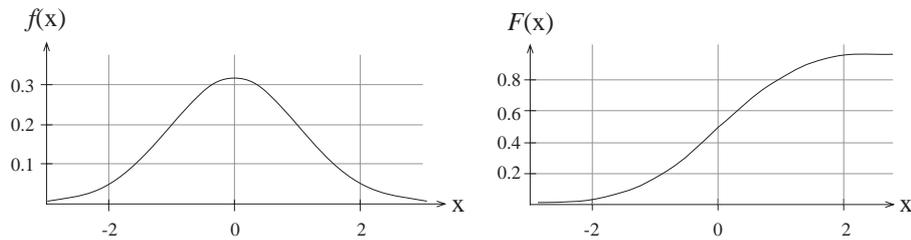


Abb D.3 Dichte und Verteilung der Standard-Normalverteilung

Eine weitere, für die Warteschlangentheorie nicht unwichtige Verteilung ist die Gammaverteilung. Aus ihr können u.a. die Erlangverteilung und die Exponentialverteilung abgeleitet werden. Wichtige Merkmale zu dieser und anderen kontinuierlichen Verteilungen sind in Tabelle D.2 und D.3 aufgelistet.

Neben diskreten und kontinuierlichen Verteilungen gibt es auch sogenannte *gemischte Verteilungen*. Ein Beispiel dafür ist die Verteilung der Lebensdauer von Glühlampen, da eine positive Wahrscheinlichkeit existiert, daß eine Glühlampe bereits schadhaft ausgeliefert wird. Weiters kann es passieren, daß die Glühlampe schon nach einigen Betriebsstunden ausfällt. Je länger die Lampe brennt, desto höher die Wahrscheinlichkeit eines Ausfalls. Die Verteilungsfunktion könnte so aussehen:

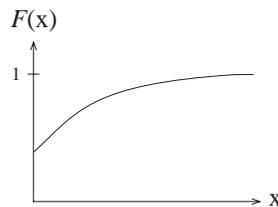


Abb D.4. Beispiel für die Verteilung einer gemischten Zufallsgröße

## D.2 Elementare Größen

Da Dichte- und Verteilungsfunktionen manchmal schwer zu handhaben bzw. überhaupt nicht zu ermitteln sind, versucht man andere Möglichkeiten zur Beschreibung von Zufallsvariablen heranzuziehen. Eine davon ist der *Erwartungswert*:

$$\mu = \mathbb{E}(X) = \begin{cases} \sum x p(x) & \text{für diskrete } X \\ \int_{-\infty}^{\infty} x f(x) dx & \text{für kontinuierliche } X \end{cases}$$

[D.9]

Der Erwartungswert der Standard Normalverteilung ist 0, der der Normalverteilung ist  $\mu$ . Eine weitere wichtige Größe ist die *Varianz*.

$$\sigma^2 = \text{Var}(X) = \mathbb{E}(X - \mathbb{E}(X))^2 = \begin{cases} \sum (x - \mathbb{E}(X))^2 p(x) & \text{für diskrete } X \\ \int_{-\infty}^{\infty} (x - \mathbb{E}(X))^2 f(x) dx & \text{für kontinuierliche } X \end{cases}$$

[D.10]

Aus der Varianz läßt sich leicht die *Streuung* ermitteln:

$$\sigma = \text{Streuung}(X) = \sqrt{\text{Var}(X)}$$

Im Gegensatz zum Erwartungswert sagen Streuung und Varianz etwas über die Form der zugrunde liegenden Verteilung aus. Je höher die Streuung/Varianz, desto breiter (und flacher) wird die Dichtefunktion und umso höher die Wahrscheinlichkeit für außenliegende Punkte. Das selbe gilt für diskrete Variablen - auch hier werden bei Erhöhung der Varianz Wahrscheinlichkeitsfunktion und Verteilungsfunktion in die Länge gezogen.

Aus Erwartungswert und Streuung läßt sich eine weitere Größe, der *Variationskoeffizient*, bestimmen

$$c = \text{Variationskoeffizient}(X) = \frac{\text{Streuung}(X)}{\mathbb{E}(X)} = \frac{\sigma}{\mu} \quad c = \text{Variationskoeffizient}(X) = \frac{\text{Streuung}(X)}{\mathbb{E}(X)} = \frac{\sigma}{\mu} \quad [\text{D.11}]$$

Ist  $c = \sigma = \sigma^2 = 0$ , so bedeutet dies, daß die Zufallsvariable mit Sicherheit einen festen Wert ( $\mu$ ) annimmt. Wird der Erwartungswert der n-ten Potenz einer Zufallsvariable  $X$  gebildet, so erhält man das n-te *Moment*  $\mathbb{E}(X^n)$  dieser Zufallsvariable  $n$ . Das erste Moment entspricht dem Erwartungswert und aus dem 2. Moment läßt sich mit dem *Verschiebungssatz*

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \quad [\text{D.12}]$$

die Varianz berechnen. Der Vorteil bei der Verwendung der Momente liegt in ihrer Berechnung. Mit Hilfe der momenterzeugenden Funktion läßt sich ein Integral bzw. eine Summe auf eine Ableitung reduzieren.

### D.3 Beziehungen zwischen mehrdimensionalen Zufallsvariablen

Bisher wurden nur einzelne Zufallsvariablen betrachtet. Allerdings gibt es Experimente, deren Ausgang von den Werten mehrerer Zufallsvariablen abhängt. Eine Möglichkeit der Beschreibung ist die *Verbundwahrscheinlichkeit*

$$\text{Pr}\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$$

für diskrete Zufallsvariablen und die *Verbundverteilung*

$$\text{Pr}\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}$$

für kontinuierliche Zufallsvariablen.

#### D.3.1 Unabhängigkeit

Die diskreten Zufallsvariablen  $X_1, X_2, \dots, X_n$  sind (statistisch) *unabhängig*, falls gilt:

$$\text{Pr}\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\} = \text{Pr}\{X_1 = x_1\} \text{Pr}\{X_2 = x_2\} \cdots \text{Pr}\{X_n = x_n\} \quad [\text{D.13}]$$

Für kontinuierliche, unabhängige Zufallsvariablen gilt ähnliches:

$$Pr\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\} = Pr\{X_1 \leq x_1\} Pr\{X_2 \leq x_2\} \cdots Pr\{X_n \leq x_n\} \quad [D.14]$$

### Beispiel D.3

Peter und Paul werfen wieder einmal eine (ideale) Münze. Bei zweimaligem Wurf kann das Experiment durch folgende Zufallsvariablen beschrieben werden:

$X_1$  ... Ergebnis beim 1. Wurf

$X_2$  ... Ergebnis beim 2. Wurf

$X_3$  ... Gesamtergebnis

Wird für "Kopf" eine 0 und für "Zahl" eine 1 vergeben, so läßt sich  $X_3$  als Summe

$$X_3 = X_1 + X_2$$

darstellen. Allerdings ist damit ein kleiner Informationsverlust verbunden, da sich die Reihenfolge von "Kopf" und "Zahl" aus  $X_3$  allein nicht eindeutig bestimmen läßt.

$X_1$  und  $X_2$  sind voneinander unabhängig, da gilt

$$Pr\{X_1 = i, X_2 = j\} = Pr\{X_1 = i\} Pr\{X_2 = j\} = \frac{1}{4} \quad \text{für } i, j \in \{0, 1\}$$

$X_1$  und  $X_3$  sind voneinander abhängig:

$$Pr\{X_1 = 1, X_3 = 2\} = Pr\{X_1 = 1\} Pr\{X_2 = 1\} = \frac{1}{4}$$

und

$$\begin{aligned} Pr\{X_1 = 1\} &= \frac{1}{2} \\ Pr\{X_3 = 2\} &= Pr\{X_1 = 1, X_2 = 1\} = \frac{1}{4} \\ Pr\{X_1 = 1\} Pr\{X_3 = 2\} &= \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \end{aligned}$$

stimmen nicht überein.

