

Institut für Nachrichtenvermittlung und Datenverarbeitung
Universität Stuttgart
Prof. Dr.-Ing. P. Kühn

37. Bericht über verkehrstheoretische Arbeiten

Verkehrsanalyse von Vermittlungs-Steuerungen
mit taktgesteuerter
Ein-/Ausgabe und Prioritäten

von
Herbert Jans

Institute of Switching and Data Technics
University of Stuttgart
Prof. Dr.-Ing. P. Kühn

37th Report on Studies in Congestion Theory

Traffic Analysis of Switching System Control
Structures with
Clocked I/O and Priorities

by
Herbert Jans

© 1983 Institut für Nachrichtenvermittlung und Datenverarbeitung Universität Stuttgart

Druck: E. Kurz & Co.

ISBN 3-922403-47-6

SUMMARY

Chapter 1 (pp. 1-13)

This report deals with the analytical investigation of a switching system control structure with partly centralized/decentralized control. Today a number of electronic switching systems with this type of control structure has been developed and is introduced into the public telephone networks or into private branch exchanges.

In the first section of this chapter some general aspects of stored program controlled (SPC) switching systems are discussed concerning basic tasks, special features, grade of control centralization and interprocessor communication in connection with input/output mechanisms. Theoretical traffic aspects for the central control part are discussed as applied to the basic control structure of a typical SPC system (Fig. 1.1).

A short survey of the report concludes Chapter 1.

Chapters 2 and 3 (pp. 14-44)

These chapters deal with some fundamentals of teletraffic and queueing theory. Modelling aspects and analysis methods are described. After defining the components of a queueing model (arrival and service processes including batch processes), a comparison between interrupt-driven and clock-driven I/O-schemes is given. Fig. 2.3 shows the advantage of the clocked scheme at higher traffic intensities which is usually used in SPC-systems.

The analysis methods described in Chapter 3 are subdivided into exact and approximate analytical methods with or without Markov property and experimental methods like system measurements and simulations. A definition of characteristic values and two theorems (Little's law and a work conservation law) being applied in the following part of this report conclude Chapter 3.

Chapter 4 (pp. 45-124)

In this chapter, a three-level queueing model for the central control part is investigated (Fig. 4.1). This model is based on the fact that the central processor in connection with the I/O-transfer scheme forms the system bottleneck. The clocked I/O-transfer initialized by a hardware clock and supervised by the central processor is modelled by the switches. For the exchange of messages from and commands to the decentralized control units, a constant I/O-overhead time is assumed. Only one class of requests is taken into consideration in this model.

Because of the infinite queue capacities all waiting requests can be transferred from the input buffer into the work queue and also out of the output buffer at each I/O-event. With this generalization and with renewal assumptions at the clock instants, the analysis is based on the decomposition method, whereby the complex system is split up into three subsystems. The different subsystems are analyzed as isolated parts and the results for the complex system can be obtained by linearly adding up the particular results for the subsystems.

The main analysis steps for the three subsystems are briefly discussed below.

The first subsystem is a queueing model of type $GI/C/\infty$, general arrival process and clocked service. Based on renewal theory assumptions, an imbedded Markov chain approach with regeneration points at the clock instants is carried out. Considering a test customer arriving at arbitrary time within a clock period, it is shown that the mean waiting time is, regardless of the arrival process type, half the time of the clock period (with the exception of a deterministic arrival process type). The output process is a clocked batch process. Because of the clocked service, the batch distribution is identical to the distribution of the number of arrivals during a clock period. From renewal theory it is well known that for a stationary process the mean number of arrivals is the product of the arrival rate and the time interval or $E[K] = \lambda T$. The variance $Var[K]$ is equal to the variance of the

number of renewals in a clock period (Eq. (4.19)). Results for the special cases $E_k/C/\infty$ and $H_2/C/\infty$ are derived in detail.

The second subsystem can be characterized by the short notation $C^{[X]}/G/1$. The clocked arrival process $C^{[X]}$ is given by the output process of the first subsystem. A fixed overhead time T_{OH} is assumed for the execution of the I/O-program. Three different analysis methods are presented:

- Imbedded Markov chain approach with regeneration points just prior to the clock period
- Analysis by application of Lindley's integral equation.
- Transformation into a queueing model of type GI/G/1 and using approximation formulas derived by Krämer and Langenbach-Belz.

For the Markov chain approach, an interrupted service will be continued after the I/O-overhead phase; the analysis is greatly enhanced by using the forward recurrence time of a stationary renewal process. Knowing the state probabilities just prior to the clock instant, the state probabilities at an arbitrary time instant can be calculated. From this state probabilities, the characteristic values like mean queue size, mean waiting time, mean flow time, and probability of waiting are obtained. The key formula for the second and third analysis method is Eq. (4.47) by which the mean total waiting time (including the portion caused by the I/O-overhead is calculated. Besides the mean $E[K]$ of the batch size, only the mean number $E[X,T]$ in the system just prior to the clock instant has to be taken into account.

Lindley's integral equation describes the waiting time distribution function for the general model GI/G/1. A direct application to models with batch arrivals such as $C^{[X]}/G/1$ is not possible. Therefore, an equivalent model $C/G_x/1$ with single arrivals has been defined. G_x takes the service of the whole arriving batch (super request) of size $K>0$ into account. From the system point of view, the waiting time w_{01} of the first request of a batch in the original system $C^{[X]}/G/1$ is equal to the unfinished work seen by the super request in the equivalent system $C/G_x/1$. The distri-

bution function of this unfinished work can be expressed by Lindley's integral equation given in Eq. (4.52). Because of the clocked arrival process, the equation can be solved by a time discretization method whereby all probability density functions (pdf's) of continuous time are approximated by discrete pdf's with the elementary time quantum Δt . A closed-form solution for the generating function $G_i(z,T)$ of the probabilities for the unfinished work quanta is derived and contains the roots $z_v, v = 0,1,2,\dots,l-1$, of the equation $z^l - G_m(z) = 0$ inside and on the unit circle (Eq. (4.65)). By differentiation and setting $z = 1$, the mean number $E[I,T]$ of unfinished work quanta and, finally, from Eq. (4.90) the mean number $E[X,T]$ in the system are obtained.

Section 4.3.3 describes the GI/G/1 approximation. For more details see [45].

The third subsystem in Section 4.4 is similar to the first one. From the analysis by a Markov chain the mean waiting time and queue length are given. Also, the first two moments of the batch output process are derived.

In the last part of this chapter, some results are reported and validated with computer simulations, especially for the flow time in the second and for the coefficient of variation in the third subsystem.

Chapter 5 (pp. 125-176).

According to the various call processing programs in SPC-systems, different priority classes are implemented in real systems. This is taken into account in Chapter 5 where a three-level queueing model similarly structured to that one in Chapter 4 but extended to an arbitrary number P of priority classes is considered. An exact analysis for the special case $C^{[X]}/M/1$ with two preemptive priority classes, different exponential service time distributions for each class, and constant I/O-overhead is carried out in the first section. The analysis is based on a two-dimensional state description and an imbedded Markov chain with regeneration points at the clock instants. Results from this investigation are used

to validate those results obtained by the approximation of Section 5.2.2.

The main principles of the approximation are:

- a) Superposition of individual class processes for composite single-class systems of type $C^{[x]}/G/1$ and description of the substitute processes by at least two parameters such as mean and variance of batch sizes and service times (Sections 5.2.2.2 and 5.2.2.3).
- b) Isolated analysis of the composite single-class systems of type $C^{[x]}/G/1$.
- c) Approximate calculation of the mean waiting time for each priority class by application of a conservation law.

1. Composite single-class systems

In the given system $C^{[x]}/G/1$ with P priority classes the i -th class is considered (Fig. 5.2). To evaluate the mean waiting time of this class, two composite single-class systems, indicated by A and B , are defined. The first one is originated from all classes 1 to $i-1$, the second one from 1 to i , respectively. According to the preemptive priority schedule, all other classes ($i+1$ to P) do not influence the waiting time of the i -th class. Because of the synchronous clocked batch arrivals, the substitute batch distributions for the composite systems are found by a convolution of all concerned component batch distributions with mean $E[K]$ and variance $Var[K]$ according to Eq. (5.32a) and Eq. (5.33). The pdf's of the substitute service processes are given by the weighted sum of the pdf's of the individual service processes. For the first two moments see Eq. (5.35).

2. Approximation for the queueing model $C^{[x]}/G/1$

With the two-parameter description, all values are known for the approximate analysis of the composite single-class system using the second or third analysis method described in Section 4.3 for the second subsystem.

3. Approximation by applying a conservation law

To find the mean waiting time for the different priority classes, a conservation law is used which is exactly valid for the system $M/G/1$ preemptive-resume but an approximation for the system under study (see Section 5.2.2.1).

In addition to the two-dimensional state description (Section 5.2.1), the third subsystem is analyzed in Section 5.3 for two priority classes. In Section 5.4.3, an approximation for the mean waiting time of the i -th class is carried out by definition of an effective service time T_{eff}^i for the considered class.

