



Universität Stuttgart

Institut für Nachrichtenvermittlung und Datenverarbeitung

Prof. Dr.-Ing. P. Kühn

46. Bericht über verkehrstheoretische Arbeiten

ZEITDISKRETE ANALYSE

VERKEHRSTHEORETISCHER MODELLE

IN RECHNER- UND KOMMUNIKATIONSSYSTEMEN

von

Phuoc TRAN-GIA

1988

© 1989 Institut für Nachrichtenvermittlung und Datenverarbeitung Universität Stuttgart

Druck: E. Kurz & Co., Druckerei + Reprografie GmbH.

ISBN 3-922403-56-5



University of Stuttgart

Institute of Communications Switching and Data Technics

Prof. Dr.-Ing. P. Kühn

46th Report on Studies in Congestion Theory

**DISCRETE-TIME ANALYSIS
OF PERFORMANCE MODELS
IN COMPUTER AND COMMUNICATION SYSTEMS**

by

Phuoc TRAN-GIA

1988

ABSTRACT

In the course of modeling modern computer and communication systems, discrete-time models and model components play an increasingly important role. On the one hand, new system structures and principles often employ discrete or discretized basic time and data units. Examples are the concept of cells in asynchronous wide-area networks or time slots in high-speed local and metropolitan area networks. On the other hand, system parameters and input values are often based on measured data, which are given in the form of histograms. They are discrete-time by nature. These facts lead to the development of discrete-time models in performance analyses of computer and communication systems, which can be frequently observed in the recent literature.

This report deals with analysis methods for discrete-time performance models in computer and communication systems and particularly, the use of discrete transform techniques in this class of analyses.

CHAPTER 1 INTRODUCTION

General modeling and analysis aspects in performance investigations of computer and communication systems are discussed, where attentions are devoted to discrete-time model components and to relationships between continuous- and discrete-time modeling. An outline of the objectives of the report is given.

CHAPTER 2 THEORETICAL ASPECTS IN DISCRETE-TIME MODEL ANALYSIS

The first part of this chapter deals with general assumptions and the characterization of discrete-time model components. Subsequently, transform techniques used later in this report are formally introduced, i.e. the Laplace transform for continuous-time density functions, the Z-transform and the discrete Fourier transform (DFT) applied to probability mass functions in the discrete time domain. Section 2.3 presents discrete random

variables and their distributions as well as discretization methods to approximate continuous-time processes.

Section 2.4 gives an outline of the theory of point processes and renewal processes of discrete-time types, where basic relationships are derived for use in the subsequent chapters.

The concept of cepstrum, which is often employed in the theory of homomorphic systems and signal processing, is introduced in section 2.5, where aspects of its application in the discrete-time queueing analysis of performance models are motivated.

CHAPTER 3 DISCRETE-TIME ANALYSIS OF THE GI/G/1 QUEUEING SYSTEM

To illustrate the variety of discrete-time analysis approaches, which operate in both, the time domain and the transform domain, and to show the application of transform techniques in these investigations, the analysis of the basic queueing system of type G/G/1 with discrete-time arrival and service processes is presented in this chapter.

Section 3.1 deals with general model assumptions of the class of discrete-time GI/G/1 models, while in section 3.2, a close-form solution of the special case GEOM/G/1 is derived. Basic equations for the GI/G/1 analysis are figured out in section 3.3, whereby attentions are devoted to the discrete-time representation of the well-known Lindley integral equation. Versions of this equations are given in discrete-time domain as well as in the Z-transform domain, which form the basic requirements for all analysis techniques and algorithms discussed later in the chapter.

Section 3.4 provides a survey on analysis methods for the waiting time distribution of the general class of GI/G/1 systems, starting with a presentation of iteration methods using discrete convolution algorithms in time domain. Subsequently, some major methods in the transform domain are described. Besides of the discussion of analysis techniques using polynomial factorizations and poles and

zeros allocations, the derivation is concentrated on an analysis technique employing concepts of signal processing theory, in particular the separation method in transform domain using the Cepstrum operator.

Based on the waiting time analysis of the GI/G/1 system, a new algorithm for the calculation of the idle period and the inter-departure distribution function of the output process is presented in section 3.5. Finally, numerical examples are depicted and discussed in section 3.6.

CHAPTER 4 ANALYSIS EXAMPLES FOR DISCRETE-TIME MODELS OF HIGHER COMPLEXITY

In this chapter the analysis of two discrete-time models of higher complexity is presented, in order to illustrate the capability, the efficiency as well as the methodological limit of discrete-time analysis techniques.

Section 4.1 presents a discrete-time modeling and analysis approach for a general class of overload control strategies in communication switching systems. The basic model is of type G/G/1 with bounded delay. The overload control is based on a throttling mechanism for new arriving requests when the load status of the system is above a predefined threshold. The indicator for overload conditions is chosen to be the amount of unfinished work instead of the number of calls in the system.

The input process modeling the subscriber requests is assumed to be general and is considered in section 4.1 under stationary as well as non-stationary conditions. The main overload control mechanism is modelled by means of a feedback path of the G/G/1 system, in conjunction with a workload-controlled acceptance scheme. The analysis method, which is developed in the discrete-time domain, is described in details. It allows the use of efficient discrete transform algorithms. To discuss dimensioning aspects of the overload control strategy numerical results are

given for different types of input processes and overload threshold parameters.

The class of polling systems with generally distributed input processes and finite capacity of waiting places is the subject of section 4.2. These models are often used e.g. in performance investigations of local area networks with token-passing medium access principles. The approximate analysis is done in time domain, based on the direct evaluation of discrete convolution operation taking advantages of fast convolution algorithms, e.g. the fast Fourier transform (FFT). Numerical examples are shown to illustrate the influence of the input process and the storage capacity on the main performance measures, e.g. the waiting time and the blocking probability of arriving packets or messages in the system.

CHAPTER 5 CONCLUSION

Concluding remarks on the use and significant properties of discrete-time analysis techniques are given and major results of the report are summarized.

INHALTSVERZEICHNIS

Literaturverzeichnis

Verzeichnis der wichtigsten Abkürzungen und Formelzeichen

1.	EINLEITUNG	1
1.1	Verfahren zur Leistungsuntersuchung von Rechner- und Kommunikationssystemen	1
1.2	Analyseverfahren	4
1.3	Übersicht über die Arbeit	7
2.	THEORETISCHE GRUNDLAGEN ZUR ZEITDISKRETEN MODELLANALYSE	10
2.1	Voraussetzungen	10
2.2	Transformationsmethoden	12
2.2.1	Laplace- und Laplace-Stieltjes-Transformationen	12
2.2.2	Z-Transformation und erzeugende Funktion	13
	a) Die Z-Transformation	13
	b) Die erzeugende Funktion	14
2.2.3	Diskrete Fourier-Transformation und schnelle Fourier-Transformation	15
	a) Definition der diskreten Fourier-Transformation	15
	b) Zusammenhang zwischen DFT und Z-Transformation	16
	c) Schnelle Fourier-Transformation und Algorithmen	16
2.3	Zeitdiskrete Zufallsvariablen und Verteilungen	17
2.3.1	Diskrete Verteilungen	17
2.3.2	Diskretisierung kontinuierlicher Verteilungsfunktionen	19
	a) Diskretisierungsmethode	19
	b) Diskretisierungsgenauigkeit	20
2.3.3	Grundfunktionen mit zwei diskreten Zufallsvariablen	22
	a) Summation	22
	b) Differenzbildung	22
	c) Minimumbildung	23

2.4	Zeitdiskrete Erneuerungsprozesse	24
2.4.1	Zeitkontinuierliche Punktprozesse und Erneuerungsprozesse	24
2.4.2	Definition zeitdiskreter Punkt- und Erneuerungsprozesse	26
2.4.3	Rekurrenzzeitverteilung zeitdiskreter Erneuerungsprozesse	27
	a) Beobachtung vor den diskreten Zeitpunkten	27
	b) Beobachtung nach den diskreten Zeitpunkten	30
	c) Rückwärts-Rekurrenzzeit-Verteilung	30
	d) Zeitdiskrete Erneuerungsprozesse mit gedächtnisloser Eigenschaft	31
2.4.4	Anzahl von Ereignissen in einem beliebig verteilten Intervall	32
2.5	Das Konzept des komplexen Cepstrums	34
2.5.1	Homomorphe Systeme und Transformationen	34
2.5.2	Das komplexe Cepstrum	35
	a) Homomorphe Systeme für gefaltete Signale	35
	b) Definition des komplexen Cepstrums	36
	c) Eigenschaften des komplexen Cepstrums	38
	d) Beispiel einer direkten Berechnung des komplexen Cepstrums	41
	e) Numerische Berechnung des komplexen Cepstrums	42
3.	ZEITDISKRETE ANALYSE DES WARTESCHLANGENSYSTEMS GI/G/1	44
3.1	Modellbeschreibung und Voraussetzungen	44
3.2	Direkte Lösung des GEOM(1)/G/1-Systems durch eingebettete Markoff-Kette	48
3.2.1	Systemparameter und Markoff-Kette	48
3.2.2	Zustandswahrscheinlichkeiten der Markoff-Kette	49
3.2.3	Wartezeitverteilung	52
3.3	Die Lindley-Integralgleichung im diskreten Zeitbereich	54
3.3.1	Allgemeines	54
3.3.2	Herleitung der Lindley-Integralgleichung im diskreten Zeitbereich	55
3.3.3	Charakteristische Gleichung im transformierten Bereich	60

3.3.4	Direkte Auflösung am Beispiel des Systems GEOM(m)/GEOM(m)/1	62
3.4	Algorithmen für Wartezeitverteilung des zeitdis- kreten GI/G/1-Systems	65
3.4.1	Algorithmus im diskreten Zeitbereich	65
3.4.2	Algorithmen im transformierten Bereich	67
	a) Grundprinzip	67
	b) Algorithmus mit Polynom-Faktorisierung	68
	c) Algorithmus mit Cepstrum-Bildung	70
3.4.4	Mittelwerte und weitere Systemcharakteristiken	75
	a) Wartewahrscheinlichkeit	75
	b) Wartezeit bezüglich wartender Anforderungen	75
	c) Systembelegung und mittlere Warteschlangen- länge	75
3.5	Verteilung der Freiperioden und des Ausgangsprozesses	77
3.5.1	Allgemeines	77
3.5.2	Verteilung der Freiperioden	78
3.5.3	Verteilung des Ausgangsprozesses	79
3.6	Ergebnisse	82
3.6.1	Systeme mit beliebig verteilten Zwischen- ankunfts- und Bedienzeiten	82
3.6.2	Systeme mit negativ-binomial verteilten Zwischenankunfts- und Bedienzeiten	84
	a) Wartezeitverteilung	84
	b) Freiperiode	84
	c) Ausgangsprozesse	87
3.6.3	Approximative Berechnung für zeitkontinuierliche Systeme	90
4.	BEISPIELE FÜR ZEITDISKRETE ANALYSE KOMPLEXER MODELLE	92
4.1	Zeitdiskrete Analyse einer Überlastabwehrstrategie in rechnergesteuerten Vermittlungssystemen	92
4.1.1	Modellbildung von Überlastabwehrstrategien	92
4.1.2	Lastgesteuertes Überlastabwehrverfahren und Modellbildung	93
4.1.3	Zeitdiskreter Algorithmus zur Modellanalyse	94
4.1.4	Ergebnisse	99
	a) Stationäres Verhalten der Überlastabwehr- strategie	100
	b) Systemantwort auf instationäre Überlastungen	101

4.2	Systeme mit zyklischen Abfertigungsmechanismen	103
4.2.1	Die Klasse von Polling-Systemen	103
4.2.2	Modellbeschreibung und Voraussetzungen	105
4.2.3	Zeitdiskrete Modellanalyse	106
	a) Voraussetzungen und Bezeichnungen	106
	b) Markoff-Kette und Zustandswahrscheinlichkeiten	107
	c) Zykluszeitanalyse	110
	d) Algorithmus zur iterativen Analyse	111
4.2.4	Systemcharakteristiken	112
	a) Blockierungswahrscheinlichkeit	112
	b) Zeitgemittelte Zustandswahrscheinlichkeiten	113
	c) Mittlere Wartezeit	116
4.2.5	Ergebnisse und Validierung	117
5.	ZUSAMMENFASSUNG	122

LITERATURVERZEICHNIS

- /1/ Achilles D., "Die Fourier-Transformation in der Signalverarbeitung", Springer 1978.
- /2/ Ackroyd M.H., "Computing the Waiting Time Distribution for the G/G/1 Queue by Signal Processing Methods", IEEE COM-28 (1980) 52-58.
- /3/ Ackroyd M.H., "Iterative Computation of the M/G/1 Queue Length Distribution via the Discrete Fourier Transform", IEEE COM-28(1980) 1929-1932.
- /4/ Ackroyd M.H., "M/M/1 Transient State Occupancy Probabilities Via the Discrete Fourier Transform", IEEE COM-30 (1982) 557-559.
- /5/ Bednar J.B., Watt T.L., "Calculating the Complex Cepstrum Without Phase Unwrapping or Integration", IEEE ASSP-33 (1985) 1014-1017.
- /6/ Bharath-Kumar K., "Discrete-Time Queueing Systems and their Networks", IEEE COM-28(1980) 260-263.
- /7/ Böpple M., "Zeitdiskrete Analyse von Warteschlangensystemen mit zyklischer Abfertigung und allgemeinen Ankunftsprozessen", Studienarbeit, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1986.
- /8/ Bogert B.P., Healy M.J.R., Tukey J.W., "The Quefrency Alalysis of Time Series for Echoes: Cepstrum, Pseudo-Autocovariance, Cross-Cepstrum, and Saphe Cracking", Proc. Symp. Time Series Analysis, Ed.: M. Rosenblatt, Wiley 1963, pp. 209-243.
- /9/ Burke P.J., "The Output of a Queueing System", Operations Research 4(1956) 699-704.
- /10/ Burke P.J., "Output Processes and Tandem Queues", Proc. Symp. on Comp. Comm. Networks and Teletraffic, NewYork 1972, Polytechn. Press of the PIB, 22(1972) 419-428.
- /11/ Brigham E.O., "The Fast Fourier Transform", Prentice-Hall 1974.
- /12/ Bux W., Truong H.L., "Mean-Delay Approximation for Cyclic-Service Queueing Systems", Perf. Eval. 3(1981) 187-196.
- /13/ Cavers J.K., "On the Fast Fourier Transform Inversion of Probability Generating Functions", J.Inst. Maths. Application, 22(1978) 275-282.

- /14/ Cohen J.W., "The Single Server Queue", North-Holland Publ. Co. 1969.
- /15/ Cooley J.W., Tukey J.W., "An Algorithm for the Machine Calculation of Complex Fourier Series", Math. Computation 19(1965) 297-301.
- /16/ Cooper R.B., Murray G., "Queues served in Cyclic Order", Bell Syst. Tech. J. 48 (1969) 675-689.
- /17/ Cooper R.B., "Queues served in Cyclic Order: Waiting Times", Bell Syst. Tech. J. 49 (1970) 399-413.
- /18/ Cox D.R., "Erneuerungstheorie", Oldenbourg 1966.
- /19/ Cox D.R., Miller H.D., "The Theory of Stochastic Processes", Chapman and Hall 1965.
- /20/ Cox D.R., Lewis P.A.W., "The Statistical Analysis of Series of Events", Methuen 1966.
- /21/ Daisenberger G., Oehlerich J., Wegmann G., "STATOR - Statistical Overload Regulation - and TAIL - Time Account Input Limitation - Two Concepts for Overload Regulation in SPC Systems", Proc. 11th Int. Teletraffic Congr. (ITC), pp. 2.1B-4, Kyoto 1985.
- /22/ Daley D.J., "Notes on Queueing Output Processes", Math. Methods in Queueing Theory, Springer, 1974, 351-358.
- /23/ Doetsch G., "Handbuch der Laplace-Transformation ; Band I", Birkhäuser Stuttgart/Basel 1950.
- /24/ Eisenberg M., "Queues with Periodic Service and Changeover Times", Operations Research 20 (1972) 440-451.
- /25/ Feller W., "An Introduction to Probability Theory and its Applications"; Vol.II, 2nd edition, Wiley 1971.
- /26/ Grassmann W.K., Jain J.L., "Numerical Solution of the Waiting Time of the GI/G/1 Queue", preprint, Univ. of Saskatchewan, 1985.
- /27/ Halfin S., "Delays in Queues, Properties and Approximations", Proc. 11th Int. Teletraffic Congr. (ITC), pp. 1.4-3, Kyoto 1985.
- /28/ Hashida O., "Analysis of Multiqueue", Rev. El. Commun. Lab. 20(1972) 189-199.
- /29/ Heffes H., "On the Output of a GI/M/N Queueing System with Interrupted Poisson Input", Operations Research 24(1976)3, 530-542.

- /30/ Henrici P., "Fast Fourier Methods in Computational Complex Analysis", Siam Review, 21(1979) 481-527.
- /31/ Hsu J., Burke P.J., "Behavior of Tandem Buffers with Geometric Input and Markovian Output", IEEE COM-24(1976) 358-361.
- /32/ Jans H., "Verkehrsanalyse von Vermittlungssteuerungen mit taktgesteuerter Ein-/Ausgabe und Prioritäten", 37. Bericht über verkehrstheoretische Arbeiten, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1983.
- /33/ Jury E.I., "Theory and Application of the Z-Transform Method", Wiley 1964.
- /34/ Kingman J.F.C., "Some Inequalities for the Queue GI/G/1", Biometrika 49(1962) 315-324.
- /35/ Kingman J.F.C., "Inequalities in the Theory of Queues", J. Roy. Stat. Soc. B32(1970) 102-110.
- /36/ Kleinrock L., "Queueing Systems", Vol.I: Theory, Vol.II: Computer Applications", Wiley 1975.
- /37/ Kobayashi H., Konheim A.G., "Queueing Models for Computer Communications System Analysis", IEEE Trans.Comm., 25(1977) 2-29.
- /38/ Kobayashi H., "Modelling and Analysis : An Introduction to System Performance Evaluation Methodology", Addison Wesley 1978.
- /39/ Kobayashi H., "Stochastic Modelling: Queueing Models ; Discrete-Time Queueing Systems", in : Part II, Louchard G., Latouche G. (eds.), "Probability Theory and Computer Science", Academic Press 1983.
- /40/ Konheim A.G., "The Stationary Waiting Time Distribution for a Single Server Queue", J.Soc.Indust. Appl. Math., 13(1965) 966-975.
- /41/ Konheim A.G., "An Elementary Solution of the Queueing System GI/G/1", SIAM J. Comp., 4(1975) 540-545.
- /42/ Kosten L., "Stochastic Theory of Service Systems", Pergamon Press 1973.
- /43/ Krämer W., Langenbach M., "Approximate Formulae for the Delay in the Queueing System GI/G/1", Proc. 8th ITC, Melbourne, 1976.

- /44/ Kühn P., "Tabellen für Wartesysteme", Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1976.
- /45/ Kühn P., "Approximate Analysis of General Queueing Networks by Decomposition", IEEE Trans.Comm. 27(1979) 113-126.
- /46/ Kühn P.J., "Multiqueue Systems with Nonexhaustive Cyclic Service", Bell Syst. Tech. J. 58 (1979) 671-699.
- /47/ Kühn P., "Analyse zufallsabhängiger Prozesse in Systemen zur Nachrichtenvermittlung und Nachrichtenverarbeitung", 30. Bericht über verkehrstheoretische Arbeiten, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1981.
- /48/ Leibowitz M.A., "An Approximate Method for Treating a Class of Multiqueue Problems", IBM J. Res. Develop. 5(1961) 204-209.
- /49/ Lindley D.V., "The Theory of Queues with a Single Server", Proc. of the Cambridge Philosophical Society, 48(1952) 277-289.
- /50/ Makino T., "On a Study of Output Distribution", J. of the Oper. Res. Soc. of Japan 8(1966)3, 109-133.
- /51/ Marshall K.T., "Some Inequalities in Queueing", Operations Research 16(1968) 651-665.
- /52/ Meisling T., "Discrete-Time Queueing Theory", Operations Research 6(1958) 96-105.
- /53/ Morrison J.A., "Two Discrete-Time Queues in Tandem", IEEE COM-27(1979) 563-573.
- /54/ Nussbaumer H.J., "Fast Fourier Transform and Convolution Algorithms", Springer 1982.
- /55/ Oppenheim A.V., Schafer R.W., "Digital Signal Processing", Prentice-Hall 1975.
- /56/ Ott T. J., "On the Stationary Waiting Time Distribution in the GI/G/1 Queue", Bell Communications Research, 1985.
- /57/ Pack C.D., "The Output of an M/D/1 Queue", Operations Research 23(1975)4, 750-760.
- /58/ Ponstein J., "Theory and Numerical Solution of a Discrete Queueing Problem", Statistica Neerlandica, 20(1974) 139-152.

- /59/ Papoulis A., "Probability, Random Variables and Stochastic Processes", McGraw-Hill 1965.
- /60/ Pujolle G., Claude J.P., Seret D., "A Discrete Queueing System with a Product Form Solution", Proc. Int. Seminar on Comp. Networking and Perf. Evaluation, pp.3.4, Tokyo 1985.
- /61/ Rathgeb E., "Algorithmen zur numerischen Analyse des zeitdiskreten GI/G/l-Systems im Frequenzbereich", Studienarbeit, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1985.
- /62/ Reich E., "Waiting Times when Queues are Tandem", A.M.S., 1957.
- /63/ Schmidt D., "Iterative Algorithmen zur numerischen Analyse des zeitdiskreten GI/G/l-Systems im Zeitbereich", Studienarbeit, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1985.
- /64/ Seelen L.P., Tijms H.C., Van Hoorn M.H., "Tables for Multi-Server Queues", North-Holland Publ. Co. 1985.
- /65/ Sieglen R., "Zeitdiskrete Analyse einer Überlastabwehrstrategie in Vermittlungssystemen", Studienarbeit, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1986.
- /66/ Singleton R.C., "On Computing the Fast Fourier Transform", Comm. ACM, 10(1967) 647-654.
- /67/ Smith W.L., "On the Distribution of Queueing Times", Proc. of the Cambridge Philosophical Society, 49(1953) 449-461.
- /68/ Takagi H., Kleinrock L., "Analysis of Polling Systems", Japan Science Inst. Research Report, 1985.
- /69/ Tran-Gia P., "Überlastprobleme in rechnergesteuerten Fernsprechvermittlungssystemen - Modellbildung und Analyse", 36. Bericht über verkehrstheoretische Arbeiten, Institut für Nachrichtenvermittlung und Datenverarbeitung, Universität Stuttgart, 1982.
- /70/ Tran-Gia P., Raith T., "Multiqueue Systems with Finite Capacity and Nonexhaustive Cyclic Service", Proc. Int. Seminar on Computer Networking and Performance Evaluation, Sept. 1985, Tokyo, Japan.
- /71/ Tran-Gia P., "Discrete-Time Analysis for the Interdeparture Distribution of GI/G/l Queues", Proc. Seminar on Teletraffic Analysis and Computer Performance Evaluation, June 1986, Amsterdam.

- /72/ Tribolet J.M., "A new Phase Unwrapping Algorithm", IEEE Trans. ASSP 25(1977) 170-177.
- /73/ Tripathi S.K., Agrawala A.K., "On Characterizing the Inter-Departure Process of a Server", Perf. Eval. 2(1982) 251-255.
- /74/ Tumura Y., "On the Equilibrium Probabilities of GI/G/1", Ope. Res. Soc. of Japan, 10(1968) 93-107.
- /75/ Whitt W., "Approximations for Networks of Queues", Proc. 10th Int. Teletraffic Congr. (ITC), pp. 4.1.2, Montreal 1983.
- /76/ Whitt W., "Approximations for Departure Processes and Queues in Series", Nav. Res. Log. Quart., 31(1984) 499-521.
- /77/ Winograd S., "On Computing the discrete Fourier Transform", Math. Computation 32(1978) 175-199.

Verzeichnis der wichtigsten Abkürzungen und Formelzeichen

a) Allgemeine Abkürzungen und Symbole

A	Zufallsvariable für Zwischenankunftsabstände
B	Zufallsvariable für Bedienzeiten
DFT	<u>d</u> iskrete <u>F</u> ourier- <u>T</u> ransformation
FFT	schnelle Fourier-Transformation (<u>F</u> ast <u>F</u> ourier <u>T</u> ransform)
ZT	Z-Transformation
ZV	Zufallsvariable
$\delta(t-t_i)$	Dirac-Impuls an der Stelle $t=t_i$
$\delta(k-k_i)$	Einheitsimpuls an der Stelle $k=k_i$
Δt	Diskretisierungszeiteinheit
ρ	Verkehrsangebot bzw. Auslastung der Bedieneinheit
"**"	kontinuierliche bzw. diskrete Faltungsoperation

Notation bzgl. einer zeitkontinuierlichen Zufallsvariablen A :

A	Zufallsvariable, hier z.B. zeitbezogen
A*	ZV für die Vorwärts- bzw. Rückwärts-Rekurrenzzeit von A
A(t)	Verteilungsfunktion von A
a(t)	Verteilungsdichtefunktion der ZV A
$a_{LT}(s)$	Laplace-Transformierte von a(t) bzw. Laplace-Stieltjes-Transformierte von A(t)
$E[A^n]$	n-tes gewöhnliches Moment der ZV A
$EA=E[A]$	Mittelwert von A
c_A	Variationskoeffizient von A

Notation bzgl. einer zeitdiskreten Zufallsvariablen X :

X	Zufallsvariable, hier z.B. zeitbezogen
X*	ZV für die Vorwärts- bzw. Rückwärts-Rekurrenzzeit von X
x(k)	Verteilung der ZV X
$x_{CEP}(k)$	komplexes Cepstrum von x(k)
$x_{DFT}(k)$	diskrete Fourier-Transformierte von x(k)
$x_{EF}(z)$	erzeugende Funktion von x(k)

$x_{ZT}(z)$	Z-Transformierte von $x(k)$
$E[X^n]$	n-tes gewöhnliches Moment von X
$EX=E[X]$	Mittelwert von X
C_X	Variationskoeffizient von X

Kurznotation für Verteilungen und Verteilungsfunktionen

BER	Bernoulli-Verteilung
BIN	binomiale Verteilung
D	deterministische Verteilungsfunktion
E_n	Erlang-Verteilungsfunktion n-ter Ordnung
GEOM(m)	um m Stellen verschobene geometrische Verteilung
H_n	hyperexponentielle Verteilungsfunktion n-ter Ordnung
NEGBIN	negativ-binomiale Verteilung
POIS	Poisson-Verteilung

b) Zusätzliche Abkürzungen zu Kapitel 3

A_n	ZV für die Zeitdauer zwischen den Ankunftszeitpunkten der n-ten und der (n+1)-ten Anforderung
B_n	ZV für die Bedienzeit der n-ten Anforderung
$c(k)$	Systemfunktion
D	ZV für die Zwischenendigungsabstände
F	ZV für die Durchlaufzeit von Anforderungen
I	ZV für die Dauer von Freiperioden
P_w	Wartewahrscheinlichkeit
P_1	Wahrscheinlichkeit, daß ein Ausgangsereignis am Ende einer Betriebsperiode liegt
Q	ZV für die Anzahl wartender Anforderungen
q_{ij}	Übergangswahrscheinlichkeiten der eingebetteten Markoff-Kette
\underline{Q}	Übergangsmatrix der eingebetteten Markoff-Kette
R	ZV für die Anzahl R von eingetroffenen Anforderungen während einer Bedienzeit
S	ZV für die Dauer der Betriebsperioden
$S_{ZT}(z)$	charakteristische Funktion des zeitdiskreten GI/G/1-Systems im Z-Bereich

U_n^-	ZV für die Restarbeit im System unmittelbar vor dem Ankunftszeitpunkt der n-ten Anforderung
U_n^+	ZV für die Restarbeit im System unmittelbar nach dem Ankunftszeitpunkt der n-ten Anforderung
U_{n+1}^v	ZV für die virtuelle Restarbeit im System unmittelbar vor dem Ankunftszeitpunkt der (n+1)-ten Anforderung
W	ZV für die Wartezeit bezüglich aller Anforderungen
$W_{>0}$	ZV für die Wartezeit bezüglich wartender Anforderungen
X	ZV für die Systembelegung
X_n	ZV für die Anzahl von Anforderungen nach dem n-ten Bedienende
π_0	Operator im diskreten Zeitbereich gemäß Gl. (3.31)
π^m	Operator im diskreten Zeitbereich gemäß Gl. (4.14)

c) Zusätzliche Abkürzungen zu Kapitel 4.1

B_A	Blockierungswahrscheinlichkeit einer beliebig herausgegriffenen Anforderung
$B_{A,n}$	Blockierungswahrscheinlichkeit der n-ten Anforderung
L	Schwellenwert der Überlastabwehrstrategie
U_n	ZV für die Restarbeit im System unmittelbar vor dem Ankunftszeitpunkt der n-ten Anforderung

d) Zusätzliche Abkürzungen zu Kapitel 4.2

A	ZV für die Zwischenankunftszeit des Ankunftsprozesses an einer Warteschlange
B_A	Blockierungswahrscheinlichkeit für Anforderungen
$B(k)$	bedingte Blockierungswahrscheinlichkeit
C	ZV für die Zykluszeit
C_0	Zykluszeit ohne Bedienung einer Anforderung in der betrachteten Warteschlange
C_1	Zykluszeit mit Bedienung einer Anforderung in der betrachteten Warteschlange

C_S	ZV für ein Zykluszeit-Segment
g	Anzahl der Warteschlangen des Polling-Systems
G_i	ZV für die Anzahl ankommender Anforderungen während einer bedingten Zykluszeit des Typs i (C_i), $i=0,1$
O	ZV für die Umschaltzeit
P_j^*	Wahrscheinlichkeit, daß der zufällige Beobachtungszeitpunkt in einer bedingten Zykluszeit vom Typ C_j liegt
S	Kapazität einer Warteschlange
t^*	zufälliger Beobachtungszeitpunkt
X_n	ZV für die Anzahl von Anforderungen in der betrachteten Warteschlange unmittelbar vor dem n -ten Abfragezeitpunkt
X_n^+	ZV für die Anzahl von Anforderungen unmittelbar nach dem n -ten Abfragezeitpunkt
X^*	ZV für den Zustand zum zufälligen Beobachtungszeitpunkt
X, X^+	ZV für den Zustand unmittelbar vor bzw. nach dem vorausgegangenen Regenerationszeitpunkt
W_A	mittlere Wartezeit angenommener Anforderungen
$\rho = g \cdot \frac{EB}{EA}$	Verkehrsangebot

1. EINLEITUNG

1.1 Verfahren zur Leistungsuntersuchung von Rechner- und Kommunikationssystemen

Bedingt durch den Einsatz moderner Mikrorechner-Komponenten in Kommunikationssystemen sowie durch die Entwicklung von Rechnernetzen entstehen immer leistungsfähigere und komplexere Systeme, sowohl in der Hard- als auch in der Software, aus struktureller und aus steuerungstechnischer Sicht. Im Gegensatz zu früheren Systemgenerationen, deren Leistung oft erst nach ihrer Inbetriebnahme ermittelt wurde, muß bei modernen Systemen - wegen ihrer funktionellen Komplexität und ihrer immer vielfältigeren Leistungsmerkmale - die Leistungsfähigkeit vor der Systemeinführung untersucht und nachgewiesen werden. Dies sind im einzelnen:

- Die funktionelle Leistungsfähigkeit:
Dazu gehören u.a. die Verklemmungs- bzw. Widerspruchsfreiheit der implementierten Kommunikationsprotokolle im System, die einwandfreie Zusammenarbeit der aktivierten Prozesse sowie die funktionelle Einhaltung definierter, systemspezifischer Leistungsmerkmale.

- Die Leistungsfähigkeit des Systems unter Lastbedingung:
Hier sind zu erwähnen: die Funktionsfähigkeit und die Einhaltung vorgegebener Grenzwerte für die Blockierungswahrscheinlichkeit, Durchlauf- und Wartezeiten, Verarbeitungsgüte usw. unter Nennbelastung, die Einhaltung festgelegter Verkehrsgüte unter Überlast, die Überlebensfähigkeit des Systems bei extremen Lastspitzen, etc.

Bedingt durch die große Anzahl von angeschlossenen Teilnehmern und in Echtzeit zusammenarbeitenden Systemkomponenten sowie von parallel ablaufenden Prozessen kann das Ablaufgeschehen in einem Rechner- bzw. Kommunikationssystem mit Hilfe von zufallsabhängigen Prozessen beschrieben werden. Für die Untersuchung derartiger Prozesse werden Methoden der Stochastik, insbesondere der Verkehrstheorie angewandt.

Bild 1.1 gibt eine Übersicht über Methoden zur Leistungsbeurteilung von Rechner- und Kommunikationssystemen. Im Falle eines bereits in der Betriebsphase befindlichen Systems kann die Funktionsfähigkeit durch Messungen untersucht werden, wobei realistische Belastungsprofile von angeschlossenen Teilnehmern und peripheren Einrichtungen für Meßzwecke benutzt werden können.