### D.3.2 Bedingte Wahrscheinlichkeit

Bei manchen Experimenten kann es wichtig sein, den Wert einer Zufallsvariablen  $X_1$  von den Werten anderer Zufallsvariablen  $X_2, \dots, X_n$  abhängig zu machen. Die Wahrscheinlichkeit für  $X_1 = x_1$ , falls die Bedingungen  $X_2 = x_2, \dots, X_n = x_n$  bereits erfüllt sind, lautet

$$Pr\{X_1 = x_1 \mid X_2 = x_2, \dots, X_n = x_n\} = \frac{Pr\{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}}{Pr\{X_2 = x_2, \dots, X_n = x_n\}} \quad [D.15]$$

Diese Wahrscheinlichkeit heißt *bedingte Wahrscheinlichkeit*. Für kontinuierliche Zufallsvariablen wird die bedingte Wahrscheinlichkeit folgendermaßen definiert

$$Pr\{X_1 \leq x_1 \mid X_2 \leq x_2, \dots, X_n \leq x_n\} = \frac{Pr\{X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n\}}{Pr\{X_2 \leq x_2, \dots, X_n \leq x_n\}} \quad [\text{D.16}]$$

#### Beispiel D.4

Im Anschluß an Beispiel D.3 fragen wir uns, wie hoch die Wahrscheinlichkeit für  $X_3 = 1$  ist, falls  $X_1$  bereits den Wert 0 angenommen hat. Es ergibt sich

$$Pr\{X_3 = 1 \mid X_1 = 0\} = \frac{Pr\{X_1 = 0, X_3 = 1\}}{Pr\{X_1 = 0\}} = \frac{1/4}{1/2} = \frac{1}{2}$$

### D.3.3 Zentraler Grenzwertsatz

Bezeichnen  $X_1, X_2, \dots, X_n$  unabhängige und identisch verteilte Zufallsvariablen mit Erwartungswert  $\bar{X}$  und Varianz  $\sigma_X^2$ , so gilt für deren Summe

$$S_n = \sum_{i=1}^n X_i$$

mit Erwartungswert  $n\mu$  und Varianz  $n\sigma_X^2$ , daß sich die Verteilung von  $\frac{S_n - n\mu}{\sqrt{n}\sigma_X}$  der Standard Normalverteilung nähert.

#### Beispiel D.5

Bei genauer Betrachtung von Beispiel D.2 fällt auf, daß  $X_3$  anders verteilt ist als  $X_1$  und  $X_2$ . Abbildung D.5 verdeutlicht, daß der Erwartungswert von  $X_1$  und  $X_2$  gleich  $1/2$ , der von  $X_3$  gleich 1 ist. Die Varianz im ersten Fall ist  $1/4$  und im zweiten Fall  $1/2$ .

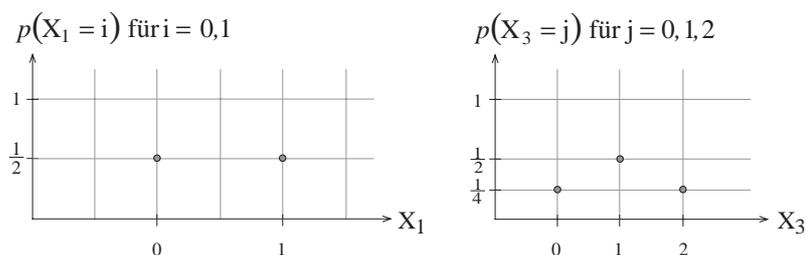


Abb D.5 Wahrscheinlichkeitsfunktion für  $X_1$  und  $X_3$ .

Je mehr unabhängig identische Zufallsvariablen man also addiert, desto stärker nähert sich die Wahrscheinlichkeits- bzw. Dichtefunktion der Dichtefunktion der Normalverteilung.

Diese Methode des Aufsummierens gleichverteilter Zufallsvariablen kann für die Erzeugung normalverteilter Zufallsgrößen verwendet werden.

## D.4 Mehrdimensionale elementare Größen

Im vorigen Abschnitt wurde u.a. der Sachverhalt der Abhängigkeit erläutert. Eines der Ziele dieses Abschnittes ist es, ein Maß für diese Abhängigkeit zu finden. Ein solches Maß ist die *Kovarianz*.

$$\text{Cov}(X, Y) = \mathbb{E}(X - \mathbb{E}(X))(Y - \mathbb{E}(Y)) \quad [\text{D.17}]$$

$$= \begin{cases} \sum_x \sum_y (x - \mathbb{E}(X))(y - \mathbb{E}(Y)) p(x, y) & \text{für diskrete } X, Y \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mathbb{E}(X))(y - \mathbb{E}(Y)) f(x, y) dx dy & \text{für kontinuierliche } X, Y \end{cases}$$

Ein Spezialfall der Kovarianz ist die Varianz

$$\text{Cov}(X, X) = \text{Var}(X) = \sigma_X^2$$

Bei unabhängigen  $X, Y$  ist

$$\text{Cov}(X, Y) = 0$$

Wird die Kovarianz normiert, so erhält man den *Korrelationskoeffizienten*.

$$\text{Cor}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} =$$

[D.18]

Für unabhängige  $X, Y$  gilt

$$\text{Cor}(X, Y) = 0$$

Man bezeichnet  $X$  und  $Y$  in diesem Fall als *unkorreliert*. Zu beachten ist, daß unabhängige Zufallsvariablen immer unkorreliert sind, aber nicht umgekehrt.

Mit diesen neuen Größen können wir nun auch die Summen von Erwartungswerten und Varianzen bilden:

$$\mathbb{E}\left(\sum_{i=1}^n k_i X_i\right) = \sum_{i=1}^n k_i \mathbb{E}(X_i) \quad \text{mit } k_i \text{ beliebig konstant} \quad [\text{D.19}]$$

$$\mathbb{E}\left(\sum_{i=1}^n k X_i\right) = k \sum_{i=1}^n \mathbb{E}(X_i) \quad \text{mit } k \text{ beliebig konstant}$$

$$\text{Var}\left(\sum_{i=1}^n k_i X_i\right) = \sum_{i=1}^n k_i^2 \text{Var}(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n k_i k_j \text{Cov}(X_i, X_j)$$

[D.20]

Für unabhängige  $X_i, i=1\dots n$  ergibt sich zusätzlich:

$$\mathbb{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbb{E}(X_i)$$

[D.21]

$$\text{Var}\left(\sum_{i=1}^n k_i X_i\right) = \sum_{i=1}^n k_i^2 \text{Var}(X_i) \quad [\text{D.22}]$$

$$\text{Var}\left(\sum_{i=1}^n k X_i\right) = k^2 \sum_{i=1}^n \text{Var}(X_i)$$

## D.5 Statistische Ungleichungen und Gesetze

Eine der bekanntesten Methoden, Aussagen über den Bereich der Realisierung einer Zufallsvariablen zu machen, ist die *Chebyshev'sche Ungleichung*:

$$\text{Pr}\{|X - \mu| \geq x\} \leq \frac{\sigma_X^2}{x^2} \quad [\text{D.23}]$$

Aus dieser Gleichung folgt u.a. auch das *Gesetz der großen Zahlen*. Es besagt, daß sich der Mittelwert einer Summe unabhängig identisch verteilter Zufallsvariablen dem Erwartungswert nähert, wenn  $n$  gegen unendlich geht.