In the last part of this section, some numerical results for the subsystems and for the complex total system are reported and validated by exact results or computer simulations (200 000 arrival events, 95% confidence intervals) otherwise. Fig. 5.7 indicates the influence of the arrival process types in the first subsystem on the mean flow times in the second one. The typical priority behavior is depicted. The comparison with exact or simulation results generally shows a good accuracy. This is also the case for the results in Fig. 5.8 where the influence of the service process types on the mean flow times $f_{1,2}^*$ ($f_{1,2}^*$ includes the first and second subsystem) can be seen. The curves show that a minimum clock period T_{min}^i can be derived for each class (Eq. 5.43). Depending on the system and traffic parameters, an optimum choice of the clock period can be defined. Since the optimum values are different for each class, a compromise must be found. Further parameter studies and validations, e.g., the influence of the overhead time in Fig. 5.9, are considered. Also, a comparison between the clocked scheme and the interrupted scheme is given in Fig. 5.11. It is shown that the clocked scheme has a better performance for higher traffic intensities whereas the interrupt scheme is useful for low traffic rates only.

Fig. 5.12 demonstrates the mean response time of the complex system. It should be emphasized that the results are calculated by the sum of the individual ones for the subsystems analyzed in "isolation". In the region of saturation, the flow times of the second subsystem dominate. This indicates that the central control forms the system bottleneck.

Chapter 6 (pp. 177-178)

Concluding remarks about the traffic behavior of the switching system with partly centralized/decentralized control are given with special regards to the central control. The investigation has shown the influence of the main system and traffic parameters on the waiting and response times and can be used for system dimensioning and performance evaluation. An application of the queueing model for the traffic analysis in real-time processing systems or for a computer network operating according to a "mail box" scheme is possible. In this network the complex model represents one station where the input and output buffers stand for the "mail boxes".

INHALTSVERZEICHNIS

LITERATURVERZEICHNIS

BEZEICHNUNGEN

Seite

1. EINLEITUNG 1

1.1 Rechnergesteuerte Vermittlungssysteme 1

1.1.1 Grundaufgaben eines Vermittlungssystems 1

1.1.2 Spezielle Leistungsmerkmale rechnergesteuerter Vermittlungssysteme 3

1.1.3 Steuerungsprinzipien 4

1.1.4 Interprozessor-Kommunikation und Ein-/Ausgabe-Mechanismen 6

1.1.5 Verkehrstheoretische Aspekte der zentralen Steuereinheit 8

1.1.6 Problemstellungen bei der verkehrstheoretischen Untersuchung und Leistungsbewertung 11

1.2 Übersicht über die Arbeit 12

2. MODELLBILDUNG UND PRINZIPIELLE ERGEBNISSE 14

2.1 Systematik verkehrstheoretischer Modelle 14

2.1.1 Klassifizierung von Bedienungssystemen und Kurznotation 14

2.2 Komponenten eines verkehrstheoretischen Modells 16

2.2.1 Der Ankunftsprozeß 16

2.2.2 Der Bedienungs- oder Endeprozeß 16

2.2.3 Gruppenankunft bzw. Gruppenbedienung 17

2.2.4 Die Struktur des Bedienungssystems 18

2.2.5 Die Organisation von Bedienungssystemen 19

2.3 Modellierung unterschiedlicher Ein-/Ausgabe-Mechanismen 20

2.3.1 Interruptgesteuerte Ein-/Ausgabe 21

2.3.2 Taktgesteuerte Ein-/Ausgabe 22

2.3.3 Prinzipieller Leistungsvergleich zwischen interruptgesteuerter und taktgesteuerter Eingabe ... 24

2.4 Anwendung von Modellen in der Vermittlungstechnik ... 25

3. METHODEN DER MODELLANALYSE 27

3.1 Analytische Methoden 27

3.1.1 Exakte analytische Methoden 27

3.1.1.1 Die "Gedächtnislosigkeit" Markoff'scher Prozesse 28

3.1.1.2 Analyse Markoff'scher Bedienungssysteme 29

3.1.1.3 Analyse Nichtmarkoff'scher Prozesse 32

3.1.2 Approximative analytische Methoden 37

3.1.2.1 Zerlegungsmethoden 37

3.1.2.2 Momentenapproximation 37

3.1.2.3 Erneuerungsannahmen 38

3.2 Experimentelle Methoden 38

3.2.1 Messung an Systemen 38

3.2.2 Simulative Methoden 39

3.3 Charakteristische Größen 41

3.3.1 Kenngrößen für Systemeigenschaften 41

3.3.2 Das Little'sche Theorem 43

3.3.3 Erhaltungssätze für Bedienungssysteme 44

3.4 Zusammenfassung 44

4. ANALYSE DES DREISTUFIGEN WARTEMODELLS MIT TAKTGESTEUERTER EIN-/AUSGABE, VERWALTUNGSZEIT UND EINER ANFORDERUNGSKLASSE 45

4.1 Struktur und Betriebsweise 45

4.2 Das erste Teilmodell GI/C/∞ 49

4.2.1 Berechnung der Gruppengrößen-Verteilung 51

4.2.2 Mittelwert und Varianz der Gruppengrößen-Verteilung 57

4.3 Das zweite Teilmodell C^[x]/G/1 63

4.3.1 Analyse mit Hilfe einer eingebetteten Markoff-Kette kurz vor den Taktzeitpunkten 65

4.3.1.1 Bestimmung der Zustandswahrscheinlichkeiten . 66

4.3.1.2 Berechnung der charakteristischen Größen	69	5.2.1.1 Berechnung der Wahrscheinlichkeiten $d_{m,n}(t)$.	132
4.3.2 Analyse mit Hilfe der Lindley'schen Integral- methode	75	5.2.1.2 Bestimmung der Zustandswahrscheinlichkeiten kurz vor den Taktzeitpunkten	135
4.3.2.1 Bestimmung der Restarbeit im System kurz vor den Taktzeitpunkten	75	5.2.1.3 Berechnung der charakteristischen Größen	144
4.3.2.2 Lösungsansatz durch Zeitdiskretisierung	78	5.2.2 Approximative Berechnung des Modells $C^{[x]}/G/1$ mit P Prioritätsklassen	148
4.3.2.3 Berechnung der erzeugenden Funktion $G_m(z)$...	83	5.2.2.1 Ein Erhaltungssatz für Prioritätssysteme	148
4.3.2.4 Zeitdiskretisierung verschiedener Bedienungs- VF	86	5.2.2.2 Gruppengrößen-Verteilung des Ersatzankunfts- prozesses	152
4.3.2.5 Lösung der charakteristischen Gleichung $z^l - G_m(z) = 0$	92	5.2.2.3 Bedienungs-VF des Ersatzbedienungsprozesses .	153
4.3.2.6 Berechnung der charakteristischen Größen	94	5.2.2.4 Algorithmus	155
4.3.3 Approximative Berechnung durch Rückführung auf das Modell des Typs GI/G/1	98	5.3 Das dritte Teilmodell mit zwei Prioritätsklassen	156
4.3.3.1 Der Ersatzankunftsprozeß GI	99	5.3.1 Berechnung der charakteristischen Größen bei Verwendung einer eingebetteten Markoff-Kette ..	157
4.3.3.2 Der Ersatzbedienungsprozeß G^*	100	5.3.2 Berechnung der Ausgangsprozesse	159
4.3.3.3 Berechnung des Angebotes ρ^*	100	5.4 Ergebnisdiskussion und Simulationsvergleich	161
4.3.3.4 Berechnung der charakteristischen Größen	101	5.4.1 Numerische Ergebnisse für das erste Teilmodell	161
4.4 Das dritte Teilmodell GO/C/ ∞	102	5.4.2 Numerische Ergebnisse für das zweite Teilmodell	162
4.4.1 Bestimmung der Zustandswahrscheinlichkeiten zu einem beliebigen Zeitpunkt bei Verwendung einer eingebetteten Markoff-Kette	103	5.4.3 Numerische Ergebnisse für das dritte Teilmodell	168
4.4.2 Berechnung der charakteristischen Größen	106	5.4.4 Mittlere Gesamtdurchlaufzeiten	169
4.4.3 Berechnung des Ausgangsprozesses $C_D^{[x]}$	107	6. ZUSAMMENFASSUNG	177
4.5 Ergebnisdiskussion und Simulationsvergleich	108	ANHANG 1 Erzeugende Funktion und ihre Eigenschaften	179
4.5.1 Numerische Ergebnisse für das erste Teilmodell GI/C/ ∞	108	ANHANG 2 Bestimmung der Übergangswahrscheinlichkeiten $q_{jx}(T)$	180
4.5.2 Numerische Ergebnisse für das zweite Teilmodell $C^{[x]}/G/1$	109	ANHANG 3 Bestimmung von Integralen für das 4. Kapitel	181
4.5.3 Numerische Ergebnisse für das dritte Teilmodell GO/C/ ∞	115	ANHANG 4 Bestimmung von Integralen für das 5. Kapitel	182
4.5.4 Mittlere Gesamtdurchlaufzeit	117		
5. ANALYSE DES DREISTUFIGEN WARTEMODELLS MIT TAKTGESTEUERTER EIN-/AUSGABE, VERWALTUNGSZEIT UND P PRIORITÄTSKLASSEN ...	125		
5.1 Struktur und Betriebsweise	125		
5.2 Das zweite Teilmodell $C^{[x]}/G/1$ mit P Prioritäts- klassen	128		
5.2.1 Analyse mit Hilfe einer eingebetteten Markoff- Kette kurz vor den Taktzeitpunkten (2 Priori- tätsklassen, exponentiell verteilte Bedienungs- zeiten)	131		