Falls ein Prototyp des zu untersuchenden Systems verfügbar ist, die anzuschließenden Teilnehmer bzw. peripheren Einrichtungen jedoch noch nicht oder nur teilweise vorhanden sind, kann die Leistung des Systems unter realen Lastbedingungen z.B. mit Hilfe von Umweltsimulationseinrichtungen untersucht werden. Dabei werden Teilnehmer- und periphere Prozesse durch Zustandsautomaten zeittreu und unter Berücksichtigung des realen Teilnehmerverhaltens nachgebildet, mit denen die Teilnehmer-System-Interaktion

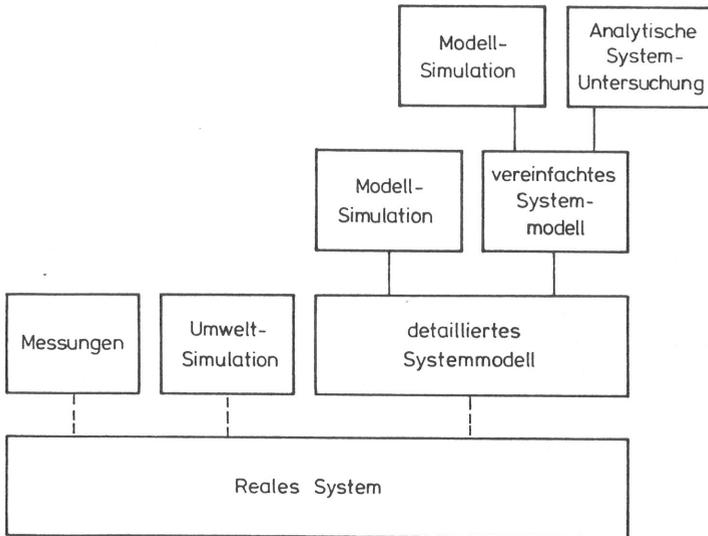


Bild 1.1 Methoden zur Leistungsuntersuchung von Rechner- und Kommunikationssystemen

realistisch simuliert und die Leistungsfähigkeit des Systems unter beliebig vorgebbaren Belastungen getestet werden kann.

In der Vor- oder Entwicklungsphase eines Systems werden Leistungsuntersuchungen häufig anhand von Systemmodellen, insbesondere von Verkehrsmodellen, durchgeführt. In einem verkehrstheoretischen Modell wird das Ablaufgeschehen im realen System in allen seinen funktionellen und zeitlichen Zusammenhängen und Interaktionen der Systemkomponenten mit Hilfe weniger abstrakter Modellelemente beschrieben. Die Systemstruktur und -organisation und das dynamische, in Echtzeit ablaufende Systemgeschehen werden auf entsprechende Modellstrukturen und -komponenten abgebildet.

Wie in Bild 1.1 dargestellt wird, kann die Modellbildung auf unterschiedlichen Abstraktionsebenen bzw. Modellierungstiefen geschehen. In einem detaillierten Systemmodell sind Modellkomponenten noch sehr systemnah. Alle strukturellen Komponenten und Ablaufsteuerungen in realen Systemen werden hier noch detailgetreu modelliert. Dies führt zu Verkehrsmodellen höherer Komplexität, die in der Regel nur mit Hilfe der Systemsimulation untersucht werden können.

Da in der Vorentwicklungsphase häufig mehrere strukturelle Alternativen verglichen werden müssen, führt die Modellsimulation auf dieser Abstraktionsebene häufig zu sehr großem Rechenaufwand. Die Leistungsuntersuchung kann in diesem Fall mit einem Systemmodell höherer Abstraktion durchgeführt werden, wobei die wesentlichen, für die Untersuchung relevanten Systemmerkmale noch im Modell enthalten sein müssen. Auf dieser Abstraktionsebene existieren oft analytisch exakte und approximative Methoden, die Untersuchungen hinreichend großer Parameterbereiche mit günstigerem Rechenaufwand ermöglichen.

Die Parametrisierung von Modellkomponenten erfolgt oft anhand von Messungen an im Betrieb befindlichen Systemen bzw. Prototypen. Meßergebnisse werden häufig in der Form von Histogrammen gewonnen, d.h. die betreffenden Modellkomponenten liegen für weitere Unter-

suchungen bereits in einer zeitdiskreten stochastischen Beschreibungsform vor. Darüberhinaus enthalten Verkehrsmodelle vielfach Komponenten, die sich zeitdiskret exakt charakterisieren lassen. Beispiele dafür sind Taktsteuerungen in Betriebssystem-Modellierungen und Übertragungsdauern von Nachrichtenpaketen, die aus einer ganzzahligen Anzahl von Bits oder Bytes bestehen.

In den meisten Modelluntersuchungen in der Literatur werden Verfahren, die im kontinuierlichen Zeitbereich operieren, angewendet. Der Grund dafür liegt nicht zuletzt darin, daß die hier verfügbaren Analysemethoden bekannter und in einigen Fällen vollständiger entwickelt sind.

Die Leistungsuntersuchungen der Klasse von Systemen mit zeitdiskreten Komponenten im kontinuierlichen Zeitbereich werden i.a. an äquivalenten zeitkontinuierlichen Modellen vorgenommen. Dabei werden z.B. zeit- und wertdiskrete Verteilungsfunktionen mit Hilfe von bekannten kontinuierlichen Verteilungsfunktionen approximiert, die eine Übereinstimmung einer Anzahl von Momenten (z.B. Mittelwert und Variationskoeffizient) mit den zeitdiskreten Verteilungsfunktionen aufweisen.

Als Alternative bieten sich Analyseverfahren an, die direkt im diskreten Zeitbereich operieren, unter unmittelbarer Einbeziehung von zeitdiskreten Modellkomponenten, deren Charakterisierung in zeitdiskreter Form vorliegt. Die Erörterung dieser Klasse von analytischen Methoden und Ansätze ist das Ziel der vorliegenden Arbeit.

1.2 Analyseverfahren

Es existiert ein breites Spektrum von Analyseverfahren für verkehrstheoretische Modelle, die in einschlägiger Fachliteratur /14,36,38,42,47/ ausführlich und systematisch dargestellt wurden.

Im folgenden wird lediglich auf einige Grundprinzipien und grundlegende analytische Ansätze eingegangen, die später zur Erläuterung der zeitdiskreten Analyseverfahren benötigt werden.

Betrachtet werde nun ein stochastischer Zustandsprozeß $X(t)$, der das Ablaufgeschehen einer zufallsabhängigen Prozeßgröße in einem Verkehrsmodell beschreibt. Dieser Prozeß wird zunächst speziell zu den Zeitpunkten $t_n, n=0,1,\dots$ ($t_0 < t_1 < t_2 < \dots$) beobachtet. Beispiele für die betrachtete Prozeßgröße sind: die Anzahl von Anforderungen im System, die Menge unerledigter Arbeit im System oder die verbliebene Bedien- oder Wartezeit einer speziellen Anforderung.

Bei der Modellanalyse wird häufig zunächst $X(t_n)$ bestimmt - wobei beispielsweise $X(t_0)$ als Anfangsbedingung vorgegeben wird -, und darauf aufbauend werden weitere Systemcharakteristiken berechnet. Die meisten Verfahren basieren auf einer Berechnungsvorschrift V , die oft wie folgt angegeben werden kann:

$$X(t_{n+1}) = V(X(t_n)). \quad (1.1)$$

Die Vorschrift V in Gl.(1.1) kann je nach Analysemethode die Form einer Matrizen-Multiplikation, eine Funktionalgleichung usw. annehmen.

Bezüglich des Prozesses $X(t)$ wird oft anhand der Eigenschaft der Gedächtnislosigkeit (Markoffsche Eigenschaft) eine Klassifizierung vorgenommen. Ein Prozeß $X(t)$ besitzt zum Zeitpunkt t_n die Markoff-Eigenschaft, falls seine vom Prozeßzeitpunkt t_n aus beobachtete zukünftige Entwicklung nur vom Zustand $X(t_n) = x_n$ zum Zeitpunkt t_n und nicht von der vergangenen Entwicklung ($t_{n-1}, t_{n-2}, \dots, t_0$) abhängig ist, d.h.

$$\begin{aligned} P(X(t_{n+1})=x_{n+1} | X(t_n)=x_n, X(t_{n-1})=x_{n-1}, \dots, X(t_0)=x_0) \\ = P(X(t_{n+1})=x_{n+1} | X(t_n)=x_n). \end{aligned} \quad (1.2)$$

Einige häufig anzutreffende Analysemethoden werden nachfolgend kurz diskutiert. Wie in späteren Kapiteln erläutert werden wird, sind diese Methoden sowohl für zeitkontinuierliche als auch für zeitdiskrete Analysen anwendbar.

- Analyse mit eingebetteter Markoff-Kette

Diese Methode ist geeignet für die Zustandsanalyse bei Prozessen, bei denen die Markoffsche Eigenschaft nur bei einer Anzahl von Punkten auf der Zeitachse besteht. In diesen sog. Regenerationszeitpunkten, welche die eingebettete Markoff-Kette bilden, verliert der Prozeß sein Gedächtnis ; die Prozeßentwicklung ist von dort aus unabhängig von der Vergangenheit.

Die oben erwähnte Prozeßgröße $X(t)$ sei nun die Anzahl von Anforderungen im System zum Zeitpunkt t . Der Zustandsprozeß besitzt zum Regenerationszeitpunkt t_n den Zustandswahrscheinlichkeitsvektor

$$\underline{p}^{(n)} = \{ p^{(n)}(k) \} \quad k=0,1,\dots \quad (1.3a)$$

mit

$$p^{(n)}(k) = P(X(t_n) = k). \quad (1.3b)$$

Läßt sich eine Übergangswahrscheinlichkeitsmatrix \underline{Q} finden, die die Beziehung zwischen den Zustandswahrscheinlichkeitsvektoren zweier beliebiger aufeinanderfolgender Regenerationszeitpunkte t_n und t_{n+1} herstellt, d.h.

$$\underline{Q} = \{ q_{ij} \} \quad (1.4a)$$

mit

$$q_{ij} = P(X(t_{n+1}) = j \mid X(t_n) = i), \quad i, j=0,1,2,\dots, \quad (1.4b)$$

so erhält man eine rekursive Berechnungsvorschrift gemäß Gl.(1.1) wie folgt

$$\underline{p}^{(n+1)} = \underline{Q} \cdot \underline{p}^{(n)} . \quad (1.5)$$

Anhand der Gl.(1.5) können Markoff-Ketten-Wahrscheinlichkeitsvektoren sukzessiv berechnet werden, falls ein Anfangsvektor, z.B. $\underline{p}^{(0)}$, bekannt ist. Im stationären Zustand des Prozesses, wobei der Wahrscheinlichkeitsvektor unabhängig vom Zeitindex ist, d.h.

$$\underline{p}^{(n+1)} = \underline{p}^{(n)} = \underline{p} , \quad (1.6)$$

läßt sich die Zustandsanalyse auf eine Eigenvektorbestimmung folgender Gleichung zurückführen:

$$\underline{p} = \underline{Q} \cdot \underline{p} . \quad (1.7)$$

- Lindley-Integralmethode

Wie in Kap.3 ausführlich hergeleitet werden wird, kann für die Klasse der allgemeinen einstufigen Wartesysteme als Berechnungsvorschrift gemäß Gl.(1.1) eine Integralgleichung vom Wiener-Hopf-Typ aufgestellt werden. Dabei wird der Prozeß der unerledigten Arbeit im System $U(t_n)$ zu den Ankunftszeitpunkten t_n der Anforderungen betrachtet. Die entstehende Gleichung wird als Lindley-Integralgleichung /49/ bezeichnet (vgl. Kap. 3.3).

1.3 Übersicht über die Arbeit

Anschließend an die Einleitung werden im 2. Kapitel einige methodische Aspekte zur zeitdiskreten Analyse von verkehrstheoretischen Modellen erörtert, die in den späteren Kapiteln benötigt werden. Dabei werden in 2.2 Transformationsmethoden und Algorithmen (z.B. erzeugende Funktion, Z-Transformation, diskrete Fourier-Transformation, schnelle Fourier-Transformation, etc.) definiert und ihre Eigenschaften sowie Anwendungsgesichtspunkte in der Verkehrsanalyse diskutiert. In 2.3 werden zeitdiskrete und diskretisierte Zufallsvariablen und ihre Verteilungen behandelt. Eine systematische Behandlung der Klasse von Erneuerungsprozessen im diskreten Zeitbereich findet sich in 2.4.

In der Analyse zeitdiskreter Systeme im transformierten Bereich wird u.a. das aus der Signalverarbeitungstheorie stammende Konzept des komplexen Cepstrums angewendet. Nach einer kurzen Abhandlung von Grundbegriffen und Eigenschaften der homomorphen Systemtheorie wird das Konzept des komplexen Cepstrums in 2.5 eingeführt.

In Kapitel 3 wird versucht, das Spektrum der Methoden der zeitdiskreten Analyse von Verkehrsmodellen am Beispiel des einstufigen Warteschlangensystems GI/G/1 aufzuzeigen. Unterkapitel 3.2 zeigt die Analyse dieses Systems für den speziellen Fall des geometrischen Ankunftsprozesses mit der in der zeitkontinuierlichen Analyse oft verwendeten Methode der eingebetteten Markoff-Kette. In Kap. 3.3 werden GI/G/1-Systeme mit allgemeinen Zwischenankunfts- und Bedienprozessen betrachtet. Eine der Lindley-Integralgleichung entsprechende Beziehung im diskreten Zeitbereich wird hergeleitet und im Zeit- und Frequenzbereich angegeben. Die direkte Auflösung der Gleichung wird am Beispiel eines Systems mit (verschobenen) geometrischen Ankunfts- und Bedienprozessen gezeigt.

Für GI/G/1-Systeme, die einer geschlossenen analytischen Lösung der Lindley-Integralgleichung nicht zugänglich sind, existieren numerische Algorithmen zur Berechnung der Wartezeitverteilung. Diese Algorithmen werden in Kap. 3.4 ausführlich diskutiert. Dies sind im einzelnen :

- Algorithmus im Zeitbereich : mit iterativer Faltung
- Algorithmus im transformierten Bereich :
 - o Separation mit expliziter Berechnung von Nullstellen und Polen der charakteristischen Funktion mit Hilfe der Polynom-Faktorisierung.
 - o Separation unter Berücksichtigung der Phasen-Eigenschaften und Anwendung des Cepstrum-Konzepts.

Eine neue zeitdiskrete Methode zur Berechnung von Verteilungen der Freiperioden und des Ausgangsprozesses zeitdiskreter GI/G/1-Systeme wird in Kap. 3.5 vorgestellt.

Die Anwendung zeitdiskreter Analysemethoden in komplexen Leistungsuntersuchungen von Rechner- und Kommunikationssystemen wird in Kapitel 4 anhand von zwei Modellbeispielen erörtert. In Kap. 4.1 wird ein lastgesteuerter Überlastabwehrmechanismus in Vermittlungssystemen untersucht und in Kap. 4.2 eine approximative zeitdiskrete Analyse für Polling-Systeme mit endlichem Warteraum und allgemeinen Ankunftsprozessen vorgestellt.

Die Klasse von Analyseverfahren im diskreten Zeitbereich stellt eine bedeutende und bisher in der Literatur ungenügend beachtete Alternative zu den zeitkontinuierlichen Methoden dar. Sie ist in erster Linie für solche Modelluntersuchungen gut geeignet, bei denen die Modellkomponenten bereits in zeitdiskreter Form vorliegen. Ferner können zeitdiskrete Analysemethoden zur Untersuchung zeitkontinuierlicher Modelle, die keiner exakten Analyse zugänglich sind, approximativ eingesetzt werden. Durch den Einsatz effizienter numerischer Algorithmen zur Durchführung von Transformationen weist die zeitdiskrete Klasse konzeptionelle Vorteile auf, welche insbesondere bei Anwendungen zum Tragen kommen, bei denen Transformationen erforderlich sind. Dies ist häufig der Fall, wenn z.B. Ergebnisse nur im transformierten Bereich geschlossen erhältlich sind und daraus Systemcharakteristiken mit Hilfe von Transformationen ermittelt werden müssen.

2. THEORETISCHE GRUNDLAGEN ZUR ZEITDISKRETEN MODELLEANALYSE

2.1 Voraussetzungen

Zur Beschreibung von zufallsabhängigen Modellelementen, z.B. Zwischenankunfts- oder Bedienzeiten, werden in den hier behandelten zeitdiskreten Modelluntersuchungen Zufallsvariablen betrachtet, deren Realisierungen ganzzahlige Vielfache einer modelleinheitlichen Zeiteinheit Δt betragen. Durch diese Annahme können Ereignisse im Modell (z.B. Ankunftsereignis, Bedienende, ...) nur zu den diskreten Zeitpunkten auf einer mit Δt diskretisierten Zeitachse (Zeitgitter) auftreten. Für den Fall, daß zu einem Zeitpunkt Ankunfts- und Bedienende-Ereignisse gleichzeitig eintreffen, wird vorausgesetzt, daß die Bedienende-Ereignisse vorrangig behandelt werden. Die Reihenfolge der Abarbeitung innerhalb eines Ereignistyps ist abhängig vom jeweiligen Modell festzulegen.

Wie in Bild 2.1 an einem Beispiel aufgezeigt, kann eine derartige zeitdiskrete Zufallsvariable (ZV) A in unterschiedlichen Formen beschrieben werden :

- Verteilungsfunktion : Die Verteilungsfunktion (Bild 2.1a)

$$A(t) = P(A \leq t) \quad (2.1a)$$

ist treppenförmig; die Stufenhöhen entsprechen den Verteilungswerten.

- Verteilungsdichtefunktion : Gemäß der Definition kann die zugehörige Verteilungsdichtefunktion als Summe gewichteter Dirac-Impulse angegeben werden (vgl. Bild 2.1b)

$$a(t) = \frac{dA(t)}{dt} = \sum_{i=-\infty}^{+\infty} P(A=t_i) \delta(t-t_i), \quad (2.1b)$$

wobei $\delta(t-t_i)$ den Dirac-Impuls an der Stelle $t=t_i$ bedeutet.

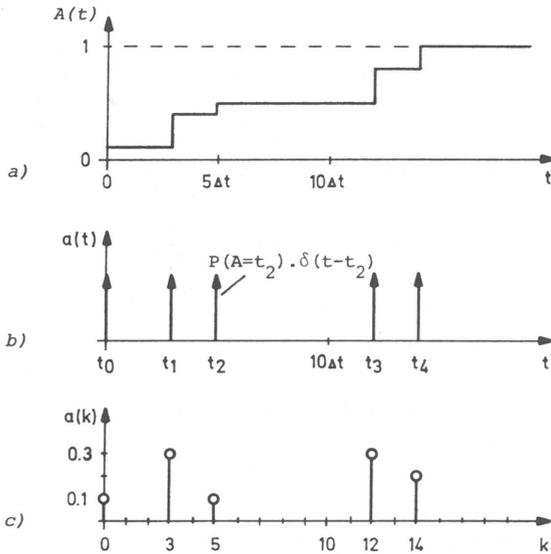


Bild 2.1 Beschreibungsformen diskreter Zufallsvariablen
a) Verteilungsfunktion
b) Verteilungsdichtefunktion
c) Verteilung

- Verteilung : Da nach der oben erwähnten Annahme die Realisierungen $t_i, i=-\infty, \dots, +\infty$, ganzzahlige Vielfache von Δt sind, kann die ZV A mit der Verteilung

$$a(k) = P(A=k \cdot \Delta t), k=-\infty, \dots, +\infty \quad (2.1c)$$

charakterisiert werden. Zur Vereinfachung der Notation werden im folgenden alle zeitbezogenen Größen auf Δt normiert angegeben. Es wird beispielsweise die normierte Form $A = k$ an Stelle von $A = k \cdot \Delta t$ geschrieben.

2.2 Transformationsmethoden

Die in den analytischen Untersuchungen in den Kapiteln 3 und 4 angewendeten Transformationsmethoden werden im folgenden zusammengefaßt dargestellt. Die aufgeführten Definitionsgleichungen gelten für allgemeine Zufallsvariablen. Sie werden hier jedoch für den praktisch bedeutsamen Bereich nicht-negativer ZV angegeben.

2.2.1 Laplace- und Laplace-Stieltjes-Transformationen

Für Analysemethoden im kontinuierlichen Zeitbereich spielt die Laplace- bzw. die Laplace-Stieltjes-Transformation eine zentrale Rolle. $A(t)$ und $a(t)$ seien die Verteilungsfunktion und die Verteilungsdichtefunktion einer nicht-negativen ZV A . Die Laplace-Stieltjes-Transformierte von $A(t)$ bzw. die Laplace-Transformierte von $a(t)$ wird definiert als /23,59/ :

$$\begin{aligned} a_{LT}(s) &= \text{LST}\{A(t)\} = \int_0^{\infty} e^{-st} dA(t) \\ &= \text{LT}\{a(t)\} = \int_0^{\infty} e^{-st} a(t) dt . \end{aligned} \quad (2.2)$$

Ist die ZV A vom zeitdiskreten Typ (vgl. Kap.2.1), d.h.

$$a(t) = \sum_{k=0}^{\infty} a(k) \delta(t-k\Delta t),$$

so ergibt sich für die Laplace-Transformierte

$$a_{LT}(s) = \sum_{k=0}^{\infty} a(k) e^{-k \cdot \Delta t \cdot s} . \quad (2.3)$$

Dieser Zusammenhang deutet auf die in Kap. 2.2.2 eingeführten Transformationsmethoden hin, die für die Analyse im diskretisierten Zeitbereich besser geeignet sind.

Eine häufig benutzte Eigenschaft der Laplace-Transformation, welche die später eingeführten Transformationsmethoden (Z-Transfor-

mation, erzeugende Funktion, diskrete Fourier-Transformation,...) ebenfalls besitzen, ist die Umwandlung der Faltungsoperationen im Ursprungsbereich in Multiplikationen im transformierten Bereich. Für eine Summe A von zwei voneinander unabhängigen Zufallsvariablen A_1 und A_2 gilt :

$$a_{LT}(s) = LT\{a(t)\} = LT\{a_1(t) * a_2(t)\} = a_{1,LT}(s) \cdot a_{2,LT}(s), \quad (2.4)$$

wobei "*" den Faltungsoperator im kontinuierlichen Bereich bezeichnet.

2.2.2 Z-Transformation und erzeugende Funktion

Die Z-Transformation gehört zu den meistbenutzten Werkzeugen der zeitdiskreten Systemanalyse. Ausführliche Abhandlungen über die Z-Transformation finden sich z.B. in /33,55/. In diesem Kapitel wird nur ein kurzer Abriß der wichtigsten Eigenschaften dieser Transformationsmethode hinsichtlich ihrer Anwendung auf zeitdiskrete Verteilungen gegeben.

a) Die Z-Transformation

Die Z-Transformierte einer Zeitfolge $x(k)$, $k=-\infty, \dots, +\infty$, wird definiert als

$$x_{ZT}(z) = Z\{x(k)\} = \sum_{-\infty}^{+\infty} x(k) z^{-k}. \quad (2.5)$$

Ist die Zeitfolge $x(k)$ die Verteilung einer nicht-negativen, zeitdiskreten ZV X , d.h.

$$0 \leq x(k) \leq 1, \quad k=0, 1, \dots \quad \text{und} \quad \sum_{k=0}^{\infty} x(k) = 1, \quad (2.6)$$

so erhält man folgende einseitige Z-Transformation

$$x_{ZT}(z) = \sum_{k=0}^{\infty} x(k) z^{-k}. \quad (2.7)$$

Anhand der Eigenschaft gemäß Gl.(2.6) ist erkennbar, daß $x_{ZT}(z)$ in Gl.(2.7) im Bereich $|z| \gg 1$ konvergiert. Ist die Verteilung $x(k)$ endlich, so kann $x_{ZT}(z)$ als Polynom in $1/z$ aufgefaßt werden. Der Konvergenzbereich von $x_{ZT}(z)$ erstreckt sich auf $|z| > 0$.

Die Z-Rücktransformation ist gegeben durch das Integral

$$x(k) = \frac{1}{2\pi i} \oint_C x_{ZT}(z) z^{k-1} dz, \quad i = \sqrt{-1}. \quad (2.8)$$

Dabei ist C eine im Konvergenzbereich von $x_{ZT}(z)$ befindliche, geschlossene Kontur, die gegen den Uhrzeigersinn verläuft und den Ursprung ($z=0$) einschließt.

Analog zu Gl.(2.4) gilt für die Summe X zweier voneinander unabhängiger, diskreter Zufallsvariablen X_1 und X_2 :

$$x_{ZT}(z) = Z\{x(k)\} = Z\{x_1(k) * x_2(k)\} = x_{1,ZT}(z) \cdot x_{2,ZT}(z), \quad (2.9)$$

wobei "*" hier die diskrete Faltung bedeutet.

b) Die erzeugende Funktion

Häufig wird in der Wahrscheinlichkeitsrechnung an Stelle der Z-Transformation die erzeugende Funktion - bzw. wahrscheinlichkeitserzeugende Funktion - verwendet, die wie folgt definiert wird

$$x_{EF}(z) = EF\{x(k)\} = \sum_{k=0}^{+\infty} x(k) z^k, \quad (2.10)$$

wobei $x(k)$ die Verteilung einer nicht-negativen, diskreten Zufallsvariablen X darstellt. Da der Unterschied zwischen den Definitionsgleichungen (2.5) und (2.10) nur in den Termen z^k bzw. z^{-k} besteht, lassen sich prinzipiell die Eigenschaften der erzeugenden Funktion aus denen der Z-Transformation ableiten, indem z durch $1/z$ ersetzt wird.

Die Verteilung $x(k)$ kann gemäß folgender inversen Transformation aus der erzeugenden Funktion berechnet werden :

$$x(k) = \frac{1}{k!} \left. \frac{d^k x_{EF}(z)}{dz^k} \right|_{z=0}, \quad k=0,1,\dots \quad (2.11)$$

Entsprechend Gl.(2.4) und Gl.(2.9) lautet die erzeugende Funktion der Summe X von zwei voneinander unabhängigen, diskreten Zufallsvariablen X_1 und X_2 :

$$x_{EF}(z) = EF\{x(k)\} = EF\{x_1(k) * x_2(k)\} = x_{1,EF}(z) \cdot x_{2,EF}(z). \quad (2.12)$$

2.2.3 Diskrete Fourier-Transformation und schnelle Fourier-Transformation

a) Definition der diskreten Fourier-Transformation

Die diskrete Fourier-Transformation (DFT) für eine endliche Folge $x(k)$, $k=0,1,\dots,N-1$, wird definiert durch :

$$x_{DFT}(n) = DFT\{x(k)\} = \sum_{k=0}^{N-1} x(k) e^{-i \frac{2\pi}{N} nk}, \quad n=0,\dots,N-1, \quad i=\sqrt{-1}. \quad (2.13)$$

Obwohl die Folgen $x(k)$ im Zeitbereich und $x_{DFT}(n)$ im transformierten Bereich prinzipiell komplexwertig sein können, sind die Zeitfolgen in dieser Arbeit oft Verteilungen und dementsprechend rein reell.

Die Rücktransformation lautet :

$$x(k) = DFT^{-1}\{x_{DFT}(n)\} = \frac{1}{N} \sum_{n=0}^{N-1} x_{DFT}(n) e^{+i \frac{2\pi}{N} kn}, \quad k=0,\dots,N-1, \quad i=\sqrt{-1}. \quad (2.14)$$

Analog zu Gl.(2.4) gilt für eine Summe X von zwei voneinander unabhängigen, diskreten Zufallsvariablen X_1 und X_2 :

$$\begin{aligned}x_{\text{DFT}}(n) &= \text{DFT}\{x(k)\} = \text{DFT}\{x_1(k) * x_2(k)\} \\ &= x_{1,\text{DFT}}(n) \cdot x_{2,\text{DFT}}(n).\end{aligned}\tag{2.15}$$

b) Zusammenhang zwischen DFT und Z-Transformation

Betrachtet werde weiterhin die Klasse endlicher Verteilungen $x(k)$, $k=0, \dots, N-1$. Es kann gezeigt werden (siehe /55/), daß sich die Z-Transformation gemäß

$$x_{\text{ZT}}(z) = \sum_{k=0}^{N-1} x(k) z^{-k}$$

mit den folgenden auf dem Einheitskreis liegenden Abtastwerten

$$x(n) = x_{\text{ZT}}\left(e^{n \frac{2\pi i}{N}}\right), \quad n=0, 1, \dots, N-1, \quad i=\sqrt{-1}\tag{2.16}$$

vollständig beschreiben läßt. Ein Vergleich mit der Definition der diskreten Fourier-Transformation zeigt, daß - für die Klasse endlicher Verteilungen - an Stelle der Z-Transformation die diskrete Fourier-Transformation eingesetzt werden kann. Entsprechend kann die inverse Z-Transformation durch die diskrete Fourier-Rücktransformation (DFT^{-1}) ersetzt werden. Dies erlaubt den Einsatz effektiver Verfahren und Algorithmen zur Auswertung der diskreten Fourier-Transformation, die in der Signalverarbeitung entwickelt wurden.

c) Schnelle Fourier-Transformation und Algorithmen

Aus den Definitionen der diskreten Fourier-Transformation DFT bzw. DFT^{-1} ist erkennbar, daß der numerische Aufwand zur Auswertung einer DFT $O(N^2)$ entspricht, wobei N^2 auf die Anzahl der komplexen Multiplikationen hinweist. Zur Reduzierung dieses Aufwands wurden durch geeignete Zusammenfassung von mehrfach verwendeten Teilsummen effektive Algorithmen zur Auswertung der DFT und der inversen DFT entwickelt. Diese Klasse von Algorithmen, die ihren

Anfang in der bahnbrechenden Arbeit von J.W. Cooley und J.W. Tukey /15/ finden und den numerischen Aufwand von $O(N^2)$ auf $O(N \cdot \log N)$ reduzieren, wird unter dem Begriff "schnelle Fourier-Transformation" (FFT : Fast Fourier Transform) zusammengefaßt (vgl./1,11,54, 55/). Eine Übersicht über Anwendungen von FFT-Algorithmen findet man z.B. in /13,30/.

Es existiert eine Reihe von Algorithmen, z.B. die Polynomial-Transformation /54/, der Winogradsche Algorithmus /77/, mit denen sich der Rechenaufwand zur Auswertung der DFT weiter reduzieren läßt. Auf diese Klassen von Algorithmen wird in der vorliegenden Arbeit nicht eingegangen.

2.3 Zeitdiskrete Zufallsvariablen und Verteilungen

2.3.1 Diskrete Verteilungen

Einige wichtige Verteilungstypen, die in den späteren Kapiteln angewendet werden, sind in Tabelle 2.1 aufgelistet. Hinsichtlich der Anwendung dieser Verteilungen sollen hier folgende Eigenschaften und Zusammenhänge erwähnt werden :

- Betrachtet man an jedem Punkt der diskretisierten Zeitachse einen Bernoulli-Versuch, bei dem $(1-q)$ die Wahrscheinlichkeit für das Auftreten einer Anforderung am beobachteten Zeitpunkt ist, so spricht man vom Bernoulli-Ankunftsprozeß. Die dazugehörige Zwischenankunftsverteilung ist vom Typ GEOM(1). Bei diesem Ankunftsverkehr beträgt der Abstand zwischen zwei Anforderungen mindestens eine Zeiteinheit. Dieser Ankunftsverkehr wird in zeitdiskreten Analysen häufig benutzt /6,37,39/.
- Vergleicht man die Z-Transformierte der negativ-binomialen Verteilung mit der geometrischen Verteilung GEOM(0), so kann die negativ-binomiale Verteilung als eine y -fache Faltung der GEOM(0)-Verteilung mit sich selbst interpretiert werden, wobei anzumerken ist, daß y nicht ganzzahlig sein muß.

Verteilungstyp	Verteilung $x(k)=P(X=k)$	Z-Transformierte $x_{ZT}(z)$	Mittelwert EX	Variationskoeffizient c_x
Bernoulli BER	$x(0)=q$ $x(1)=1-q$ $0 \leq q \leq 1$ 1-q: Erfolgswahrscheinlichk.	$q + (1-q)z^{-1}$	$(1-q)$	$\sqrt{\frac{q}{1-q}}$
Geometrisch (um m Stellen verschoben) GEOM(m)	$(1-q)q^{k-m}$ $0 \leq q \leq 1$ $m \geq 0, k \geq m$	$\frac{(1-q)z^{-m}}{1-qz^{-1}}$	$\frac{q}{1-q} + m$	$\frac{\sqrt{q}}{q+m(1-q)}$
Binomial BIN	$\binom{n}{k}(1-q)^k q^{n-k}$ $0 \leq q \leq 1, 0 \leq k \leq n$	$(q + (1-q)z^{-1})^n$	$n(1-q)$	$\sqrt{\frac{q}{n(1-q)}}$
Poisson POIS	$\frac{y^k}{k!} e^{-y}$ $k \geq 0, y$ reell	$e^{y(z^{-1}-1)}$	y	$\frac{1}{\sqrt{y}}$
Negativbinomial NEGBIN	$\binom{y+k-1}{k} (1-q)^y q^k$ $= \binom{-y}{k} (1-q)^y (-q)^k$ $0 \leq q \leq 1, y$ reell	$\left(\frac{1-q}{1-qz^{-1}}\right)^y$	$\frac{yq}{1-q}$	$\frac{1}{\sqrt{yq}}$

Tabelle 2.1 Einige diskrete Verteilungen

- Die Parameter der negativ-binomialen Verteilung lassen sich aus dem vorgegebenen Mittelwert EX und dem Variationskoeffizienten c_x wie folgt bestimmen

$$q = 1 - \frac{1}{EX \cdot c_x^2}, \quad y = \frac{EX}{EX \cdot c_x^2 - 1}, \quad (2.17a)$$

wobei

$$EX \cdot c_x^2 > 1. \quad (2.17b)$$

Dies bedeutet, daß sich mit der negativ-binomialen Verteilung theoretisch zweiparametrische Verteilungen mit beliebig vorge-

gegebenem Mittelwert EX und Variationskoeffizienten $c_X \geq 0$ erzeugen lassen, wobei die Bedingung in Gl.(2.17b) erfüllt sein muß. Diese Eigenschaft ist sehr vorteilhaft in Untersuchungen, in denen eine systematische parametrische Analyse - z.B. bzgl. des Variationskoeffizienten - durchzuführen ist. Für die Verwendung dieses Verteilungstyps in Kap. 3 ergeben sich jedoch in bestimmten Parameterbereichen sehr breite Verteilungen, die numerisch schwierig zu handhaben sind. Die daraus resultierenden Beschränkungen des Parameterbereiches werden in /61/ ausführlich diskutiert.