Untersucht man einander paarweise ausschließende Ereignisse  $\{H_i\}$ , die vereinigt den Ereignisraum ergeben, so kann die Wahrscheinlichkeit eines Ereignisses  $E$  mit Hilfe des *Satzes der totalen Wahrscheinlichkeit* berechnet werden.

$$\text{Pr}\{E\} = \sum_{i=1}^n \text{Pr}\{E | H_i\} \text{Pr}\{H_i\} \quad [\text{D.24}]$$

Daraus läßt sich die *Bayes'sche Formel* ableiten

$$\text{Pr}\{H_i | E\} = \frac{\text{Pr}\{E | H_i\} \text{Pr}\{H_i\}}{\sum_j \text{Pr}\{E | H_j\} \text{Pr}\{H_j\}} \quad [\text{D.25}]$$

Dabei wird  $\text{Pr}\{H_i | E\}$  als *a posteriori Wahrscheinlichkeit* bezeichnet.

| Name      | Wahrscheinlichkeitsfunktion<br>$p(x)$   | Parameter                              | Erwartungswert<br>$E(X)$ | Varianz<br>$\sigma^2 = E(X - E(X))^2$ |
|-----------|---|--|--------------------------|---------------------------------------|
| Bernoulli | $p(x) = \begin{cases} p & \text{für } x = 1 \\ 1 - p & \text{für } x = 0 \end{cases}$ | $0 \leq p \leq 1$                      | $p$                      | $p(1 - p)$                            |
| Binom.    | $p(x) = \binom{n}{x} p^x (1 - p)^{n-x}$<br>mit $x = 0, 1, \dots, n$                   | $n = 1, 2, \dots$<br>$0 \leq p \leq 1$ | $np$                     | $np(1 - p)$                           |
| Poisson   | $p(x) = \frac{e^{-\lambda} \lambda^x}{x!}$<br>mit $x = 0, 1, 2, \dots$                | $\lambda > 0$                          | $\lambda$                | $\lambda$                             |
| Geom.     | $p(x) = p(1 - p)^{x-1}$<br>mit $x = 1, 2, \dots$                                      | $0 \leq p \leq 1$                      | $\frac{1}{p}$            | $\frac{1-p}{p^2}$                     |
| Neg. Bin. | $p(x) = \binom{k+x-1}{x} p^k (1-p)^x$<br>mit $x = 0, 1, 2, \dots$                     | $k > 0$<br>$0 \leq p \leq 1$           | $\frac{k(1-p)}{p}$       | $\frac{k(1-p)}{p^2}$                  |

Tab D.1 Die wichtigsten Verteilungen

| Name             | Dichtefunktion   | Parameter                                | Erwartungswert                | Varianz  |
|------------------|--|--|-------------------------------|--|
| Uniform          | $f(x) = \frac{1}{b-a}$<br>mit $a \leq x \leq b$  | $-\infty < a < b < \infty$               | $\frac{a+b}{2}$               | $\sigma^2 = \frac{(b-a)^2}{12}$                        |
| Normal           | $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{(x-\mu)}{\sigma}\right]^2}$<br>mit $-\infty < x < \infty$ | $-\infty < \mu < \infty$<br>$\sigma > 0$ | $\mu$                         | $\sigma^2$   |
| Chi <sup>2</sup> | $f(x) = \frac{1}{2^{n/2} \Gamma(n/2)} x^{(n/2)-1} e^{-x/2}$<br>mit $x > 0$   | $n = 1, 2, \dots$                        | $n$                           | $2n$   |
| Beta             | $f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (x-1)^{\beta-1}$<br>mit $0 < x < 1$        | $\alpha, \beta > 0$                      | $\frac{\alpha}{\alpha+\beta}$ | $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ |

Tab D.2 Die wichtigsten kontinuierlichen Verteilungen

| Name        | Dichtefunktion  | Parameter   | Erwartungswert                        | Varianz  |
|-------------|---|---|---------------------------------------|--|
| Exponential | $f(x) = \theta e^{-\theta x}$<br>mit $x > 0$  | $\theta > 0$  | $\frac{1}{\theta}$                    | $\sigma^2 = \frac{1}{\theta^2}$  |
| Gamma       | $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}$<br>mit $x > 0$                          | $\alpha, \beta > 0$   | $\alpha\beta$                         | $\alpha\beta^2$  |
| Erlang-k    | $f(x) = \frac{(\theta k)^k}{(k-1)!} x^{k-1} e^{-k\theta x}$<br>mit $x > 0$                                      | $\theta > 0$<br>$k = 1, 2, \dots$                           | $\frac{1}{\theta}$                    | $\frac{1}{k\theta^2}$  |
| Hyperexp.   | $f(x) = \sum_{i=1}^k \alpha_i \mu_i e^{-\mu_i x}$<br>mit $x > 0$  | $\mu_i > 0; \alpha_i \geq 0$<br>$\sum_{i=1}^k \alpha_i = 1$ | $\sum_{i=1}^k \frac{\alpha_i}{\mu_i}$ | $2 \sum_{i=1}^k \frac{\alpha_i}{\mu_i^2} - \left(\sum_{i=1}^k \frac{\alpha_i}{\mu_i}\right)^2$ |
| Hypoexp.    | $f(x) = \sum_{i=1}^k \alpha_i \mu_i e^{-\mu_i x}$<br>mit $\alpha_i = \prod_{j=1}^k \frac{\mu_j}{\mu_j - \mu_i}$ | $\mu_i > 0$   | $\sum_{i=1}^k \frac{1}{\mu_i}$        | $\sum_{i=1}^k \frac{1}{\mu_i^2}$   |

Tab D.3 Die wichtigsten kontinuierlichen Verteilungen

Literaturhinweise

Die einschlägige Literatur bietet eine gute Einführung. Ein Beispiel dafür ist [WONN72]. Hier wird, ausgehend von Ereignissen mittels graphischer Darstellungen der Weg zum Verständnis der Statistik gebnet. An ein ähnliches Konzept hält sich [TRIV82], der zusätzlich noch viele Beispiele aus der Informatik beinhaltet.

Die wichtigsten Zusammenhänge sind in [BOLC89] Kapitel 2.1 und [KLEI74] Appendix 2 kurz, aber genau dargestellt.

Auch in [ALLE78] in den Kapiteln 2 und 3 ist eine detaillierte Beschreibung zu finden.

## E. Stochastische Prozesse und stochastische Ketten

Ein stochastischer Prozeß ist die mathematische Abstraktion eines empirischen Prozesses. Dieser wird durch eine Familie von Zufallsvariablen  $\{X(t) : t \in T\}$ <sup>①</sup> dargestellt. Die Indizes der Zufallsvariablen bilden den Parameterraum  $T$ .  $T$  bezeichnet zumeist ein Zeitintervall und  $X(t)$  eine Beobachtung zum Zeitpunkt  $t$ .

### Beispiel E.1

Peter und Paul werfen eine ideale (unverfälschte, nicht manipulierte) Münze. Bei "Kopf" zahlt Peter an Paul 1 Schilling, bei "Zahl" ist es umgekehrt. Den Betrag, den Paul nach  $n$  Münzwürfen verdient hat, können wir graphisch anschaulich darstellen:

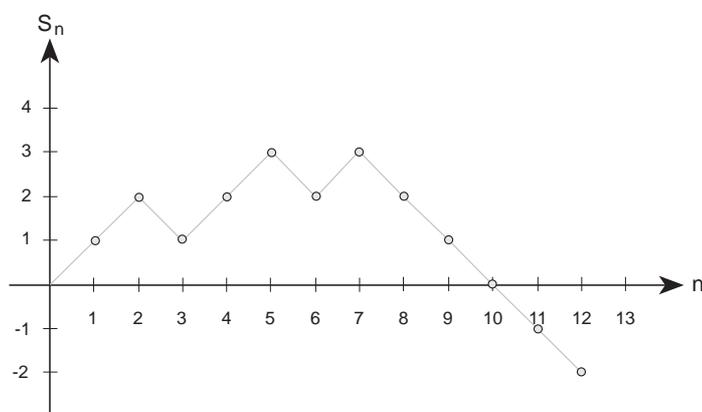


Abb E.1 Darstellung eines möglichen Spielverlaufs  
(Die Verbindungen dienen nur der optischen Untermalung)

Alle hier möglichen Verdienste von Paul stellen einen *stochastischen Prozeß*  $\{S_n : n \in \mathbb{N}_0\}$  dar.<sup>①</sup>

Im Beispiel E.1 könnte  $S_n$  als eine Summe von  $n$  Zufallsvariablen dargestellt werden, d.h.  $S_n = \sum_{i=1}^n E_i = E_1 + E_2 + \dots + E_n$ . Dabei kann jedes  $E_i$  die Werte  $(+1)$  oder  $(-1)$  annehmen (der Wertebereich von  $E_i$  ist also  $\{+1, -1\}$ , d.h.  $E_i \in \{+1, -1\}$ ). Weiters ist aus vorigem Beispiel folgendes zu erkennen:

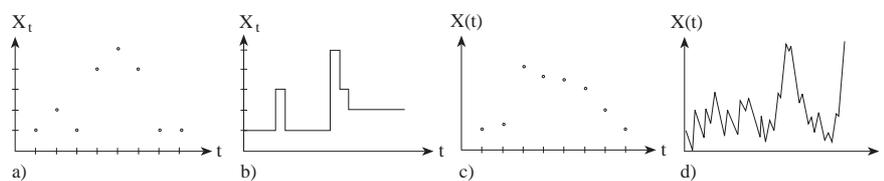
- Wie schon anfangs erwähnt, besteht ein stochastischer Prozeß aus einer Familie von Zufallsvariablen, die wir im folgenden mit  $X_t$ ,  $t \in T$  bezeichnen wollen.
- Der Parameterraum  $T$  ist *diskret* d.h.  $T = \{0, 1, 2, 3, \dots, n\}$ . Weiters ist er *endlich*, da anzunehmen ist, daß Peter und Paul das Spiel irgendwann beenden werden. Ein *kontinuierlicher* Parameterraum wird durch seine Schranken dargestellt, so ist  $T = \{t : -\infty < t < +\infty\}$  ein Beispiel für einen kontinuierlichen, unendlichen Parameterraum.
- Die Distanzen zwischen den einzelnen Zeitpunkten müssen nicht unbedingt gleich (äquidistant) sein. Es ist bei obigem Beispiel auch nur schwer vorstellbar, daß die Münze exakt jede Minute geworfen wird. Wir werden auch in Zukunft die Zeitpunkte mit  $0, 1, 2, \dots$  bezeichnen, ohne daß dadurch Äquidistanz der Zeitpunkte impliziert wird.

<sup>①</sup> Das Symbol  $\in$  gibt an, daß  $t$  ein Element der Menge  $T$  ist.

**Definition E.1** Ein stochastischer Prozeß ist eine Zusammenfassung von Zufallsvariablen, die alle über demselben Wahrscheinlichkeitsraum definiert sind. Werden die einzelnen Zufallsvariablen durch einen Parameter  $t$  gekennzeichnet, so ist der betrachtete stochastische Prozeß die Menge  $\{X(t) : t \in T\}$ . Dabei durchläuft  $t$  einen nichtleeren Parameterraum  $T$ , der für eine Folge von Zeitpunkten (diskret) oder für ein Zeitintervall (kontinuierlich) steht (endlich oder unendlich).

**Definition E.2** Eine stochastische Kette ist ein stochastischer Prozeß, bei dem die Wertebereiche der Zufallsvariablen  $X_t$ ,  $t \in T$  diskret sind. Die Verteilungen der  $X_t$  zu jedem Zeitpunkt  $t$  müssen im Gegensatz zu den Wertebereichen nicht die selben sein.

Abb E.2 Beispiele für stochastische Prozesse



In Abbildung E.2 sind einige Beispiele für stochastische Prozesse mit

- a) diskretem Parameter- & Zustandsraum
  - b) kontinuierlichem Parameter- & diskretem Zustandsraum
  - c) diskretem Parameter- & kontinuierlichem Zustandsraum
  - d) kontinuierlichem Parameter- & Zustandsraum
- dargestellt.

### Literaturhinweise

Eine sehr detaillierte Beschreibung der stochastischen Prozesse und der stochastischen Ketten sind in [STOC78] Kapitel 2 und 1, ebenso in [TAKA59] zu finden.

Eine kurze Erklärung ohne mathematische Umschweife bietet [GROS85] Kapitel 1.10.

Eine sehr verständliche Einführung gibt [TRIV82] Kapitel 6.

Eine kurze Darstellung bietet [LANG92] Kapitel 7.1.2. Es werden auch der Poisson-Prozeß und der Geburten-Todes-Prozeß behandelt.

② Die Menge  $n_0$  gibt die Menge der natürlichen Zahlen  $n$  plus Element  $\{0\}$  an.

$$n_0 = n \cup \{0\}$$

## F. Markovprozesse und Markovketten

Markovprozesse sind stochastische Prozesse, die die *Markoveigenschaft* erfüllen. Diese besagt, daß die Zukunft des Systems nur von der Gegenwart, nicht aber von der Vergangenheit abhängen darf. Der Markovprozeß ist also "gedächtnislos".

**Definition F.1** Ein stochastischer Prozeß  $\{X(t) : t \in T\}$  heißt *Markovprozeß*, falls für eine beliebige Auswahl von  $n$  Werten  $t_1 < t_2 < \dots < t_n$  aus dem Parameterraum und für eine beliebige Auswahl von  $n$  Zuständen  $\{x_1, x_2, \dots, x_n\}$  gilt:

$$Pr\{X(t_n) \leq x_n | X(t_1) = x_1, \dots, X(t_{n-1}) = x_{n-1}\} = Pr\{X(t_n) \leq x_n | X(t_{n-1}) = x_{n-1}\} \quad [F.1]$$

**Definition F.2** Ein Markovprozeß heißt *Markovkette*, wenn sein Zustandsraum diskret ist. Es handelt sich also um eine stochastische Kette  $\{X_t : t \in T\}$ , für die, bei einer beliebigen Auswahl von  $n$  Zuständen  $\{x_1, \dots, x_n\}$ , folgendes gilt:

$$Pr\{X_n = x_n | X_1 = x_1, \dots, X_{n-1} = x_{n-1}\} = Pr\{X_n = x_n | X_{n-1} = x_{n-1}\} \quad [F.2]$$

### F.1 Markovketten mit diskretem Parameterraum

Falls die *Übergangswahrscheinlichkeiten*  $Pr\{X_n = j | X_{n-1} = i\}$  unabhängig von  $n$  sind, heißt die Markovkette *homogen*. Wir können also vereinfachend definieren:

$$p_{ij} = Pr\{X_n = j | X_{n-1} = i\} \quad [F.3]$$

Daraus läßt sich eine *Übergangsmatrix* (der 1-stufigen Übergangswahrscheinlichkeiten) erstellen.

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{12} & \dots \\ p_{21} & p_{22} & \dots \\ \cdot & \cdot & \dots \end{pmatrix} \quad [F.4]$$

Die Wahrscheinlichkeit, in  $m$  Schritten vom Zustand  $i$  in den Zustand  $j$  zu gelangen ist definiert als

$$p_{ij}^{(m)} = Pr\{X_{n+m} = j | X_n = i\} \quad [F.5]$$

und kann mit Hilfe der *Chapman - Kolmogorov - Gleichung* (für Markovketten mit diskretem Parameterraum)

$$p_{ij}^{(m)} = \sum_k p_{ik}^{(m-1)} p_{kj} \quad m = 2, 3, \dots \quad [F.6]$$

errechnet werden. Um vom Zustand  $i$  zum Zustand  $j$  in  $m$  Schritten zu gelangen, müssen zuerst der Weg von  $i$  nach  $k$  in  $m-1$  Schritten und dann der verbleibende Weg von  $k$  nach  $j$  in einem einzigen Schritt zurückgelegt werden. Da die beiden letztgenannten Ereignisse unabhängig sind, ergibt das Produkt ihrer Wahrscheinlichkeiten

die Wahrscheinlichkeit eines möglichen Weges. Werden die Wahrscheinlichkeiten aller möglichen Wege summiert, so erhält man  $p_{ij}^{(m)}$ . Jetzt ist auch absehbar, daß  $p_{ij}^{(m)}$  nicht unbedingt gleich  $p_{ij}^m$  sein muß.

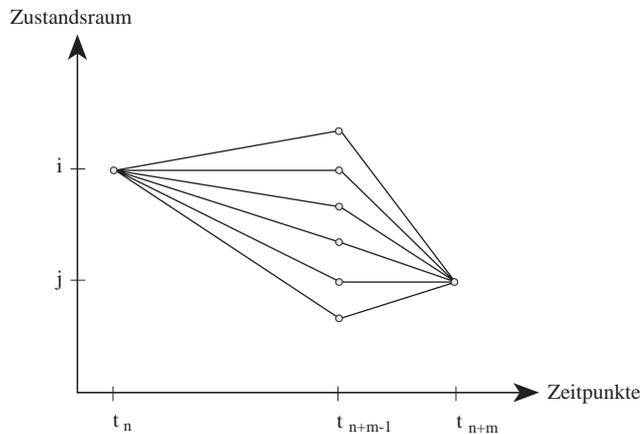


Abb F.1 Um vom Zustand  $i$  im Zeitpunkt  $t_n$  zum Zustand  $j$  im Zeitpunkt  $t_{n+m}$  zu gelangen, kann man im Zeitpunkt  $t_{n+m-1}$  alle möglichen Zustände des Zustandsraumes durchlaufen.