LITERATURVERZEICHNIS

- [1] Gerke, P.R. Rechnergesteuerte Vermittlungssysteme. Springer-Verlag, Berlin/Heidelberg/New York, 1972.
- [2] - Special Issues BSTJ on No. 1 ESS 43 (1964); No. 4 ESS, BSTJ 56 (1977), Common Channel Interoffice Signalling (CCIS), BSTJ 57 (1978).
- [3] Oden, H. Nachrichtenvermittlung. R. Oldenburg Verlag, München/Wien, 1975.
- [4] - System 12 - Das Digital-Vermittlungssystem. Technische Beschreibung. Standard Elektrik Lorenz, 1981.
- [5] Besier, H. Digitale Vermittlungstechnik. Heuer, P. R. Oldenburg Verlag, München/Wien, 1981. Kettler, G.
- [6] Gerke, P.R. Neue Kommunikationsnetze. Springer-Verlag, Berlin/Heidelberg/New York, 1982.
- [7] - Queuing Standardization Conference Report. May 11 (1971).
- [8] Herzog, U. Priority Models for Communication Processors Including System Overhead. 8. ITC, Melbourne (1976), Congressbook, S. 623/1-7.
- [9] Langenbach-Belz, M. Getaktete Wartesysteme bei Rechnern und zentralgesteuerten Nachrichtenvermittlungsanlagen. Dissertationsschrift, Universität Stuttgart, 1973.
- [10] Schwaertzel, H.G. Serving Strategies of Batch Arrivals in Common Control Switching Systems. NTZ 27 (1974)9, S. 341-345.
- [11] Weisschuh, H. Investigations on the Traffic Behavior of the Common Control in SPC Switching Systems. Wizgall, M. 8. ITC, Melbourne (1976), Congressbook, S. 612/1-8.
- [12] Weisschuh, H. Entwicklung der Steuersoftware für eine rechnergesteuerte Vermittlungsstelle. Dissertationsschrift, Universität Stuttgart, 1977.
- [13] Kühn, P. Analysis of Switching System Control Structures by Decomposition. 9. ITC, Torremolinos (1979).
- [14] Fujiki, M. Murao, Yo. Queuing Model with Regular Service Interruptions. 8. ITC, Melbourne (1976), Congressbook, S. 232/1-6.
- [15] Langenbach-Belz, M. Vergleich zweier Warteschlangenmodelle für Realzeit-Rechnersysteme mit Interrupt- bzw. Takt-gesteuerter Übernahme von Anforderungen aus der Peripherie. Lecture Notes in Computer Science, No. 1. Springer-Verlag, Berlin/Heidelberg/New York (1973), S. 304-313.
- [16] Jans, H. Kühn, P. A Priority Queueing Model with Synchronous Batch Arrivals and General Service Times. ORSA/TIMS Joint Nat. Meeting, Detroit (1982).
- [17] Kühn, P. Analyse zufallsabhängiger Prozesse in Systemen zur Nachrichtenvermittlung und Nachrichtenverarbeitung. Habilitationsschrift, Universität Stuttgart, 1981.
- [18] Kleinrock, L. Queuing Systems. Vol. I: Theory, Vol. II: Computer Applications. John Wiley & Sons, New York/London/Sydney/Toronto, 1975, 1976.
- [19] Cooper, R.B. Introduction to Queuing Theory. The Mac Millon Comp., New York/London, 1972.
- [20] Gross, D. Harris, C.M. Fundamentals of Queueing Theory. John Wiley & Sons, New York/London/Sydney/Toronto, 1974.
- [21] Cox, D.R. Miller, H.D. The Theory of Stochastic Processes. Chapman and Hall Ltd., London, 1965.
- [22] Cohen, J.W. The Single Server Queue. North-Holland, Amsterdam, 1982.
- [23] Kosten, L. Stochastic Theory of Service Systems. International Series of Monoqr. in Pure and Appl. Math., Vol. 103. Pergamon Press, Oxford/New York/Toronto/Sydney/Braunschweig, 1973.
- [24] Kendall, D.G. Stochastic Processes Occurring in the Theory of Queues and their Analysis by the Method of Imbedded Markov Chains. Ann. Math. Statistics 24 (1953), S. 338-354.

- [25] Cox, D.R. A Use Complex Probabilities in the Theory of Stochastic Processes. Proc. Cambr. Phil. Soc. 51 (1955), S. 313-319.
- [26] Cox, D.R. The Analysis of Non-Markovian Stochastic Processes by the Inclusion of Supplementary Variables. Proc. Cambr. Phil. Soc. 51 (1955), S. 433-441.
- [27] Lindley, D.V. The Theory of Queues with a Single Server. Proc. Cambr. Phil. Soc. 48 (1952), S. 277-289.
- [28] Cobham, A. Priority Assignment in Waiting Line Problems. Opns. Res. 2 (1954), S. 70-76.
- [29] Herzog, U. Verkehrsfluß in Datennetzen. Habilitationsschrift, Universität Stuttgart, 1973.
- [30] Kühn, P. Analysis of Complex Queuing Networks by Decomposition. 8. ITC, Melbourne (1976), Congressbook, S. 236/1-8.
- [31] Kühn, P. Approximate Analysis of General Queuing Networks by Decomposition. IEEE Transact. COM-27 (1979)1, S. 113-126.
- [32] Kühn, P. Über die Berechnung der Wartezeiten in Vermittlungs- und Rechnersystemen. Dissertationsschrift, Universität Stuttgart, 1972.
- [33] Kampe, G. Kühn, P. Graded Delay Systems with Infinite or Finite Source Traffic and Exponential or Constant Holding Time. Part I: Theory. AEÜ 32 (1978), S. 8-18. Part II: Numerical Results and Validation. AEÜ 32 (1978), S. 59-61.
- [34] Law, A.M. Kelton, W.D. Simulation Modeling and Analysis. Mc Graw-Hill Book Comp., New York, 1982.
- [35] Schäfer, U. Simulation von Wartenetzen mit getakteter Übergabe. Diplomarbeit Nr. 33, FG Nachrichtentechnik, Universität Siegen, 1982.
- [36] Little, J.D.C. A Proof of the Queueing Formula $L=\lambda W$. Opns. Res. 9 (1961), S. 383-387.
- [37] Tran-Gia, P. Jans, H. Clocked Event Transfer Protocol in Distributed Processing Systems - A Performance Analysis. 6. Int. Conf. on Comp. Com., London (1982), Congressbook, S. 769-774.

- [38] Dor, N.M. Guide to the Length of Buffer Storage Required for Random (Poisson) Input and Constant Output Rates. IEEE Transact. Electr. Comp. EC-16 (1967), S. 683-684.
- [39] Chu, W.W. Buffer Behavior for Poisson Arrivals and Multiple Synchronous Constant Outputs. IEEE Transact. Comp. C-19 (1970), S. 530-534.
- [40] Maritsas, D.G. Hartley, M.G. Buffer Length for Erlang-Input and Constant Removal Rate. IEEE Transact. Comp. C-19 (1970), S. 839-843.
- [41] Kühn, P. Tabellen für Wartesysteme. Institut für Nachrichtenvermittlung und Datenverarbeitung. Universität Stuttgart, 1976.
- [42] Cox, D.R. Erneuerungstheorie. R. Oldenburg Verlag, München/Wien, 1966.
- [43] Törnig, W. Numerische Mathematik für Ingenieure und Physiker, Band 1,2. Springer-Verlag, Berlin/Heidelberg/New York, 1979.
- [44] Burke, P.J. Delays in Single-Server Queues with Batch Input. Opns. Res. 23 (1975), S. 830-833.
- [45] Krämer, W. Langenbach-Belz, M. Approximate Formulae for the Delay in the Queuing System GI/G/1. 8. ITC, Melbourne (1976), Congressbook, S. 235/1-8.
- [46] Gradshteyn, I.S. Ryzhik, I.M. Table of Integrals, Series and Products. Academic Press, New York/London/Toronto/Sydney/San Francisco, 1980, S. 318, 948, 1058.
- [47] Gröbner, W. Hofreiter, N. Integraltafel, Teil II: Bestimmte Integrale. Springer-Verlag, Wien/New York, 1973, S. 11.
- [48] Herzog, U. Woo, L. Chandy, K.M. Solution of Queueing Problems by a Recursive Technique. IBM J. Res. Dev. 19 (1975)3, S. 295-300.
- [49] Conway, R.W. Maxwell, W.L. Miller, L.W. Theory of Scheduling. Addison-Wesley, Reading, Mass., 1967.