2.3.2 Diskretisierung kontinuierlicher Verteilungsfunktionen

In zeitdiskreten Modelluntersuchungen, bei denen zeitkontinuierliche Verteilungsfunktionen vorgegeben sind, sowie bei Leistungs- und Genauigkeitsvergleichen zeitkontinuierlicher und -diskreter Analyseverfahren, wobei der gleiche Verteilungstyp genommen werden muß, werden häufig zeitkontinuierliche Verteilungsfunktionen durch diskrete Verteilungen approximativ ersetzt. Im folgenden wird die in dieser Arbeit verwendete Approximationsmethode vorgestellt und die Approximationsgenauigkeit diskutiert.

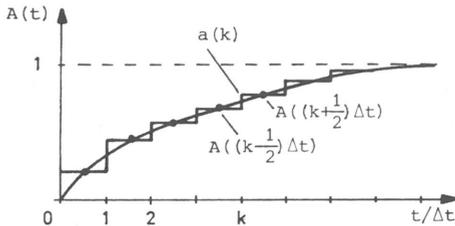


Bild 2.2 Diskretisierung zeitkontinuierlicher Verteilungsfunktionen

a) Diskretisierungsmethode

Wie in Bild 2.2 schematisch dargestellt, kann eine zeitkontinuierliche Verteilungsfunktion $A(t)$ durch eine zeitdiskrete Verteilung $a(k)$ wie folgt approximiert werden (vgl./32,63/) :

$$a(0) = A\left(\frac{\Delta t}{2}\right)$$
$$a(k) = A\left(\left(k+\frac{1}{2}\right)\Delta t\right) - A\left(\left(k-\frac{1}{2}\right)\Delta t\right), \quad k=1,2,\dots \quad (2.18)$$

wobei Δt die in Kap. 2.1 eingeführte Diskretisierungszeiteinheit bedeutet. In /32/ finden sich Ausdrücke der diskreten Approximation wichtiger Verteilungsfunktionsstypen unterschiedlicher Variationskoeffizientenbereiche.

b) Diskretisierungsgenauigkeit

Die Genauigkeit der in Gl. (2.18) angegebenen Diskretisierungsverfahren hängt offensichtlich von der Diskretisierungszeiteinheit und vom Typ sowie vom Verlauf der Verteilungsfunktion ab. Im folgenden wird am Beispiel der negativ-exponentiellen Verteilungsfunktion

$$A(t) = 1 - e^{-\lambda t}, \quad EA = 1/\lambda, \quad c_A = 1. \quad (2.19)$$

die Approximationsgenauigkeit untersucht. Die diskrete Approximation gemäß Gl. (2.18) lautet

$$a(0) = 1 - p$$
$$a(k) = p^{2k+1} - p^{2k-1}, \quad k=1,2,\dots \quad (2.20a)$$

wobei

$$p = e^{-\lambda \Delta t / 2} \quad (2.20b)$$

Der Mittelwert und der Variationskoeffizient der diskreten Approximation lauten

$$EA_{\text{Diskr.}} = \frac{p}{1-p^2} \cdot \Delta t, \quad c_{A,\text{Diskr.}} = \sqrt{\frac{1-p+p^2}{p}}. \quad (2.21)$$

Der Betrag des relativen Diskretisierungsfehlers für den Mittelwert bzw. für das Quadrat des Variationskoeffizienten errechnet sich zu:

$$\Delta_E = \left| \frac{EA_{\text{Diskr.}} - EA}{EA} \right| = \left| \frac{\lambda \cdot \Delta t \cdot p}{1 - p^2} - 1 \right| \quad (2.22a)$$

$$\Delta_{c^2} = \left| \frac{c_{A, \text{Diskr.}}^2 - c_A^2}{c_A^2} \right| = \frac{(1 - p)^2}{p} \quad (2.22b)$$

Bild 2.3 zeigt die in Gl.(2.22a) und Gl.(2.22b) angegebenen Diskretisierungsfehler als Funktion der Diskretisierungsauflösung. Eine feine Auflösung garantiert zwar geringe Diskretisierungsfehler, führt jedoch zu Verteilungen mit größeren Wertebereichen, die in der Analyse mit höherem numerischen Aufwand verbunden sind.

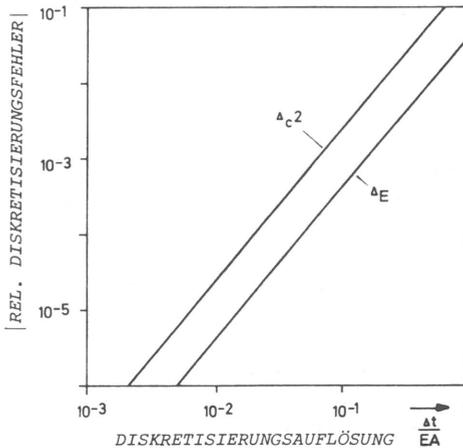


Bild 2.3 Diskretisierungsgenauigkeit am Beispiel der negativ-exponentiellen Verteilungsfunktion

2.3.3 Grundfunktionen mit zwei diskreten Zufallsvariablen

Gegeben seien zwei statistisch voneinander unabhängige, diskrete Zufallsvariablen A und B. Im folgenden werden einige häufig angewendeten Funktionen dieser Zufallsvariablen und ihre Verteilungen hergeleitet.

a) Summation

Die Verteilung $c(k)$ der Summe $C = A + B$ errechnet sich zu

$$c(k) = P(C = k) = P(A+B=k) = \sum_{j=-\infty}^{\infty} P(A=k-j \mid B=j) \cdot P(B=j). \quad (2.23a)$$

Da A und B voneinander unabhängig sind, erhält man

$$c(k) = \sum_{j=-\infty}^{\infty} a(k-j) \cdot b(j) = a(k) * b(k), \quad (2.23b)$$

wobei "*" den diskreten Faltungsoperator bedeutet. Im Z-Bereich ergibt sich

$$c_{ZT}(z) = a_{ZT}(z) \cdot b_{ZT}(z). \quad (2.23c)$$

Die Vollständigkeitsbedingung für die Verteilung $c(k)$ läßt sich anhand der Gl. (2.5) und Gl. (2.23c) einfach nachprüfen :

$$\sum_{k=-\infty}^{\infty} c(k) = c_{ZT}(1) = a_{ZT}(1) \cdot b_{ZT}(1) = 1.$$

b) Differenzbildung

Die Verteilung $c(k)$ der Differenz $C = B-A$ kann wie folgt berechnet werden

$$\begin{aligned} c(k) &= P(C = k) = P(B-A=k) = \sum_{j=-\infty}^{\infty} P(B=k+j \mid A=j) \cdot P(A=j) \\ &= \sum_{j=-\infty}^{\infty} b(k+j) \cdot a(j) = b(k) * a(-k). \end{aligned} \quad (2.24a)$$

Die Verteilung der Differenz ist gemäß Gl. (2.24a) die Kreuzkorrelation der Verteilungen $b(k)$ und $a(k)$. Im Z -Bereich erhält man

$$c_{ZT}(z) = b_{ZT}(z) \cdot a_{ZT}(1/z). \quad (2.24b)$$

Ähnlich wie bei der Summation kann ebenfalls anhand der Gl.(2.5) verifiziert werden, daß $c(k)$ die Vollständigkeitsbedingung erfüllt.

c) Minimumbildung

Die Verteilung $c(k)$ des Minimums

$$C = \min(A, B)$$

kann durch die Betrachtung folgender Beziehung berechnet werden :

$$P(C > k) = P(A > k) \cdot P(B > k),$$

d.h.

$$1 - C(k) = (1 - A(k)) \cdot (1 - B(k))$$

oder

$$C(k) = \sum_{j=-\infty}^k c(j) = A(k) + B(k) - A(k) \cdot B(k) . \quad (2.25)$$

Ein Punktprozeß wird als gewöhnlicher Erneuerungsprozeß (bzw. rekurrenter Prozeß) bezeichnet, falls die Intervalle A_i unabhängig voneinander und identisch verteilt sind, d.h.

$$A_i(t) = A(t) \quad , \quad i=0,1,\dots, \quad t \geq 0. \quad (2.26)$$

Für die Vorwärts-Rekurrenzzeit A^* , die als die Zeitdauer von einem zufälligen Beobachtungszeitpunkt t^* bis zur nächsten Anforderung (Ereigniszeitpunkt) definiert wird, errechnet sich die Verteilungsdichtefunktion $a^*(t)$ (siehe /18,19,20/) zu

$$a^*(t) = \frac{1}{EA} (1 - A(t)) \quad (2.27a)$$

oder nach der Laplace-Transformation :

$$a_{LT}^*(s) = \frac{1}{s EA} (1 - a_{LT}(s)) \quad (2.27b)$$

Daraus ergeben sich die gewöhnlichen Momente der Vorwärts-Rekurrenzzeit

$$E[A^{*k}] = \frac{E[A^{k+1}]}{(k+1) EA} \quad (2.28)$$

und speziell für den Mittelwert

$$EA^* = \frac{EA}{2} (c_A^2 + 1) \quad (2.29)$$

Häufig wird in der Erneuerungstheorie neben der Vorwärts-Rekurrenzzeit die Rückwärts-Rekurrenzzeit in Betracht gezogen. Dies ist das Intervall zwischen dem Beobachtungszeitpunkt und dem letzten Ereigniszeitpunkt. Da Erneuerungsprozesse bzgl. des Zeitparameters reversibel sind /18,19/, erhält man dieselbe Verteilungsfunktion (Gl.(2.27a)) für die Rückwärts-Rekurrenzzeit.

2.4.2 Definition zeitdiskreter Punkt- und Erneuerungsprozesse

Der Zeitparameter wird nun mit Δt diskretisiert. Ein zeitdiskreter Punktprozeß ist eine Folge von Ereigniszeitpunkten, die auf der diskretisierten Zeitachse liegen. Die Zwischenankunftsabstände A_i (Zeitintervall zwischen den Ereigniszeitpunkten t_{i-1} und t_i) sind dementsprechend zeitdiskreter Natur mit den zugehörigen Verteilungen $a_i(k)$, $k=0,1,\dots$

Ein zeitdiskreter Punktprozeß wird als gewöhnlicher, zeitdiskreter Erneuerungsprozeß bezeichnet, falls die Intervalle A_i unabhängig voneinander und identisch verteilt sind, d.h.

$$a_i(k) = a(k) \quad , \quad i=0,1,\dots \quad k=0,1,\dots \quad (2.30)$$

Da $a(0) = 0$ nicht vorausgesetzt wurde, sind in der hier diskutierten Betrachtungsweise Fälle enthalten, in denen mehrere Ereignisse auf einem diskreten Zeitpunkt liegen. Die Klasse der hier betrachteten Prozesse umfaßt daher auch Gruppenankunftsprozesse mit geometrisch verteilter Gruppengröße. Es kann z.B. gezeigt werden, daß die folgenden zeitdiskreten Erneuerungsprozesse P_1 und P_2 identisch sind :

- Prozeß P_1

Erneuerungsprozeß mit Einzelankünften und folgender Verteilung des Zwischenankunftsabstands A_1

$$a_1(k) = P(A_1=k) \quad , \quad k=0,1,\dots \quad (2.31a)$$

- Prozeß P_2

Gruppenankunftsprozeß mit folgender Verteilung des Zwischenankunftsabstands A_2

$$a_2(k) = P(A_2=k) = \frac{a_1(k)}{1 - a_1(0)} \quad , \quad k=1,2,\dots \quad (2.31b)$$

und geometrisch verteilter Gruppengröße G

$$g(i) = P(G = i) = (1 - a_1(0)) a_1(0)^{i-1} \quad , \quad i=1,2,\dots \quad (2.31c)$$

2.4.3 Rekurrenzzeitverteilung zeitdiskreter Erneuerungsprozesse

Analog zur Definition im kontinuierlichen Zeitbereich (vgl. Gl. (2.27a,b)) wird die Vorwärts-Rekurrenzzeit A^* im diskreten Zeitbereich als die Zeitspanne von einem zufälligen Beobachtungszeitpunkt t^* bis zum nächsten Ereigniszeitpunkt definiert. Da hier im diskretisierten Zeitbereich operiert wird, darf der Prozeß auch nur zu den diskreten Zeitpunkten observiert werden.

Prinzipiell sind für die Konstruktion der Vorwärts- sowie Rückwärts-Rekurrenzzeit im diskreten Zeitbereich beide Festlegungen möglich, ob der Beobachtungszeitpunkt unmittelbar vor oder nach einem diskreten Zeitpunkt liegt. Diese zwei alternativen Betrachtungsweisen führen zu unterschiedlichen Rekurrenzzeit-Verteilungen, die je nach Anwendungsfall zutreffen und eingesetzt werden (vgl. Kapitel 4).

a) Beobachtung vor den diskreten Zeitpunkten

Bild 2.5 zeigt einen zeitdiskreten Erneuerungsprozeß, der zu einem zufälligen Zeitpunkt t^* von einem außenstehenden, unabhängigen Beobachter betrachtet wird. Es wird hier vorausgesetzt, daß der Beobachtungszeitpunkt t^* unmittelbar vor einem diskreten Zeitpunkt liegt. Trifft zum gleichen Zeitpunkt ein Ereignis des zu beobachtenden Prozesses ein, so hat die Dauer der Vorwärts-Rekurrenz-

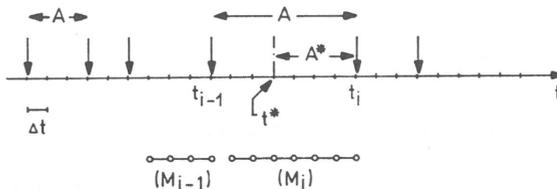


Bild 2.5 Zeitdiskrete Erneuerungsprozesse
 t^* zufälliger Beobachtungszeitpunkt
 (M_i) Menge möglicher Beobachtungszeitpunkte, für die das nächste Ereignis zum Zeitpunkt t_i eintritt

zeit den Wert Null. Im folgenden wird die Herleitung der Verteilung $a^*(k)$ der Vorwärts-Rekurrenzzeit detailliert dargestellt.

Die Verteilung $a^*(k)$ wird zunächst als totale Wahrscheinlichkeit aufgefaßt

$$a^*(k) = \sum_{n=k+1}^{\infty} a^*(k|n) p(n), \quad (2.32)$$

wobei

$p(n)$ die Wahrscheinlichkeit ist, daß ein Zwischenankunftsintervall der Länge n beobachtet wird
 $a^*(k|n) = P(A^* = k | t^* \text{ liegt in einem Zwischenankunftsintervall der Länge } n).$

Die Wahrscheinlichkeit $p(n)$ läßt sich aus folgenden Feststellungen ermitteln :

- $p(n)$ ist proportional zur Auftrittswahrscheinlichkeit $a(n)$
- $p(n)$ ist proportional zur Länge n , da zu ersehen ist, daß aus der Sicht eines außenstehenden, unabhängigen Beobachters ein langes Intervall wahrscheinlicher als ein kurzes ange-troffen wird.

Dies führt zu

$$p(n) = K \cdot n \cdot a(n), \quad (2.33a)$$

wobei K eine anhand der Normierungsbedingung zu ermittelnde Konstante ist

$$\sum_{n=0}^{\infty} p(n) = K \sum_{n=0}^{\infty} n a(n) = K EA = 1 \quad (2.33b)$$

oder

$$p(n) = n a(n) / EA. \quad (2.33c)$$

Unter der Bedingung, daß das Zwischenankunftsintervall n Zeiteinheiten lang ist, existieren n Beobachtungszeitpunkte (vgl. Menge (M_i) in Bild 2.5), die gleichwahrscheinlich angetroffen werden, d.h.

$$a^*(k|n) = \begin{cases} 1/n & k = 0, 1, \dots, n-1 \\ 0 & \text{sonst.} \end{cases} \quad (2.34)$$

Schließlich erhält man aus den Gleichungen (2.32), (2.33c) und (2.34) für die Vorwärts-Rekurrenzzeit-Verteilung

$$a^*(k) = \frac{1}{EA} \left(1 - \sum_{i=0}^k a(i) \right), \quad k = 0, 1, \dots \quad (2.35)$$

oder nach der Z-Transformation

$$a_{ZT}^*(z) = \frac{1 - a_{ZT}(z)}{EA (1 - z^{-1})} \quad (2.36)$$

Der Mittelwert ergibt sich zu

$$EA^* = \frac{EA}{2} (c_A^2 + 1) - \frac{1}{2} \quad (2.37)$$

Ein Vergleich mit der Beziehung für den Mittelwert der Vorwärts-Rekurrenzzeit kontinuierlicher Erneuerungsprozesse (Gl. (2.29)) zeigt eine normierte Differenz von $1/2$. Dieser Unterschied kommt dadurch zustande, daß die Dauer der Vorwärts-Rekurrenzzeit bei einem angetroffenen Zwischenankunftsintervall der Länge n zwischen $0 \cdot \Delta t$ und $(n-1) \cdot \Delta t$ gleichverteilt mit dem Mittelwert $(n-1) \cdot \Delta t / 2$ variiert (siehe Bild 2.5).

Die Berechnung der Verteilung $a^*(k)$ zeitdiskreter Erneuerungsprozesse kann rekursiv erfolgen :

$$\begin{aligned} a^*(0) &= \frac{1}{EA} (1 - a(0)) \\ a^*(k) &= a^*(k-1) - \frac{1}{EA} a(k), \quad k=1, 2, \dots \end{aligned} \quad (2.38)$$

b) Beobachtung nach den diskreten Zeitpunkten

Für die Konstruktion der Vorwärts-Rekurrenzzeit wird nun festgelegt, daß der Beobachtungszeitpunkt t^* unmittelbar nach einem diskreten Zeitpunkt liegt. Die Herleitung der Vorwärts-Rekurrenzzeit-Verteilung erfolgt analog zum vorausgegangenen Unterkapitel.

Die bedingte Verteilung aus Gl.(2.34) lautet nun

$$a^*(k|n) = \begin{cases} 1/n & k = 1, \dots, n \\ 0 & \text{sonst} \end{cases} \quad (2.39)$$

Man erhält für die Vorwärts-Rekurrenzzeit-Verteilung

$$a^*(k) = \sum_{n=k}^{\infty} a^*(k|n) \cdot p(n) = \frac{1}{EA} \left(1 - \sum_{i=0}^{k-1} a(i) \right), \quad k = 1, 2, \dots \quad (2.40)$$

oder nach der Z-Transformation

$$a_{ZT}^*(z) = \frac{z^{-1} (1 - a_{ZT}(z))}{EA (1 - z^{-1})} \quad (2.41)$$

Der Mittelwert errechnet sich nun zu

$$EA^* = \frac{EA}{2} (c_A^2 + 1) + \frac{1}{2} \quad (2.42)$$

Ein Vergleich von Gl.(2.40) mit (2.36) zeigt, daß die Vorwärts-Rekurrenzzeit nach Gl.(2.40) die um Δt verschobene Verteilung nach Gl.(2.36) darstellt.

c) Rückwärts-Rekurrenzzeit-Verteilung

Ähnlich wie bei zeitkontinuierlichen Erneuerungsprozessen wird im diskreten Zeitbereich häufig die Rückwärts-Rekurrenzzeit benötigt. Bei der Konstruktion des Rückwärts-Rekurrenzintervalls wird festgelegt, daß der Beobachtungszeitpunkt für die Rückwärts-Rekurrenzzeit unmittelbar nach einem diskreten Zeitpunkt liegt,

falls der Beobachtungszeitpunkt für die Vorwärts-Rekurrenzzeit unmittelbar vor einem diskreten Zeitpunkt vereinbart wurde, und umgekehrt. Die Rückwärts-Rekurrenzzeit ist dann die Zeitspanne zwischen dem so definierten Beobachtungszeitpunkt und dem letzten Ereigniszeitpunkt. Unter dieser Voraussetzung ist die zeitdiskrete Rückwärts-Rekurrenzzeit-Verteilung jeweils identisch mit der Verteilung der Vorwärts-Rekurrenzzeit, die in Gl.(2.35) und (2.40) angegeben ist.

d) Zeitdiskrete Erneuerungsprozesse mit gedächtnisloser Eigenschaft

Bei Erneuerungsprozessen mit gedächtnisloser Eigenschaft ist die Vorwärts-Rekurrenzzeit-Verteilung identisch mit der Zwischenankunftsverteilung /18,19,20/. Dies gilt auch für zeitdiskrete Erneuerungsprozesse. Abhängig von der Vereinbarung bzgl. der Lage des Beobachtungszeitpunktes, d.h. abhängig davon, ob dieser unmittelbar vor oder nach einem diskreten Zeitpunkt liegt, können verschiedene Prozeßtypen, welche die gedächtnislose Eigenschaft (Markoffsche Eigenschaft) besitzen, gefunden werden.

- Beobachtung vor den diskreten Zeitpunkten :

Aus Tabelle 2.1 und Gl.(2.36) kann gezeigt werden, daß Erneuerungsprozesse mit GEOM(0)-Zwischenankunftsverteilung die Markoffsche Eigenschaft aufweisen:

$$a_{ZT}^*(z) = \frac{1 - a_{ZT}(z)}{EA(1 - z^{-1})} = \frac{1 - \frac{1 - q}{1 - qz^{-1}}}{\frac{q}{1 - q}(1 - z^{-1})} = a_{ZT}(z). \quad (2.43)$$

- Beobachtung nach den diskreten Zeitpunkten :

In diesem Falle erhält man aus Tabelle 2.1 und Gl.(2.41) den Bernoulli-Ankunftsprozeß (vgl. Kap. 2.3.1), dessen Zwischenankunftsabstände GEOM(1)-verteilt sind, als gedächtnislosen Erneuerungsprozeß

$$a_{zT}^*(z) = \frac{z^{-1}(1 - a_{zT}(z))}{EA(1 - z^{-1})} = \frac{z^{-1}(1 - \frac{1-q}{1-qz^{-1}})}{\frac{q}{1-q}(1 - z^{-1})} = a_{zT}(z). \quad (2.44)$$

Der Bernoulli-Ankunftsprozeß wird in der Literatur /6,37,39,53/ sehr häufig zur Charakterisierung gedächtnisloser Verkehrsströme, insbesondere als zum Poisson-Prozeß analogen Prozeß im diskreten Zeitbereich, angewendet.

2.4.4 Anzahl von Ereignissen in einem beliebig verteilten Intervall

Betrachtet werde ein zeitdiskreter Erneuerungsprozeß mit der ZV A für Zwischenankunftsabstände. Der Prozeß wird während eines Intervalls mit der zeitdiskreten ZV V und der Verteilung $v(m)$, $m=0,1,\dots$, beobachtet. Es wird vorausgesetzt, daß die zwei Endpunkte des Intervalls V jeweils unmittelbar vor einem diskreten Zeitpunkt (Bild 2.6) liegen. Die Verteilung $r(j)$ der Anzahl R von Ereignissen (Erneuerungen), die während des zufälligen Beobachtungsintervalls eintreffen, wird im folgenden hergeleitet.

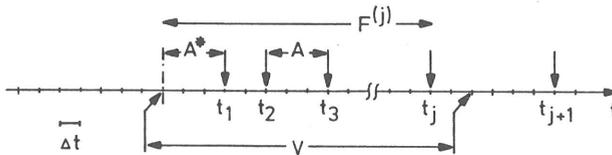


Bild 2.6 Anzahl von Ereignissen in einem zufälligen Intervall

Die Verteilung $r(j)$ wird zunächst als totale Wahrscheinlichkeit formuliert:

$$r(j) = \sum_{m=0}^{\infty} P(R=j|V=m) \cdot P(V=m) = \sum_{m=0}^{\infty} r(j|m) v(m). \quad (2.45)$$

Bezeichnet man $F^{(j)}$ als die ZV für die Zeitspanne zwischen dem Anfang des Beobachtungsintervalls und dem j -ten Ereignis (vgl. Bild 2.6), so gilt

$$F^{(j)} = A^* + \underbrace{A + \dots + A}_{(j-1)\text{-mal}}$$

bzw.

$$f^{(j)}(k) = a^*(k) * \underbrace{a(k) * \dots * a(k)}_{(j-1)\text{-mal}}. \quad (2.46)$$

Die bedingte Verteilung $r(j|m)$ aus Gl.(2.45) errechnet sich damit zu

$$\begin{aligned} r(j|m) &= P(F^{(j)} < m \leq F^{(j+1)}) \\ &= P(F^{(j)} < m) - P(F^{(j+1)} < m). \end{aligned} \quad (2.47a)$$

Unter Berücksichtigung des Sonderfalls, daß ein Intervall V der Länge $m=0$ keine Ereignisse enthält, erhält man schließlich :

$$\begin{aligned} r(j|0) = \delta(j) &= \begin{cases} 1 & j=0 \\ 0 & \text{sonst} \end{cases}, \quad m=0 \\ r(j|m) &= \sum_{i=0}^{m-1} (f^{(j)}(i) - f^{(j+1)}(i)), \quad m=1,2,\dots \end{aligned} \quad (2.47b)$$

Aus Gl. (2.45) und Gl. (2.47) erhält man die Verteilung der Anzahl von Ereignissen während des Beobachtungsintervalls V

$$\begin{aligned} r(j) &= v(0) \delta(j) \\ &+ \sum_{m=1}^{\infty} v(m) \sum_{i=0}^{m-1} (f^{(j)}(i) - f^{(j+1)}(i)), \quad j=0,1,\dots \end{aligned} \quad (2.48)$$

2.5 Das Konzept des komplexen Cepstrums

In der Analyse zeitdiskreter Systeme im transformierten Bereich, z.B. bei Systemen vom Typ GI/G/1 /2,61,63,71/, wird u.a. das aus der Theorie der Signalverarbeitung stammende Konzept des komplexen Cepstrums /8,55/ angewendet. Dieses Konzept läßt sich mit Elementen der homomorphen Systemtheorie /55/ systematisch beschreiben. Vor der Einführung und Erläuterung des Cepstrum-Konzeptes werden deshalb einige Eigenschaften der Klasse homomorpher Transformationen und Systeme diskutiert.

2.5.1 Homomorphe Systeme und Transformationen

Betrachtet werde ein System, das die Transformation T ausführt (s. Bild 2.7a). Die Verknüpfungen (I) und (O) werden jeweils für die Menge der Eingangssignale x_i bzw. der Ausgangssignale y_i definiert. Ferner wird eine skalare Größe c mit den Eingangssignalen durch den Operator (i) und mit den Ausgangssignalen durch den Operator (o) verknüpft. Das System bzw. die Transformation wird als homomorph bezeichnet, falls

$$T(x_1(I) x_2) = T(x_1) (O) T(x_2) = y_1 (O) y_2 \quad (2.49a)$$

$$T(c (i) x_1) = c (o) T(x_1) = c (o) y_1 \quad (2.49b)$$

Wie aus den Gleichungen (2.49a,b) zu ersehen ist, stellt die homomorphe Eigenschaft eine Verallgemeinerung des Linearität-Begriffes in der Systemtheorie dar. Ein Beispiel homomorpher Systeme wird in Bild 2.7b gezeigt, wobei

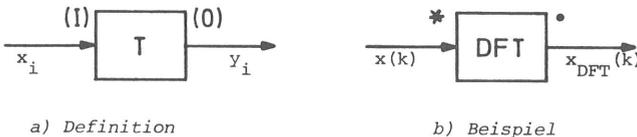


Bild 2.7 Homomorphe Systeme

- Eingangs- und Ausgangssignale : Vektoren von komplexen Zahlen
- Transformation T : Diskrete Fourier-Transformation (DFT, vgl Gl.(2.13))
- Eingangsverknüpfung : diskrete Faltung ("*")
- Ausgangsverknüpfung : elementenweise Multiplikation von Vektoren
- Eingangs- und Ausgangsverknüpfung mit skalaren Größen : Multiplikation.

2.5.2 Das komplexe Cepstrum

a) Homomorphe Systeme für gefaltete Signale

Werden homomorphe Systeme in Serie geschaltet, so ist das gesamte System ebenfalls homomorph. Diese Eigenschaft spielt eine bedeutende Rolle in der Signalverarbeitung und führt zu kanonischen Strukturen homomorpher Systeme.

Betrachtet werde nun die in Bild 2.8 dargestellte kanonische Struktur homomorpher Systeme (vgl./55/), die folgende Eigenschaften aufweist:

- Die Klasse $x(k)$ zeitdiskreter Eingangssignale sowie die Klasse $y(k)$ zeitdiskreter Ausgangssignale (z.B. Abtastwerte in Übertragungssystemen, Wahrscheinlichkeits-Verteilungen, usw.) werden mit der diskreten Faltungsoperation verknüpft.
- Durch die homomorphe Transformation T, über deren Aufbau im nächsten Unterkapitel ausführlich diskutiert wird, werden Faltungsoperationen in Additionen umgewandelt. Dabei wird aus einem Eingangssignal $x(k)$ das zugehörige komplexe Cepstrum $x_{\text{CEP}}(k)$ gebildet.
- Das innere System führt im Cepstrum-Bereich Transformationen durch, bei denen die Addition als Verknüpfungsvorschrift erhalten bleibt.

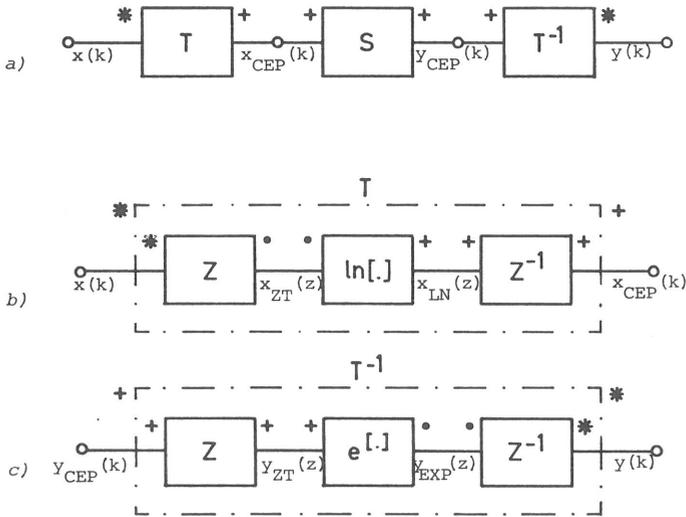


Bild 2.8 Kanonische Form eines homomorphen Systems für gefaltete Signale

- a) Das gesamte System
- b) Teilsystem zur Transformation
- c) Teilsystem zur Rücktransformation

- Die Rücktransformation T^{-1} führt die inverse Abbildung vom Cepstrum-Bereich zum Zeitbereich durch.

Wie zu erkennen ist, besteht das Ziel der vorgestellten Struktur darin, mittels homomorpher Transformationen faltungsbehaftete Operationen im Ursprungsbereich (diskreter Zeitbereich) durch Additionen im transformierten Bereich (Cepstrum-Bereich) zu ersetzen.

b) Definition des komplexen Cepstrums

Die Bildung des komplexen Cepstrums und die zugehörige inverse Transformation werden in Bild 2.8b und Bild 2.8c schematisch dargestellt. Dabei setzen sich die Transformation T und die Rücktransformation T^{-1} aus jeweils drei Teiltransformationen zusammen. Die Cepstrum-Bildung T erfolgt in folgenden drei Schritten :

- Z-Transformation
(Umwandlung der Faltungsoperation in Multiplikation)

$$x_{ZT}(z) = Z \{x(k)\} . \quad (2.50a)$$

- Bildung des komplexen Logarithmus
(Umwandlung der Multiplikation in Addition)

$$x_{LN}(z) = \ln(x_{ZT}(z)) . \quad (2.50b)$$

Das Attribut "komplex" des Cepstrums bezieht sich auf die Anwendung des komplexen Logarithmus und besagt nicht, daß das komplexe Cepstrum prinzipiell komplexwertig sein muß. In der Literatur werden außer dem hier definierten komplexen Cepstrum ebenfalls Phasen-Cepstrum sowie Betrags-Cepstrum angewendet.