Eine Markovkette ist *irreduzibel*, falls man von jedem Zustand  $i$  ausgehend zu jedem anderen Zustand  $j$  in endlich vielen Schritten gelangen kann.

Definition F.3 Eine Markovkette heißt *irreduzibel*, wenn es für jedes Paar von Zuständen  $i$  und  $j$  eine ganze Zahl  $m$  gibt, für die gilt:

$$p_{ij}^{(m)} > 0. \quad [F.7]$$

Bezeichnet  $f_i^{(m)}$  die Wahrscheinlichkeit, vom Zustand  $i$  aus nach  $m$  Schritten zum ersten Mal in den Zustand  $i$  zurückzukehren,

$$f_i^{(m)} = Pr\{X_m = i, X_k \neq i \text{ für } k = 1, 2, \dots, m-1 | X_0 = i\} \quad [F.8]$$

dann ist die Wahrscheinlichkeit, überhaupt zu  $i$  zurückzukehren, gegeben durch

$$f_i = \sum_{m=1}^{\infty} f_i^{(m)} \quad [9]$$

Definition F.4 Falls  $f_i < 1$  ist, heißt der Zustand  $i$  *transient*. Falls  $f_i = 1$  ist, heißt der Zustand  $i$  *rekurrent*. In diesem Fall ist die *mittlere Rekurrenzzeit* von  $i$ , das ist die mittlere Zeit, um zu  $i$  zurückzukehren, gegeben durch

$$m_i = \sum_{n=1}^{\infty} n f_i^{(n)} \quad [F.10]$$

Ist  $m_i = \infty$ , dann heißt der Zustand  $i$  *nullrekurrent*. Ist  $m_i < \infty$ , dann heißt der Zustand  $i$  *positiv rekurrent*.

Definition F.5 Man spricht von einem *absorbierenden* Zustand  $i$ , falls  $p_{ii} = 1$  ist. Sobald ein solcher Zustand betreten wird, kann er nicht mehr verlassen werden.

Definition F.6 Eine Markovkette heißt *aperiodisch*, falls der größte gemeinsame Teiler der ganzzahligen Menge

$$\left\{ n \in \mathbb{N} \mid p_{ii}^{(n)} > 0 \right\} \quad \text{für alle Zustände } i$$

gleich 1 ist. Ist der größte gemeinsame Teiler größer als 1, so heißt die Markovkette *periodisch*.

Um die obige Definition zu illustrieren, betrachten wir das folgende Beispiel:

Beispiel F.1

Ein Prozeß kehrt in 5 und 7 Schritten zu einem bestimmten Zustand zurück. Der Prozeß ist *aperiodisch*, da eine Rückkehr zu den Zeitpunkten 5,7,10,12,14,15,17,19,20,21,22,24,25,26,27, ... erfolgt.

Betrachten wir einen anderen Prozeß, der nach 3 und 6 Schritten in einen bestimmten Zustand zurückkehrt. Er ist *periodisch*, da eine Rückkehr zu den Zeitpunkten 3, 6, 9, 12, ... erfolgt.

Definition F.7 Eine aperiodische, irreduzible und positiv rekurrente Kette heißt *ergodisch*.

(Siehe auch Abbildung F.3)

Definition F.8 Für eine irreduzible, aperiodische Markovkette existieren immer die Grenzwahrscheinlichkeiten

$$\pi_j = \lim_{n \rightarrow \infty} Pr\{X_n = j\} \quad \text{für alle } j \quad \text{[F.11]}$$

Diese Grenzwahrscheinlichkeiten sind völlig unabhängig von ihrer Anfangsverteilung. Sind alle Zustände transient oder nullrekurrent, so ist  $\pi_j = 0$  für alle  $j$  und es existiert keine stationäre Verteilung. Sind hingegen alle Zustände positiv rekurrent (d.h. ergodisch), dann ist  $\pi_j > 0$  für alle  $j$  und  $\{\pi_j\}$  bezeichnet eine Verteilung mit  $\pi_j = 1/m_j$ . ( $m_j$  bezeichnet die mittlere Rekurrenzzeit, s.v.)

Die Grenzverteilung ist die einzige Lösung des folgenden Gleichungssystems:

$$\begin{aligned} \pi_j &= \sum_{i=0}^{\infty} \pi_i P_{ij} \\ \sum_{i=0}^{\infty} \pi_i &= 1 \end{aligned} \quad \text{[F.12]}$$

Zum vorangegangenen Text abschließend noch ein Beispiel:

Beispiel F.2

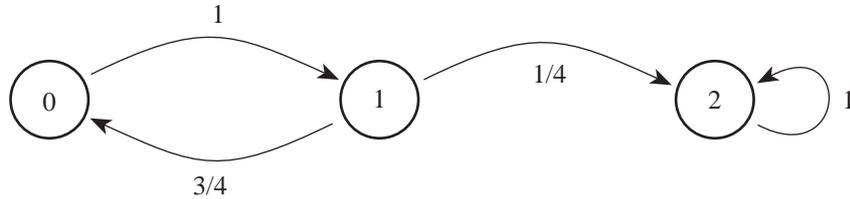


Abb F.2 Zustandsübergangsdiagramm einer endlichen Markovkette

Man erkennt in Abbildung F.2, daß der Zustand 2 von 1 (und damit auch von 0 aus) erreicht werden kann, aber nicht umgekehrt. Die Markovkette ist also nicht irreduzibel. Die Zustände 0 und 1 sind periodisch, da

$$p_{00}^{(n)} = p_{11}^{(n)} > 0 \quad n = 2, 4, 6, 8$$

Der Zustand 2 hingegen ist aperiodisch. Die Wahrscheinlichkeit für die erste Rückkehr in den Zustand 0 in zwei Schritten ist  $f_0^{(2)} = 3/4$ . Für alle  $n \neq 2$  ist  $f_0^{(n)} = 0$ , denn um in mehr als zwei Schritten in den Zustand 0 zurückzukehren, muß der Zustand 0 passiert werden. Also gilt:

$$f_0 = \sum_{n=1}^{\infty} f_0^{(n)} = \frac{3}{4} < 1$$

Das selbe gilt für den Zustand 1 ( $f_0 = f_1$ ). Die Zustände 0 und 1 sind also transient. Der Zustand 2 hingegen ist positiv rekurrent und absorbierend.