BEZEICHNUNGEN

Die folgende Zusammenstellung umfaßt nur die wichtigsten Bezeichnungen. Alle weiteren modellspezifischen Bezeichnungen werden in den einzelnen Kapiteln definiert. Bei der Betrachtung des mehrklassigen Modells im Kap. 5 kennzeichnet der Index i die Größen der i-ten Prioritätsklasse.

1. Parameter und Abkürzungen

- λ Mittlere Ankunftsrate
- ϵ, μ Mittlere Bedien- oder Enderaten
- ρ λ/μ
Angebot
- ZV Zufallsvariable
- VF Verteilungsfunktion
- VDF Verteilungsdichtefunktion
- LT Laplace-Transformation
- LST Laplace-Stieltjes-Transformation

2. Zufallsvariable, Verteilungen, Transformationen und Momente

- T_A Ankunftsabstand (interarrival time)
- T_H Bedienungsdauer (holding time)
- K Gruppengröße
- $A(t) = P\{T_A \leq t\}$
VF T_A der Ankunftsabstände
- $a(t) = dA(t)/dt$
VDF der Ankunftsabstände
- $H(t) = P\{T_H \leq t\}$
VF T_H der Bedienungsdauern
- $h(t) = dH(t)/dt$
VDF der Bedienungsdauern
- $\Phi_A(t) = LT\{a(t)\} = LST\{A(t)\}$
Laplace-Transformierte von $a(t)$ bzw. Laplace-Stieltjes-Transformierte von $A(t)$
- $\Phi_H(t) = LT\{h(t)\} = LST\{H(t)\}$
- G_k Erzeugende Funktion von g_k
- $t_A = E[T_A]$
Mittlerer Ankunftsabstand

- $h = E[T_H]$
Mittlere Bedienungsdauer
- $h^{(2)} = E[T_H^2]$
2-tes gewöhnliches Moment der ZV T_H

3. Charakteristische Größen

- Ω Mittlere Warteschlangenlänge
- w Mittlere Wartezeit (bezogen auf alle eintreffenden Anforderungen)
- t_f Mittlere Durchlaufzeit
- w, w_0 Wartewahrscheinlichkeiten

1. EINLEITUNG

1.1 Rechnergesteuerte Vermittlungssysteme

Seit der Inbetriebnahme des ersten serienmäßig gefertigten rechnergesteuerten Vermittlungssystems (Typ No. 1 ESS) im Jahre 1965 in Succasunna, New Jersey, USA, hat die Entwicklung solcher Systeme große Fortschritte erzielt. Erst die Bereitstellung leistungsfähiger Mikroprozessoren und hochintegrierter Schaltkreise ermöglicht einen wirtschaftlichen Einsatz rechnergesteuerter Vermittlungssysteme mit gespeichertem Programm (im engl. als stored program controlled, SPC, bezeichnet) und gestattet damit die Einführung spezieller Leistungsmerkmale, die nur bei SPC-Systemen mit vertretbarem Aufwand realisierbar sind.

Im folgenden werden die Aufgaben - unterteilt in Grundaufgaben bzw. spezielle Leistungsmerkmale - und die Steuerungsprinzipien rechnergesteuerter Vermittlungssysteme kurz dargestellt. Ferner wird auf Ein-/Ausgabe-Mechanismen und die prioritätsgesteuerte Bearbeitung der vermittlungstechnischen Funktionen näher eingegangen, da die Leistungsfähigkeit eines Systems hiervon stark beeinflusst wird.

1.1.1 Grundaufgaben eines Vermittlungssystems

Die Grundaufgaben sind solche, die unabhängig von der technischen Realisierung von jedem Vermittlungssystem durchgeführt werden müssen. Sie lassen sich grob einteilen in die Teilaufgaben

- Vermitteln
- Sichern
- Betreiben.

Unter Vermitteln versteht man den Aufbau, das Halten und den Abbau einer Verbindung. Das Vermittlungssystem stellt hierzu Vermittlungseinrichtungen zur Verfügung, die aufgrund von Steuerinformationen (Wählziffern, interne Signale) eine zeitweilige Ver-

bindung zwischen dem rufenden Teilnehmer A (Tln. A) und dem gerufenen Tln. B herstellen. Damit der rufende Tln. über den Zustand des Vermittlungsvorganges informiert wird, sendet das Vermittlungssystem Hörtöne aus. Diese sind:

- Wählton zur Wahlaufforderung
- Freiton (zeigt Frei-Zustand des Tln. B an)
- Belegt- oder Besetztton (zeigt Gassen-besetzt oder Teilnehmer-besetzt an).

Aus der Sicht des Vermittlungssystems müssen die folgenden Grundfunktionen ausgeführt und die entsprechenden Einrichtungen zur Verfügung gestellt werden:

- Erkennen des Verbindungswunsches (Tln. A hebt ab), z. B. durch wiederholte Abtastung des Teilnehmerzustandes.
- Wahlaufforderung durch Senden des Wähltons und Bereitstellung von Einrichtungen zur Erkennung, Aufnahme und Verarbeitung der Wählziffern (Zielinformation).
- Weiterleitung der in der Ursprungs-Vermittlungsstelle nicht verwendeten Wählziffern zu anderen Vermittlungsstellen bei Extern- bzw. Fernverkehr mit Hilfe von Signalisierungseinrichtungen.
- Sprechwegedurchschaltung im Durchschaltenetz der Vermittlungsstelle des Tln. A.
- Anschalten von Frei-/Belegtton.
- Auslösung der Verbindung (Tln. A und/oder Tln. B legt auf).
- Gebührenerfassung.

In der Regel steuert ein Vermittlungssystem gleichzeitig eine große Anzahl von Verbindungen, die sich aufgrund ihres zufälligen Beginns in unterschiedlichen Phasen befinden. Die Verarbeitung der verschiedenen Steuerungssignale, die von den aktiven Verbindungen erzeugt werden und in ihrer zeitlichen Aufeinanderfolge statistisch verteilt eintreffen, unterliegt teilweise strengen Realzeit-Bedingungen. Zur Einhaltung dieser Forderungen müssen die Systeme sorgfältig geplant, verschiedene Entwurfskonzepte

modellmäßig auf ihre Leistungsfähigkeit hin untersucht und gegeneinander abgewogen und schließlich das real aufgebaute System in der Praxis überwacht und wenn erforderlich modifiziert werden.

Unter Sichern versteht man alle Maßnahmen, die zum automatischen Aufrechterhalten der Betriebsbereitschaft beitragen. Hierzu gehören u. a.

- Fehlererkennung
- Fehlerlokalisierung
- Automatische Fehlerbeseitigung oder Fehlerabschaltung
- Ersatzschaltung
- Wiederanlauf abgeschalteter Teilsysteme.

Auf die Problematik, die sich durch die geforderte hohe Verfügbarkeit (ca. 99.992 % oder 2 Stunden Totalausfall in 30 Jahren) ergibt und die in der Praxis verwendeten Methoden zur Erhöhung der Betriebssicherheit, wie z. B. duplizieren und Einfügen von Redundanz, soll hier nicht näher eingegangen werden. In jedem Falle stellen alle Sicherungsmaßnahmen eine zusätzliche Belastung für die Steuerung des Vermittlungssystems dar.

Das Betreiben umfaßt z. B. die Aufgaben

- Vergabe von Teilnehmerberechtigungen
- Gebührenerfassung
- Messung von Verkehr und Dienstgüte
- Planung und Dimensionierung.

Hinzu kommen weitere Aufgaben, die sich teilweise mit den Prüf- und Wartungsaufgaben überschneiden.

1.1.2 Spezielle Leistungsmerkmale rechnergesteuerter Vermittlungssysteme

Erst durch den Einsatz rechnergesteuerter Vermittlungssysteme können spezielle Leistungsmerkmale (Facilities) allen Teilnehmern wirtschaftlich angeboten werden. Die dem Teilnehmer zugute kommenden Leistungen sind z. B.:

- Kurzwahl oder Verbindung ohne Wahl
- Anrufumleitung, Rufweiserschaltung oder Anrufwiederholung
- Ruhe vor dem Telefon, "Anklopfen" während des Gesprächs, Warten auf Freiwerden des Tln. B, Gebührenübernahme durch Tln. B oder Sperre für abgehende Verkehrsarten (z. B. Ferngespräche).

Die Aufzählung erhebt keinen Anspruch auf Vollständigkeit. Es sei besonders vermerkt, daß die Maßnahmen eine zusätzliche Kommunikation zwischen Teilnehmer und Vermittlungssystem erfordern, da der Teilnehmer per Tastenwahl auf spezielle Speicher des Vermittlungssystems zugreifen kann und die von ihm gewünschten Leistungen kennzeichnet bzw. ändert. Erst ein rechnergesteuertes System kann die hierzu notwendigen Einrichtungen mit vertretbarem Aufwand zur Verfügung stellen. Die neue Technik bietet jedoch auch dem Betreiber wesentliche Vorteile wie Betriebserleichterungen für die Verwaltung, Rationalisierung der Wartung, des Betriebs und der Erfassung von Störungen bis hin zur automatischen Ersatzschaltung von Funktionseinheiten.