- inverse Z-Transformation

$$x_{CEP}(k) = Z^{-1}\{x_{LN}(z)\} . \quad (2.50c)$$

Gl.(2.50c) setzt voraus, daß $x_{LN}(z)$ eine gültige Z-Transformierte darstellt.

Entsprechend erfolgt die inverse Transformation zur Cepstrumbildung T^{-1} (vgl. Bild 2.8c):

- Z-Transformation

$$Y_{ZT}(z) = Z \{y_{CEP}(k)\} \quad (2.51a)$$

- Exponentialfunktion

$$Y_{EXP}(z) = e^{y_{ZT}(z)} \quad (2.51b)$$

- inverse Z-Transformation

$$y(k) = Z^{-1}\{y_{EXP}(z)\} \quad (2.51c)$$

c) Eigenschaften des komplexen Cepstrums

Im folgenden werden einige Eigenschaften des komplexen Cepstrums erläutert, die für die im Kapitel 3 vorgestellte Analyse von Bedeutung sind. Dabei wird speziell die häufig anzutreffende Klasse von Signalen bzw. Zeitfolgen, die rationale Funktionen als Z-Transformierte besitzen, betrachtet. Für diese Klasse lautet im Z-Bereich (vgl./55/) :

$$x_{ZT}^*(z) = \frac{A \cdot z^r \cdot \prod_{n=1}^{n_i} (1 - z^{-1} q_{in}) \cdot \prod_{n=1}^{n_o} (1 - z/q_{on})}{\prod_{n=1}^{m_i} (1 - z^{-1} p_{in}) \prod_{n=1}^{m_o} (1 - z/p_{on})}, \quad (2.52)$$

wobei

- q_{in} und q_{on} Nullstellen innerhalb bzw. außerhalb des Einheitskreises und
- p_{in} und p_{on} Pole innerhalb bzw. außerhalb des Einheitskreises sind.

Bei der anschließenden Logarithmus-Bildung der Funktion $x_{ZT}^*(z)$, bei der A positiv sein muß, ist der Term z^r für die weitere Berechnung (Z-Rücktransformation) im allgemeinen schwierig /55/ zu handhaben. Da der Term z^r jedoch in der ursprünglichen Folge $x^*(k)$ lediglich eine Verschiebung um r Stellen bedeutet, wird zur Vereinfachung der Cepstrum-Bildung häufig statt $x^*(k)$ die abgeänderte Folge $x(k)$, die durch eine entsprechende Verschiebung von $x^*(k)$ und Betragsbildung von A entsteht, genommen.

Die weiteren Schritte zur Berechnung des komplexen Cepstrums können nun gemäß Gl. (2.50a,b) erfolgen (vgl. Bild 2.8b) :

$$x_{ZT}(z) = |A| \frac{\prod_{n=1}^{n_i} (1 - z^{-1} q_{in}) \cdot \prod_{n=1}^{n_o} (1 - z/q_{on})}{\prod_{n=1}^{m_i} (1 - z^{-1} p_{in}) \prod_{n=1}^{m_o} (1 - z/p_{on})}, \quad (2.53a)$$

$$\begin{aligned}
 x_{LN}(z) = & \ln|A| + \sum_{n=1}^{n_i} \ln(1 - z^{-1}q_{in}) + \sum_{n=1}^{n_o} \ln(1 - z/q_{on}) \\
 & - \sum_{n=1}^{m_i} \ln(1 - z^{-1}p_{in}) - \sum_{n=1}^{m_o} \ln(1 - z/p_{on})
 \end{aligned}
 \tag{2.53b}$$

Da $|q_{in}|, |p_{in}| < 1$ und $|q_{on}|, |p_{on}| > 1$, kann $x_{LN}(z)$ durch die Reihenentwicklungen

$$\ln(1 - uz^{-1}) = - \sum_{k=1}^{\infty} \frac{u^k}{k} z^{-k}, \quad |u| < 1, \quad |z| > |u|,$$

(2.54)

$$\ln(1 - z/v) = - \sum_{k=1}^{\infty} \frac{v^{-k}}{k} z^k, \quad |v| > 1, \quad |z| < |v|$$

wie folgt angegeben werden :

$$\begin{aligned}
 x_{LN}(z) = & \sum_{k=-\infty}^{-1} \left(\sum_{n=1}^{n_o} \frac{q_{on}^k}{k} - \sum_{n=1}^{m_o} \frac{p_{on}^k}{k} \right) z^{-k} \\
 & + \ln|A| \\
 & + \sum_{k=1}^{\infty} \left(\sum_{n=1}^{m_i} \frac{p_{in}^k}{k} - \sum_{n=1}^{n_i} \frac{q_{in}^k}{k} \right) z^{-k}.
 \end{aligned}
 \tag{2.55}$$

Die Funktion $x_{LN}(z)$ in Gl. (2.55) liegt bereits in einer ähnlichen Form wie bei der Definition der Z-Transformation vor. Mit den Gleichungen (2.8), (2.50 c), (2.55) sowie dem Cauchyschen Integralsatz und dem Residuensatz läßt sich das komplexe Cepstrum ablesen

$$x_{CEP}(k) = \begin{cases} \sum_{n=1}^{n_o} \frac{q_{on}^k}{k} - \sum_{n=1}^{m_o} \frac{p_{on}^k}{k} & k < 0 \\ \ln|A| & k = 0 \\ \sum_{n=1}^{m_i} \frac{p_{in}^k}{k} - \sum_{n=1}^{n_i} \frac{q_{in}^k}{k} & k > 0. \end{cases}
 \tag{2.56}$$

Aus Gl. (2.56) lassen sich folgende Eigenschaften des komplexen Cepstrum feststellen :

- Ist $x(k)$ eine Minimalphasen-Folge, d.h. liegen alle Pole und Nullstellen von $x_{ZT}(z)$ innerhalb des Einheitskreises, dann existiert $x_{CEP}(k)$ nur für $k \geq 0$.
- Ist $x(k)$ eine Maximalphasen-Folge, d.h. liegen alle Pole und Nullstellen von $x_{ZT}(z)$ außerhalb des Einheitskreises, dann existiert $x_{CEP}(k)$ nur für $k \leq 0$.
- Das komplexe Cepstrum nimmt mit steigendem k mindestens so schnell wie die Funktion $1/k$ ab.
- Das komplexe Cepstrum begrenzter Folgen im Zeitbereich kann unendlich lang sein.

Die oben aufgezeigte Eigenschaft der räumlichen Separation von Minimal- und Maximalphasenanteilen im Cepstrum-Bereich spielt eine wichtige Rolle in der zeitdiskreten Analyse des GI/G/1-Systems (vgl. Kap. 3). Diese Eigenschaft wird daher näher erörtert.

Betrachtet man das Faltungsprodukt einer Maximal- und einer Minimalphasenfolge

$$x(k) = x^{(MAX)}(k) * x^{(MIN)}(k) \quad (2.57a)$$

d.h.

$$x_{ZT}(z) = x_{ZT}^{(MAX)}(z) \cdot x_{ZT}^{(MIN)}(z), \quad (2.57b)$$

so gilt im Cepstrum-Bereich

$$x_{CEP}(k) = \begin{cases} x_{CEP}^{(MAX)}(k) & k < 0 \\ x_{CEP}^{(MAX)}(0) + x_{CEP}^{(MIN)}(0) & k = 0 \\ x_{CEP}^{(MIN)}(k) & k > 0 \end{cases} \quad (2.58)$$

d) Beispiel einer direkten Berechnung des komplexen Cepstrums

Als Beispiel wird die negativ-binomiale Verteilung genommen (vgl. Tabelle 2.1) mit

$$x(k) = \binom{y+k-1}{k} (1-q)^y q^k, \quad 0 \leq q < 1, \quad y \text{ reell}, \quad (2.59a)$$

$$x_{zT}(z) = \left(\frac{1-q}{1-qz^{-1}} \right)^y \quad (2.59b)$$

Zur Berechnung des komplexen Cepstrums wird zunächst der Logarithmus gemäß Gl.(2.50b) gebildet

$$x_{LN}(z) = y \cdot \ln(1-q) - y \cdot \ln(1-qz^{-1})$$

Läßt man den Sonderfall $q=1$ außer Betracht, für den $x(k)=0$ ($k < \infty$) und $x_{LN}(z)$ nicht definiert sind, kann das komplexe Cepstrum wie folgt berechnet werden ($q < 1$) :

$$x_{LN}(z) = y \cdot \ln(1-q) + y \sum_{k=1}^{\infty} \frac{q^k}{k} z^{-k},$$

$$x_{CEP}(k) = z^{-1}\{x_{LN}(z)\} = \begin{cases} 0 & k < 0 \\ y \cdot \ln(1-q) & k = 0 \\ y \frac{q^k}{k} & k > 0 \end{cases} \quad (2.60)$$

Aus Gl. (2.60) ist erkennbar, daß das komplexe Cepstrum der negativ-binomialen Verteilung nur Minimalphasenanteile besitzt. Dies entspricht der Lage der Pole von $x_{zT}(z)$ in Gl. (2.59b).

e) Numerische Berechnung des komplexen Cepstrums

Wie in Kap. 2.2.3 bereits erörtert wurde, kann die Z-Transformierte einer endlichen Zeitfolge durch die diskrete Fourier-Transformation (DFT) ersetzt werden (Gl.(2.16)). Dementsprechend können bei der numerischen Auswertung des komplexen Cepstrums die benötigten Z-Transformationen in Gl. (2.50a) und Gl.(2.51a) durch die DFT, die inversen Z-Transformationen in Gl.(2.50c) und Gl. (2.51c) durch die inverse DFT approximiert werden. Diese Transformationen können dann mittels effektiver Algorithmen zur schnellen Fourier-Transformation (FFT) durchgeführt werden (siehe Kap. 2.2.3).

Ausgehend von einer begrenzten Folge der Länge N

$$x(k), k=0,1,\dots,N-1,$$

können die approximativen Funktionen zur numerischen Cepstrumbildung entsprechend den Gleichungen (2.50a-c) folgendermaßen angegeben werden :

$$x_{ZT}(z) : x_{DFT}(n) = DFT\{x(k)\}, n=0,1,\dots,N-1 \quad (2.61a)$$

$$x_{LN}(z) : x_{LN,Appr.}(n) = \ln x_{DFT}(n), n=0,1,\dots,N-1 \quad (2.61b)$$

$$x_{CEP}(k) : x_{CEP,Appr.}(k) = DFT^{-1}\{x_{LN,Appr.}(n)\}, k=0,1,\dots,N-1. \quad (2.61c)$$

Analog lauten die approximativen Funktionen für die inverse Transformation zur Cepstrumbildung gemäß Gl.(2.51a-c) :

$$y_{ZT}(z) : y_{DFT}(n) = DFT\{y_{CEP}(k)\}, n=0,1,\dots,N-1 \quad (2.62a)$$

$$y_{EXP}(z) : y_{EXP,Appr.}(n) = e^{y_{DFT}(n)}, n=0,1,\dots,N-1 \quad (2.62b)$$

$$y(k) : y_{Appr.}(k) = DFT^{-1}\{y_{EXP,Appr.}(n)\}, k=0,1,\dots,N-1 \quad (2.62c)$$

Bei der Auswertung der Gleichungen (2.61a-c) treten folgende Problematiken auf (vgl. /55,61/), die zu berücksichtigen sind:

- Phasen-Entflechtung (phase unwrapping):

Bei der Bildung des komplexen Logarithmus nach Gl.(2.61b) können Phasenanteile mehrdeutig werden, da sie Sprünge um $k \cdot 2\pi$ aufweisen können. Diese Sprünge sind noch im entstehenden Phasenverlauf, der sich aus eine Reihe von N Phasenhauptwerten zusammensetzt, enthalten. Die Ermittlung des tatsächlichen Verlaufs der Phase von $x_{LN}(n)$ kann mittels eines Phasen-Entflechtungsalgorithmus (phase unwrapping algorithm) erfolgen (siehe z.B. Tribollet /72/, Oppenheim /55/). Es soll an dieser Stelle darauf hingewiesen werden, daß Algorithmen zur Berechnung des komplexen Cepstrums, die ohne Phasen-Entflechtung auskommen, existieren (vgl. /5/).

- Überlappung der Cepstrum-Komponenten (aliasing) :

Es kann gezeigt werden, daß die in Gl.(2.61c) angegebene approximative Funktion $x_{CEP,Appr.}(k)$ eine durch periodische Überlappung modifizierte Funktion des exakten komplexen Cepstrums $x_{CEP}(k)$ darstellt:

$$x_{CEP,Appr.}(k) = \sum_{j=-\infty}^{\infty} x_{CEP}(k+jN) . \quad (2.63)$$

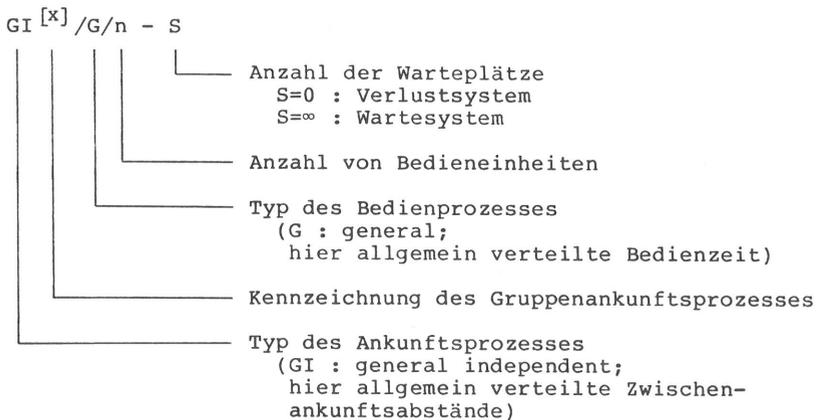
Durch die oben erwähnte Eigenschaft des komplexen Cepstrums, nach der $x_{CEP}(k)$ mit steigendem k mindestens so schnell wie $1/k$ abnimmt, können mit Hilfe einer Vergrößerung von N, d.h. einer Verlängerung der Folge $x(k)$ durch Hinzufügen von Nullelementen, die überlagerten Anteile $x_{CEP}(k+jN)$, $j \neq 0$, verringert und dadurch die Approximationsgenauigkeit von $x_{CEP,Appr.}(k)$ erhöht werden. Diese Verlängerung von $x(k)$ entspricht einer feineren Abtastung von $x_{ZT}(z)$ auf dem Einheitskreis (vgl. Gl.(2.16)) durch $x_{DFT}(n)$ (Gl.(2.61a)).

3. ZEITDISKRETE ANALYSE DES WARTESCHLANGENSYSTEMS GI/G/1

In diesem Kapitel wird die Methodik der zeitdiskreten Analyse von Verkehrsmodellen am Beispiel des einstufigen Warteschlangensystems GI/G/1 erörtert.

3.1 Modellbeschreibung und Voraussetzungen

Die Kennzeichnung von einstufigen Warteschlangensystemen erfolgt hier in Erweiterung der von Kendall (1954) eingeführten Kurznotation:



In diesem Kapitel wird das Wartesystem vom Typ GI/G/1 ($S=\infty$) im diskreten Zeitbereich untersucht. Die zufallsabhängigen, zeitdiskreten Prozesse sind wie folgt definiert:

- Ankunftsprozeß : Die Zwischenankunftszeiten mit der diskreten Zufallsvariablen (ZV) A sind allgemein verteilt mit einer Verteilung der Länge n_A

$$a(k) = P(A = k) , k = 0, 1, \dots, n_A - 1. \quad (3.1a)$$

Da im allgemeinen Fall die Wahrscheinlichkeit $a(0)$ existiert, sind Ankunftsabstände der Länge $k=0$ zulässig. Dies bedeutet implizit die Einbeziehung von Gruppenankunftsprozessen mit geo-

metrisch verteilter Gruppengröße und allgemein verteilten Zwischenankunftszeiten in der Analyse (vgl. Gl.(2.31a-c)).

- Bedienprozeß : Die Bedienzeit mit der diskreten ZV B ist allgemein verteilt mit einer Verteilung der Länge n_B

$$b(k) = P(B = k) , k = 0, 1, \dots, n_B - 1. \quad (3.1b)$$

Es werde der Wartebetrieb mit unbegrenzter Warteschlangenlänge betrachtet. Während bei der Wartezeitanalyse im Kap. 3.4 die Abfertigungsdisziplin FIFO (first-in, first-out) vorausgesetzt wird, ist die Analyse der Freiperiode im Kap. 3.5.2 sowie des Ausgangsprozesses im Kap. 3.5.3 für alle arbeitserhaltenden (vgl. /36/: work conservative) Abfertigungsdisziplinen gültig - z.B. LIFO: last-in, first-out, RANDOM, usw. -.

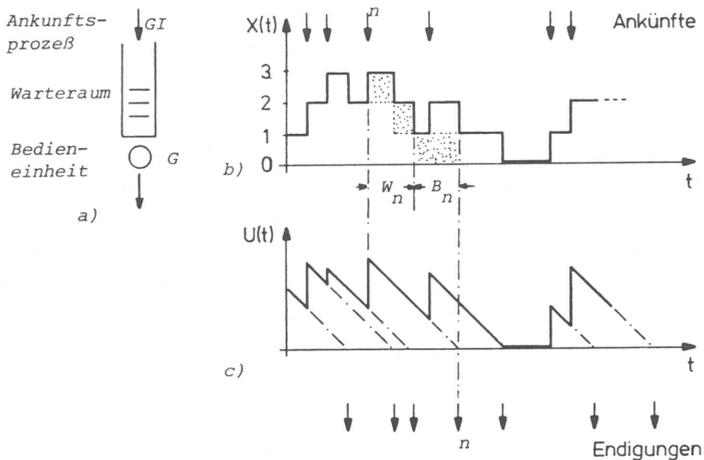


Bild 3.1 Struktur und Darstellungsformen des Zustandsprozesses eines Warteschlangensystems vom Typ GI/G/1

- a) Struktur
- b) Prozeßverlauf der Anzahl von Anforderungen im System
- c) Prozeßverlauf der Restarbeit im System

Da hier der Wartebetrieb vorausgesetzt wird, lautet die Stabilitätsbedingung des Systems

$$\rho = \frac{EB}{EA} < 1, \quad (3.2)$$

wobei ρ die Auslastung der Bedieneinheit bzw. das Verkehrsangebot im System bedeutet.

Die Struktur und ein typischer Prozeßverlauf im System GI/G/1 werden in Bild 3.1 dargestellt, wobei zwei in der Systemanalyse häufig verwendete Darstellungsformen des Zustandsprozesses illustriert sind:

- Anzahl $X(t)$ der Anforderungen im System

Der Systemzustand wird hier durch die zufallsabhängige Anzahl der zu einem Zeitpunkt t im System befindlichen Anforderungen charakterisiert (vgl. Bild 3.1b). Diese Zustandsbeschreibung wird z.B. bei der Methode der eingebetteten Markoff-Kette bzw. bei Analyseverfahren homogener Markoff-Prozesse angewendet.

- Restarbeit $U(t)$ im System

Bei dieser Beschreibungsform wird die zufallsabhängige Summe der Bedien- bzw. Restbedienzeiten aller zum Zeitpunkt t im System befindlichen Anforderungen betrachtet. Dies wird im folgenden als Restarbeit bezeichnet (vgl. Bild 3.1c). Die Restarbeit $U(t)$ erhöht sich bei der Ankunft einer Anforderung um die Bedienzeit dieser Anforderung. Solange $U(t)$ positiv ist, nimmt sie linear mit der Zeit ab. Diese Betrachtung eignet sich insbesondere in Wartezeitanalysen von Systemen mit FIFO Abfertigung, da in diesem Falle die von einer Anforderung zum Ankunftszeitpunkt ange-troffene Restarbeit genau der Wartezeit entspricht, die diese Anforderung erfährt (s. Wartezeit W_n der n -ten Anforderung in Bild 3.1). Diese Zustandsbeschreibung wird hauptsächlich bei Verfahren benutzt, die auf der Lindley-Integralgleichung basieren (vgl. Kap. 3.3).

Es existiert eine umfangreiche Literatur zur analytischen Behandlung des GI/G/1-Systems. Die meisten Arbeiten untersuchten das System im kontinuierlichen Zeitbereich, basierend auf der Lindley'schen Integralgleichung /14,25,34,35,36,49,56,67/. Für die approximative Berechnung von Mittelwerten wurden einige Beziehungen in /43/ entwickelt. Tabellen für Systemcharakteristiken des GI/G/1-Systems finden sich in /44,64/.

Einige Untersuchungen betrachteten das System GI/G/1 im diskreten Zeitbereich, wobei Separationsmethoden zur Gewinnung der Wartezeitverteilung vorgestellt wurden /2,26,37,39,40,41,58/. Konheim /41/ und Ponstein /58/ setzten Ankunfts- und Bedienprozesse gemäß Gl. (3.1a,b) voraus, wobei die Separation mittels direkter Berechnung der Nullstellen bzw. Polynom-Faktorisierung durchgeführt wurden.

Ein Separationsansatz, bei dem die Trennung im transformierten Bereich mit dem in der Signalverarbeitung häufig verwendeten Konzept des komplexen Cepstrums erfolgte, wurde in Ackroyd /2/ vorgestellt. Dabei werden die Pole und Nullstellen der zu trennenden charakteristischen Funktion des Systems GI/G/1 nicht explizit berechnet. Die Separation wurde bei diesem Verfahren durch die Betrachtung von Maximal- und Minimalphasenanteilen dieser Funktion bewerkstelligt.

3.2 Direkte Lösung des GEOM(1)/G/1-Systems durch eingebettete Markoff-Kette

Ähnlich wie im kontinuierlichen Zeitbereich kann bei zeitdiskreten Warteschlangensystemen, bei denen der Zustandsprozeß an bestimmten Zeitpunkten die gedächtnislose Eigenschaft aufweist, die Methode der eingebetteten Markoff-Kette zur Modellanalyse /36/ herangezogen werden. Dies ist z.B. zutreffend bei folgenden Systemtypen :

- GEOM(0)/G/1 bzw. GEOM(1)/G/1 (Bernoulli-Ankunftsprozeß)
- GI/GEOM(0)/1 bzw. GI/GEOM(1)/1 (Bernoulli-Bedienprozeß).

In diesem Unterkapitel wird anhand des GEOM(1)/G/1-Systems diese Analysemethode näher erläutert.

3.2.1 Systemparameter und Markoff-Kette

Für das hier betrachtete System vom Typ GEOM(1)/G/1, d.h. der Ankunftsprozeß ist ein Bernoulli-Prozeß, lautet die Zwischenankunftsverteilung:

$$a(k) = (1-q_A) q_A^{k-1}, \quad k=1,2,\dots, \quad EA = \frac{1}{1-q_A}, \quad (3.3a)$$

wobei EA auf Δt normiert angegeben wird und

$$a_{zT}(z) = \frac{1-q_A}{z-q_A}. \quad (3.3b)$$

Eine Stichprobe des Prozeßverlaufs zeigt Bild 3.2, wobei zur Zustandsbeschreibung die Anzahl der Anforderungen im System zum betrachteten Zeitpunkt dient. Es ist ersichtlich, daß unmittelbar nach einem Bedienende der Prozeß die Eigenschaft der Gedächtnislosigkeit aufweist, da

- der Bernoulli-Prozeß gedächtnislos ist, falls die Beobachtungzeitpunkte unmittelbar nach den diskreten Zeitpunkten liegen
- die gedächtnisbehaftete Bedienzeit gerade beendet ist.

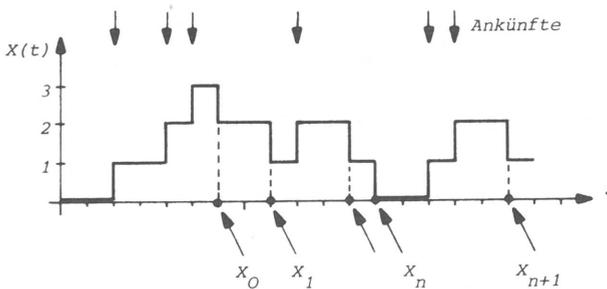


Bild 3.2 Stichprobe des Prozeßverlaufs im GEOM(1)/G/1-System

In diese Zeitpunkte, die Regenerationszeitpunkte genannt werden, wird eine Markoff-Kette eingebettet, wobei X_n die ZV für die Anzahl von Anforderungen nach dem n-ten Bedienende darstellt (s. Bild 3.2). Die folgende Analyse der Markoff-Kette erfolgt analog zur Untersuchung des zeitkontinuierlichen M/G/1-Systems /14,36/.

3.2.2 Zustandswahrscheinlichkeiten der Markoff-Kette

Zur weiteren Berechnung wird die Verteilung $r(k)$ der Anzahl R von eingetroffenen Anforderungen während einer Bedienzeit mit der Verteilung $b(k)$ benötigt. Unter Berücksichtigung der Eigenschaften des Bernoulli-Ankunftsprozesses erhält man

$$r(k) = \sum_{m=k}^{\infty} b(m) \binom{m}{k} (1-q_A)^k q_A^{m-k} \quad , k=0, 1, \dots, \quad (3.4)$$

oder nach der Z-Transformation

$$r_{ZT}(z) = b_{ZT}\left(\frac{z}{1-q_A+q_A z}\right) = b_{ZT}\left(\frac{EA \cdot z}{1+(EA-1)z}\right). \quad (3.5)$$

Ferner gilt nach der Ableitung von Gl.(3.5)

$$ER = - \left. \frac{dr_{ZT}(z)}{dz} \right|_{z=1} = \rho. \quad (3.6)$$

Betrachtet man zwei aufeinanderfolgende Regenerationszeitpunkte X_n und X_{n+1} mit den Zustandswahrscheinlichkeiten

$$x_n(k) = P(X_n=k), \quad x_{n+1}(k) = P(X_{n+1}=k), \quad k=0,1,\dots, \quad (3.7)$$

so kann das Übergangsverhalten des Systems durch die Übergangsmatrix

$$\underline{Q} = \{q_{ij}\} \quad i,j = 0,1,\dots \quad (3.8a)$$

mit

$$q_{ij} = P(X_{n+1}=j | X_n=i) \quad (3.8b)$$

vollständig beschrieben werden.

Die Übergangswahrscheinlichkeiten können mit den in Gl.(3.4) berechneten Wahrscheinlichkeiten $r(k)$ wie folgt angegeben werden:

$$q_{ij} = \begin{cases} r(j) & i=0 \\ r(j-i+1) & i>0, j \geq i-1 \\ 0 & \text{sonst} \end{cases} \quad (3.9)$$

Die Zustandswahrscheinlichkeiten $x_{n+1}(j)$, $j=0,1,\dots$, können demgemäß aus den Zustandswahrscheinlichkeiten am vorausgegangenen Regenerationszeitpunkt anhand folgender Zustandsgleichungen berechnet werden:

$$x_{n+1}(j) = x_n(0) \cdot r(j) + \sum_{i=1}^{j+1} x_n(i) \cdot r(j-i+1), \quad j=0,1,\dots \quad (3.10)$$

Für ein im eingeschwungenen Zustand befindliches System mit den stationären Zustandswahrscheinlichkeiten

$$x(j) = \lim_{n \rightarrow \infty} x_n(j), \quad j=0,1,\dots \quad (3.11)$$

erhält man aus Gl.(3.10) folgendes Gleichungssystem zur Bestimmung der Markoff-Ketten-Wahrscheinlichkeiten des GEOM(1)/G/1-Systems:

$$x(j) = x(0) \cdot r(j) + \sum_{i=1}^{j+1} x(i) \cdot r(j-i+1), \quad j=0,1,\dots \quad (3.12)$$

Die Lösung erfolgt hier im transformierten Bereich. Die Z-Transformation der Gl.(3.12) lautet :

$$x_{ZT}(z) = x(0) \frac{r_{ZT}(z)(1-z)}{1 - z r_{ZT}(z)} . \quad (3.13)$$

Der verbleibende unbekanntete Term $x(0)$ errechnet sich durch Einsetzen von $z = 1$ in Gl.(3.13) und unter Berücksichtigung des Mittelwerts ER in Gl.(3.6) zu

$$x(0) = 1 - \rho . \quad (3.14)$$

Schließlich ergeben sich aus Gl.(3.5), (3.13) und (3.14) die Zustandswahrscheinlichkeiten der eingebetteten Markoff-Kette im Z-Bereich :

$$x_{ZT}(z) = (1-\rho) \frac{(1-z) b_{ZT}\left(\frac{z}{1-q_A+q_A z}\right)}{1 - z b_{ZT}\left(\frac{z}{1-q_A+q_A z}\right)} . \quad (3.15)$$

Gl.(3.15) entspricht einer zeitdiskreten Form der Pollaczek-Khintchine Formel /14,36/ für Zustandswahrscheinlichkeiten des zeitkontinuierlichen M/G/1-Systems.

Zunächst gilt diese Beziehung nur an den Regenerationszeitpunkten der eingebetteten Markoff-Kette (s. Bild 3.2). Es kann jedoch speziell für das GEOM(1)/G/1-System nachgewiesen werden (vgl./39/), daß die in Gl.(3.15) angegebenen stationären Zustandswahrscheinlichkeiten für beliebige Prozeß- bzw. Beobachtungszeitpunkte gelten, die unmittelbar nach den diskreten Zeitpunkten liegen.

3.2.3 Wartezeitverteilung

Die Wartezeitverteilung wird nachfolgend aus den Zustandswahrscheinlichkeiten hergeleitet. Bezeichnet man W und F als Zufallsvariablen für die Wartezeit bzw. Durchlaufzeit von Anforderungen im System, so gilt:

$$F = W + B \quad (3.16a)$$

bzw. nach Gl.(2.23a,b)

$$f(k) = w(k) * b(k) \quad (3.16b)$$

$$f_{zT}(z) = w_{zT}(z) \cdot b_{zT}(z) \quad (3.16c)$$

Setzt man die Abfertigungsdisziplin FIFO (first-in, first-out) voraus, so entspricht die Anzahl der eintreffenden Anforderungen während der Durchlaufzeit einer zufällig beobachteten Testanforderung exakt der Anzahl von Anforderungen im System unmittelbar nach der Bedienzeit dieser Testanforderung, d.h. diese Anzahl entspricht dem Systemzustand an den Regenerationszeitpunkten. Dies bedeutet, daß die Verteilung der Anzahl der während einer Durchlaufzeit eintreffenden Anforderungen den Markoff-Ketten-Zustandswahrscheinlichkeiten entspricht, oder analog zur Gl.(3.4)

$$x(k) = \sum_{m=k}^{\infty} f(m) \binom{m}{k} (1-q_A)^k q_A^{m-k}, k=0,1,\dots \quad (3.17)$$

oder nach der Z-Transformation

$$x_{zT}(z) = f_{zT}\left(\frac{z}{1-q_A+q_A z}\right) \quad (3.18)$$

Vergleicht man Gl.(3.18) mit Gl.(3.15), so ergibt sich

$$f_{zT}(z) = (1-\rho) \frac{(1-z) b_{zT}(z)}{1 - q_A z - (1-q_A)z b_{zT}(z)} \quad (3.19)$$

Aus Gl.(3.16c) und Gl.(3.19) erhält man die Wartezeitverteilung von Anforderungen in einem System vom Typ GEOM(1)/G/1

$$w_{ZT}(z) = \frac{(1-\rho) \cdot (1-z)}{1 - q_A z - (1-q_A)z b_{ZT}(z)} \cdot \quad (3.20)$$

Gl.(3.20) kann als eine zeitdiskrete Form der Pollaczek-Khintchine Formel /14,36/ für die Wartezeitverteilungsfunktion des zeitkontinuierlichen M/G/1-Systems interpretiert werden.

Für den speziellen Fall des GEOM(1)/GEOM(1)/1-Systems mit der Bedienzeitverteilung :

$$b(k) = (1-q_B) q_B^{k-1} \quad , \quad k=1,2,\dots, \quad EB = \frac{1}{1-q_B} \quad (3.21a)$$

$$b_{ZT}(z) = \frac{1-q_B}{z-q_B} \quad , \quad (3.21b)$$

kann aus Gl.(3.20) die Wartezeitverteilung wie folgt angegeben werden

$$w_{ZT}(z) = \frac{(1-\rho) \cdot (z-q_B)}{q_A z - q_B} = \frac{(1-\rho) \cdot ((z-1)EB+1)}{z\rho(EA-1) - (EB-1)} \cdot \quad (3.22)$$

Dieses System kann als ein zeitdiskretes Äquivalent des M/M/1-Systems betrachtet werden (vgl./6,37,39/).

3.3 Die Lindley-Integralgleichung im diskreten Zeitbereich

3.3.1 Allgemeines

In diesem Unterkapitel werden GI/G/1-Systeme mit allgemeinen Zwischenankunfts- und Bedienprozessen betrachtet, für die keine geschlossene analytische Lösung existiert. In der Analyse dieses Systemtyps wird häufig zur Charakterisierung des Zustandsprozesses die Restarbeit im System (vgl. Kap. 3.1) genommen. Dies führt im allgemeinen zu einer von Lindley /14,49/ eingeführten Integralgleichung, die in der Analyse des GI/G/1-Wartesystems im kontinuierlichen Zeitbereich eine zentrale Rolle spielt.