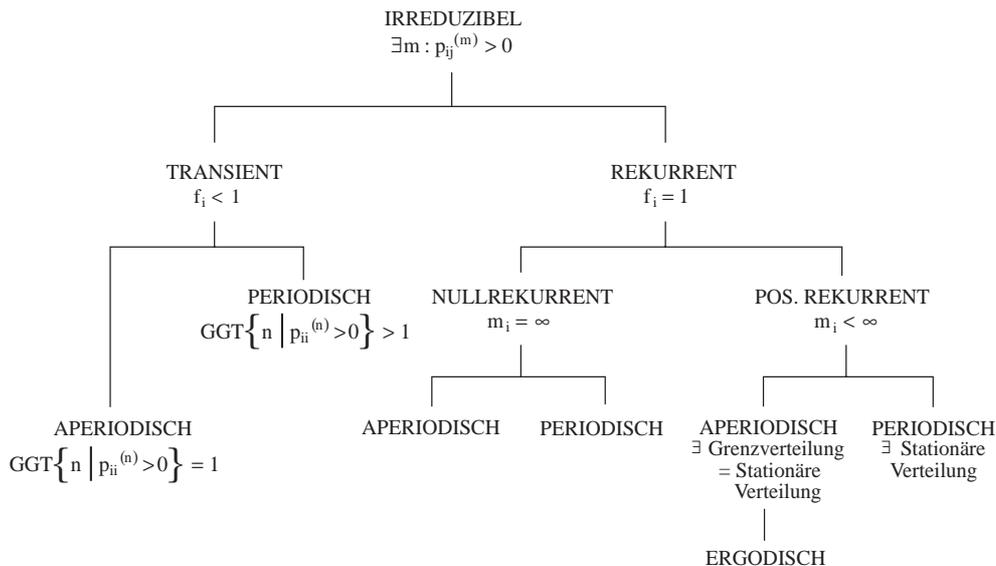


Abb F.3 Einteilung der irreduziblen Markovketten

**Bemerkung F.1** Da ein Markovprozeß laut Definition "gedächtnislos" ist, hat die Zeit, in der in einem bestimmten Zustand verweilt wird, auf den Prozeß keine Auswirkungen. Bei manchen Warteschlangenmodellen kann diese Einschränkung jedoch nicht beibehalten werden (z.B. M/G/1). Die Verweilzeit in einem bestimmten Zustand muß also bekannt sein. Der hier zugrunde liegende stochastische Prozeß ist nicht "gedächtnislos", daher handelt es sich auch um keinen Markovprozeß.

Betrachtet man den stochastischen Prozeß lediglich zu den sogenannten *Regenerationspunkten* (das sind die Zeitpunkte, zu denen ein Zustandsübergang erfolgt) so hängt die zukünftige Entwicklung dieses Prozesses nur von dem soeben erreichten Zustand ab. Die Markoveigenschaft trifft zu und in den ursprünglichen stochastischen Prozeß wird eine Markovkette mit diskretem Parameterraum eingebettet. Diese *eingebetteten Markovketten* sind ein wichtiges Instrument der Warteschlangentheorie.

## F.2 Markovketten mit kontinuierlichem Parameterraum

Die Analyse von Markovketten mit kontinuierlichem Parameterraum erfolgt genauso wie bei Markovketten mit diskretem Parameterraum. Der einzige Unterschied besteht darin, daß ein Zustandswechsel zu jedem Zeitpunkt innerhalb eines bestimmten Zeitabschnittes erfolgen kann.

Das Verhalten einer Markovkette mit kontinuierlichem Parameterraum läßt sich komplett aus den Anfangswahrscheinlichkeiten

$$p_j(0) = Pr\{X_0 = j\} \quad \text{für alle } j \quad [F.13]$$

und den Übergangswahrscheinlichkeiten

$$p_{ij}(s, t) = Pr\{X_t = j \mid X_s = i\} \quad [F.14]$$

bestimmen. Dabei ist die Übergangswahrscheinlichkeit  $p_{ij}(s, t)$  definiert als die Wahrscheinlichkeit, um vom Zustand  $i$  in den Zustand  $j$  innerhalb eines Zeitraumes  $s$  bis  $t$  zu gelangen.

Eine Grundlage zur analytischen Behandlung von Markovketten mit kontinuierlichem Parameterraum bildet die Chapman-Kolmogorov Gleichung. Für beliebige Zeitpunkte  $t > u > s \geq 0$  gilt

$$p_{ij}(s, t) = \sum_r p_{ir}(s, u) p_{rj}(u, t) \quad [F.15]$$

Daraus können die *Kolmogorov-Vorwärts-Gleichung*

$$\frac{\partial p_{ij}(s, t)}{\partial t} = -q_j(t) p_{ij}(s, t) + \sum_{r \neq j} p_{ir}(s, t) q_{rj}(t) \quad [F.16]$$

und die *Kolmogorov-Rückwärts-Gleichung*

$$\frac{\partial p_{ij}(s, t)}{\partial t} = q_i(s) p_{ij}(s, t) - \sum_{r \neq j} q_{ir}(s) p_{rj}(s, t) \quad [\text{F.17}]$$

abgeleitet werden. Die beiden nicht-negativen kontinuierlichen Funktionen  $q_j(t)$  und  $q_{ij}(t)$  sind wie folgt definiert:

$$\begin{aligned} q_{ij}(t) &= \lim_{h \rightarrow 0} \frac{p_{ij}(t, t+h) - p_{ij}(t, t)}{h} \\ q_j(t) &= \lim_{h \rightarrow 0} \frac{1 - p_{jj}(t, t+h)}{h} \end{aligned} \quad [\text{F.18}]$$

Wenn man eine homogene Markovkette unterstellt, d.h.

$$\begin{aligned} q_{ij}(t) &= q_{ij} & q_j(t) &= q_j \\ p_{ij}(t, t+h) &= p_{ij}(h) \end{aligned}$$

so läßt sich Gleichung [F.16] wie folgt anschreiben

$$\frac{\partial p_{ij}(s, t)}{\partial t} = -q_j p_{ij}(t-s) + \sum_{r \neq j} p_{ir}(t-s) q_{rj} \quad [\text{F.19}]$$

Mit Hilfe des Satzes der totalen Wahrscheinlichkeit läßt sich folgender Zusammenhang herstellen:

$$p_j(t) = \sum_i p_{ij}(t-s) p_i(s) \quad [\text{F.20}]$$

Wird Gleichung [F.19] genauso mit  $p_i(s)$  multipliziert und über alle  $i$  summiert, so ergibt sich

$$\frac{dp_j(t)}{dt} = -q_j p_j(t) + \sum_{r \neq j} p_r(t) q_{rj} \quad [\text{F.21}]$$

Bei irreduziblen (jeder Zustand kann von jedem anderen Zustand erreicht werden) Markovketten mit kontinuierlichem Parameterraum existieren immer die Grenzwahrscheinlichkeiten

$$p_j = \lim_{t \rightarrow \infty} p_j(t) \quad [\text{F.22}]$$

Wie bei den Markovketten mit diskretem Parameterraum sind diese Grenzwahrscheinlichkeiten unabhängig von ihrer Anfangsverteilung.

Da sich diese stationären Wahrscheinlichkeiten mit der Zeit nicht ändern, gilt

$$\frac{dp_j}{dt} = 0$$

Nun kann Gleichung [F.2] zu

$$0 = -q_j p_j + \sum_{r \neq j} p_r q_{rj} \quad [F.23]$$

vereinfacht werden. Zusätzlich muß folgende Bedingung erfüllt sein:

$$\sum_j p_j = 1 \quad [F.24]$$

Dabei stellt  $q_{ij}$  die stationäre Übergangsrate dar. Die  $q_i$  lassen sich wie folgt berechnen:

$$q_i = \sum_{j \neq i} q_{ij} \quad [F.25]$$

Das Gleichungssystem [F.23], [F.24] ist ein wichtiges Instrument in der Warteschlangentheorie. Es ermöglicht u.a. die Berechnung der Zustandswahrscheinlichkeiten von Geburten/Todesprozessen.

### F.3 Verwendung der Matrizenrechnung

Mit Hilfe der Matrizenrechnung können viele der bisher angeführten Gleichungen stark vereinfacht werden. Weiters können die bei der Analyse auftretenden Gleichungssysteme leichter gelöst werden.

Den Ausgangspunkt bildet die Gleichung [F.21]. Sie lautet in Matrixnotation

$$\mathbf{p}'(t) = \mathbf{p}(t) \mathbf{Q} \quad [F.26]$$

wobei

$$\mathbf{p}(t) = (p_0(t), p_1(t), \dots)$$

$$\mathbf{p}'(t) = \left( \frac{dp_0(t)}{dt}, \frac{dp_1(t)}{dt}, \dots \right)$$

$$\mathbf{Q} = \begin{pmatrix} -q_0 & q_{01} & q_{02} & \dots \\ q_{10} & -q_1 & q_{12} & \dots \\ q_{20} & q_{21} & -q_2 & \dots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

Bei Betrachtung der stationären Wahrscheinlichkeiten ergibt sich analog zu Gleichung Nummer [F.23]

$$\mathbf{0} = \mathbf{p} \mathbf{Q} \quad \text{mit } \mathbf{p} = (p_0, p_1, p_2, \dots) \quad [\text{F.27}]$$

Obige Gleichung gilt für Markovketten mit diskretem wie auch kontinuierlichem Parameterraum. Bei Markovketten mit diskretem Parameterraum läßt sich Matrix  $\mathbf{Q}$  aus der Matrix  $\mathbf{P}$  (siehe [F.4]) ermitteln

$$\mathbf{Q} = \mathbf{P} - \mathbf{I} \quad [\text{F.28}]$$

Die Elemente des Vektors lassen sich mit Gleichung [F.11] ermitteln.