1.1.3 Steuerungsprinzipien

Alle zum Auf- und Abbau von Verbindungen eintreffenden Signale müssen von der Steuerung eines Vermittlungssystems aufgenommen, verarbeitet und wenn erforderlich, entsprechende Befehle generiert und weitergeleitet werden. Abhängig von der Systemarchitektur kann die Intelligenz (Funktionen, Daten) unterschiedlich innerhalb des Systems verteilt werden. Hieraus ergeben sich prinzipiell drei verschiedene Steuerungsprinzipien:

- Zentrale Steuerung
- Dezentrale Steuerung
- Teilzentrale oder gemischt zentral/dezentrale Steuerung.

Im folgenden werden die wichtigsten Eigenschaften der verschiedenen Steuerungsprinzipien kurz diskutiert:

a) Zentrale Steuerung

Bei zentraler Steuerung werden alle vermittlungstechnischen Funktionen durch ein oder wenige, gleichartige Universal-Steuerwerke durchgeführt. Dies bedeutet, daß sich alle Vermittlungsprogramme und Zustandsspeicher (Abbild der aktuellen Zustände der Teilnehmer und der übrigen peripheren Einrichtungen) im zentralen Speicher befinden müssen. Dieses Prinzip erfordert wegen des intensiven Steuerdatentransports zwischen Peripherie und zentraler Steuerung (Signale, Befehle) einen leistungsfähigen Ein-/Ausgabe-Mechanismus. Weitere Nachteile sind die großen Speicher- und Rechnerkapazitäten, die notwendige Duplizierung der gesamten Steuerung und die kapazitätsmäßige Konzipierung für einen möglichen Endausbau. Der Vorteil ist jedoch ein konfliktfreier Zugriff auf alle zentralisierten, vermittlungstechnischen Einrichtungen.

b) Dezentrale Steuerung

Dieses Steuerungsprinzip liegt der klassischen, direkt gesteuerten Vermittlungstechnik zugrunde [1,2,3]. Die einzelnen Betriebsmittel besitzen eigene Steuereinrichtungen, in denen ihr Zustand abgespeichert ist und die untereinander direkt Steuerinformationen zum Auf- und Abbau der Verbindung austauschen können. Durch die Entwicklung von Mikrorechnern und anderen hochintegrierten Schaltkreisen wurde es möglich, den Zentralisierungsgrad bei rechnergesteuerten Vermittlungssystemen zu verringern und die Intelligenz entsprechend den Steuerungsaufgaben auf mehrere Teilsteuerungen nach dem Funktions- oder Lastteilungsprinzip (function sharing, load sharing) zu verteilen. Die Dezentralisierung erlaubt damit einen besseren modularen Aufbau in Hardware und Software. Durch die Festlegung von Schnittstellen ist eine Flexibilität gegenüber technologischen Fortschritten gewährleistet. Weitere Vorteile gegenüber der zentralen Steuerung sind die Zuverlässigkeit, die geringere Wirkungsbreite bei Teilausfällen und die bedarfsmäßige Ausbaufähigkeit durch Hinzufügen weiterer Module. Nachteilig ist die gegenseitige Blockierung während des Informationsaustausches bei steigender Anzahl von Prozessoren, die an einem gemeinsamen Bussystem angeschlossen sind. Da an der Abwicklung einer Verbin-

dung mehrere Module beteiligt sind, ist eine leistungsfähige Interprozessor-Kommunikation für das hohe Steuerdaten-Verkehrsvolumen zwingend notwendig.

c) Teilzentrale oder gemischt zentral/dezentrale Steuerung

Diese Mischform zwischen zentraler und dezentraler Steuerung ist das Ergebnis von Erfahrungen in der Entwicklung und im praktischen Einsatz von rechnergesteuerten Vermittlungssystemen. Man versucht die Vorteile beider Steuerungsprinzipien zu nutzen. Durch die Verwendung peripherer Gruppen-Steuerwerke können leitungsnahe (insbesondere mit der Signalisierung zusammenhängende) Teilfunktionen wie z. B. Teilnehmerzustands- bzw. Leitungszustandsüberwachung und Verwaltung von peripheren Betriebsmitteln (Wahlsätze, Konzentrationskoppelnetz) eigenständig dezentral durchgeführt werden. Die hierdurch entlastete zentrale Steuerung übernimmt die übergeordneten rufbezogenen Funktionen, semipermanente Daten sowie die Verwaltung zentraler Betriebsmittel (z. B. zentrales Durchschaltenetz). Der Datenaustausch zwischen den einzelnen dezentralen Steuerwerken und der zentralen Steuerung erfolgt z. B. über ein gemeinsames Bussystem oder über das ohnehin vorhandene Durchschaltenetz mittels fest zugeordneter bzw. vermittelter Kanäle [4]. Die im Bild 1.1 dargestellte Grobstruktur eines gemischt zentral/dezentral gesteuerten Vermittlungssystems mit gemeinsamen Bussystem, die in den meisten bestehenden Systemen Anwendung findet, liegt den folgenden Betrachtungen und den in dieser Arbeit durchgeführten verkehrstheoretischen Untersuchungen zugrunde. Zur weiteren Diskussion siehe auch [5,6].

1.1.4 Interprozessor-Kommunikation und Ein-/Ausgabe-Mechanismen

Der Umfang des Steuerdatenaustausches zwischen den verteilten Steuermodulen (Interprozessor-Kommunikation) ist direkt abhängig vom Zentralisierungsgrad und bestimmt maßgeblich die Leistungsfähigkeit des Systems. Bei der dezentralen Steuerung erfolgt der Datenaustausch z. B. über ein gemeinsames Bussystem mittels eines Übertragungs-Protokolls unter Festlegung einer ge-

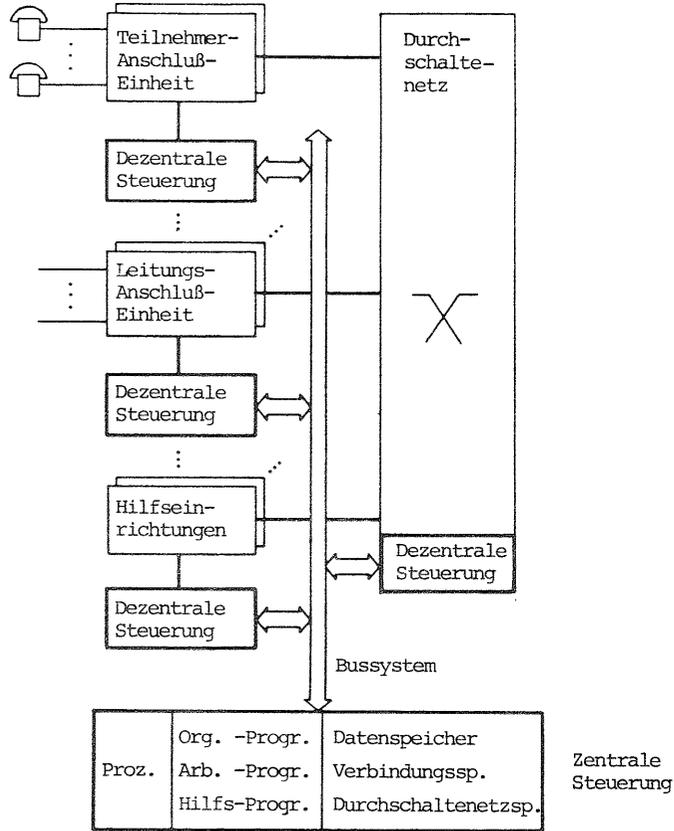


Bild 1.1. Grobstruktur eines gemischt zentral/dezentral gesteuerten Vermittlungssystems

eigneten Zuteilungsmethode. Dagegen wird bei der gemischt zentral/dezentralen Struktur sowohl der Zugriff auf das Bussystem als auch der Datentransfer häufig von der zentralen Steuerung übernommen und benötigt einen nicht zu vernachlässigenden Verwaltungsaufwand (Overhead). Für die Ein-/Ausgabe zwischen der zentralen Steuerung und den dezentralen Steuerungen stehen hierbei prinzipiell zwei verschiedene Mechanismen zur Verfügung:

- ereignisgesteuerte (interruptgesteuerte) Ein-/Ausgabe
- programmgesteuerte Ein-/Ausgabe.

Bei der interruptgesteuerten Methode erzeugt jede Einzelmeldung einen Interrupt, der das laufende Programm direkt unterbricht und die Steuereinheit durch Aktivierung des Ein-/Ausgabeprogrammes zur Übernahme veranlaßt. Dieses Verfahren gestattet zwar eine quasi-simultane Bedienung verschiedener Ein-/Ausgabe-Schnittstellen, jedoch steigt der Verwaltungsaufwand mit steigender Belastung (Datenrate) so stark an, daß die Wartezeiten der zu bearbeitenden Anforderungen sehr ungünstig beeinflusst werden.