Bezeichnet man $A(t)$ bzw. $B(t)$ als die Verteilungsfunktionen der Zwischenankunftszeit und Bedienzeit, $a(t)$ bzw. $b(t)$ als die zugehörigen Verteilungsdichtefunktionen, so gilt unter stationären Bedingungen für die Wartezeitverteilungsfunktion $W(t)$ bzw. für die Verteilungsdichtefunktion $w(t)$ folgende funktionale Beziehung:

$$W(t) = \begin{cases} 0 & t < 0 \\ W(t) * c(t) & t \geq 0 \end{cases} \quad (3.23)$$

wobei

$$c(t) = a(-t) * b(t) \quad (3.24)$$

und "*" den Faltungsoperator im kontinuierlichen Zeitbereich kennzeichnet. Da in $c(t)$ alle zufallsabhängigen Prozesse des Systems GI/G/1 parametrisch enthalten sind, wird diese Funktion häufig als Systemfunktion bezeichnet.

Gl.(3.23) gibt die Lindleysche Integralgleichung zur Bestimmung der Wartezeitverteilungsfunktionen des GI/G/1-Wartesystems im kontinuierlichen Zeitbereich wieder. Es ist ersichtlich, daß die Integralgleichung nach Lindley eine leicht modifizierte Form der in der mathematischen Physik und stochastischen Systemtheorie bekannten und oft anzutreffenden Wiener-Hopf-Integralgleichung darstellt. Bildet man die Ableitung beider Seiten der Gl.(3.23), so

erhält man nach kurzer Umformung folgende Form für die Wartezeit-Verteilungsdichtefunktion :

$$w(t) = \begin{cases} 0 & t < 0 \\ \delta(t) \int_{-\infty}^{0^+} (w(u) * c(u)) du & t = 0 \\ w(t) * c(t) & t > 0 \end{cases} \quad (3.25)$$

oder in einer kompakteren Notation /2,36/

$$w(t) = \pi_0(w(t) * c(t)) , \quad (3.26)$$

wobei der Operator π_0 im kontinuierlichen Zeitbereich wie folgt definiert wird:

$$\pi_0(f(t)) = \begin{cases} 0 & t < 0 \\ \delta(t) \int_{-\infty}^{0^+} f(u) \cdot du & t = 0 \\ f(t) & t > 0 \end{cases} . \quad (3.27)$$

3.3.2 Herleitung der Lindley-Integralgleichung im diskreten Zeitbereich

Aus der für zeitkontinuierliche Systeme angegebenen Form nach Gl. (3.23) lassen sich prinzipiell entsprechende Ausdrücke für zeitdiskrete GI/G/1-Systeme herleiten. Dies kann z.B. unter Anwendung der Darstellung von Dichtefunktionen mit gewichteten Dirac-Funktionen (vgl. Gl.(2.1b)) und ihrer Ausblendeigenschaft geschehen.

Im folgenden wird jedoch wegen der Übersichtlichkeit und der in späteren Kapiteln benötigten Zufallsvariablen eine analoge Form im diskreten Zeitbereich für die Lindleysche Integralgleichung direkt hergeleitet.

Bild 3.3 zeigt eine Stichprobe des Prozeßverlaufs der Restarbeit im System GI/G/1. Wie in Kapitel 3.1 dargestellt wurde, versteht

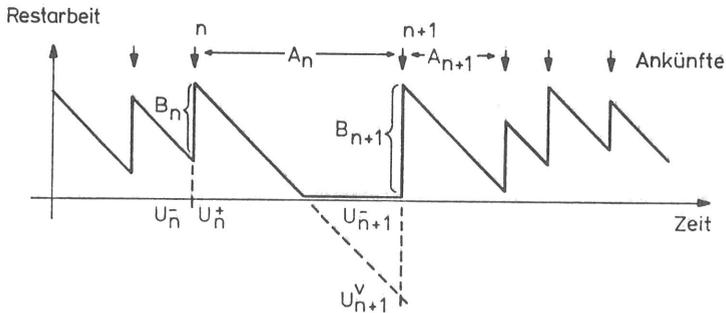


Bild 3.3 Stichprobe des Prozeßverlaufs der Restarbeit im GI/G/1-System

man unter Restarbeit die Summe von Bedien- und Restbedienzeiten aller zum betrachteten Zeitpunkt im System befindlichen Anforderungen. Da die Bedienzeitverteilung als zeitdiskret vorausgesetzt worden ist, wird die Restarbeit ebenfalls diskreter Natur, bestehend aus einer ganzzahligen Anzahl von Arbeitseinheiten. Ist die Bedieneinheit belegt, so wird pro diskretisierter Zeiteinheit genau eine Arbeitseinheit abgefertigt. Dies führt zu einem stufenförmig abfallenden Verlauf der Restarbeit im System, der aus Gründen der Übersichtlichkeit in Bild 3.3 geradlinig dargestellt wird.

Folgende Zufallsvariablen (ZV) werden eingeführt (vgl. Bild 3.3):

- A_n ZV für die Zeitdauer zwischen den Ankunftszeitpunkten der n-ten und der (n+1)-ten Anforderung
- B_n ZV für die Bedienzeit der n-ten Anforderung
- U_n^- ZV für die Restarbeit im System unmittelbar vor dem Ankunftszeitpunkt der n-ten Anforderung
- U_n^+ ZV für die Restarbeit im System unmittelbar nach dem Ankunftszeitpunkt der n-ten Anforderung
- U_{n+1}^v ZV für die virtuelle Restarbeit im System unmittelbar vor dem Ankunftszeitpunkt der (n+1)-ten Anforderung.

Die Betrachtung der virtuellen Restarbeit wird hier zur Vereinfachung der Herleitung eingeführt. Die virtuelle Restarbeit wird für

den theoretischen Fall definiert, in dem die Bedieneinheit während eines Zwischenankunftsintervalls stetig weiter bedienen würde, selbst wenn keine Anforderungen im System vorhanden sind. Demgemäß kann die virtuelle Restarbeit negative Werte annehmen (s. Bild 3.3).

Die Restarbeit wird nach der Ankunft der n-ten Anforderung um die Bedienzeit dieser Anforderung erhöht, d.h.

$$U_n^+ = U_n^- + B_n . \quad (3.28a)$$

Gemäß Gl.(2.23b) für die Summe zweier unabhängiger Zufallsvariablen erhält man

$$u_n^+(k) = u_n^-(k) * b_n(k) , \quad (3.28b)$$

wobei das Symbol "*" hier für die diskrete Faltungsoperation verwendet wird. Aus Bild 3.3 ist ersichtlich, daß

$$U_{n+1}^v = U_n^+ - A_n . \quad (3.29a)$$

Nach Gl.(2.24a) für die Differenz zweier unabhängiger Zufallsvariablen kann die Verteilung der virtuellen Restarbeit wie folgt angegeben werden :

$$u_{n+1}^v(k) = u_n^+(k) * a_n(-k) . \quad (3.29b)$$

Eine Beziehung zwischen der Restarbeit und der virtuellen Restarbeit kann hergestellt werden, indem man berücksichtigt, daß in einem leeren System die virtuelle Restarbeit weiter abnimmt, während die reale Restarbeit den Wert Null behält, d.h. :

$$U_{n+1}^- = \max(0, U_{n+1}^v) \quad (3.30a)$$

oder

$$u_{n+1}^-(k) = \pi_0(u_{n+1}^v(k)) , \quad (3.30b)$$

wobei π_m hier ein im diskreten Zeitbereich definierter Operator ist, welcher der Definition nach Gl.(3.27) entspricht

$$\pi_m(x(k)) = \begin{cases} 0 & k < m \\ \sum_{i=-\infty}^m x(i) & k = m \\ x(k) & k > m \end{cases} \quad (3.31)$$

Aus (3.28), (3.29) und (3.30) erhält man die allgemeine diskrete Form der in (3.23) angegebenen Lindleyschen Integralgleichung:

$$\begin{aligned} u_{n+1}^-(k) &= \pi_0(u_n^-(k) * a_n(-k) * b_n(k)) \\ &= \pi_0(u_n^-(k) * c_n(k)) \end{aligned} \quad (3.32)$$

oder

$$w_{n+1}(k) = \pi_0(w_n(k) * c_n(k)) \quad (3.33a)$$

wobei

$$c_n(k) = a_n(-k) * b_n(k) \quad (3.33b)$$

die zeitdiskrete Systemfunktion bildet.

Es soll an dieser Stelle darauf hingewiesen werden, daß die Beziehung (3.33a) für die Wartezeitverteilungen zweier aufeinanderfolgender Anforderungen in der Analyse weitgehend allgemeiner GI/G/1-Systeme, in denen Zwischenankunfts- und Bedienzeiten anforderungsabhängig sind - d.h. sie sind von der Ordnungsnummer der ankommenden Anforderung abhängig -, angewendet werden kann. Der aus diesem Sachverhalt entstehende Algorithmus im Zeitbereich wird im Kapitel 3.4 diskutiert.

Setzt man voraus, daß Zwischenankunftsabstände sowie Bedienzeiten unabhängig und identisch verteilt sind (Gl.(3.1a,b)), d.h.:

$$A_n = A \quad , \quad B_n = B \quad \text{für alle } n, \quad (3.34a)$$

und ferner, daß sich das System im eingeschwungenen Zustand befindet, mit

$$W = \lim_{n \rightarrow \infty} W_n, \quad (3.34b)$$

so ergibt sich aus Gl.(3.33a) und (3.34a,b) eine der Lindley-Integralgleichung ähnliche Beziehung für stationäre, zeitdiskrete GI/G/1-Systeme

$$w(k) = \pi_0(w(k) * c(k)), \quad \text{mit } c(k) = a(-k) * b(k). \quad (3.35)$$

Aus Gl.(3.35), die für die Wartezeitverteilung angegeben wird, kann durch eine Summenbildung eine Beziehung für die Wartezeitverteilungsfunktion gewonnen werden. Für $k \geq 0$ gilt:

$$\begin{aligned} W(k) &= \sum_{i=0}^k w(i) = \sum_{i=-\infty}^k c(i) * w(i) = \sum_{i=-\infty}^k \sum_{j=-\infty}^{\infty} c(j) \cdot w(i-j) \\ &= \sum_{j=-\infty}^{\infty} c(j) W(k-j), \quad k=0,1,\dots \end{aligned}$$

oder

$$W(k) = \begin{cases} 0 & k < 0 \\ c(k) * W(k) & k \geq 0 \end{cases} \quad (3.36)$$

An dieser Stelle soll erwähnt werden, daß die Wartewahrscheinlichkeit P_W für Anforderungen im System sich wie folgt angeben läßt:

$$\begin{aligned} P_W &= P(\text{eine Anforderung muß vor der Bedienphase warten}) \\ &= P(W > 0) = 1 - W(0) = 1 - w(0) \end{aligned} \quad (3.37)$$

Für ein stabiles System, d.h.

$$\frac{EB}{EA} < 1$$

gilt folgende Beziehung

$$0 \leq P_W < 1 \quad \text{oder} \quad 0 < W(0) \leq 1 . \quad (3.38)$$

3.3.3 Charakteristische Gleichung im transformierten Bereich

Wie in Bild 3.4 dargestellt, entsteht die Wartezeitverteilung $W(k)$ durch die diskrete Faltung $c(k) * W(k)$, nach der die Anteile im negativen Zeitbereich weggelassen werden. Führt man die Funktion $W^-(k)$, $k < 0$, für diese Anteile ein, so ergibt sich :

$$W^-(k) + W(k) = c(k) * W(k) \quad (3.39)$$

oder im Z-Bereich

$$W_{ZT}^-(z) + W_{ZT}(z) = c_{ZT}(z) \cdot W_{ZT}(z)$$

bzw.

$$\frac{W_{ZT}^-(z)}{W_{ZT}(z)} = c_{ZT}(z) - 1 .$$

Nach formaler Anwendung der Z-Transformation auf $W(k)$

$$W_{ZT}(z) = \frac{W_{ZT}(z)}{1 - z^{-1}} \quad (3.40)$$

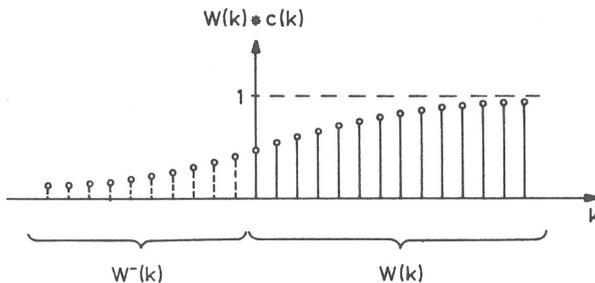


Bild 3.4 Struktur der Wartezeitverteilungsfunktion $W(k)$ und der Hilfsfunktion $W^-(k)$

erhält man die charakteristische Gleichung zur Wartezeitanalyse des zeitdiskreten GI/G/1-Systems

$$W_{ZT}^-(z) \cdot \frac{1}{w_{ZT}(z)} = \frac{c_{ZT}(z) - 1}{1 - z^{-1}} . \quad (3.41)$$

Die rechte Seite der Gl.(3.41)

$$S_{ZT}(z) = \frac{c_{ZT}(z) - 1}{1 - z^{-1}} \quad (3.42)$$

wird im folgenden als charakteristische Funktion bezeichnet.

Die charakteristische Gleichung nach (3.41), die oft auch im Bereich der erzeugenden Funktionen /39,41/ angegeben wird, spielt bei den meisten Algorithmen und Analyseverfahren zur Auflösung der zeitdiskreten Form der Lindley-Integralgleichung eine zentrale Rolle. Einige wichtige Eigenschaften der charakteristischen Gleichung bzw. der charakteristischen Funktion werden nachfolgend erläutert.

- Die Funktion $W_{ZT}^-(z)$ in der charakteristischen Gleichung ist die Z-Transformierte einer Folge mit positiven, auf der negativen Zeitachse nicht ansteigenden Koeffizienten. Nach dem Eneström-Kekeya-Theorem (vgl. /2/) liegen alle Nullstellen dieser Funktion außerhalb des Einheitskreises. Ferner, da $W^-(k)$ eine linksseitige Folge ist (s. Bild 3.4) mit $0 \leq W^-(k) < 1$, $k < 0$, kann gezeigt werden (s. /55/), daß der Konvergenzbereich der Funktion $W_{ZT}^-(z)$ sich auf den abgeschlossenen Einheitskreis erstreckt, d.h. alle ihre Pole müssen außerhalb des Einheitskreises liegen. Die Funktion $W_{ZT}^-(z)$ stellt daher die Z-Transformierte einer Maximalphasen-Folge dar (vgl. Kap. 2.5.2).
- Die Funktion $w_{ZT}(z)$ ist die Z-Transformierte einer Verteilung. Infolgedessen konvergiert diese Funktion außerhalb und auf dem Einheitskreis, d.h. alle Pole von $w_{ZT}(z)$ bzw. alle Nullstellen

des Termes $1/w_{ZT}(z)$ der charakteristischen Gleichung liegen innerhalb des Einheitskreises.

- Es gilt nach der Stabilitätsbedingung (3.38b) und dem Grenzwertsatz der Z-Transformation /55/

$$0 < \lim_{z \rightarrow \infty} w_{ZT}(z) \leq 1 \quad . \quad (3.43)$$

3.3.4 Direkte Auflösung am Beispiel des Systems GEOM(m)/GEOM(m)/1

Bei einigen Systemen kann aufgrund der oben aufgeführten Eigenschaften die Wartezeitverteilung direkt aus der charakteristischen Gleichung hergeleitet werden. Anhand des Systems GEOM(m)/GEOM(m)/1 wird dieser Lösungsweg aufgezeigt.

Bei den verschobenen geometrischen Verteilungen der Zwischenankunftsabstände und Bedienzeiten wird die Verschiebung m einheitlich für beide Prozesse gewählt, d.h.

$$a(k) = (1 - q_A) q_A^{k-m}, \quad k \geq m, \quad EA = m + \frac{q_A}{1 - q_A} \quad (3.44a)$$

$$b(k) = (1 - q_B) q_B^{k-m}, \quad k \geq m, \quad EB = m + \frac{q_B}{1 - q_B} \quad (3.44b)$$

oder im Z-Bereich

$$a_{ZT}(z) = \frac{(1 - q_A) z^{-m}}{1 - q_A z^{-1}} \quad (3.45a)$$

$$b_{ZT}(z) = \frac{(1 - q_B) z^{-m}}{1 - q_B z^{-1}} \quad . \quad (3.45b)$$

Das System ist stabil für

$$\rho = \frac{EB}{EA} < 1 \quad \text{oder} \quad \frac{q_B}{q_A} < 1 . \quad (3.46)$$

Aus Gl.(3.33b) ergibt sich für die Systemfunktion

$$c_{zT}(z) = a_{zT}(z^{-1}) \cdot b_{zT}(z) = \frac{(1-q_A)(1-q_B)}{(1-q_A z)(1-q_B z^{-1})} . \quad (3.47)$$

Die charakteristische Gleichung des Systems lautet gemäß Gl.(3.42)

$$S_{zT}(z) = \frac{q_A z - q_B}{(1-q_A z)(1-q_B z^{-1})} . \quad (3.48a)$$

oder in einer nach Gl.(2.53a) für die Erkennung der Lage von Polen und Nullstellen günstigeren Form

$$S_{zT}(z) = q_A \frac{z(1 - \frac{q_B}{q_A} z^{-1})}{(1-q_A z)(1-q_B z^{-1})} = W_{zT}^-(z) \cdot \frac{1}{w_{zT}(z)} . \quad (3.48b)$$

Der Term $(1-q_A z)$ entspricht einem außerhalb des Einheitskreises liegenden Pol, welcher der Maximalphasen-Folge $W^-(k)$ angehört. Aus Gl.(3.48b) kann die Wartezeitverteilung wie folgt geschrieben werden :

$$w_{zT}(z) = K_0 \frac{1}{q_A} z^{K_0-1} \frac{1 - q_B z^{-1}}{1 - \frac{q_B}{q_A} z^{-1}} . \quad (3.49)$$

Aufgrund der Normierungsbedingung

$$w_{zT}(1) = 1$$

errechnet sich K_0 zu :

$$K_0 = \frac{q_A - q_B}{1 - q_B} \quad (3.50a)$$

und aus Gl.(3.43)

$$\lim_{z \rightarrow \infty} w_{ZT}(z) > 0$$

ergibt sich für K_1

$$K_1 = 1 \quad (3.50b)$$

Aus den Gleichungen (3.49), (3.50a) und (3.50b) erhält man die Wartezeitverteilung für das System GEOM(m)/GEOM(m)/1 im Z-Bereich:

$$w_{ZT}(z) = \frac{q_A - q_B}{1 - q_B} \cdot \frac{z - q_B}{zq_A - q_B} \quad (3.51)$$

Für den speziellen Fall $m=1$ erhält man mit

$$(1-\rho) = 1 - \frac{EB}{EA} = \frac{q_A - q_B}{1 - q_B}$$

die in Gl.(3.22) angegebene Wartezeitverteilung des Systems GEOM(1)/GEOM(1)/1, die bereits in Kap. 3.2 mit der Methode der eingebetteten Markoff-Kette berechnet wurde :

$$w_{ZT}(z) = \frac{(1-\rho) (z-q_B)}{q_A z - q_B} \quad (3.52)$$

3.4 Algorithmen für die Wartezeitverteilung des zeitdiskreten GI/G/1-Systems

Für GI/G/1-Systeme, die einer geschlossenen analytischen Lösung der Lindleyschen Integralgleichung nicht zugänglich sind, existieren numerische Algorithmen zur Berechnung der Wartezeitverteilung, die in diesem Unterkapitel behandelt werden.

3.4.1 Algorithmus im diskreten Zeitbereich

Aus der allgemeinen Form der Lindleyschen Integralgleichung im diskreten Zeitbereich gemäß Gl.(3.33a,b)

$$w_{n+1}(k) = \pi_0(w_n(k) * a_n(-k) * b_n(k)) * c_n(k) \quad (3.53)$$

kann auf sukzessive Weise die angetroffene Restarbeit bzw. die Wartezeitverteilung der (n+1)-ten Anforderung aus der Wartezeitverteilung der n-ten Anforderung berechnet werden, wobei die Zwischenankunftsverteilung sowie die Bedienzeitverteilung anforderungsindividuell gewählt werden können. Dies führt zu einem iterativen

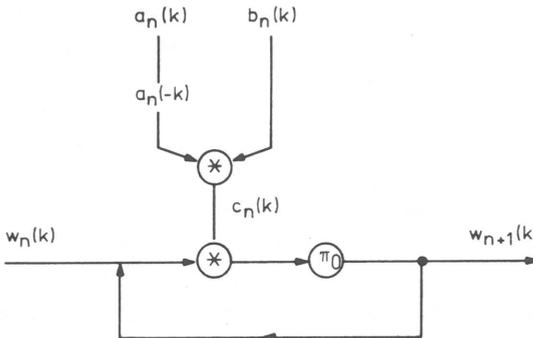


Bild 3.5 Schematische Darstellung des Algorithmus zur Berechnung der GI/G/1-Wartezeitverteilung im Zeitbereich

Algorithmus zur Berechnung der Wartezeitverteilung des GI/G/1-Systems im Zeitbereich /2,63/, der in Bild 3.5 schematisch dargestellt wird.

Die Einzelschritte des Algorithmus sind :

- i) Berechnung der Systemfunktion $c(k)$, falls die Verteilungen der Zwischenankunfts- und Bedienzeiten anforderungsunabhängig sind.
- ii) Initialisierung der Wartezeitverteilung $w_0(k)$. Dies kann im Prinzip mit einer beliebigen Verteilung geschehen. Zweckmäßigerweise wird jedoch die Iteration z.B. durch die Betrachtung eines leeren Systems beim Eintreffen der 0-ten Anforderung initiiert, d.h.

$$w_0(k) = \delta(k) = \begin{cases} 1 & k=0 \\ 0 & \text{sonst} \end{cases} . \quad (3.54)$$

- iii) Berechnung der Systemfunktion $c_n(k)$, falls die Zwischenankunfts- und Bedienverteilungen anforderungsabhängig sind.
- iv) Durchführung der Faltungsoperation zur Berechnung der Wartezeitverteilung der $(i+1)$ -ten Anforderung aus der i -ten Anforderung gemäß Gl.(3.53).
- v) Durchführung der π_0 -Operation gemäß Gl.(3.31)
- vi) Wiederholung der Schritte iii), iv) und v) bis zur Konvergenz der Wartezeitverteilung. Als Konvergenzkriterium kann z.B. die Differenz zwischen den Mittelwerten zweier bzw. mehrerer sukzessiv ermittelter Wartezeitverteilungen genommen werden.

Die hier benötigten Faltungsoperationen können z.B. mit Hilfe der diskreten Fourier-Transformation (DFT) bzw. der entsprechenden Al-

gorithmen der schnellen Fourier-Transformation (FFT) - vgl. Kap. 2.2.3 - effektiv durchgeführt werden.

Sind die Zwischenankunftsabstände und Bedienzeiten anforderungsabhängig, werden pro Iterationszyklus zwei Faltungsoperationen benötigt. Eine Faltungsoperation pro Iterationsschritt kann dagegen durch eine Vorausberechnung der Systemfunktion eingespart werden, falls die Verteilungen $a(k)$ und $b(k)$ anforderungsunabhängig sind.

Im Vergleich zu den Algorithmen im transformierten Bereich ist der hier vorgestellte Algorithmus im diskreten Zeitbereich relativ rechenzeitintensiv, falls die Systemfunktion eine große Anzahl von Verteilungswerten umfaßt. Ein Vorteil dieses Algorithmus liegt jedoch darin, daß das Konvergenzverhalten hinsichtlich der Typen von Verteilungsfunktionen sehr robust ist. Ferner, wie erwähnt wurde, kann der Algorithmus im Zeitbereich, im Gegensatz zu den Methoden im transformierten Bereich, für Systeme mit anforderungsabhängigen Zwischenankunfts- und Bedienzeiten angewendet werden.

3.4.2 Algorithmen im transformierten Bereich

a) Grundprinzip

Algorithmen für die Berechnung der Wartezeitverteilung von zeitdiskreten GI/G/1-Systemen im transformierten Bereich basieren fast ausschließlich auf der charakteristischen Gleichung gemäß (3.41) für Z-Transformierte

$$w_{ZT}^-(z) \cdot \frac{1}{w_{ZT}(z)} = \frac{c_{ZT}(z) - 1}{1 - z^{-1}} \quad (3.55a)$$

bzw. für erzeugende Funktionen

$$w_{EF}^-(z) \cdot \frac{1}{w_{EF}(z)} = \frac{c_{EF}(z) - 1}{1 - z} \quad (3.55b)$$

Unter Berücksichtigung der Eigenschaften bzgl. der Lage von Polen und Nullstellen der Terme auf der rechten Seite der charakteristischen Gleichung kann die Wartezeitverteilung $w_{ZT}(z)$ von $\bar{W}_{ZT}(z)$ (bzw. $w_{EF}(z)$ von $\bar{W}_{EF}(z)$) im transformierten Bereich getrennt werden. Zwei prinzipielle Möglichkeiten sind :

- Separationsmethode durch explizite Berechnung von Nullstellen bzw. Polen mit Hilfe der Polynom-Faktorisierung (vgl./41,58/)
- Separationsmethode unter Berücksichtigung der Phasen-Eigenschaften der zu trennenden Terme der charakteristischen Funktion. Dabei kann z.B. das Cepstrum-Konzept (vgl. Kap. 2.5) herangezogen werden /2,55/.

Bei den nachfolgend vorgestellten Algorithmen wird vorausgesetzt, daß die Zwischenankunfts- und Bedienzeitverteilungen von endlicher Länge sind (vgl. Gl.(3.1a,b)) :

$$a(k) = P(A = k), \quad k=0,1,\dots,n_A-1, \quad n_A < \infty, \quad (3.56a)$$

$$b(k) = P(B = k), \quad k=0,1,\dots,n_B-1, \quad n_B < \infty. \quad (3.56b)$$

b) Algorithmus mit Polynom-Faktorisierung

Der Algorithmus nach Konheim /41/ setzt ferner voraus, daß der Ankunftsprozeß ein Einzelankunftsprozeß ist, d.h.

$$a(0) = 0. \quad (3.57)$$

Da

$$c_{EF}(z) - 1 \Big|_{z=1} = 0 \quad \text{und} \quad \frac{d}{dz}(c_{EF}(z) - 1) \Big|_{z=1} = EB - EA < 0, \quad (3.58)$$

ist die Nullstelle $z=1$ des Zählerpolynoms der charakteristischen Funktion im Bereich der erzeugenden Funktionen

$$S_{EF}(z) = W_{EF}^-(z) \cdot \frac{1}{w_{EF}(z)} = \frac{c_{EF}(z) - 1}{1 - z} \quad (3.59)$$

eine einfache Nullstelle. Zusammen mit der Voraussetzung in Gl. (3.56a,b) ist ersichtlich, daß die charakteristische Funktion als Polynom geschrieben werden kann. Dies führt zum folgenden Faktorisierungsansatz

$$S_{EF}(z) = S_{EF}^+(z) \cdot S_{EF}^-(z) \quad (3.60)$$

mit

$$S_{EF}^+(z) = K_1 \prod_{n=1}^{n_0} (z - q_{on}), \quad |q_{on}| > 1 \quad (3.61a)$$

und

$$S_{EF}^-(z) = K_2 z^{-K_3} \prod_{n=1}^{n_i} (z - q_{in}), \quad |q_{in}| < 1, \quad (3.61b)$$

wobei q_{in} , $n=1, \dots, n_i$ und q_{on} , $n=1, \dots, n_0$ die Nullstellen des Polynoms $S_{EF}(z)$ innerhalb bzw. außerhalb des Einheitskreises darstellen.

Die Konstanten K_1 und K_2 müssen der Normierungsbedingung genügen

$$S_{EF}^+(1) = 1. \quad (3.62)$$

In Konheim /41/ wird gezeigt, daß sich die erzeugende Funktion der Wartezeitverteilung wie folgt angeben läßt

$$w_{EF}(z) = \frac{1}{S_{EF}^+(z)}. \quad (3.63)$$

Die Anwendung dieser Methode setzt eine explizite Berechnung von Nullstellen der charakteristischen Funktion, z.B. mit Hilfe von Algorithmen zur Polynom-Faktorisierung, voraus. Für GI/G/1-Systeme

mit relativ langen Zwischenankunfts- und Bedienzeitverteilungen ist der damit verbundene numerische Aufwand ziemlich hoch. Ein weiterer Nachteil des Verfahrens besteht darin, daß die Lösung gemäß Gl.(3.63) im transformierten Bereich angegeben wird. Benötigt man die Wartezeitverteilung im Zeitbereich oder den Mittelwert bzw. den Variationskoeffizienten der Verteilung, müssen weitere Algorithmen zur Rücktransformation bzw. Differentiation /30/ herangezogen werden.

Für Systeme mit geringerer Anzahl von Verteilungswerten stellt der Algorithmus mit Polynom-Faktorisierung jedoch eine effektive Methode zur Wartezeitanalyse dar.

c) Algorithmus mit Cepstrum-Bildung

Die Separationsmethode im Cepstrum-Bereich /2/ geht von der charakteristischen Gleichung im Z-Bereich aus

$$S_{ZT}(z) = W_{ZT}^-(z) \cdot \frac{1}{w_{ZT}(z)} = \frac{c_{ZT}(z) - 1}{1 - z^{-1}} \quad (3.64)$$

Durch die Voraussetzung der endlichen Zwischenankunfts- und Bedienzeitverteilungen können diese Verteilungen sowie die Systemfunktion im Z-Bereich als Polynome geschrieben werden :

$$a_{ZT}(z) = \sum_{k=0}^{n_A-1} a(k) z^{-k} \quad (3.65a)$$

$$b_{ZT}(z) = \sum_{k=0}^{n_B-1} b(k) z^{-k} \quad (3.65b)$$

$$c_{ZT}(z) = a_{ZT}(1/z) \cdot b_{ZT}(z) = \sum_{k=-(n_A-1)}^{n_B-1} c(k) z^{-k} \quad (3.66)$$

Ähnlich wie im vorausgegangenen Kapitel 3.4.2b kann hier gezeigt werden, daß die Funktion $c_{ZT}(z) - 1$ eine einfache Nullstelle an

der Stelle $z = 1$ besitzt, d.h. die charakteristische Funktion $S_{ZT}(z)$ hat keine Pole und kann deshalb als Polynom in $1/z$ formuliert werden.

Ferner ist nach Gl.(3.39) - vgl. Bild 3.4 -

$$W^-(k) + W(k) = c(k) * W(k) .$$

Unter Berücksichtigung der endlichen Länge der Systemfunktion $c(k)$ nach Gl.(3.66) läßt sich die Z-Transformierte $W_{ZT}^-(z)$ ebenfalls als Polynom der Form

$$W_{ZT}^-(z) = \sum_{k=-(n_A-1)}^{-1} W^-(k) z^{-k} = \sum_{k=1}^{n_A-1} W^-(-k) z^k \quad (3.67)$$

angeben. Dies bedeutet, daß der Term $W_{ZT}^-(z)$ keine Pole besitzt.

Es kann nun mit folgenden Schritten gezeigt werden, daß die zwei multiplikativen Terme $W_{ZT}^-(z)$ und $1/w_{ZT}(z)$ der charakteristischen Funktion jeweils die Z-Transformierte einer Maximalphasen- bzw. einer Minimalphasenfolge darstellen :

- i) Die charakteristische Funktion $S_{ZT}(z)$ nach Gl.(3.64) besitzt keine Pole
- ii) Wie oben gezeigt wurde, besitzt der Term $W_{ZT}^-(z)$ keine Pole. Da $W^-(-k)$ in Gl.(3.39) eine Folge mit nicht-abnehmenden Koeffizienten ist (vgl. Bild 3.4), kann ferner nach dem Theorem von Eneström und Kakeya (vgl./2/) nachgewiesen werden, daß alle Nullstellen von $W_{ZT}^-(z)$ außerhalb des Einheitskreises liegen. Dies bedeutet, daß $W_{ZT}^-(z)$ die Z-Transformierte einer Maximalphasenfolge ist.
- iii) Da $w_{ZT}(z)$ die Z-Transformierte einer Verteilung ist und für $|z| > 1$ konvergiert, liegen alle Pole von $w_{ZT}(z)$ innerhalb des Einheitskreises. Dies bedeutet, daß der Term $1/w_{ZT}(z)$ nur Nullstellen innerhalb des Einheitskreises besitzt.

- iv) Aus i) und ii) ist ersichtlich, daß $1/w_{ZT}(z)$ keine Pole besitzt. Mit iii) folgt, daß $1/w_{ZT}(z)$ die Z-Transformierte einer Minimalphasenfolge darstellt.

Zur Berechnung der Wartezeitverteilung $w(k)$ kann nun die Separationsmethode von Maximal- und Minimalphasenfolgen durch Cepstrumbildung (vgl. Gl.(2.57a,b) und (2.58)) auf die charakteristische Funktion in Gl.(3.64) angewendet werden. Die zur Cepstrumbildung benötigte Z-Transformation bzw. -Rücktransformation kann wiederum anhand der diskreten Fourier-Transformation (DFT) bzw. der entsprechenden Algorithmen der schnellen Fourier-Transformation (FFT) - vgl. Kap. 2.2.3 - bewerkstelligt werden.