#### Literaturhinweise

Eine kurze Zusammenstellung des Materials bieten [LANG92] Kapitel 7, [ALLE78] Kapitel 4, sowie [GROS74] im Anhang.

Etwas genauer behandelt werden die Markovprozesse in [GROS85] Kapitel 1.10. Dort werden auch die eingebetteten Markovketten (Kapitel 1.10.4) und das Langzeitverhalten von Markovprozessen beschrieben. Die eingebetteten Markovketten werden auch von [COOP72] Kapitel 5-5 im Zuge des M/G/1-Modells kurz angeschnitten.

Sehr verständlich ist die Darstellung in [KLEI75] Kapitel 2 und [TRIV82] Kapitel 7 und 8. Eine detaillierte Beschreibung mit mehreren Beispielen befindet sich in [STOCH78] Kapitel 1 und [FELL68] Kapitel 15 und 16.

[HORN92] bringt eine mathematische Kurzdarstellung über Markovketten in diskreter und stetiger Zeit. Auch finden Erneuerungsprozesse und Warteschlangen Eingang.

Besonders erwähnenswert ist [IRAN88] Kapitel 8, mit seiner einleuchtenden Erklärung der Markovketten, ohne dabei Bezug auf die Theorie der Warteschlangen zu nehmen.

---

## G. Bibliographie

- [AGRA85] METAMODELLING -  
A STUDY OF APPROXIMATIONS IN QUEUEING MODELS  
Agrawal S.C.  
MIT-Press 1985
- [ALLE78] PROBABILITY, STATISTICS AND QUEUEING THEORY  
Allen Arnold O.  
Academic Press 1978
- [BART89] MATHEMATISCHE FORMELN, 22. Auflage  
Bartsch Hans Jochen  
VEB Fachbuchverlag Leipzig 1988
- [BHAT68] A STUDY OF THE QUEUEING SYSTEMS M/G/1 AND GI/M/1  
aus LECTURE NOTES IN ECONOMICS AND MATHEMATICAL  
Bhat U.N.  
Springer Verlag 1968
- [BOLC89] LEISTUNGSBEWERTUNG VON RECHENSYSTEMEN MITTELS  
ANALYTISCHER WARTESCHLANGENMODELLE  
Bolch G.  
Teubner 1989
- [COOP72] INTRODUCTION TO QUEUEING THEORY  
Cooper Robert B.  
Mac Millan Company 1972 (3.Auflage 1990 CeePress Books)
- [COX61] QUEUES  
Cox D.R. and Smith W.L.  
Chapman and Hall 1961
- [DOR188] MATHEMATIK FÜR WIRTSCHAFTSINFORMATIKER, BAND I  
Dorninger D., Karigl G.  
Springer Verlag 1988
- [DOR288] MATHEMATIK FÜR WIRTSCHAFTSINFORMATIKER, BAND II  
Dorninger D., Karigl G.  
Springer Verlag 1988
- [FELL68] AN INTRODUCTION TO PROBABILITY THEORY AND ITS  
APPLICATIONS  
Feller W.  
Wiley Interscience 1968
-

- 
- [GROS74] FUNDAMENTALS OF QUEUEING THEORY  
Gross D., Harris C.M.  
Wiley Interscience 1974
- [GROS85] FUNDAMENTALS OF QUEUEING THEORY, 2nd Edition  
Gross D., Harris C.M.  
Wiley Interscience 1985
- [HAIG81] APPLIED PROBABILITY  
Haight F.A..  
Plenum Press 1981
- [HORN92] STOCHASTISCHE PROZESSE, ZUFÄLLIGE PROZESSE  
Skripten zur Vorlesung, Hornik K.  
Technische Universität Wien 1992
- [IRAN88] BASIC STOCHASTIC PROCESSES  
Iranpour R., Chagon P.  
Macmillan Publishing 1988
- [JAIN91] THE ART OF COMPUTER SYSTEMS PERFORMANCE ANALYSIS  
Jain Rai  
Wiley 1991
- [JAIS68] PRIORITY QUEUES  
Jaiswal N.K.  
Academic Press 1968
- [KRAM93] ESTIMATING PARAMETERS OF COX DISTRIBUTIONS  
Kramer Manfred 1993  
aus  
MESSUNG, MODELLIERUNG UND BEWERTUNG VON RECHEN-  
UND KOMMUNIKATIONSSYSTEMEN  
Herausgeber Walke, Spaniol  
Springer Verlag 1993
- [KLEI75] QUEUEING SYSTEMS THEORY, VOL. 1  
Kleinrock Leonard  
Wiley Interscience 1975
- [KLEI76] QUEUEING SYSTEMS COMPUTER APPLICATIONS, VOL.2  
Kleinrock Leonard  
Wiley Interscience 1976
-

- 
- [LANG92] LEISTUNGSANALYSE VON RECHENSYSTEMEN  
Langendörfer  
Hanser Verlag 1992
- [MEYE71] WARTESYSTEME MIT VARIABLEN BEARBEITUNGSRATE  
Meyer Karl Heinz F.  
aus LECTURE NOTES IN ECONOMICS AND MATHEMATICAL  
SYSTEMS  
Springer Verlag 1971
- [MINO93] BROADBAND NETWORK ANALYSIS AND DESIGN  
Minoli Daniel  
Artech House 1993
- [MORS58] QUEUES, INVENTORIES & MAINTENANCE  
Morse Philip M.  
Wiley 1958
- [PRAB65] QUEUES AND INVENTORIES  
Prabhu N.U.  
Wiley 1965
- [RIO62] STOCHASTIC SERVICE SYSTEMS  
Riordan John  
WILEY 1962
- [SAAT61] ELEMENTS OF QUEUEING THEORY  
Saaty Thomas L.  
Mc Graw Hill 1961
- [SCHA73] WARTESCHLANGEN  
Schassberger R.  
Springer Verlag 1973
- [SCHM87] APPROXIMATION VON EMPIRISCHEN  
VERTEILUNGSFUNKTIONEN MIT ERLANGVERTEILUNGEN  
UND COXVERTEILUNGEN  
Schmickler Leonhard 1987  
aus  
MESSUNG, MODELLIERUNG UND BEWERTUNG VON RECHEN-  
UND KOMMUNIKATIONSSYSTEMEN  
Herausgeber Paterok  
Springer Verlag 1987
-

- [SCHM89] ERWEITERUNG DES VERFAHRENS MEDA  
ZUR ANALYTISCHEN BESCHREIBUNG EMPIRISCHER  
VERTEILUNGSFUNKTIONEN  
Schmickler Leonhard 1989  
aus  
MESSUNG, MODELLIERUNG UND BEWERTUNG VON RECHEN-  
UND KOMMUNIKATIONSSYSTEMEN  
Herausgeber Stiege, Lie  
Springer Verlag 1989
- [SCHW93] NUMERISCHE MATHEMATIK  
Schwarz H.R..  
Teubner 1993
- [STOC78] MARKOFFKETTEN-STOCHASTISCHE PROZESSE-  
WARTESCHLANGEN  
Lindenberg, Nuske, Schriever  
de Gruyter Lehrbuch 1978
- [TAKA58] STOCHASTIC PROCESSES  
Takács L.  
Wiley Interscience 1958
- [WAAL87] PERFORMANCE ANALYSIS & OPTIMAL CONTROL OF AN  
M/M/1/K QUEUEING SYSTEM WITH IMPATIENT CUSTOMERS  
de Waal 1987  
aus  
MESSUNG, MODELLIERUNG UND BEWERTUNG VON RECHEN-  
UND KOMMUNIKATIONSSYSTEMEN  
Herausgeber Paterok  
Springer Verlag 1987
- [WONN72] INTRODUCTORY STATISTICS  
Wonnacott Thomas H., Wonnacott Ronald J.  
Wiley Interscience 1972
-