Bei der programmgesteuerten Ein-/Ausgabe geht die gesamte Initiative von der Steuereinheit (Rechner) aus, die den Datentransfer aktiviert. Sie veranlaßt per Programm von sich aus die Übernahme von Anforderungen bzw. die Ausgabe von Befehlen. Die zeitlichen Abstände der "Abfragezeitpunkte" sind bei bestehenden Systemen häufig konstant, die Aktivierung zum Transfer erfolgt in der Regel hardwaremäßig mit Hilfe einer Realzeituhr. Dieses Verfahren wird deshalb auch als taktgesteuerte Ein-/Ausgabe bezeichnet. Der Vorteil des Verfahrens besteht darin, daß der Verwaltungsaufwand nahezu lastunabhängig ist, da das Ein-/Ausgabeprogramm nur zu den Taktzeitpunkten aufgerufen wird. Alle eintreffenden Anforderungen innerhalb einer Taktperiode müssen jedoch bis zum Ende des laufenden Taktintervalls zwischengespeichert werden.

Das Ablaufgeschehen für die taktgesteuerte Ein-/Ausgabe unter Einbeziehung von Prioritätsklassen für die Bearbeitung der unterschiedlichen Anforderungen durch die zentrale Steuereinheit zeigt das Bild 1.3 im nächsten Abschnitt.

1.1.5 Verkehrstheoretische Aspekte der zentralen Steuereinheit

Alle Steuerungsaufgaben in Vermittlungssystemen sind ihrer Natur nach Realzeit-Anforderungen und müssen deshalb mehr oder weniger dringend bearbeitet werden, um vorgeschriebene Reaktionszeiten einhalten zu können. Diese Forderung verlangt neben der beschriebenen Hardware-Modularisierung eine geeignete Software-Struktur für Vermittlungs- und Betriebsaufgaben sowohl im dezentralen als auch im zentralen Steuerungsbereich. Bei der im Bild 1.1 gegebenen

Struktur stellt die Interprozessor-Kommunikation zwischen der zentralen Steuerung und dezentralen Steuerungen über das gemeinsame Bussystem und die Bearbeitung der übergeordneten Steuerungsfunktionen durch die zentrale Steuerung den Systemengpaß dar und bestimmt damit maßgeblich die Leistungsfähigkeit des gesamten Systems. Hieraus ergibt sich die Notwendigkeit neben einer leistungsfähigen Ein-/Ausgabe der Anforderungen von und zur zentralen Steuerung eine prioritätsgesteuerte Abfertigungsstrategie für die unterschiedlichen Anforderungsklassen einzurichten.

Die Gesamtheit der erforderlichen Programme, die als "Systemsoftware" bezeichnet wird, läßt sich grob in drei Klassen einteilen:

- Organisationsprogramme
- Arbeitsprogramme
- Hilfsprogramme.

Die Organisationsprogramme ("Betriebssystem") steuern und überwachen den Ablauf der Arbeitsprogramme, die die eigentlichen Vermittlungsaufgaben durchführen. Die Hilfsprogramme stehen hierbei für wiederkehrende Funktionen zur Verfügung. Die Organisationsprogramme, denen hardwaremäßig verschiedene Prioritäten zugeteilt sind, lassen sich prinzipiell in drei Klassen einteilen:

1. Fehler-/Alarmbehandlungsprogramme
2. Ein-/Ausgabeprogramme für die Interprozessor-Kommunikation
3. Vermittlungsprogramme.

Die Klasse 1 hat die höchste Priorität. Der Aufruf der einzelnen Organisationsprogramme erfolgt hardwaremäßig mit Hilfe eines Unterbrechungssystems grundsätzlich durch Unterbrechungssignale. Während die Fehler-/Alarmbehandlungsprogramme wegen ihrer Dringlichkeit interruptgesteuert aufgerufen werden, erfolgt der Aufruf der Ein-/Ausgabeprogramme für die Interprozessor-Kommunikation zwischen den dezentralen und zentralen Steuereinheiten taktgesteuert, insbesondere bei Systemen mit größeren Datenraten. Den zeitlichen Programmablauf unter den genannten Voraussetzungen zeigt das Bild 1.2.

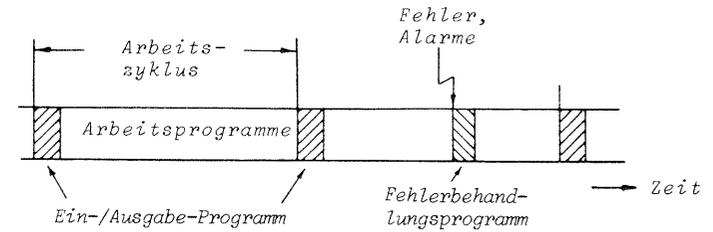


Bild 1.2. Zeitlicher Programmablauf der zentralen Steuereinheit

Die Anzahl der eintreffenden Fehler-/Alarmmeldungen, d. h. die Ankunftsrate der höchsten interruptgesteuerten Prioritätsklasse, ist voraussetzungsgemäß so niedrig, daß sie für die verkehrstheoretische Untersuchung der zentralen Steuereinheit unbedeutend ist. Berücksichtigt man nun noch, daß die Vermittlungsprogramme ihrerseits verschiedenen Prioritätsklassen zugeordnet sind, so ergibt sich z. B. der folgende zeitliche Verlauf für den Systemzustand bei taktgesteuerter Ein-/Ausgabe und Bearbeitung nach unterbrechenden Prioritäten. X gibt die zufällige Anzahl von Anforderungen im System an.

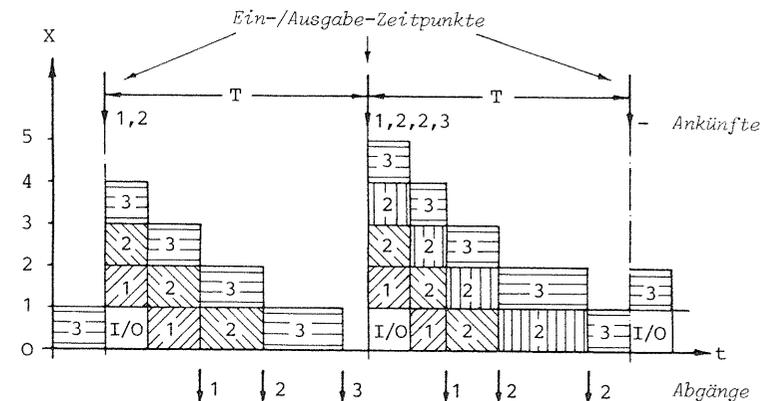


Bild 1.3. Zeitlicher Verlauf des Systemzustandes bei taktgesteuerter Ein-/Ausgabe und Bearbeitung nach unterbrechenden Prioritäten

Den im Bild 1.3 angenommenen drei Prioritätsklassen könnten z. B. die folgenden vermittlungstechnischen Funktionen zugeordnet werden:

1. Rufverarbeitung (Wählziffern, Signalisierung, zeitkritische Signale)
2. Rufannahme
3. Auslösen.

Die Bearbeitung der Programme nach Beendigung des Ein-/Ausgabe-Programms erfolgt hier nach unterbrechenden Prioritäten, d. h. die Bearbeitung eines durch die Ein-/Ausgabe unterbrochenen Programms wird erst dann fortgesetzt, wenn sich keine Anforderungen höherer Priorität mehr im System befinden.

Im Abschnitt 2.3 werden für die interrupt- und taktgesteuerte Ein-/Ausgabe Warteschlangenmodelle unter Einbeziehung von Prioritäten vorgestellt und leistungsmäßig miteinander verglichen. Es wird gezeigt, daß die taktgesteuerte Ein-/Ausgabe der interrupt-gesteuerten Methode bei höheren Datenraten (Verkehrsangebot) überlegen ist. Bei den im Rahmen dieser Arbeit untersuchten Modellen wird stets die taktgesteuerte Ein-/Ausgabe vorausgesetzt.

1.1.6 Problemstellungen bei der verkehrstheoretischen Untersuchung und Leistungsbewertung

Hat man bei der Entwicklung eines rechnergesteuerten Vermittlungssystems die Hard- und Softwarestruktur festgelegt, so erfolgt die Untersuchung des gesamten Systems nach verkehrstheoretischen Gesichtspunkten. Dazu werden die einzelnen Funktions- oder Steuereinheiten hinsichtlich der

- Belastung
- Wartezeiten
- Durchlaufzeiten
- Verluste usw.

untersucht. Mit Hilfe der Methoden der Verkehrstheorie können diese Kenngrößen durch Abbildung der Teilsysteme auf Modelle und durch

Beschreibung des zeitlichen Systemverhaltens mittels stochastischer Prozesse ermittelt werden. Hierdurch lassen sich frühzeitig verschiedene Entwicklungskonzepte miteinander vergleichen, Systemengpässe ermitteln und Fehlplanungen weitgehend vermeiden. Auf die Problematik der Modellbildung und auf die Methoden der Modellanalyse wird in Kap. 2 und 3 näher eingegangen.