Wie in Bild 3.6 schematisch dargestellt, können die einzelnen Schritte des Algorithmus mit Cepstrumbildung /2/ wie folgt zusammengefaßt werden :

- i) Berechnung der diskreten Fourier-Transformation (DFT) der charakteristischen Funktion

$$S_{DFT}(n) = DFT\{C_1(k)\} \quad (3.68)$$

mit

$$C_1(k) = \sum_{j=-(n_A-1)}^k c_1(j) \quad \text{und} \quad c_1(j) = c(j) - \delta(j). \quad (3.69)$$

- ii) Berechnung des komplexen Logarithmus

$$S_{LN}(n) = \ln S_{DFT}(n) \quad (3.70)$$

und des komplexen Cepstrums der charakteristischen Funktion

$$S_{CEP}(k) = DFT^{-1}\{S_{LN}(n)\}. \quad (3.71)$$

iii) Separation der Maximal- und Minimalphasenanteile. Das Teil-Cepstrum

$$S_{\text{CEP}}^+(k) = \begin{cases} S_{\text{CEP}}(k) & k \geq 0 \\ 0 & k < 0 \end{cases} \quad (3.72)$$

stellt den zum $1/w_{\text{ZT}}(z)$ gehörenden Anteil, der jedoch wegen der Überlappung von Amplitudenanteilen von Maximal- und Minimalphasenfolgen an der Stelle $k=0$ (vgl. Gl.(2.58)) noch unnormiert ist.

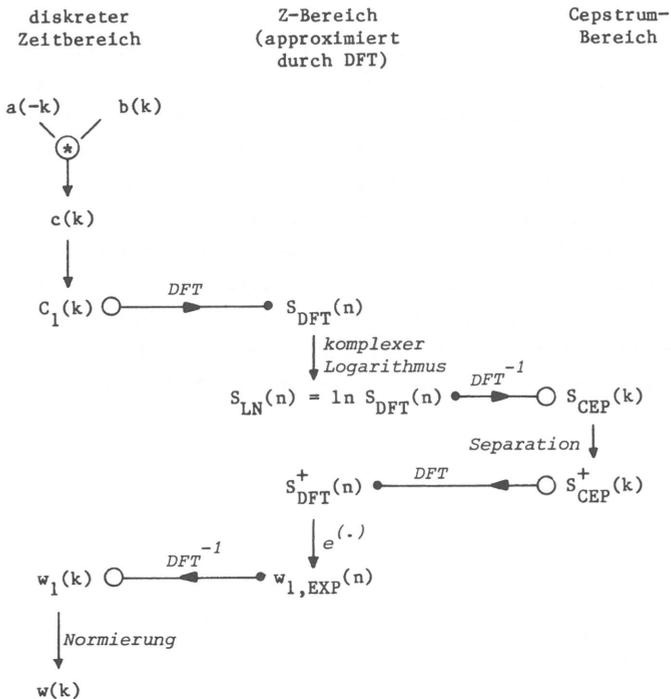


Bild 3.6 Schematische Darstellung des Algorithmus zur Wartezeitanalyse des zeitdiskreten GI/G/1-Systems mittels Cepstrum-Bildung

- iv) Berechnung der unnormierten Wartezeitverteilung $w_1(k)$ mittels inverser Cepstrum-Bildung nach Gl.(2.51a-c)

$$S_{\text{DFT}}^+(n) = \text{DFT}\{S_{\text{CEP}}^+(k)\} \quad (3.73)$$

$$w_{1,\text{EXP}}(n) = \frac{1}{S_{\text{EXP}}^+(n)} = e^{-S_{\text{DFT}}^+(n)} \quad (3.74)$$

$$w_1(k) = \text{DFT}^{-1}\{w_{1,\text{EXP}}(n)\}. \quad (3.75)$$

- v) Berechnung der Wartezeitverteilung $w(k)$ durch Normierung von $w_1(k)$.

Der hier behandelte Algorithmus mit Cepstrum-Bildung stellt eine effektive Methode zur Berechnung der Wartezeitverteilung des zeitdiskreten GI/G/1-Systems dar. Der Algorithmus ist besonders gut geeignet für Systeme, bei denen die Anzahl von Verteilungswerten der Zwischenankunfts- und Bedienzeiten groß ist. Dies ist begründet durch die Anwendung von effektiven Algorithmen zur schnellen Fourier-Transformation (FFT) /1,30/. Dabei müssen die in Kapitel 2.5.2e diskutierten numerischen Probleme, die bei der Berechnung des komplexen Cepstrums auftreten, z.B. die Überlappung von Cepstrum-Komponenten (aliasing) oder die Entflechtung von Phasenverläufen des Cepstrums (phase unwrapping /72/), berücksichtigt werden.

3.4.4 Mittelwerte und weitere Systemcharakteristiken

Einige wichtige Systemcharakteristiken, die direkt aus der Wartezeitverteilung $w(k)$ der Anforderungen in einem stationären GI/G/1-System gewonnen werden können, werden im folgenden diskutiert.

a) Wartewahrscheinlichkeit

Wie anhand von Gl.(3.37) bereits diskutiert, ist die Wahrscheinlichkeit dafür, daß eine ankommende Anforderung warten muß, d.h. eine positive Wartezeit erfährt, identisch mit der Wahrscheinlichkeit, daß unmittelbar vor dem Ankunftszeitpunkt der Anforderung die Bedieneinheit belegt ist

$$P_W = \sum_{k=1}^{\infty} u^-(k) = 1 - u^-(0) = 1 - w(0) . \quad (3.76)$$

b) Wartezeit bezüglich wartender Anforderungen

Aus der Wartezeitverteilung $w(k)$ der ZV W für die Wartezeit einer beliebigen Anforderung läßt sich die Verteilung $w_{>0}(k)$ der Wartezeit $W_{>0}$ von Anforderungen, die warten müssen, wie folgt berechnen :

$$w_{>0}(k) = P(W_{>0} = k) = \begin{cases} 0 & k \leq 0 \\ w(k)/P_W & k > 0 \end{cases} . \quad (3.77)$$

Zwischen den Mittelwerten von W und $W_{>0}$ besteht infolgedessen der Zusammenhang:

$$EW = EW_{>0} \cdot P_W . \quad (3.78)$$

c) Systembelegung und mittlere Warteschlangenlänge

Aus den mittleren Wartezeiten können durch Anwendung des Little'schen Theorems die Mittelwerte der Warteschlangenlänge (ZV Q) und der Systembelegung (ZV X) berechnet werden.

Betrachtet man ein beliebiges System mit der mittleren Zwischenankunftszeit EA, der mittleren Aufenthaltsdauer ET von Anforderungen und der mittleren Systembelegung EX - d.h. die mittlere Anzahl von Anforderungen im System -, so gilt die Little'sche Mittelwertformel:

$$EX = \frac{ET}{EA} . \quad (3.79)$$

Für das gesamte GI/G/1-System mit der mittleren Aufenthaltsdauer EW + EB und der mittleren Zwischenankunftszeit EA (Ankunftsrate $1/EA$), so ergibt sich die mittlere Systembelegung EX gemäß Gl. (3.79) zu

$$EX = \frac{EW + EB}{EA} = \frac{EW}{EA} + \rho . \quad (3.80)$$

Wird dagegen bei der Anwendung der Little-Formel nur die Warteschlange als System mit der mittleren Aufenthaltsdauer $EW_{>0}$ und der mittleren Zwischenankunftszeit EA/P_W , (Ankunftsrate P_W/EA) betrachtet, errechnet sich die mittlere Warteschlangenlänge zu

$$EQ = \frac{EW_{>0} \cdot P_W}{EA} = \frac{EW}{EA} = EX - \rho . \quad (3.81)$$

Weitere interessierende Größen und Verteilungen, z.B. die Freiperiode der Bedieneinheit und der Ausgangsprozess des zeitdiskreten GI/G/1-Systems, werden im nächsten Unterkapitel hergeleitet.

3.5 Verteilung der Freiperioden und des Ausgangsprozesses

3.5.1 Allgemeines

Der Ausgangsprozess eines GI/G/1-Systems gehört i.a. nicht zur Klasse von Erneuerungsprozessen. Nur in einzelnen speziellen Fällen ist es möglich, diesen Ausgangsprozess, welcher sich aus aufeinanderfolgenden, z.T. voneinander abhängigen Zwischenendigungsabständen zusammensetzt, vollständig zu beschreiben /9,10,29,57,62/.

In Burke /9,10/ wurde gezeigt, daß der Ausgangsprozess des zeitkontinuierlichen M/M/n-Systems ein Poisson-Prozeß ist. Ein Beweis dieses Sachverhalts, bei dem die zeitreversible Eigenschaft von stationären Markoff-Prozessen herangezogen wird, findet sich in Reich /62/. Eine ausführliche Analyse des M/D/1-Systems im kontinuierlichen Zeitbereich wurde von Pack /57/ vorgestellt. Heffes /29/ untersuchte den Ausgangsstrom des GI/M/1-Systems mit unterbrochenen Poisson-Prozessen (IPP: interrupted Poisson process) als Ankunftsprozeß.

Außer diesen Ausnahmen beschränken sich die meisten Untersuchungen in der Literatur auf die Ermittlung von Verteilungsfunktionen der Zwischenendigungsabstände - im folgenden auch als Ausgangsverteilungsfunktionen bezeichnet - und ihrer Momente. In Marshall /51/ wurden Beziehungen für die ersten zwei Momente der Ausgangsverteilung angegeben. Momente dieser Verteilungsfunktion bei zeitkontinuierlichen Systemen der Klassen M/G/1, $E_k/M/1$ und $E_2/E_2/1$ wurden in Makino /50/ hergeleitet. Eine Übersicht über Untersuchungen der Ausgangsprozesse findet sich in Daley /22/.

In diesem Unterkapitel wird eine Methode im diskreten Zeitbereich zur Berechnung von Verteilungen der Freiperioden und der Ausgangsprozesse zeitdiskreter GI/G/1-Systeme vorgestellt. Für die Analyse werden folgende Zufallsvariablen (ZV) definiert :

- I ZV für die Dauer von Freiperioden
- D ZV für die Zwischenendigungsabstände

3.5.2 Verteilung der Freiperioden

Bild 3.7 zeigt eine Stichprobe des Prozeßverlaufs der Restarbeit in einem eingeschwungenen GI/G/1-System, wobei der Zusammenhang zwischen der virtuellen Restarbeit U^V und der Freiperiode I verdeutlicht wird. Es ist ersichtlich, daß die Länge einer Freiperiode identisch ist mit dem negativen Anteil der virtuellen Restarbeit. Dabei wird vorausgesetzt, daß die Dauer einer Freiperiode eine positive Größe ist. Die Verteilung der Freiperiode kann nun aus der stationären Verteilung der virtuellen Restarbeit

$$u^V(k) = w(k) * c(k) \quad , k \geq -(n_A - 1) \quad , \quad (3.82)$$

gewonnen werden:

$$\begin{aligned} i(k) &= P(I = k) = P(U^V = -k) \\ &= K \cdot u^V(-k), \quad k=1, \dots, n_A - 1, \end{aligned} \quad (3.83a)$$

wobei K eine Normierungskonstante darstellt:

$$K^{-1} = \sum_{j=1}^{n_A - 1} u^V(-j) \quad . \quad (3.83b)$$

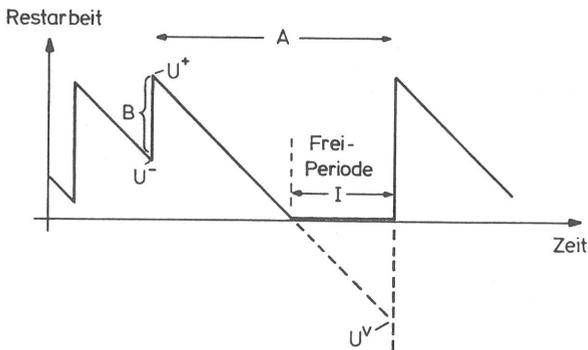


Bild 3.7 Virtuelle Restarbeit und Freiperiode

Man erhält schließlich für die Verteilung der Freiperiode eines GI/G/1-Systems:

$$i(k) = \frac{u^{\vee}(-k)}{n_A - 1} , k=1,2,\dots,n_A-1. \quad (3.84)$$

$$\sum_{j=1} u^{\vee}(-j)$$

3.5.3 Verteilung des Ausgangsprozesses

Bild 3.8 zeigt einen Prozeßverlauf der Restarbeit, wobei aufeinanderfolgende Betriebs- (S) und Freiperioden (I) dargestellt werden. Die Endigungen kennzeichnen Ereigniszeitpunkte des Ausgangsprozesses.

Zwei unterschiedliche Typen von Ausgangsereignissen können unterschieden werden:

- Ausgangsereignis vom Typ 1: Die Endigung ist am Ende einer Betriebsperiode. Die Zwischenendigungszeit D, d.h. das Zeitintervall bis zur nächsten Endigung, ist die Summe einer Freiperiode I und einer Bedienzeit B.
- Ausgangsereignis vom Typ 2: Die Endigung findet während einer Betriebsperiode statt. Die Zwischenendigungszeit D in diesem Fall beträgt nur eine Bedienzeit B.

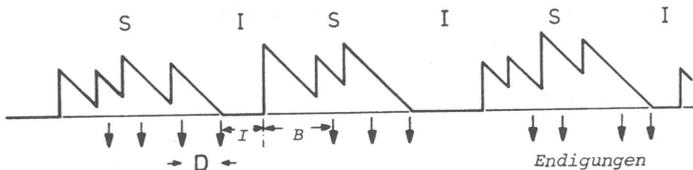


Bild 3.8 Stichprobe des Ausgangsprozesses
(S : Betriebsperiode, I : Freiperiode, B : Bedienzeit)

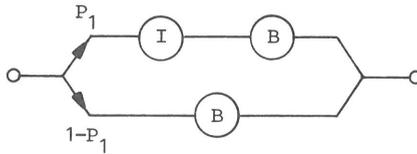


Bild 3.9 Phasendiagramm der Zwischenendigungszeit des Ausgangsprozesses

Die entsprechende Phasendarstellung für die Verteilung des Ausgangsprozesses zeigt Bild 3.9. Die Wahrscheinlichkeit P_1 , daß ein Ausgangsereignis vom Typ 1 ist, läßt sich aus folgenden Schritten bestimmen:

- i) Wird der Prozeß der Restarbeit (s. Bild 3.8) über eine hinreichend lange Zeitspanne T_0 beobachtet, so ist die Anzahl der Betriebsperioden während dieser Zeitspanne

$$\frac{T_0}{EI + ES} \cdot$$

Die Anzahl von ankommenden Ereignissen (bzw. von Ausgangsereignissen) während der Zeit T_0 beträgt

$$\frac{T_0}{EA} \cdot$$

Daraus ergibt sich die mittlere Anzahl von Ausgangsereignissen pro Betriebsperiode

$$n_0 = \frac{EI + ES}{EA} \quad (3.85a)$$

- ii) Die Wahrscheinlichkeit P_1 lautet

$$P_1 = \frac{1}{n_0} = \frac{EA}{EI + ES} \quad (3.85b)$$

wobei der Mittelwert ES der Betriebsperiode noch zu bestimmen ist.

iii) Andererseits erhält man von der Auslastung ρ des Systems

$$\rho = \frac{EB}{EA} = \frac{ES}{EI + ES} . \quad (3.85c)$$

iv) Durch Einsetzen von ES aus iii) in ii) ergibt sich schließlich für die Wahrscheinlichkeit P_1

$$P_1 = \frac{EA - EB}{EI} . \quad (3.86)$$

Aus der Phasendarstellung der Zwischenendigungszeit D in Bild 3.9 läßt sich die Verteilung des Ausgangsprozesses des GI/G/1-Systems wie folgt angeben

$$d(k) = P_1(i(k) * b(k)) + (1 - P_1) b(k), \\ k=0, 1, \dots, (n_A-1) + (n_B-1), \quad (3.87)$$

oder im Z-Bereich

$$d_{zT}(z) = b_{zT}(z) (1 - P_1 + P_1 i_{zT}(z)) . \quad (3.88)$$

3.6 Ergebnisse

In diesem Unterkapitel werden Ergebnisse für charakteristische Größen zeitdiskreter GI/G/1-Systeme diskutiert. Die numerischen Berechnungen basieren auf dem in Kap. 3.4 beschriebenen Algorithmus mit Cepstrum-Bildung. Die Ergebnisse sind im folgenden auf $\Delta t = 1$ normiert.

Für Vergleichszwecke werden folgende zeitkontinuierlichen Verteilungsfunktionen - hinsichtlich verschiedener Bereiche von Variationskoeffizienten - genommen (vgl./45,47,70/):

- D : Deterministisch ($c=0$)
- E_n : Erlangsche Verteilung n-ter Ordnung ($0 < c \leq 1$)
- M : Markoff, negativ-exponentielle Verteilung ($c=1$)
- H_2 : Hyperexponentielle Verteilung 2-ter Ordnung mit Symmetrieannahme /45/ ($c > 1$)

3.6.1 Systeme mit beliebig verteilten Zwischenankunfts- und Bedienzeiten

In den meisten Anwendungen zeitdiskreter Analyseverfahren liegen Verteilungsfunktionen in der Form von Histogrammen vor, die z.B. aus Messungen gewonnen werden. Zur Illustration werden folgende einfache Verteilungen für die Zwischenankunfts- und Bedienzeiten betrachtet:

$$a(2)=25/72, \quad a(5)=22/72, \quad a(8)=25/72, \quad a(k)=0 \text{ sonst,}$$
$$b(1)=1/2, \quad b(2)=1/3, \quad b(8)=1/6, \quad b(k)=0 \text{ sonst,}$$

mit:

$$EA = 5, \quad EB = 2.5 \quad (\rho = 0.5), \quad c_A = 0.5, \quad c_B = 1.$$

Bild 3.10 zeigt die komplementären Verteilungsfunktionen für die Wartezeit bezüglich wartender Anforderungen, die Freiperiode und den Ausgangsprozess dieses Systems, wobei der treppenförmige Charakter zu erkennen ist. Da die Zwischenankunftsverteilung begrenzt ist, sind die Verteilungen der Freiperioden und der Zwischenendigungszeit ebenfalls endlich (vgl. Gl.(3.84) und (3.87)).

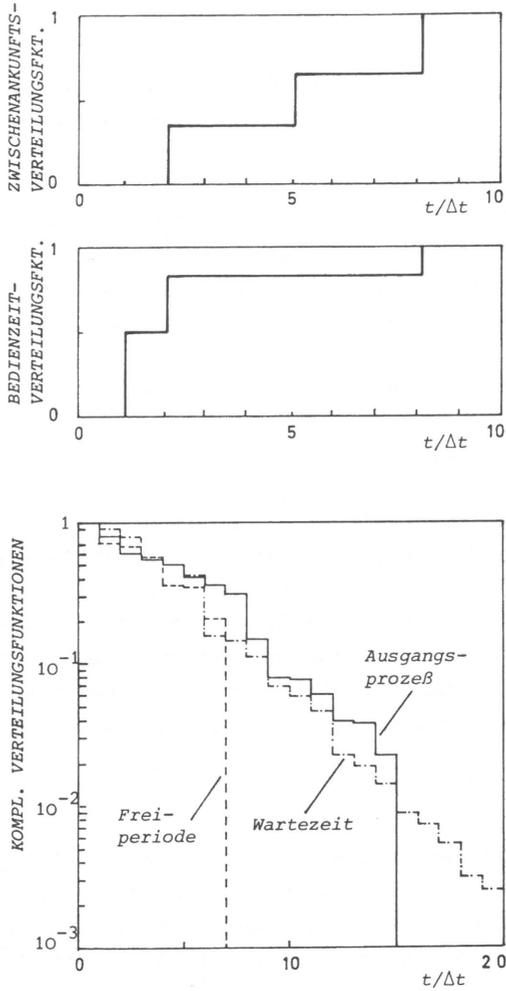


Bild 3.10 Komplementäre Verteilungsfunktionen der Wartezeit, der Freiperioden und der Zwischenendigungszeit des Ausgangsprozesses eines zeitdiskreten GI/G/1-Systems

3.6.2 Systeme mit negativ-binomial verteilten Zwischenankunfts- und Bedienzeiten

In diesem Unterkapitel wird anhand von negativ-binomial verteilten Zwischenankunfts- und Bedienzeiten eine systematische Untersuchung des zeitdiskreten GI/G/1-Systems mit zweiparametrischen Verteilungen durchgeführt (vgl./71/). Das entstandene System ist dementsprechend vom Typ NEGBIN/NEGBIN/1, welcher gemäß Gl.(2.17a,b) parametrisch festgelegt werden kann. Für numerische Zwecke werden die eingesetzten negativ-binomialen Verteilungen begrenzt (typischerweise werden Verteilungswerte kleiner als $\epsilon=10^{-9}$ vernachlässigt und die verbliebenen Verteilungswerte normiert), so daß die in Gl. (3.56a,b) diskutierten Voraussetzungen erfüllt sind.

a) Wartezeitverteilung

Die komplementären Wartezeitverteilungsfunktionen bzgl. wartender Anforderungen zeitdiskreter Systeme vom Typ NEGBIN/D/1 werden in Bild 3.11 dargestellt. Die Variationskoeffizienten sind hier so gewählt, daß die betrachteten Bedienzeiten jeweils die zeitdiskreten Äquivalente der oben erwähnten zeitkontinuierlichen Verteilungstypen Erlang 4-ter Ordnung E_4 ($c_A=0.5$), negativ-exponentiell M ($c_A=1$) und hyperexponentiell H_2 ($c_A=1.5$), darstellen.

b) Freiperiode

Die mittlere Dauer der Freiperioden wird in Bild 3.12 und Bild 3.13 als Funktion des Variationskoeffizienten der Zwischenankunftsverteilung dargestellt. Bild 3.12 zeigt den Einfluß des Bedienprozesses und Bild 3.13 der Verkehrsbelastung auf die Freiperiode.

Für den speziellen Fall $c_B = 1$ entspricht das betrachtete zeitdiskrete NEGBIN/NEGBIN/1-System dem zeitkontinuierlichen GI/M/1-System. Für GI/M/1 Systeme kann bewiesen werden (vgl. Halfin /27/), daß bei hoher Verkehrsintensität ($\rho \rightarrow 1$) die Grenzverteilung der Freiperioden mit der Rekurrenzzeit-Verteilung des Zwischenankunftsprozesses identisch ist. Dies wird in Bild 3.13 verdeutlicht.

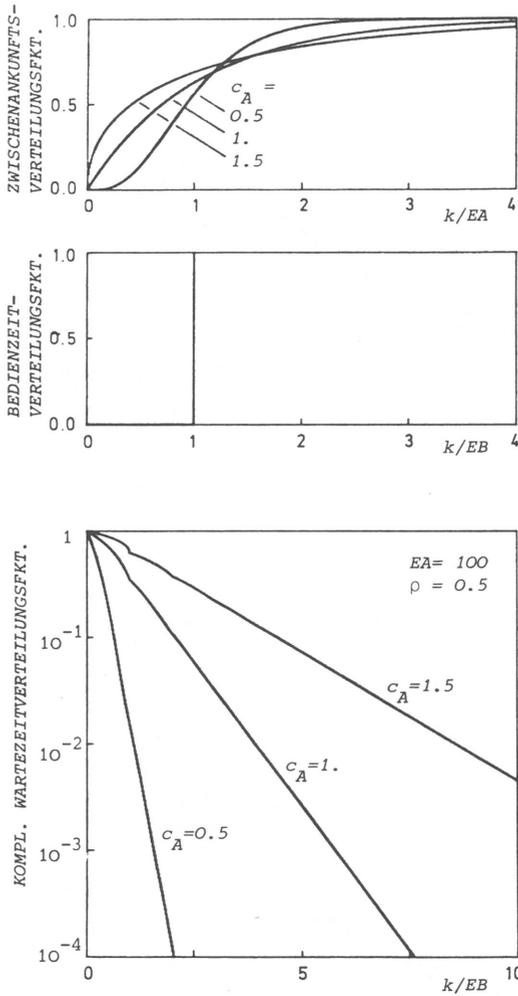


Bild 3.11 Komplementäre Wartezeitverteilungsfunktionen zeitdiskreter NEGBIN/NEGBIN/1-Systeme

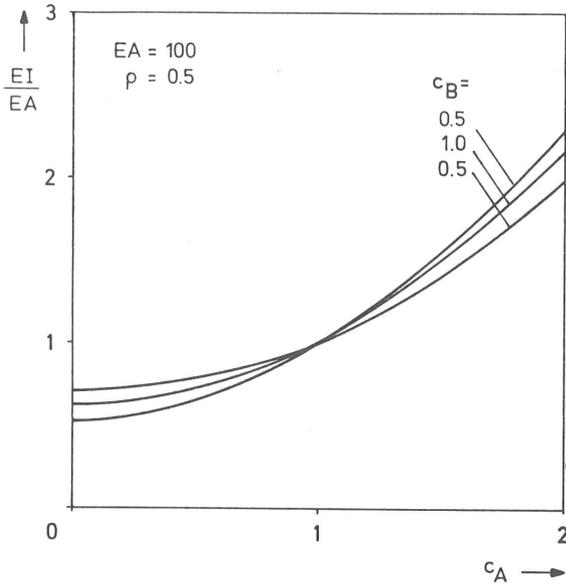


Bild 3.12

Einfluß des Ankunfts-
bzw. des Bedienpro-
zesses auf die Frei-
periode

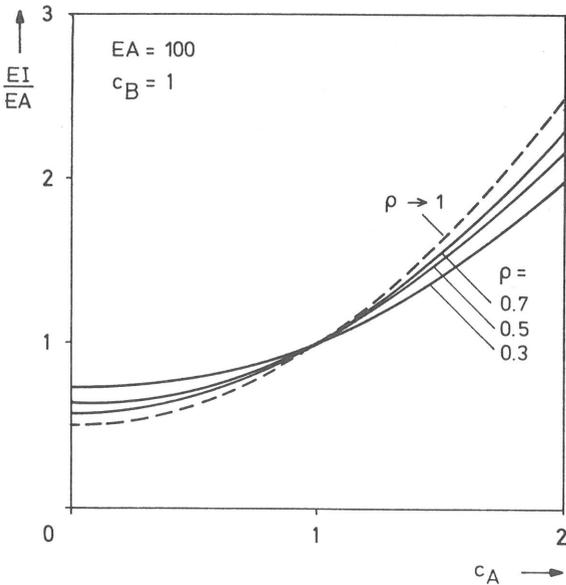


Bild 3.13

Einfluß der Verkehrs-
intensität auf die
Freiperiode

c) Ausgangsprozesse

Die komplementären Verteilungsfunktionen der Zwischenendigungszeit zeitdiskreter NEGBIN/NEGBIN/1-Systeme werden in Bild 3.14 für verschiedene Zwischenankunftsprozeßtypen gezeigt. Es ist ersichtlich, daß für größere Werte der Zeitparameter die Verteilung des Ausgangsprozesses in eine geometrische Verteilung übergeht.

Der Variationskoeffizient der Zwischenendigungszeit wird in Bild 3.15 und Bild 3.16 als Funktion des Variationskoeffizienten der Bedienzeit dargestellt. Die parametrische Beeinflussung des Ankunftsprozesses wird in Bild 3.15 verdeutlicht, während in Bild 3.16 verschiedene Zwischenankunftsprozesse mit unterschiedlichen Mittelwerten (bzw. Verkehrsangeboten $\rho = EB/EA$) und Variationskoeffizienten c_A untersucht werden.

Den Einfluß der Verkehrsintensität auf den Ausgangsprozeß zeigt Bild 3.17, wobei der Variationskoeffizient c_D der Zwischenendigungszeit in Abhängigkeit vom Variationskoeffizienten c_A der Zwischenankunftszeit und vom Verkehrsangebot dargestellt wird. Wie erwartet, geht der Ausgangsprozeß bei verschwindendem Verkehrsangebot ($\rho \rightarrow 0$) in den Eingangsprozeß über ($c_D = c_A$). In Bild 3.17 ist ebenfalls ersichtlich, daß bei starkem Verkehr, d.h. an der Stabilitätsgrenze des Systems ($\rho \rightarrow 1$), die Zwischenendigungszeit mit der Bedienzeit identisch ist. Der Grund liegt darin, daß bei diesem Grenzübergang praktisch keine Freiperioden existieren.

Die Genauigkeit der zeitdiskreten Ausgangsprozeßberechnung wird mit dem Vergleich in Tabelle 3.1 aufgezeigt, wobei Ergebnisse gemäß der Beziehung in Marshall /51/ für GI/G/1-Systeme

$$c_D^2 = c_A^2 + 2\rho^2 c_B^2 - 2\rho(1 - \rho) \frac{EW}{EB} \quad (3.89)$$

den anhand Gl.(3.87) ermittelten Ergebnissen für den Ausgangsprozeß gegenübergestellt werden.

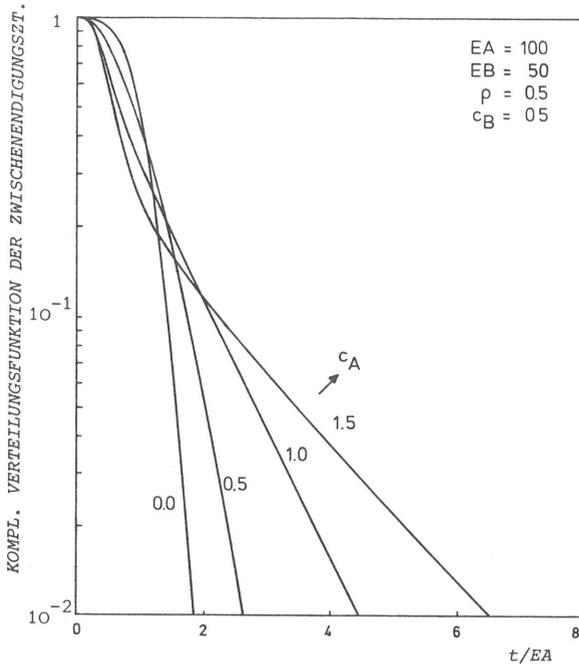


Bild 3.14

Komplementäre Verteilungsfunktion der Zwischenendigungszeit eines zeitdiskreten GI/G/1-Systems mit negativ-binomial verteilten Zwischenankunfts- und Bedienzeiten

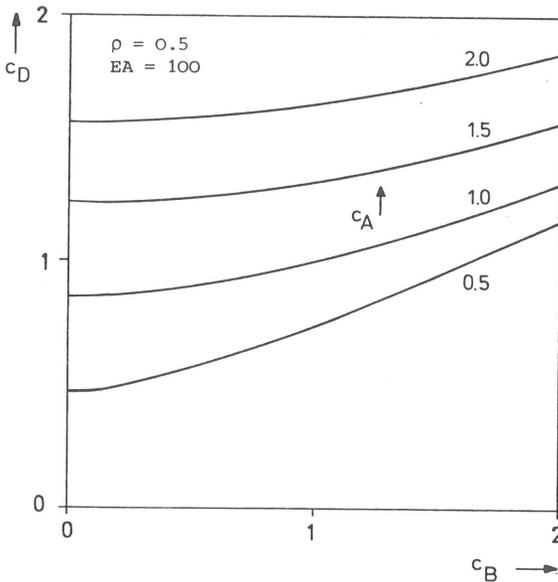


Bild 3.15

Einfluß des Ankunfts- bzw. des Bedienprozesses auf den Ausgangsprozess

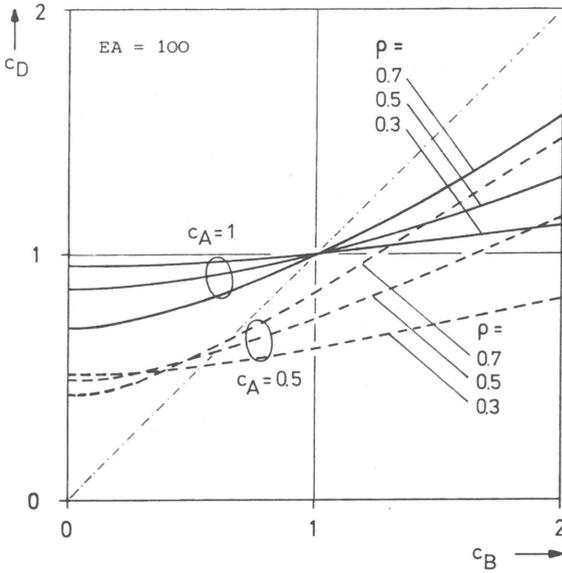


Bild 3.16

Einfluß der Verkehrsintensität bzw. des Anknftsprozesses auf den Ausgangsprozeß

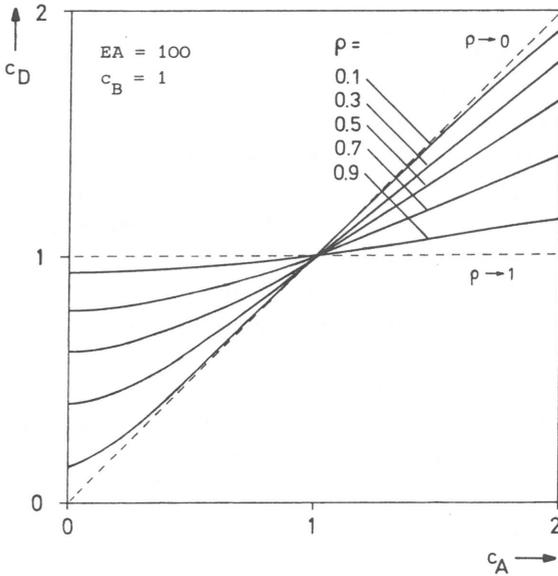


Bild 3.17

Eingang/Ausgang-Darstellung eines NEGBIN/NEGBIN/1-Systems bzgl. der Variationskoeffizienten der Verteilungen

Tabelle 3.1 *Variationskoeffizient der Zwischenendigungszeitverteilung des Ausgangsprozesses : Vergleich zwischen zeitdiskreter Analyse und Auswertung der Marshall-Formel /51/ ($\rho=0.5$, $EA=100$)*

c_D	System	$c_B=0.5$	$c_B=1.0$	$c_B=1.5$
$c_A = 0.5$	NEGBIN/NEGBIN/1 (MARSHALL)	0.5565010	0.7305068	0.9343127
		0.5565011	0.7305071	0.9343137
$c_A = 1.0$	NEGBIN/NEGBIN/1 (MARSHALL)	0.9006082	0.9994338	1.145259
		0.9006088	0.9994346	1.145261
$c_A = 1.5$	NEGBIN/NEGBIN/1 (MARSHALL)	1.259669	1.324368	1.429765
		1.259670	1.324369	1.429768

3.6.3 Approximative Berechnung für zeitkontinuierliche Systeme

Die zeitdiskrete Analyse kann approximativ zur Untersuchung zeitkontinuierlicher Systeme angewendet werden, die keinem geschlossenen exakten Analyseverfahren zugänglich sind bzw. für die eine numerische Lösung im kontinuierlichen Zeitbereich mit hohem Aufwand verbunden ist. Die Verteilungsfunktionsstypen in Tabelle 3.2 werden so gewählt, daß Vergleichswerte (z.B., /64/), die mittels zeitkontinuierlicher Analyse gewonnen werden, vorhanden sind. Die Approximationsgenauigkeit wird für GI/G/1-Systeme mit unterschiedlichen Zwischenankunfts- und Bedienprozessen verdeutlicht.