## H. Symbolverzeichnis

|                   |  |
|-------------------|--|
| $\mathbf{0}$      | Nullmatrix $\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & \end{pmatrix}$ |
| $A(x)$            | Ankunftsverteilung   |
| $a(x)$            | Dichte der Ankünfte  |
| $\bar{a}$         | Laplace-Transformierte von $a(x)$  |
| $b_n$             | Hemmfunktion, siehe 2.2.4  |
| $B(x)$            | Bedienzeitverteilung   |
| $b(x)$            | Dichte der Bedienzeit  |
| $\bar{b}$         | Laplace-Transformierte von $b(x)$  |
| $\text{Cor}(X,Y)$ | Korrelationskoeffizient zwischen Zufallsvariablen $X$ und $Y$ , siehe [D.18]   |
| $\text{Cov}(X,Y)$ | Kovarianz zweier Zufallsvariablen $X$ und $Y$ , siehe [D.17]   |
| $c$               | Variationskoeffizient $c = \frac{\sigma}{\mu}$   |
| $\gamma$          | Wahrscheinlichkeit, daß mindestens 1 Bediener im G/M/1-System beschäftigt ist, $\gamma = \bar{a}(\mu(1 - \gamma))$   |
| $d$               | Differentialquotient, z.B. $\frac{dy}{dx}$   |
| $\Delta$          | Differenzenquotient $\Delta y = y_{t+1} - y_t$ ; $D^n = (\Delta + 1)^n$  |
| $\Delta(z)$       | Deltafunktion $\Delta(z) = \begin{cases} 0 & \text{für } z \leq 0 \\ 1 & \text{für } z > 0 \end{cases}$  |
| $D$               | Differenzen-/Differentialoperator $Dy_t = y_{t+1}$ ; $Dy = y'$   |
| $E(\cdot)$        | Erwartungswert einer Zufallsvariablen, siehe [D.9]   |
| $E$               | Ereignis   |
| $\mathbf{e}$      | Spaltenvektor $\mathbf{e} = \begin{pmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$   |
| $e^x$             | Exponentialfunktion, $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$  |

---

|                |  |
|----------------|--|
| $f(x)$         | Dichtefunktion   |
| $\bar{f}(s)$   | Laplace-Transformierte einer Funktion $f(t)$   |
| $F(x)$         | Verteilungsfunktion einer kontinuierlichen Zufallsvariablen $X$ , $F(x) = Pr\{X \leq x\}$  |
| $F(z)$         | z-Transformierte einer Funktion $f_t$  |
| $\Gamma(x)$    | Gamma-Funktion $\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$   |
| GGT            | ‘Größter Gemeinsamer Teiler’   |
| $H_i$          | einander paarweise ausschließende Ereignisse   |
| K              | Anzahl der Warteplätze + Anzahl der Bediener   |
| K-m            | Anzahl der Warteplätze   |
| <b>I</b>       | Einheitsmatrix $\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$ |
| L              | mittlere Anzahl der Kunden im System, $L = \sum_{i=1}^n np_n$  |
| $L^{(A)}$      | mittlere Anzahl der Kunden im System unmittelbar vor einer Ankunft   |
| $L_q$          | mittlere Warteschlangenlänge   |
| $L_q^{(A)}$    | mittlere Warteschlangenlänge unmittelbar vor einer Ankunft   |
| $\lambda'$     | effektive Ankunftsrate $\lambda' = \lambda(1 - p_k)$   |
| $\lambda$      | mittlere Ankunftsrate $\lambda = 1/t$  |
| M              | Anzahl der Terminals bzw. Maschinen, siehe 2.2.5   |
| m              | Anzahl der Bediener  |
| $m_i$          | mittlere Rekurrenzzeit des Zustandes i, siehe [F.10]   |
| $\mu$          | mittlere Bedienrate $\mu = 1/s$  |
| N              | Zufallsvariable für Anzahl der Kunden im System  |
| n!             | n-Faktorielle, $n! = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1$   |
| $\binom{n}{k}$ | Kombinationen von n unterschiedlichen Elementen bei k Stichproben $\binom{n}{k} = \frac{n!}{k!(n-k)!}$   |

---

---

|                     |  |
|---------------------|--|
| $p(k)$              | Wahrscheinlichkeitsfunktion  |
| $\mathbf{P}$        | Vektor der Zustandswahrscheinlichkeiten $\mathbf{p} = (p_0, p_1, p_2, \dots)$  |
| $p_j$               | Wahrscheinlichkeit, daß sich der Prozeß im Zustand $j$ befindet  |
| $\pi_j$             | stationäre Wahrscheinlichkeit für den Zustand $j$  |
| $p_{ij}$            | Zustandsübergangswahrscheinlichkeit vom Zustand $i$ zum Zustand $j$  |
| $p_{ij}^{(m)}$      | Zustandsübergangswahrscheinlichkeit in $m$ Schritten   |
| $\prod$             | Produkt $\prod_{i=1}^n a_i = a_1 \cdot a_2 \cdot a_3 \cdot \dots \cdot a_{n-1} \cdot a_n$  |
| $\mathbf{P}$        | Zustandsübergangsmatrix $\mathbf{P} = (p_{ij})$  |
| $P(k)$              | Verteilungsfunktion einer diskreten Zufallsvariablen $X$ , $P(k) = Pr\{X \leq k\}$   |
| $q_a$               | Übergangswahrscheinlichkeiten des M/G/1-Modells $q_a = q_{i+1-j}$  |
| $q_b$               | Übergangswahrscheinlichkeiten des G/M/1-Modells $q_b = q_{i+1-j}$  |
| $q_{ij}$            | stationäre Übergangsrate Zustand $i \rightarrow j$   |
| $\mathbf{Q}$        | Matrix der Übergangsraten, $\mathbf{Q} = \mathbf{P} - \mathbf{I}$ ; $\mathbf{Q} = (-q_i \text{ für } i = j, q_{ij} \text{ für } i \neq j)$ |
| $r(n)$              | Verzichtsfunction, siehe 2.2.4   |
| $S$                 | Zufallsvariable Bedienzeit   |
| $S^{(n)}$           | Bedienzeit des $n$ -ten Kunden   |
| $\rho$              | Verkehrsintensität $\rho = \lambda/\mu$  |
| $\bar{s}$           | mittlere Bedienzeit  |
| $\sum$              | Summe $\sum_{i=1}^n a_i = a_1 + a_2 + \dots + a_{n-1} + a_n$   |
| $\sigma$            | Streuung einer Zufallsgröße  |
| $\sigma^2$          | Varianz einer Zufallsgröße, siehe [D.10]   |
| $\sigma_T^2$        | Varianz der Zwischenankunftszeiten   |
| $\sigma_S^2$        | Varianz der Bedienzeiten   |
| $T$                 | Zufallsvariable Zwischenankunftszeit   |
| $T^{(n)}$           | Zwischenankunftszeit des $n$ -ten Kunden   |
| $\bar{t}$           | mittlere Zwischenankunftszeit  |
| $u$                 | Ausdehnung $u = \lambda/m\mu$  |
| $\text{Var}(\cdot)$ | Varianz einer Zufallsvariablen, siehe [D.10]   |
| $W$                 | mittlere Antwortzeit / Verweilzeit   |
| $w_q$               | mittlere Wartezeit   |
| $W_q^{(n)}$         | Wartezeit der $n$ -ten Kunden  |
| $\bar{W}(s)$        | Laplace-Transformierte der Antwortzeitverteilung   |
| $\bar{W}_q(s)$      | Laplace-Transformierte der Wartezeitverteilung   |

---

|        |  |
|--------|--|
| X      | Durchsatz                                |
| $X_t$  | Zufallsvariable in diskretem Raum        |
| $X(t)$ | Zufallsvariable in kontinuierlichem Raum |

$y', y'', y''', y^{(n)}$  1., 2., 3., n. Ableitung einer Funktion  $y(x)$ ;  $y^{(n)} = \frac{d^n y}{d x^n}$

---