1.2 Übersicht über die Arbeit

Anschließend an die Einführung über rechnergesteuerte Vermittlungssysteme wird im 2. Kapitel versucht, einen prinzipiellen Überblick über die verkehrstheoretische Modellbildung zu geben. Neben der Systematik werden die Komponenten verkehrstheoretischer Modelle näher beschrieben. Weiterhin werden Modelle zur Nachbildung der unterschiedlichen Ein-/Ausgabe-Prinzipien vorgestellt, die in den Kap. 4 und 5 unter allgemeinen Voraussetzungen analysiert werden.

Das 3. Kapitel befaßt sich mit den Methoden der Modellanalyse. Es werden vorwiegend analytische und simulative Methoden diskutiert, die in Kap. 4 und 5 Anwendung finden. Hier sind insbesondere die Methode der eingebetteten Markoff-Kette, die Lindley'sche Integralmethode (beides Analyseverfahren Nichtmarkoff'scher Prozesse) und approximative analytische Methoden zu nennen.

Kap. 4 befaßt sich mit der Analyse eines dreistufigen Wartemodells mit taktgesteuerter Ein-/Ausgabe, Verwaltungszeit und einer Anforderungsklasse. Ausgehend von der Struktur und Betriebsweise wird das Gesamtmodell in drei Teilmodelle zerlegt, die nacheinander analysiert werden. Den Schwerpunkt des Kapitels stellt die Untersuchung des zweiten Teilmodells dar, für das drei Analyseverfahren vorgeschlagen werden. Die einzelnen Verfahren werden systematisch entwickelt; der Genauigkeitsnachweis der approximativ berechneten Größen erfolgt durch Simulationsvergleiche am Ende des Kapitels.

Im Kap. 5 wird ein dreistufiges Wartemodell mit taktgesteuerter

Ein-/Ausgabe und Verwaltungszeit untersucht, in dem nun mehrere Anforderungsklassen mit unterschiedlicher Dringlichkeit (Prioritäten) berücksichtigt werden. Auch hier wird das Gesamtmodell zunächst in drei Teilmodelle zerlegt. Für das zweite Teilmodell wird unter einschränkenden Voraussetzungen eine Analyse mit Hilfe einer eingebetteten Markoff-Kette vorgestellt. Aufgrund des hierbei bereits erforderlichen hohen numerischen Aufwands wird ein approximatives Berechnungsverfahren entwickelt, welches das mehrklassige Modell unter allgemeinen Voraussetzungen auf mehrere einklassige Modelle nach Kap. 4 zurückführt. Mit Hilfe eines Simulationsprogramms, das das gesamte Wartemodell unter Einbeziehung aller gegenseitigen Abhängigkeiten nachbildet, wird die Güte der Approximation gezeigt. Bei der Ergebnisdiskussion wird die Durchlaufzeit für die Anforderungen der einzelnen Klassen in Abhängigkeit von den wichtigsten Systemparametern berechnet und graphisch dargestellt. Sie stellt ein wichtiges Maß für die Leistungsbewertung der Steuereinheit eines rechnergesteuerten Vermittlungssystems dar.

Die im Rahmen dieser Arbeit entwickelten und analysierten Modelle können auch zur Nachbildung und Verkehrsanalyse des Systemverhaltens von Realzeit-Steuerungen in der Prozeßdatenverarbeitung oder bei Untersuchungen des Operations Research verwendet werden. Insbesondere können sie auch zur Analyse eines verteilten Rechnersystems, das nach dem "Mailbox-Prinzip" arbeitet, herangezogen werden. Hierunter versteht man ein System, bei dem jeder angeschlossene Rechner autonom arbeitet, während die Kommunikation untereinander über "Ein- und Ausgabebriefkästen" (mail boxes), die in den betrachteten Modellen den Ein- bzw. Ausgabewarteschlangen entsprechen, abgewickelt wird.

2. MODELLBILDUNG UND PRINZIPIELLE ERGEBNISSE

Die Modellbildung spielt bei der Untersuchung komplexer Vorgänge in vielen Bereichen der Naturwissenschaften eine entscheidende Rolle. Modelle sind dabei eine abstrakte Darstellung des realistischen Systemverhaltens. Alle logischen und zeitlichen Zusammenhänge realer Systeme werden auf äquivalente Vorgänge innerhalb der Modelle abgebildet. Abhängig von der Modellierungstiefe, d. h. von der Berücksichtigung der einzelnen Parameter des realen Systems, haben die aufgrund der Modellanalyse gewonnenen Ergebnisse einen begrenzten Gültigkeitsbereich.

In Naturwissenschaften, Technik und Ökonomie, aber auch im Transportwesen treten in großem Umfang Bedienungssituationen auf, bei denen zufällige Einflüsse eine Rolle spielen. Bei der Untersuchung derartiger Fragestellungen verwendet man Methoden der Bedienungstheorie (Verkehrstheorie). Aufgabe der Bedienungstheorie ist es, mit Hilfe verkehrstheoretischer Modelle, die Struktur und Organisationsform des jeweiligen Bedienungssystems zu optimieren.

In diesem Kapitel werden nach einer kurzen Darstellung der Systematik verkehrstheoretischer Modelle deren Komponenten näher erläutert und anschließend solche Modelle betrachtet, die im Rahmen dieser Arbeit untersucht werden.

2.1 Systematik verkehrstheoretischer Modelle

Eine systematische Klassifizierung verkehrstheoretischer Modelle ist prinzipiell möglich, wenn man die unterschiedlichen Modellkomponenten des gesamten Bedienungssystems näher betrachtet.

2.1.1 Klassifizierung von Bedienungssystemen und Kurznotation

Bedienungssysteme und die zugehörigen verkehrstheoretischen Modelle - Warteschlangenmodelle - sind charakterisiert durch

- den Ankunfts- und Bedienungsprozess (Endeprozess)
- die Struktur des Bedienungssystems
- die Organisation des Bedienungssystems.

Diese Modellkomponenten werden im Abschnitt 2.2 näher betrachtet.

Die Charakterisierung eines einstufigen Bedienungssystems mit n Bedienungseinheiten erfolgt in der Literatur nahezu einheitlich mit einer erweiterten Form der von Kendall [7] eingeführten Kurznotation:

$$GI^{[x]} / G / n - s .$$

Hierin bedeuten:

- GI Typ des Ankunftsprozesses
(GI $\hat{=}$ general independent, d. h. rekurrenter Ankunftsprozess mit allgemein verteilten und voneinander unabhängigen Ankunftsabständen)
- x Kennzeichnung eines Gruppenankunftsprozesses
- G Typ des Bedienungsprozesses
(G $\hat{=}$ general, d. h. allgemeine Bedienungsdauern)
- n Anzahl der Bedienungseinheiten
- s Anzahl der Warteplätze (für $s \rightarrow \infty$ entfällt diese Angabe)

Die folgenden, speziellen Typen für den Ankunfts- und/oder Bedienungsprozess werden in der Verkehrstheorie häufig verwendet:

- D Konstante Ankunftsabstände bzw. Bedienungsdauern
(Deterministischer Prozess)
- E_k Erlang-verteilte Ankunftsabstände bzw. Bedienungsdauern
 k -ter Ordnung
- M Exponentiell verteilte Ankunftsabstände bzw. Bedienungsdauern
(Markoff'scher Prozess)
- H_k Hyperexponentiell verteilte Ankunftsabstände bzw. Bedienungsdauern k -ter Ordnung
- C Taktmäßige Ankünfte bzw. Abfertigung.

2.2 Komponenten eines verkehrstheoretischen Modells

Die wesentlichen Komponenten eines verkehrstheoretischen Modells und deren Beschreibung durch Methoden der Wahrscheinlichkeitstheorie sollen im folgenden kurz erläutert werden.

2.2.1 Der Ankunftsprozess

Der Ankunftsprozess beschreibt die Verteilung der Ankunftsabstände von Nachrichten bzw. Befehlen (im folgenden Anforderungen genannt). Im allgemeinen Falle des rekurrenten Ankunftsprozesses sind die Ankunftsabstände T_A (interarrival times) unabhängige Zufallsvariable mit der allgemeinen Verteilungsfunktion (VF)

$$A(t) = P\{T_A \leq t\} . \quad (2.1)$$

Die Verteilungsfunktion $A(t)$ ist definiert als die Wahrscheinlichkeit, daß der Ankunftsabstand T_A höchstens gleich der beliebigen Zeit t ist.

Bezeichnet man den mittleren Ankunftsabstand mit t_A , so wird der Kehrwert λ als mittlere Ankunftsrate definiert:

$$\lambda = \frac{1}{t_A} . \quad (2.2)$$

Die mittlere Ankunftsrate entspricht der mittleren Anzahl von Anforderungen pro Zeiteinheit (Intensität).