Tabelle 3.3 zeigt einen Vergleich der zeitdiskreten Ausgangsprozeßberechnung von NEGBIN/NEGBIN/1-Systemen ($c_A=1$) mit exakten Ergebnissen, die anhand der Formel in Makino /50/

$$c_D^2 = 1 - \rho^2(1 - c_B^2) \quad (3.90)$$

für Systeme vom Typ M/G/1 berechnet werden.

Tabelle 3.2 Mittlere Wartezeit in zeitkontinuierlichen GI/G/1-Systemen : Vergleich zwischen zeitdiskreter Analyse und tabellierten Ergebnissen in /64/ ($\rho=0.5, EA=40$)

EW	Ankunfts- prozeßtyp	E_2 $c_A = 0.5$	M $c_A = 1.0$	H_2 $c_A = 2.0$
Bedienpro- zeßtyp	Analysemethode			
$c_B = 0$	diskrete Approx.	0.17672	0.50018	1.3199
	exakte Analyse /64/	0.17672	0.50000	1.3191
$c_B = 0.5$	diskrete Approx.	0.39045	0.75030	1.7647
	exakte Analyse /64/	0.39054	0.75000	1.7640
$c_B = 1.0$	diskrete Approx.	0.61812	1.0003	2.1630
	exakte Analyse /64/	0.61803	1.0000	2.1621
$c_B = 2$	diskrete Approx.	2.0938	2.5004	4.0449
	exakte Analyse /64/	2.0935	2.5000	4.0439

Tabelle 3.3 Variationskoeffizient der Zwischenendigungszeitverteilung des Ausgangsprozesses : Vergleich zwischen zeitdiskreter Analyse (NEGBIN/NEGBIN/1-System) und Auswertung der Ergebnisse in Makino /50/ ($c_A=1, EA=100$)

c_D	System	$c_B=0.5$	$c_B=1.0$	$c_B=1.5$
$\rho = 0.3$	NEGBIN/NEGBIN/1	0.96495	0.99943	1.0543
	M/G/1 (MAKINO)	0.96566	1.0	1.0547
$\rho = 0.5$	NEGBIN/NEGBIN/1	0.90061	0.99943	1.1452
	M/G/1 (MAKINO)	0.90139	1.0	1.1456
$\rho = 0.7$	NEGBIN/NEGBIN/1	0.79463	0.99959	1.2695
	M/G/1 (MAKINO)	0.79530	1.0	1.2698

4. BEISPIELE FÜR ZEITDISKRETE ANALYSE KOMPLEXER MODELLE

Ziel dieses Kapitels ist es, die Anwendung zeitdiskreter Analysemethoden in komplexen Leistungsuntersuchungen anhand einiger Modellbeispiele zu erörtern.

4.1 Zeitdiskrete Analyse einer Überlastabwehrstrategie in rechnergesteuerten Vermittlungssystemen

4.1.1 Modellbildung von Überlastabwehrstrategien

In der Konzipierung und in der Entwicklung moderner rechnergesteuerter Vermittlungssysteme wird die Überlaststeuerung zu einem der zentralen Probleme, welche die Leistungsfähigkeit eines Vermittlungssystems entscheidend beeinflussen. Dieses Problem wird, bedingt durch die Einführung neuer Dienste und Leistungsmerkmale, zunehmend komplexer. Im Zuge der technologischen Fortschritte, die zu neuartigen Systemarchitekturen führen, wird die Leistungsbeurteilung bzw. die Gewährleistung der Funktionsfähigkeit eines Vermittlungssystems in Überlastsituationen schwieriger.

Bei den Verkehrsmodellen zur Untersuchung des Überlastphänomens und der Überlastregelungsmechanismen können zwei Hauptklassen von Maßnahmen und entsprechende Modellbildungen unterschieden werden /69/ :

i) Optimierung der Ausnutzung vorhandener Kapazität:

Bei dieser Klasse von Überlastabwehrstrategien wird in Überlastsituationen die Systemkapazität vorrangig für die Behandlung bereits vorhandener Arbeiten im System eingesetzt. Dies kann z.B. durch lastgesteuerte Abfertigungsstrategien und Prioritätsgebung geschehen.

ii) Drosselung des angebotenen Verkehrs:

Zu dieser Klasse gehören Überlastabwehrmaßnahmen und Modelle, bei denen in Überlastsituationen die Annahme des ankommenden Verkehrs so kontrolliert wird, daß Überlastentwicklungen recht-

zeitig abgeschwächt bzw. vermieden werden können. In Vermittlungssystemen wird diese Klasse von Maßnahmen meist zusammen mit den unter i) klassifizierten Strategien konzipiert und implementiert.

Eine Überlastabwehrstrategie wird durch folgende Merkmale gekennzeichnet :

- Überlastindikator :
Anzeige des Gefährdungsgrades der aktuellen Überlastsituation
- Abwehrmechanismus
Maßnahme bzw. eine abgestimmte Folge von Maßnahmen, die im Überlastfall zur Überlastabwehr aktiviert werden.

In diesem Unterkapitel wird ein lastgesteuerter Überlastabwehrmechanismus mittels zeitdiskreten Analyseverfahren untersucht. Abwehrstrategien dieser Klasse werden z.B. zur Überlastregelung in peripheren Steuerungen des EWSD-Vermittlungssystems - unter der Bezeichnung TAIL (Time Account Input Limitation) - konzipiert und untersucht /21/.

4.1.2 Lastgesteuertes Überlastabwehrverfahren und Modellbildung

Bild 4.1 zeigt schematisch das Verkehrsmodell der hier untersuchten Überlastabwehrstrategie, die auf einem rückkopplungsbehafteten GI/G/1-System basiert. Der Rufankunftsverkehr wird mit einem Erneuerungsprozeß mit allgemein verteilten Zwischenankunftsabstän-

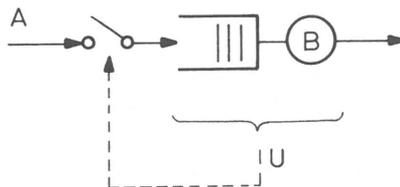


Bild 4.1 Rückkopplungsbehaftetes System zur Modellierung einer lastgesteuerten Überlastabwehrstrategie

den (ZV A) modelliert. Die Rufverarbeitung setzt sich aus Verarbeitungszeiten einzelner Steuerungsaufrufe (Teilrufe, vgl. /69/) zusammen. Der gesamte Aufwand der Steuerung zur Verarbeitung eines Rufes - der im folgenden auch als Anforderung bezeichnet wird - wird durch die allgemein verteilte ZV B charakterisiert. Es wird für die Analyse vorausgesetzt, daß die ZV A und B vom zeitdiskreten Typ sind.

Betrachtet wird nun eine Testanforderung (bzw. Testruf), die zum Ankunftszeitpunkt die Restarbeit U antrifft. Die Überlastabwehrstrategie besteht prinzipiell darin, daß die Testanforderung abgewiesen wird, falls U einen Schwellenwert L erreicht oder überschritten hat.

Es soll an dieser Stelle darauf hingewiesen werden, daß als Überlastindikator nicht die Anzahl der angetroffenen Rufe im System, sondern die Restarbeit, die der Anzahl der noch zu verarbeitenden Steuerungsaufrufe entspricht, genommen wird. Wegen dieser Rückkopplung ist das Modell in der vorgestellten allgemeinen Form keinem Analyseverfahren im kontinuierlichen Zeitbereich zugänglich.

Mit der Notation der Zufallsvariablen

U_n : ZV für die Restarbeit im System unmittelbar vor dem Ankunftszeitpunkt der n -ten Anforderung (bzw. des n -ten Rufes)

kann die Überlastabwehr- bzw. die Rufannahmestrategie wie folgt formuliert werden :

$U_n < L$ Anforderung n wird angenommen
 $U_n \geq L$ Anforderung n wird abgewiesen.

4.1.3 Zeitdiskreter Algorithmus zur Modellanalyse

Bild 4.2 zeigt eine Stichprobe des Prozeßverlaufs der Restarbeit im System. Die Zwischenankunftszeit und die Bedienzeit werden zunächst als anforderungsabhängig betrachtet

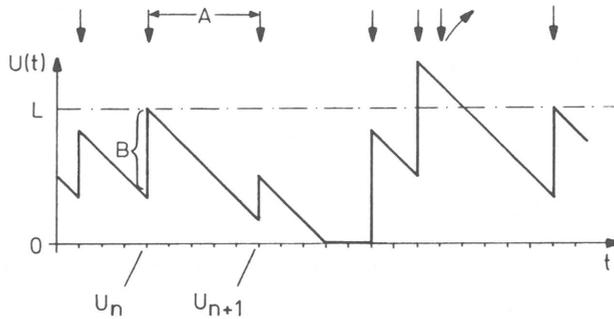


Bild 4.2 Stichprobe des Prozeßverlaufs der Restarbeit im Überlastabwehr-Modell

A_n ZV für die Zeitdauer zwischen den Ankunftszeitpunkten der n-ten und der (n+1)-ten Anforderung. Diese ZV kann z.B. zur Beschreibung von instationären Ankunftsprozessen herangezogen werden.

B_n ZV für die Bedienzeit der n-ten Anforderung. Diese Zeitspanne modelliert die gesamte Verarbeitungszeit des n-ten Rufes.

Betrachtet man als Bedingung die Annahme der n-ten Anforderung ($U_n < L$), so entstehen folgende Zufallsvariablen

$$U_{n,0} = U_n | U_n < L \quad , \quad U_{n,1} = U_n | U_n \geq L \quad (4.1)$$

mit den normierten Verteilungen

$$u_{n,0}(k) = \frac{1}{1-B_{A,n}} u_n(k) \quad , \quad k=0,1,\dots,L-1, \quad (4.2a)$$

$$u_{n,1}(k) = \frac{1}{B_{A,n}} u_n(k) \quad , \quad k=L,\dots, \infty, \quad (4.2b)$$

wobei $B_{A,n}$ die Blockierungswahrscheinlichkeit der n-ten Anforderung bedeutet :

$$B_{A,n} = \sum_{i=L}^{\infty} u_n(i) \quad (4.3)$$

Die Verteilungen in Gl.(4.2a,b) werden im später vorgestellten Algorithmus in der unnormierten Form benutzt :

$$\hat{u}_{n,0}(k) = u_n(k) \quad , \quad k=0,1,\dots,L-1, \quad (4.4a)$$

$$\hat{u}_{n,1}(k) = u_n(k) \quad , \quad k=L,\dots \quad (4.4b)$$

Führt man zusätzlich folgende ZV bzgl. der (n+1)-ten Anforderung ein

$$U_{n+1,0} = U_{n+1} | U_n < L \quad , \quad U_{n+1,1} = U_{n+1} | U_n \geq L, \quad (4.5)$$

so ergeben sich folgende Beziehungen zwischen den Zufallsvariablen und Verteilungen (vgl. Herleitung der Lindley-Integralgleichung in Kap. 3.3) :

i) $U_n < L$: die n-te Anforderung wird angenommen

$$U_{n+1,0} = U_{n,0} + B_n - A_n \quad (4.6a)$$

oder gemäß Gl.(3.32)

$$u_{n+1,0}(k) = \pi_0(u_{n,0}(k) * b_n(k) * a_n(-k)), \quad (4.6b)$$

wobei π_0 den in Gl.(3.31) definierten Operator bedeutet.

ii) $U_n \geq L$: die n-te Anforderung wird abgewiesen

$$U_{n+1,1} = U_{n,1} - A_n \quad (4.7a)$$

oder

$$u_{n+1,1}(k) = \pi_0(u_{n,1}(k) * a_n(-k)) \quad (4.7b)$$

Mit der Betrachtung der totalen Wahrscheinlichkeit erhält man :

$$u_{n+1}(k) = (1-B_{A,n}) u_{n+1,0}(k) + B_{A,n} u_{n+1,1}(k) \quad (4.8)$$

Aus Gl.(4.6b), (4.7b), (4.8) und (4.2a,b) ergibt sich die rekursive Beziehung zur Bestimmung der Restarbeit:

$$\begin{aligned}
 u_{n+1}(k) &= \pi_0(\hat{u}_{n,0}(k) * b_n(k) * a_n(-k)) + \pi_0(\hat{u}_{n,1}(k) * a_n(-k)) \\
 &= \pi_0((\hat{u}_{n,0}(k) * b_n(k) + \hat{u}_{n,1}(k)) * a_n(-k)) \quad (4.9)
 \end{aligned}$$

Aus Gl. (4.9) läßt sich auf sukzessive Weise die angetroffene Restarbeit der (n+1)-ten Anforderung aus der angetroffenen Restarbeit der n-ten Anforderung berechnen, wobei die Zwischenankunftsverteilung sowie die Bedienzeitverteilung anforderungsindividuell gewählt werden können. Dies führt zu einem iterativen Algorithmus zur Berechnung der Restarbeit im System zu den Ankunftszeitpunkten. Daraus läßt sich die Blockierungswahrscheinlichkeit der ankommenden Anforderungen bzw. der Rufe berechnen. Dieser Algorithmus im Zeitbereich (vgl. Kap.3.4.1 und Bild 3.5) wird in Bild 4.3 schematisch dargestellt.

Der hier vorgestellte Algorithmus kann in der allgemeinen Form für Systeme mit anforderungsabhängigen Zwischenankunfts- und Bedienzeiten angewendet werden. Dies bedeutet, daß Leistungsuntersuchungen sowohl unter stationären als auch unter instationären Belastungen mit diesem Algorithmus durchführbar sind.

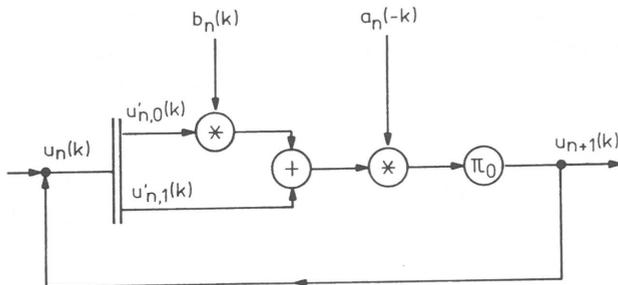


Bild 4.3 Schematische Darstellung des Algorithmus zur Analyse des Überlastabwehr-Modells

Die Schritte des Algorithmus sind im einzelnen :

- i) Initialisierung der Restarbeit $u_n(k)$, $n=0$. Dies kann z.B. durch die Betrachtung eines leeren Systems beim Eintreffen der 0-ten Anforderung geschehen (vgl. Gl.(3.54)).
- ii) Aufteilung der Verteilung $u_n(k)$ in $u_{n,0}^-(k)$ und $u_{n,1}^-(k)$ gemäß Gl.(4.4a,b).
- iii) Durchführung der Faltungs- und π_0 -Operationen gemäß Gl.(4.9) zur Berechnung der von der $(n+1)$ -ten Anforderung angetroffenen Restarbeit $u_{n+1}(k)$.
- iv) Falls Zwischenankunftsabstände und Bedienzeiten unabhängig und identisch verteilt sind, d.h.:

$$A_n = A \quad , \quad B_n = B \quad \text{für alle } n,$$

so kann die Verteilung der Restarbeit beim Ankunftszeitpunkt einer Anforderung eines Systems im eingeschwungenen Zustand

$$U = \lim_{n \rightarrow \infty} U_n$$

durch Wiederholung der Schritte ii) und iii) bis zur Konvergenz berechnet werden. Als Konvergenzkriterium wird die Differenz zwischen den Mittelwerten zweier bzw. mehrerer sukzessiv ermittelter Restarbeit-Verteilungen genommen.

Die hier benötigten Faltungsoperationen können z.B. mit der diskreten Fourier-Transformation (DFT) und entsprechenden Algorithmen der schnellen Fourier-Transformation (FFT) - vgl. Kap. 2.2.3 - durchgeführt werden.

4.1.4 Ergebnisse

Im folgenden wird - ähnlich wie im Kapitel 3.6.2 - anhand von negativ-binomial verteilten Zwischenankunfts- und Bedienzeiten eine systematische Untersuchung der parametrischen Beeinflussung der Ankunfts- und Bedienprozesse auf die Leistung der Überlastabwehrstrategie durchgeführt.

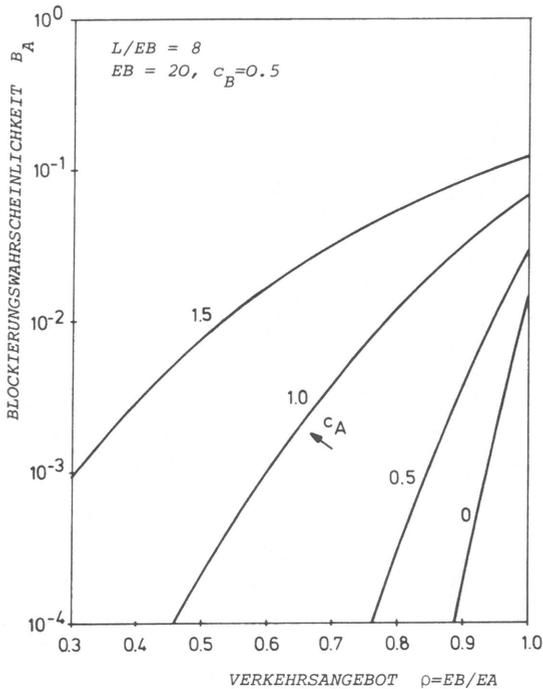


Bild 4.4 Abhängigkeit der Blockierungswahrscheinlichkeit vom Typ des Ankunftsprozesses und von der Verkehrsintensität

a) Stationäres Verhalten der Überlastabwehrstrategie

Bild 4.4 zeigt den Einfluß der Verkehrsintensität und des Ankunftsprozesses auf die Blockierungswahrscheinlichkeit. Es ist ersichtlich, daß die Blockierungswahrscheinlichkeit vom Typ bzw. Variationskoeffizienten des Ankunftsprozesses sehr stark abhängig ist. Dieser Effekt muß bei der Dimensionierung des Schwellenwerts

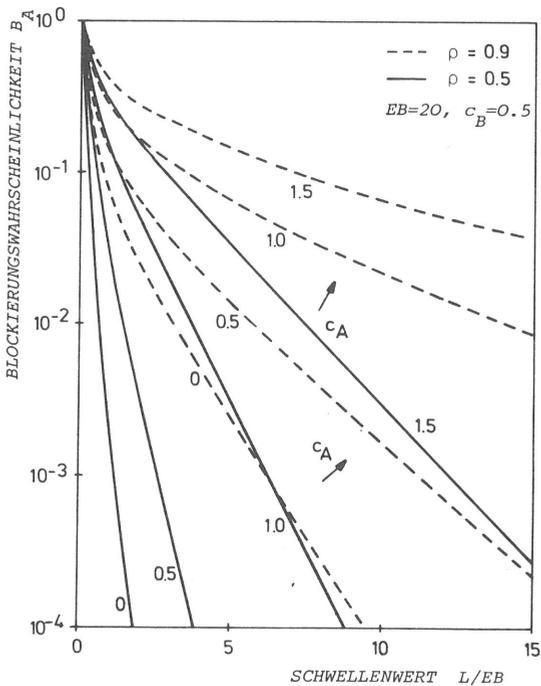


Bild 4.5 Einfluß des Ankunftsprozesses und der Dimensionierung des Schwellenwerts auf die Blockierungswahrscheinlichkeit

berücksichtigt werden, wobei der Blockierungseffekt unter normalen Lastbedingungen verschwindend klein gehalten werden muß. Dieses Verhalten ist ebenfalls in Bild 4.5 deutlich erkennbar, bei dem die Blockierungswahrscheinlichkeit als Funktion des Schwellenwerts der Überlastabwehrstrategie dargestellt wird für verschiedene Ankunftsprozeßtypen und Systembelastungen.

b) Systemantwort auf instationäre Überlastungen

Im Falle einer impulsförmigen Überlastung kann die Systemantwort durch die Beobachtung der anforderungsindividuellen Blockierungswahrscheinlichkeit $B_{A,n}$ in Gl.(4.3) registriert werden.

Der Überlastimpuls wird im folgenden durch eine Veränderung des Mittelwertes der negativ-binomial verteilten, anforderungsabhängigen Zwischenankunftszeit realisiert, wobei

$$\begin{aligned} EA_n &= 2 EB, & n=0, \dots, 14 & \quad (\rho=0.5) \\ EA_n &= EB & n=15, \dots, 39 & \quad (\rho=1) \\ EA_n &= 2 EB & n=40, \dots & \quad (\rho=0.5). \end{aligned}$$

Dies entspricht einer rechteckförmigen Überlastung mit der mittleren Dauer 25 EB. Bild 4.6 zeigt die dynamische Systemantwort auf diese instationäre Überlastung. Durch unterschiedliche Variationskoeffizienten der Zwischenankunftszeiten wird der Einfluß des Ankunftsprozesses auf den Verlauf der Blockierungswahrscheinlichkeit sichtbar. Es wird dabei vorausgesetzt, daß das System am Anfang leer ist, d.h. $B_{A,0} = 0$.

Bei den angenommenen Parametern ist ersichtlich, daß die Überlastabwehrstrategie verhältnismäßig schnell auf den Lastsprung reagiert und rasch zu den stationären Werten zurückkehrt, unabhängig vom Typ des Ankunftsprozesses.

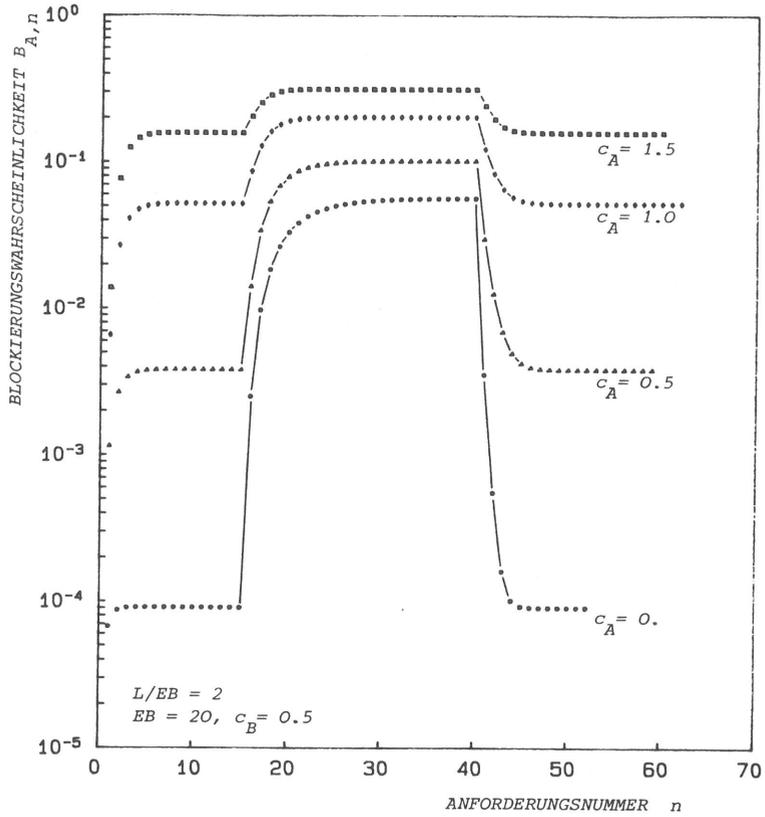


Bild 4.6 Systemantwort auf einen rechteckförmigen Überlastimpuls

4.2 Systeme mit zyklischen Abfertigungsmechanismen

4.2.1 Die Klasse von Polling-Systemen

Warteschlangensysteme mit zyklischen Abfertigungsmechanismen, die auch als Polling-Systeme bezeichnet werden, bilden eine Klasse von Verkehrsmodellen, die oft in Leistungsuntersuchungen von Rechner- und Kommunikationssystemen angewendet werden.

Ein Polling-System besteht aus einer Bedieneinheit (oder mehreren Bedieneinheiten), die eine Anzahl von Warteschlangen abfragt und anstehende Anforderungen abarbeitet. Die Ablaufsteuerung wird durch eine Abfertigungsreihenfolge und eine Bedienstrategie charakterisiert. Die Abfertigungsreihenfolge schreibt vor, wie die Warteschlangen nacheinander abgefragt werden. Diese kann z.B. zyklisch oder prioritätsgesteuert organisiert werden.

Die Bedienstrategie legt die maximale Anzahl von Anforderungen in der betreffenden Warteschlange fest, die zu einem Abfragezeitpunkt abgearbeitet werden. Beispiele dafür sind :

- Erschöpfende Bedienung : die abgefragte Warteschlange wird solange bedient, bis keine Anforderungen mehr vorhanden sind
- Gating-Bedienung : nur Anforderungen, die sich zum Abfragezeitpunkt in der betreffenden Warteschlange befinden, werden abgefertigt
- Nicht-erschöpfende Bedienung : zu einem Abfragezeitpunkt wird maximal eine bestimmte Anzahl von Anforderungen bearbeitet.

Nach der Bedienung einer Warteschlange wird bis zum Bedienungsbeginn der nächsten Warteschlange eine Verwaltungszeit benötigt. Diese Zeitspanne wird im folgenden als Umschaltzeit bezeichnet.

Die Anwendung von Polling-Systemen in Modellbildungen für Rechner- und Kommunikationssysteme erstreckt sich über eine große Anzahl von Systemen, z.B. bei Terminal-Multiplexern /48/, in Vermittlungssystemen mit teilzentralen Steuerungen /16,17,24,28,46/ und

in lokalen Rechnernetzen mit Token-Zugriffsverfahren /12,68,70/. Zur Verkehrsanalyse von Polling-Systemen existiert eine umfangreiche Literatur; eine ausführliche Übersicht findet sich in /68/.

Leibowitz /48/ untersuchte Polling-Systeme mit zyklischer Abfrage, symmetrischem Verkehr, konstanter Umschaltzeit und Gating-Bedienung. In /16,17/ wurden Systeme mit erschöpfender Bedienung behandelt, wobei die Umschaltzeit nicht berücksichtigt wurde. Die Lösung von Cooper und Murray in /16/ wurde von Hashida /28/ für nicht-verschwindende Umschaltzeiten verallgemeinert. Eine exakte Lösung für den speziellen Fall eines Systems mit zwei Warteschlangen und allgemein verteilter Umschaltzeit wurde von Eisenberg /24/ vorgestellt.

Eine approximative Lösung für Polling-Systeme mit nicht-erschöpfender Bedienung und allgemein verteilter Umschaltzeit wurde in Kühn /46/ entwickelt. Systeme mit begrenzter Warteschlangenlänge wurden in Tran-Gia und Raith /70/ untersucht, wobei ein numerischer Algorithmus im kontinuierlichen Zeitbereich vorgestellt wurde.

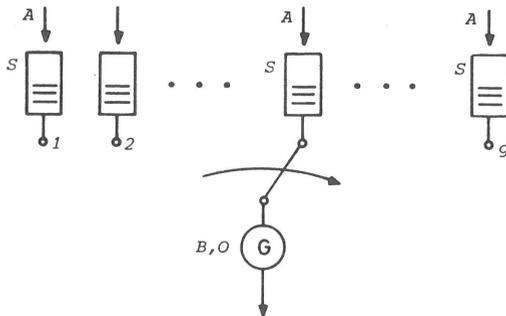


Bild 4.7 Struktur des Polling-Systems

Die meisten analytischen Arbeiten über Polling-Systeme setzen Poisson-Ankunftsverkehr voraus. Dies stellt eine gute approximative Verkehrsbeschreibung in Systemen mit einer ausreichend großen Anzahl von Verkehrsquellen dar. Realistischer jedoch läßt sich der Verkehr charakterisieren, indem Ankunftsverkehrsströme mittels Erneuerungsprozessen beschrieben werden.

Eine weitere, häufig anzutreffende Annahme in der Literatur besteht darin, daß die Kapazität der betrachteten Warteschlangen unendlich ist. Die meisten Untersuchungen basieren auf einer Analyse im transformierten Bereich, die mit Hilfe der Methode der eingebetteten Markoff-Kette durchgeführt wurden. Diese Methode führt jedoch nicht zu geschlossenen oder numerisch leicht auswertbaren Lösungen, falls Polling-Systeme mit begrenzter Warteschlangen-Kapazität Gegenstand der Untersuchung sind.

Im folgenden wird eine approximative zeitdiskrete Analyse für Polling-Systeme mit endlichem Warteraum und allgemeinen Ankunftsprozessen vorgestellt.

4.2.2 Modellbeschreibung und Voraussetzungen

Die Struktur des betrachteten Modells wurde bereits in Bild 4.7 dargestellt. Obwohl das hier vorgestellte Verfahren zur Analyse von unsymmetrischen Polling-Systemen mit warteschlangenindividuellen Belastungen angewendet werden kann, wird in diesem Unterkapitel wegen der Übersichtlichkeit der Fall symmetrischer Systeme behandelt.

Folgende Symbole und Zufallsvariablen (ZV) werden benutzt :

- g Anzahl der Warteschlangen, die zyklisch abgefragt werden
- S Kapazität einer Warteschlange
- A ZV für die Zwischenankunftszeit des Ankunftsprozesses an einer Warteschlange, Verteilung $a(k)$
- B ZV für die Bedienzeit, Verteilung $b(k)$
- O ZV für die Umschaltzeit, Verteilung $o(k)$

Die Warteschlangen werden zyklisch abgefragt und nicht-erschöpfend bedient, wobei maximal eine Anforderung pro Abfrage abgefertigt werden kann.

4.2.3 Zeitdiskrete Modellanalyse

a) Voraussetzungen und Bezeichnungen

Eine Stichprobe des Prozeßgeschehens in einer Warteschlange - z.B. in der Warteschlange j - wird in Bild 4.8 gezeigt. Bezeichnet man als Zykluszeit das Zeitintervall zwischen zwei aufeinanderfolgenden Abfragezeitpunkten bzgl. einer Warteschlange, so kann, von der Warteschlange j aus gesehen, zwischen zwei Typen bedingter Zykluszeiten /46/ unterschieden werden :

- C_0 Zykluszeit ohne Bedienung einer Anforderung in der Warteschlange j
- C_1 Zykluszeit mit Bedienung einer Anforderung in der Warteschlange j .

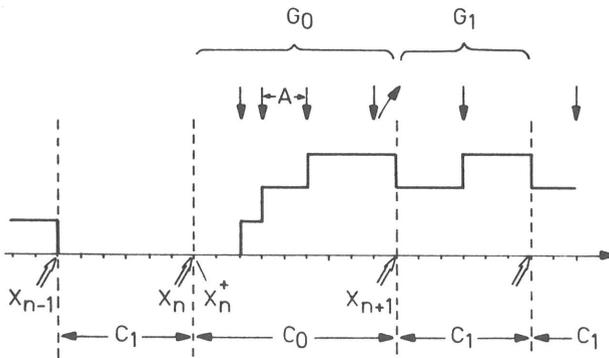


Bild 4.8 Stichprobe des Prozeßverlaufs in einer Warteschlange des Polling-Systems

Wird die Anzahl der während einer Zykluszeit ankommenden Anforderungen beobachtet, so können zwei Typen von Gruppengrößen definiert werden:

G_i ZV für die Anzahl ankommender Anforderungen in einer bedingten Zykluszeit des Typs i (C_i), $i=0,1$.

Da hier im diskreten Zeitbereich operiert wird, können zu einem diskreten Zeitpunkt verschiedene Ereignistypen auftreten. Für den Fall, daß Bedienende-, Abfrage- und Ankunftsereignisse gleichzeitig eintreffen, wird die Konvention getroffen, daß diese Ereignisse in der aufgezählten Reihenfolge behandelt werden.

b) Markoff-Kette und Zustandswahrscheinlichkeiten

Wie in Bild 4.8 dargestellt, wird der Zustandsprozeß der Warteschlange j von zwei Zufallsprozessen beeinflusst : dem Ankunftsprozeß und dem Abfrageprozeß. Unter der Annahme, daß diese Prozesse unkorreliert sind, kann im diskreten Zeitbereich approximativ zu den Zeitpunkten, die unmittelbar vor den Abfragezeitpunkten der betrachteten Warteschlange liegen, eine Markoff-Kette eingebettet werden. Folgende ZV zur Zustandsbeschreibung an den Regenerationszeitpunkten werden definiert :

X_n ZV für die Anzahl von Anforderungen in der betrachteten Warteschlange unmittelbar vor dem n -ten Abfragezeitpunkt
 X_n^+ ZV für die Anzahl von Anforderungen unmittelbar nach dem n -ten Abfragezeitpunkt.

Ähnlich wie bei der Definition bedingter Zykluszeiten, unter Berücksichtigung der Bedingung für X_n , werden folgende ZV eingeführt:

$$X_{n,0}^+ = X_n^+ | X_n = 0 \quad , \quad X_{n,1}^+ = X_n^+ | X_n > 0 \quad (4.10)$$

$$X_{n+1,0} = X_{n+1} | X_n = 0 \quad , \quad X_{n+1,1} = X_{n+1} | X_n > 0 \quad (4.11)$$

Die Beziehungen zwischen diesen Zufallsvariablen und ihren Verteilungen können wie folgt angegeben werden :

i) $\underline{X_n = 0}$

$$X_{n,0}^+ = 0 \quad (4.12a)$$

$$x_{n,0}^+(k) = \delta(k) \quad (4.12b)$$

$$X_{n+1,0} = \min(X_{n,0}^+ + G_0, S) = \min(G_0, S) \quad (4.13a)$$

$$x_{n+1,0}(k) = \pi^S(x_{n,0}^+(k) * g_0(k)) , \quad (4.13b)$$

wobei der Operator π^m wie folgt definiert wird (analog zur Definition von π_0 in Gl.(3.31)) :

$$\pi^m(x(k)) = \begin{cases} x(k) & k < m \\ \sum_{i=m}^{\infty} x(i) & k = m \\ 0 & k > m \end{cases} . \quad (4.14)$$

ii) $\underline{X_n > 0}$

$$X_{n,1}^+ = X_n - 1 \quad (4.15a)$$

$$x_{n,1}^+(k) = \begin{cases} \frac{x_n(k+1)}{P(X_n > 0)} = \frac{x_n(k+1)}{1-x_n(0)} & k=0,1,\dots,S-1 \\ 0 & \text{sonst} \end{cases} \quad (4.15b)$$

$$X_{n+1,1} = \min(X_{n,1}^+ + G_1, S) \quad (4.16a)$$

$$x_{n+1,1}(k) = \pi^S(x_{n,1}^+(k) * g_1(k)) . \quad (4.16b)$$

Nach dem Gesetz der totalen Wahrscheinlichkeit erhält man für die Zustandswahrscheinlichkeiten am (n+1)-ten Regenerationszeitpunkt :

$$x_{n+1}(k) = x_n(0) \cdot x_{n+1,0}(k) + (1-x_n(0)) \cdot x_{n+1,1}(k), \quad k=0,1,\dots,S, \quad (4.17)$$

oder mit Gl.(4.12b), (4.13b), (4.15b) und (4.16b) die Zustandsgleichungen für das Übergangsverhalten der eingebetteten Markoff-Kette :

$$x_{n+1}(k) = x_n(0) \cdot g_0(k) + \sum_{i=1}^{k+1} x_n(i) \cdot g_1(k-i+1), \quad k=0, \dots, S-1$$

$$x_{n+1}(S) = x_n(0) \sum_{i=S}^{\infty} g_0(i) + \sum_{i=1}^S x_n(i) \sum_{m=S-i+1}^{\infty} g_1(m), \quad k = S.$$

(4.18)

Befindet sich das System im eingeschwungenen Zustand, mit

$$X = \lim_{n \rightarrow \infty} X_n, \quad (4.19)$$

so kann aus Gl.(4.18) das System der stationären Zustandsgleichungen der Markoff-Kette gewonnen werden

$$x(k) = x(0) \cdot g_0(k) + \sum_{i=1}^{k+1} x(i) \cdot g_1(k-i+1), \quad k = 0, \dots, S-1,$$

$$x(S) = x(0) \sum_{i=S}^{\infty} g_0(i) + \sum_{i=1}^S x(i) \sum_{m=S-i+1}^{\infty} g_1(m), \quad k = S.$$

(4.20)

Da der Ankunftsprozeß und der Abfrageprozeß als zueinander unkorreliert vorausgesetzt wurden, können die Gruppengrößenverteilungen $g_0(k)$ und $g_1(k)$ aus der in Kapitel 2.4.4 durchgeführten Berechnung der Anzahl von Ereignissen in einem beliebig verteilten Intervall wie folgt angegeben werden :

$$g_j(k) = c_j(0) \cdot \delta(k)$$

$$+ \sum_{m=1}^{\infty} c_j(m) \sum_{i=0}^{m-1} (f^{(k)}(i) - f^{(k+1)}(i)) \quad \begin{matrix} k=0,1,\dots, \\ j=0,1, \end{matrix} \quad (4.21a)$$

wobei gemäß Gl. (2.46)

$$f^{(k)}(i) = a^*(i) * \underbrace{a(i) * \dots * a(i)}_{(k-1)\text{-mal}} . \quad (4.21b)$$

Die zur Auswertung der Gl.(4.21a) benötigten Zykluszeitverteilungen sind Gegenstand des im nächsten Unterkapitel 4.2.3c behandelten Analyseschritts.

c) Zykluszeitanalyse

Setzt man ein im eingeschwungenen Zustand befindliches, symmetrisches System voraus, so gilt die aus Gl.(4.20) berechnete Zustandsverteilung der eingebetteten Markoff-Kette für alle Warteschlangen. Die unter dieser Bedingung gemessene Zykluszeit setzt sich aus g gleichartigen Zykluszeit-Segmenten zusammen. Abhängig vom Systemzustand zum Abfragezeitpunkt besteht ein Zykluszeit-Segment bzgl. einer Warteschlange aus

- der Summe einer Umschaltzeit und einer Bedienzeit, falls mindestens eine Anforderung angetroffen wird, oder
- lediglich einer Umschaltzeit, falls die abgefragte Warteschlange leer ist.

Die Verteilung eines Zykluszeit-Segments C_S lautet dementsprechend

$$c_S(k) = x(0) \cdot o(k) + (1-x(0)) \cdot (o(k) * b(k)) \quad (4.22a)$$

oder im Z -Bereich

$$c_{S,ZT}(z) = o_{ZT}(z) (x(0) + (1-x(0)) \cdot b_{ZT}(z)) . \quad (4.22b)$$

Aus der Sicht einer Warteschlange errechnen sich die bedingten Zykluszeiten zu

$$c_{0,ZT}(z) = o_{ZT}(z) \cdot c_{S,ZT}(z)^{g-1} \quad (4.23a)$$

$$c_{1,ZT}(z) = o_{ZT}(z) \cdot b_{ZT}(z) \cdot c_{S,ZT}(z)^{g-1} . \quad (4.23b)$$

Die Zykluszeitverteilung lautet schließlich :

$$\begin{aligned} c_{ZT}(z) &= c_{S,ZT}(z)^g \\ &= x(0) c_{0,ZT}(z) + (1 - x(0)) c_{1,ZT}(z) . \end{aligned} \quad (4.24)$$

Die Auswertung der Gl.(4.23a,b) und (4.24) kann mit Hilfe der effektiven Algorithmen zur diskreten Fourier-Transformation (DFT) und entsprechender schneller Fourier-Transformation (FFT) durchgeführt werden (Kap.2.2.3).

d) Algorithmus zur iterativen Analyse

Basierend auf den Analyseschritten für Markoff-Ketten-Wahrscheinlichkeiten bzw. für Zykluszeiten, die in Kap. 4.2.3c und 4.2.3d beschrieben wurden, wird ein numerischer Algorithmus zur iterativen Analyse des Polling-Systems entwickelt /70/. Der Algorithmus besteht hauptsächlich aus einer alternierenden Iteration zwischen der Zykluszeit- und der Zustandswahrscheinlichkeit-Berechnung. Die Einzelschritte sind :

- i) Initialisierung der Zustandswahrscheinlichkeiten und Zykluszeiten.
- ii) Berechnung der bedingten Zykluszeiten aus Gl.(4.23a) und (4.23b)
- iii) Berechnung der Gruppenankunftsverteilungen aus Gl.(4.21a).
- iv) Berechnung der Zustandswahrscheinlichkeiten gemäß Gl.(4.18) bzw (4.20).
- v) Wiederholung der Schritte ii), iii) und iv) bis zur Konvergenz der Zustandswahrscheinlichkeiten bzw. der Zykluszeiten. Als Konvergenzkriterium kann z.B. die Differenz zwischen den Mittelwerten sukzessiv ermittelter Zustandsverteilungen genommen werden.

Aus den durch die Iteration gewonnenen Zykluszeitverteilungen und Zustandswahrscheinlichkeiten können Systemcharakteristiken berechnet werden.

4.2.4 Systemcharakteristiken

a) Blockierungswahrscheinlichkeit

Aufgrund der begrenzten Kapazität der Warteschlangen im System können Blockierungen auftreten. Die Wahrscheinlichkeit, daß eine Anforderung abgewiesen wird, kann aus den Zustandswahrscheinlichkeiten der Markoff-Kette hergeleitet werden.

Betrachtet wird zunächst die bedingte Blockierungswahrscheinlichkeit $B(k)$, die wie folgt definiert wird :

$B(k) = P(\text{Blockierung einer Anforderung} \mid X = k)$
: Wahrscheinlichkeit für die Abweisung einer Anforderung, falls sich zum vorausgegangenen Abfragezeitpunkt k Anforderungen in der Warteschlange befinden.

Setzt man nun den Fall $X > 0$ voraus, d.h. die darauffolgende Zykluszeit ist vom Typ C_1 und die Gruppengröße vom Typ G_1 , so kann durch die Beobachtung einer in dieser Gruppe befindlichen Testanforderung die bedingte Blockierungswahrscheinlichkeit $B(k)$ in folgenden Schritten bestimmt werden :

i) Die Wahrscheinlichkeit, daß sich die Testanforderung in einer Gruppe der Länge i befindet, ist proportional zur Gruppengröße i und zur Auftrittswahrscheinlichkeit $g_1(i)$; sie lautet

$$i \cdot g_1(i) / EG_1$$

ii) Blockierung findet statt, falls $k-1+i > S$ ist. In diesem Falle werden $k-1+i-S$ Anforderungen in der Gruppe abgewiesen. Die Wahrscheinlichkeit, daß die Testanforderung im abgewiesenen Teil der Gruppe enthalten ist, d.h. blockiert wird, ist

$$(k-1+i-S)/i$$

iii) Aus i) und ii) ergibt sich für $B(k)$, $k > 0$

$$B(k) = \sum_{i=S-k+2}^{\infty} \frac{k-1+i-S}{i} \frac{i \cdot g_1(i)}{EG_1} = \frac{1}{EG_1} \sum_{i=S-k+2}^{\infty} (k-1+i-S) \cdot g_1(i) \quad k=1, 2, \dots, S. \quad (4.25a)$$

Entsprechend erhält man für den Fall $k=0$:

$$B(0) = \sum_{i=S+1}^{\infty} \frac{i-S}{i} \frac{i \cdot g_0(i)}{EG_0} = \frac{1}{EG_0} \sum_{i=S+1}^{\infty} (i-S) g_0(i) \quad (4.25b)$$

Die Blockierungswahrscheinlichkeit B_A für Anforderungen errechnet sich schließlich aus Gl.(4.25a,b) zu :

$$\begin{aligned} B_A &= \sum_{k=0}^S x(k) \cdot B(k) \\ &= \frac{1}{EG_0} x(0) \sum_{i=S+1}^{\infty} (i-S) g_0(i) + \frac{1}{EG_1} \sum_{k=1}^S x(k) \sum_{i=S-k+2}^{\infty} (k-1+i-S) g_1(i) \end{aligned} \quad (4.26)$$

b) Zeitgemittelte Zustandswahrscheinlichkeiten

Wie bereits diskutiert wurde, gelten die im Kap. 4.2.3b berechneten Zustandswahrscheinlichkeiten der eingebetteten Markoff-Kette nur zu den Regenerationszeitpunkten, d.h. zu den diskreten Zeitpunkten, die unmittelbar vor den Abfragezeitpunkten einer Warteschlange liegen. Für die Berechnung weiterer Systemcharakteristiken, z.B. mittlerer Wartezeit angenommener Anforderungen, werden im folgenden die zeitgemittelten Zustandswahrscheinlichkeiten bestimmt. Diese stellen Zustandswahrscheinlichkeiten dar, die von einem unabhängigen außenstehenden Beobachter gesehen werden.

Betrachtet wird nun ein zufälliger Beobachtungszeitpunkt t^* , der unmittelbar vor einem diskreten Zeitpunkt im diskretisierten Zeitbereich liegt (Bild 4.9). Die Wahrscheinlichkeit P_j^* , daß t^* in einer bedingten Zykluszeit vom Typ C_j liegt ($j=0,1$), läßt sich aus

der Theorie von Semi-Markoff-Prozessen wie folgt angeben

$$P_0^* = x(0) \frac{EC_0}{EC} \quad (4.27a)$$

$$P_1^* = 1 - P_0^* = (1-x(0)) \frac{EC_1}{EC} \quad (4.27b)$$

Wird eine bedingte Zykluszeit vom Typ C_j ($j=0,1$) angetroffen, so ist das Intervall vom vorausgegangenen Abfragezeitpunkt bis zum zufälligen Beobachtungszeitpunkt genau die diskrete Rekurrenzzeit C_j^* ($j=0,1$) der bedingten Zykluszeit C_j (s. Bild 4.9) mit der in Kapitel 2.4 hergeleiteten Verteilung

$$c_j^*(k) = \frac{1}{EC_j} \left(1 - \sum_{i=0}^k c_j(i) \right), \quad k=0,1,\dots, \quad j=0,1. \quad (4.28)$$

Die Anzahl der während des Intervalls C_j^* ankommenden Anforderungen wird mit der ZV G_j^* ($j=0,1$) gekennzeichnet (s. Bild 4.9). Die zugehörige Gruppengrößenverteilung kann entsprechend Gl. (4.21a) hergeleitet werden :

$$g_j^*(k) = c_j^*(0) \cdot \delta(k) + \sum_{m=1}^{\infty} c_j^*(m) \sum_{i=0}^{m-1} (f^{(k)}(i) - f^{(k+1)}(i)), \quad k=0,\dots, \quad j=0,1. \quad (4.29)$$

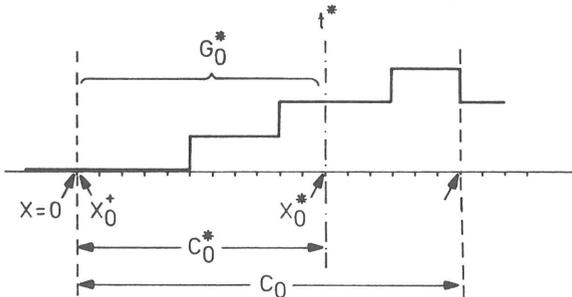


Bild 4.9 Prozeßbetrachtung zur Berechnung der zeitgemittelten Zustandswahrscheinlichkeiten

Folgende Zufallsvariablen werden für den Zustand einer Warteschlange, der durch die Anzahl wartender Anforderungen charakterisiert wird, definiert :

X^* ZV für den Zustand zum zufälligen Beobachtungszeitpunkt t^*
 X, X^+ ZV für den Zustand unmittelbar vor bzw. nach dem voraus-
gegangenen Regenerationszeitpunkt.

und, ähnlich wie bei der Analyse der Markoff-Kette (Kap. 4.2.3b),
die ZV :

$$X_0^+ = X^+ | X = 0 \quad , \quad X_1^+ = X^+ | X > 0 \quad . \quad (4.30)$$

$$X_0^* = X^* | X = 0 \quad , \quad X_1^* = X^* | X > 0 \quad . \quad (4.31)$$

Analog zu den Gleichungen (4.12), (4.13), (4.15) und (4.16) ergibt
sich für diese Zufallsvariablen und ihre Verteilungen :

i) $X = 0$

$$X_0^+ = 0 \quad (4.32a)$$

$$x_0^+(k) = \delta(k) \quad (4.32b)$$

$$X_0^* = \min(X_0^+ + G_0^*, S) = \min(G_0^*, S) \quad (4.33a)$$

$$x_0^*(k) = \pi^S(x_0^+(k) * g_0^*(k)) \quad , \quad (4.33b)$$

ii) $X > 0$

$$X_1^+ = X - 1 \quad (4.34a)$$

$$x_1^+(k) = \begin{cases} \frac{x(k+1)}{P(X>0)} = \frac{x(k+1)}{1-x(0)} & k=0,1,\dots,S-1 \\ 0 & \text{sonst} \end{cases} \quad (4.34b)$$

$$X_1^* = \min(X_1^+ + G_1^*, S) \quad (4.35a)$$

$$x_1^*(k) = \pi^S(x_1^+(k) * g_1^*(k)) . \quad (4.35b)$$

Nach dem Gesetz der totalen Wahrscheinlichkeit erhält man für die zeitgemittelten Zustandswahrscheinlichkeiten

$$x^*(k) = P_0^* x_0^*(k) + P_1^* x_1^*(k), \quad k=0,1,\dots,S. \quad (4.36)$$

oder durch Einsetzen der Gl.(4.27a,b), (4.28) und (4.32-35b) in Gl.(4.36):

$$x^*(k) = \frac{EC_0}{EC} x(0)g_0^*(k) + \frac{EC_1}{EC} \sum_{i=1}^{k+1} x(i) \cdot g_1^*(k-i+1), \quad k=0,\dots,S-1$$

$$x^*(S) = \frac{EC_0}{EC} x(0) \sum_{i=S}^{\infty} g_0^*(i) + \frac{EC_1}{EC} \sum_{i=1}^S x(i) \sum_{m=S-i+1}^{\infty} g_1^*(m), \quad k=S. \quad (4.37)$$

c) Mittlere Wartezeit

Aus den zeitgemittelten Zustandswahrscheinlichkeiten $x^*(k)$ und der Blockierungswahrscheinlichkeit B_A kann durch Anwendung des Little'schen Theorems (s.Kap. 3.4) die mittlere Wartezeit der angenommenen Anforderungen hergeleitet werden.

Betrachtet man nun als System eine Warteschlange mit der mittleren Systembelegung EX_A^* , der Ankunftsrate der nicht-blockierten Anforderungen $\frac{1}{EA}(1-B_A)$ - d.h. mit der mittleren Zwischenankunftszeit $EA/(1-B_A)$ -, so ergibt sich für die mittlere Wartezeit angenommener Anforderungen (die mittlere Aufenthaltsdauer im System) gemäß Gl.(3.79) :

$$EW_A = \frac{EX_A^* \cdot EA}{(1-B_A)} . \quad (4.38)$$

4.2.5 Ergebnisse und Validierung

Als Beispiel wird im folgenden ein symmetrisches Polling-System mit $g=5$ und $S=5$ untersucht und die Genauigkeit des vorgestellten Berechnungsverfahrens diskutiert. Die Umschaltzeit wird dabei als konstant mit dem Mittelwert $EO = EB/2$ angenommen. Die Ankunfts- und Bedienprozesse werden durch negativ-binomiale Verteilungen parametrisch festgelegt. Die approximativ gewonnenen Ergebnisse werden mit Simulationswerten validiert. Diese Simulationspunkte werden in den Diagrammen zusammen mit ihren 95%-Vertrauensintervallen eingezeichnet. Die in den Ergebnisdiagrammen angegebenen Zeiten - z.B. Zykluszeit, Wartezeit, usw. - sind auf $EB = 1$ normiert.

Der Mittelwert und der Variationskoeffizient der Zykluszeit werden in Bild 4.10 und Bild 4.11 als Funktion des Verkehrsangebots

$$\rho = g \cdot EB/EA$$

aufgetragen, für unterschiedliche Ankunfts- und Bedienprozeßtypen. In Bild 4.10 wird die Beeinflussung des Ankunftsprozesses und in Bild 4.11 die des Bedienprozesses auf die Zykluszeit untersucht. Ähnlich wie in Kap.3.6 werden die Variationskoeffizienten hier so gewählt, daß die betrachteten Ankunfts- und Bedienzeiten jeweils die zeitdiskreten Äquivalente der zeitkontinuierlichen Verteilungstypen Erlang 4-ter Ordnung E_4 ($c_A=0.5$), negativ-exponentiell M ($c_A=1$) und hyperexponentiell H_2 ($c_A=1.5$), darstellen.

Bei höheren Verkehrsintensitäten wird in Bild 4.10 das asymptotische Verhalten der Zykluszeit deutlich. Die begrenzende maximale Zykluszeit setzt sich aus der Summe von Bedien- und Umschaltzeiten aller Warteschlangen im System zusammen, d.h. der Mittelwert des maximalen Zyklus beträgt 7.5 EB. In dieser Überlastregion verringert sich ebenfalls der Variationskoeffizient wie erwartet (Bild 4.11). Geht die Verkehrsintensität gegen Null, so erhält man den aus g Umschaltzeiten bestehenden Leerzyklus (mit dem Mittelwert 2.5 EB).

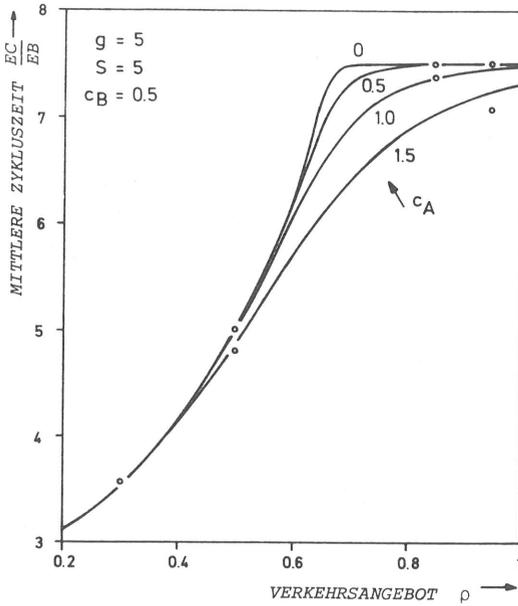


Bild 4.10

Einfluß des Ankunftsprozesses auf die mittlere Zykluszeit

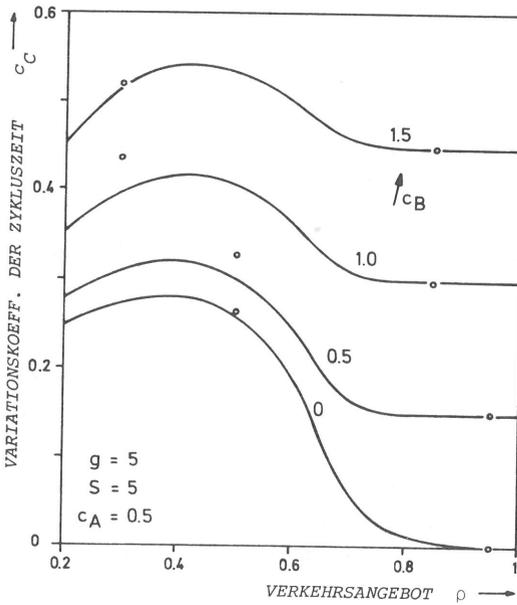


Bild 4.11

Einfluß des Bedienprozesses auf den Variationskoeffizienten der Zykluszeit

Die Abhängigkeit der Wartezeit und der Blockierungswahrscheinlichkeit der Anforderungen vom Ankunftsprozeß wird in Bild 4.12 und Bild 4.13 verdeutlicht, wobei diese Systemcharakteristiken als Funktion des Verkehrsangebots dargestellt werden. Bei verschwindendem Verkehrsangebot ist die mittlere Wartezeit identisch mit der mittleren Vorwärts-Rekurrenzzeit des Leerzyklus. Aus Bild 4.13 wird deutlich, daß die Blockierungswahrscheinlichkeit sehr drastisch mit dem Verkehrsangebot und mit dem Variationskoeffizienten der Zwischenankunftszeit ansteigt. Der Anteil des angenommenen Verkehrs verringert sich entsprechend. Dadurch läßt sich erklären, daß bei größerem Verkehrsangebot die mittlere Wartezeit angenommener Anforderungen mit steigendem c_A geringer wird.

Bild 4.14 zeigt die komplementäre Verteilungsfunktion der Zykluszeit für unterschiedliche Bedienprozeßtypen, wobei der treppenförmige Verlauf, der für zeitdiskrete Verteilungsfunktionen charakteristisch ist, deutlich zu erkennen ist.

In den Diagrammen ist zu ersehen, daß die Approximation in den untersuchten Parameterbereichen für praktische Anwendungen hinreichend genau ist. Weitere Parameteruntersuchungen zeigen jedoch, daß - bedingt durch die Unabhängigkeitsannahme bei der Zykluszeitanalyse und durch die approximative Einbettung der Markoff-Kette - die Approximationsgenauigkeit bei kürzeren Umschaltzeiten und Ankunftsprozessen mit größeren Variationskoeffizienten abnimmt (vgl. /46,70/).

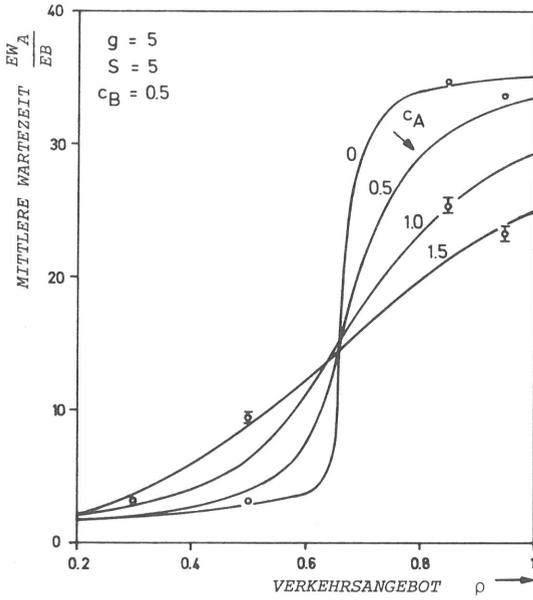


Bild 4.12

Mittlere Wartezeit
angenommener Anfor-
derungen

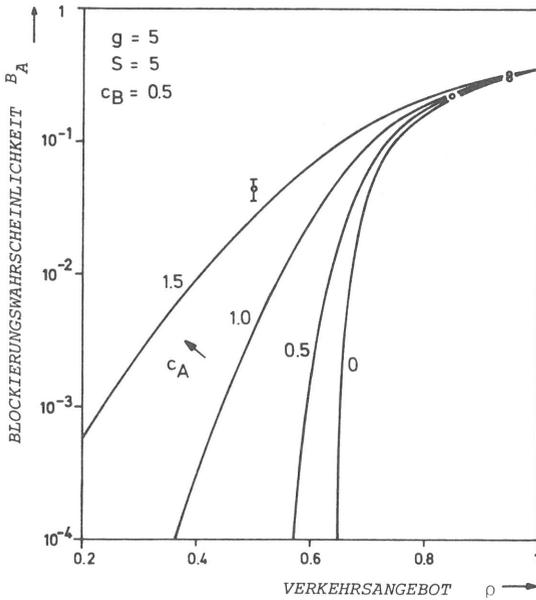


Bild 4.13

Einfluß des Ankunfts-
prozesses auf die
Blockierungswahrschein-
lichkeit

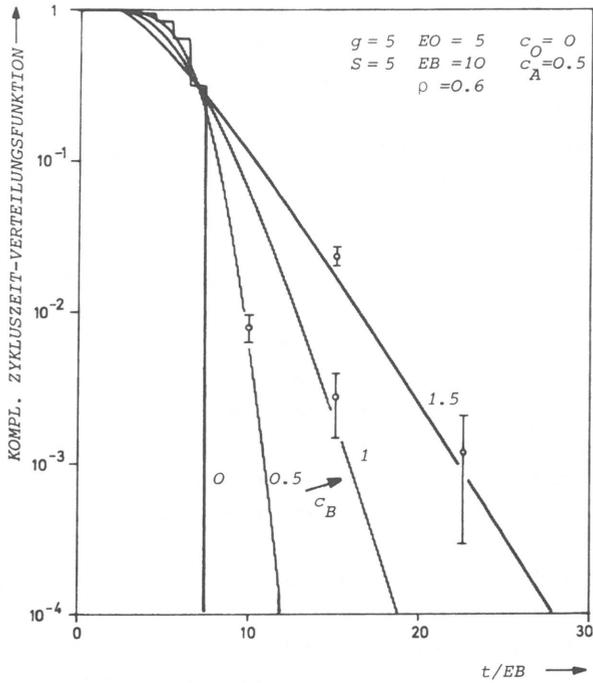


Bild 4.14 Komplementäre Verteilungsfunktion der Zykluszeit

5. ZUSAMMENFASSUNG

In der vorliegenden Arbeit wurde die Methodik und Anwendungsaspekte zur Analyse verkehrstheoretischer Modelle im diskreten Zeitbereich diskutiert. Diese Klasse von Methoden bietet ein äquivalentes Spektrum von Lösungsansätzen und ist in einigen Fällen eine methodische Ergänzung bzw. Erweiterung der zeitkontinuierlichen Analyseverfahren.

Es wurde zunächst eine Zusammenstellung der Grundlagenaspekte zur zeitdiskreten Analyse von verkehrstheoretischen Modellen gegeben. Dabei wurden benötigte Transformationsmethoden, Konzepte aus der System- und Signalverarbeitungstheorie sowie Algorithmen (z.B. diskrete Fourier-Transformation DFT, schnelle Fourier-Transformation FFT, komplexes Cepstrum) eingeführt und Anwendungsgesichtspunkte in der Verkehrstheorie diskutiert. Die Erneuerungstheorie für zeitdiskrete Prozesse wurde erörtert.

Die Methodik der zeitdiskreten Analyse von Verkehrsmodellen wurde am Beispiel des Warteschlangensystems GI/G/1 aufgezeigt, wobei exakte geschlossene Lösungen und numerische Algorithmen im Zeit- sowie transformierten Bereich vorgestellt wurden. Für diese allgemeine Klasse von Systemen wurde eine neue zeitdiskrete Methode zur Berechnung von Verteilungen der Freiperioden und des Ausgangsprozesses im diskreten Zeitbereich präsentiert.

Im abschließenden Kapitel wurde die Anwendung zeitdiskreter Methoden in der Analyse komplexer Modelle für Rechner- und Kommunikationssysteme anhand von zwei Beispielen demonstriert. Dabei wurde ein rückkopplungsbehafteter lastgesteuerter Überlastabwehrmechanismus in Vermittlungssystemen untersucht und eine zeitdiskrete Analyse für Polling-Systeme mit endlichem Warteraum und allgemeinen Ankunftsprozessen vorgestellt.

Zeitdiskrete Analysemethoden sind in erster Linie für Modelluntersuchungen gut geeignet, bei denen die Modellkomponenten und ihre Parametrisierung bereits in zeitdiskreter Form vorliegen. Wie anhand des GI/G/1-Systems gezeigt wurde, können diese Metho-

den ebenfalls approximativ zur Untersuchung zeitkontinuierlicher Systeme angewendet werden, die keiner exakten Analyse zugänglich sind.

Gegenüber den Analyseverfahren im kontinuierlichen Zeitbereich sind zeitdiskrete Methoden insbesondere in Anwendungsfällen vorteilhaft, bei denen Transformationen bzw. Faltungsoperationen numerisch notwendig sind. Dies ist z.B. der Fall, wenn Ergebnisse nur im transformierten Bereich geschlossen angebar sind und interessierende Systemcharakteristiken daraus mit Hilfe von Rücktransformationen ermittelt werden müssen. In einigen Fällen, z.B. bei der Berechnung der Wartezeitverteilung und des Ausgangsprozesses des GI/G/1-Systems, stellt die zeitdiskrete Analyse eine effektive und genaue numerische Methode dar.

Der Grund dieses konzeptionellen Vorteils liegt u.a. darin, daß hier effektive und z.T. in anderen Anwendungsgebieten (z.B. Signalverarbeitungs- und Systemtheorie) bereits optimierte Algorithmen und Verfahren im diskreten Bereich herangezogen werden können (z.B. diskrete Fourier-Transformation DFT, schnelle Fourier-Transformation FFT, Cepstrum-Verfahren). Die Anwendung zeitdiskreter Analyseverfahren zur Untersuchung der Modellbeispiele im Kapitel 4 deutet an, daß das Anwendungsspektrum zur Analyse komplexer Modelle in Rechner- und Nachrichtensystemen äußerst vielfältig ist und im Rahmen der vorliegenden Arbeit nicht erschöpfend behandelt werden konnte.