2.2.2 Der Bedienungs- oder Endeprozess

Der Bedienungs- oder Endeprozess beschreibt die Verteilung der Bedienungszeiten T_H (Bedienungsdauern, service times oder holding times) der einzelnen Anforderungen. Analog zu Gl. (2.1) gilt für voneinander unabhängige Belegungszeiten T_H die VF

$$H(t) = P\{T_H \leq t\} . \quad (2.3)$$

Bezeichnet man die mittlere Belegungsdauer mit h , dann entspricht der Kehrwert der mittleren Bedienrate μ :

$$\mu = \frac{1}{h} . \tag{2.4}$$

Die mittlere Anzahl von eintreffenden Anforderungen während der mittleren Belegungsdauer wird als Angebot ρ bezeichnet. Das Angebot ist also aus der mittleren Ankunftsrate λ und der mittleren Belegungsdauer h wie folgt berechenbar:

$$\rho = \lambda \cdot h = \frac{\lambda}{\mu} . \tag{2.5}$$

Die Markoff'schen Prozesse für den Ankunfts- und/oder Bedienungsprozeß, d. h. exponentiell verteilte Ankunftsabstände und/oder Bedienungsdauern, spielen in der Verkehrstheorie eine wichtige Rolle. Die Analyse Markoff'scher Prozesse wird deshalb im Abschnitt 3.1.1.2 genauer beschrieben.

2.2.3 Gruppenankunft bzw. Gruppenbedienung

Die oben beschriebenen Ankunfts- und Endeprozesse sind dadurch gekennzeichnet, daß innerhalb eines infinitesimal kleinen Zeitabschnittes dt höchstens ein Ankunfts- oder Endeereignis auftreten kann. Treffen in einem Bedienungssystem gleichzeitig mehrere Anforderungen ein, dann liegt ein Gruppenankunftsprozeß vor. Ein Gruppenbedienungsprozeß ist gekennzeichnet durch den gleichzeitigen Bedienungsverlauf mehrerer ankommender oder wartender Anforderungen.

Der Ankunftsabstand zwischen den Gruppen ist meistens wie bei Systemen mit Einzelankünften verteilt (M,D usw.).

Die Verteilung der Gruppengrößen ist zusätzlich anzugeben. Ist K eine diskrete Zufallsvariable, dann ist

$$g_k = P\{K = k\} \tag{2.6}$$

die Wahrscheinlichkeit, daß die ankommende Gruppe genau aus $K = k$ Anforderungen besteht, $k = 0, 1, 2, \dots$

2.2.4 Die Struktur des Bedienungssystems

Entsprechend der Struktur eines Bedienungssystems unterscheidet man Verlust-, Warte- und kombinierte Warte-Verlustsysteme. Beim Verlustsystem stehen keine ($s = 0$), beim Wartesystem unendlich viele ($s \rightarrow \infty$) und beim kombinierten Warte-Verlustsystem eine begrenzte Anzahl von Warteplätzen zur Verfügung.

Für die Untersuchung komplexer Bedienungssysteme, z. B. Systeme mit verteilter Steuerung, benötigt man häufig Modelle mit mehrstufiger Bedienung (Netzmodelle). Netzmodelle entstehen durch die Zusammenschaltung von Bedienungssystemen mit einstufiger Bedienung, bei denen die eintreffenden Anforderungen eine Bedieneinheit durchlaufen und dann das Teilsystem verlassen.

Man unterscheidet grundsätzlich zwei Typen von Netzmodellen:

- offene Warteschlangennetze
- geschlossene Warteschlangennetze.

Offene Netze haben, im Gegensatz zu geschlossenen Netzen, wenigstens einen Ein- und Ausgang zur "Außenwelt". Zur Beschreibung von Warteschlangennetzen benötigt man außer den Angaben über die Zahl der Netzknoten, der Ankunfts- und Bedienungsprozesse, Zuteilungs-, Abfertigungs- und Prioritätsstrategien auch Angaben über die Art der "Wegelenkung" (routing) der einzelnen Anforderungen innerhalb des Netzes. Die Wegelenkung kann fest, alternativ, adaptiv oder probabilistisch durchgeführt werden. Die probabilistische Wegelenkung erfolgt durch Vorgabe unabhängiger Wahrscheinlichkeiten mit Hilfe einer Verzweigungsmatrix \underline{Q} .

$\underline{Q} = (q_{ij})$ mit q_{ij} = Wahrscheinlichkeit für den Übergang einer Anforderung von Knoten i nach Knoten j , $i = 1, 2, \dots, N$, $j = 0, 1, \dots, N$ (Knoten "0" entspricht der "Außenwelt").

2.2.5 Die Organisation von Bedienungssystemen

Die Organisation des Bedienungssystems legt die Art der Einordnung neuer eintreffender Anforderungen in die Warteschlangen und die Abfertigung eintreffender und wartender Anforderungen durch die Bedienungseinheit fest. Im einzelnen lassen sich unterscheiden

- Warteschlangen-Zuteilungsstrategien
- Warteschlangen-Organisation
- Prioritätsorganisation.

a) Warteschlangen-Zuteilungsstrategien

Für die Zuteilung einer Warteschlange kommen unter anderem folgende Strategien in Betracht:

- Auswahl einer Warteschlange bei Rufankunft (queue allocation), z. B. Auswahl der Warteschlange entsprechend der Priorität der eintreffenden Anforderung.
- Prüfung des Zugangs zur ausgewählten Warteschlange (queue access), z. B. Blockierung einer Anforderung bei voll belegter Warteschlange, systemabhängige Zuweisung einer anderen Warteschlange (Überlaststeuerung).
- Auswahl einer Warteschlange nach Bedienungsende (interqueue discipline), z. B. Auswahl der Warteschlange mit höchster Priorität, zyklische Auswahl oder systemzustandsabhängige Auswahl.

b) Warteschlangen-Organisation

Die Warteschlangenorganisation beschreibt die Art und Weise der Abfertigung von wartenden Anforderungen (queue discipline). Die wichtigsten Prinzipien sind:

- FIFO first in, first out, d. h. die zuerst eingetroffene Anforderung wird zuerst bedient
- LIFO last in, first out, d. h. die zuletzt eingetroffene Anforderung wird zuerst bedient
- RANDOM die wartenden Anforderungen werden zufällig ausgesucht und bedient, unabhängig von der Ankunftsreihenfolge.

c) Prioritäts-Organisation

Aufgrund der unterschiedlichen Dringlichkeiten von Anforderungen werden diese in Prioritätsklassen eingeteilt (externe Prioritäten). Bei der Abfertigung mit Prioritäten unterscheidet man zwischen

- Unterbrechenden Prioritäten (preemptive priority)
- Nichtunterbrechenden Prioritäten (nonpreemptive priority).

Bei unterbrechenden Prioritäten wird eine in Bedienung befindliche Anforderung beim Eintreffen einer Anforderung höherer Priorität sofort unterbrochen. Die unterbrochene Anforderung wird erst dann weiter bedient, wenn keine Anforderungen höherer Priorität mehr im System sind.

Bei nichtunterbrechenden Prioritäten kann eine in Bedienung befindliche Anforderung durch das Eintreffen einer Anforderung höherer Priorität nicht unterbrochen werden. Nach jedem Bedienungsende erfolgt die Abfertigung der Anforderung mit der höchsten Priorität.

In realen Systemen benutzt man eine Kombination aus unterbrechenden und nichtunterbrechenden Prioritäten, die sog. gemischten Prioritäten.

2.3 Modellierung unterschiedlicher Ein-/Ausgabe-Mechanismen

Wie bereits in Kap. 1 beschrieben, unterscheidet man prinzipiell zwischen den beiden Ein-/Ausgabe-Mechanismen

- Interruptgesteuerte Ein-/Ausgabe
- Taktgesteuerte Ein-/Ausgabe.

Zur Modellierung dieser Mechanismen benötigt man unterschiedliche Modelle, die im folgenden kurz vorgestellt werden. Danach erfolgt ein prinzipieller Leistungsvergleich der beiden Ein-/Ausgabe-Verfahren.

2.3.1 Interruptgesteuerte Ein-/Ausgabe

Bei der interruptgesteuerten Eingabe müssen die Anforderungen nach ihrer zufälligen Entstehung in der Peripherie sofort von der Steuereinheit übernommen und in die zugehörige Programm- bzw. Arbeitswarteschlange eingespeichert werden. Zur kurzzeitigen Zwischenspeicherung steht meist ein Übernahmespeicher zur Verfügung. Für diesen unterbrechungsgesteuerten Transfer wird ein Eingabe-Programm G_0 aktiviert, welches ein gerade in Bedienung befindliches Programm sofort unterbricht. Die Bearbeitung des Eingabe-Programms, das im wesentlichen das "Retten" des Zustands des unterbrochenen Programms, die Übernahme der Anforderungen und die Wiederaufnahme der Bearbeitung des unterbrochenen Programms durchführt, benötigt eine gewisse Verwaltungszeit (overhead). Für die Ausgabe, d. h. Rückmeldungen von der Steuereinheit zur Peripherie, gelten ähnliche Überlegungen.

Ein Warteschlangenmodell zur Beschreibung der interruptgesteuerten Eingabe unter Einbeziehung von P Prioritätsklassen und Overhead zeigt Bild 2.1.

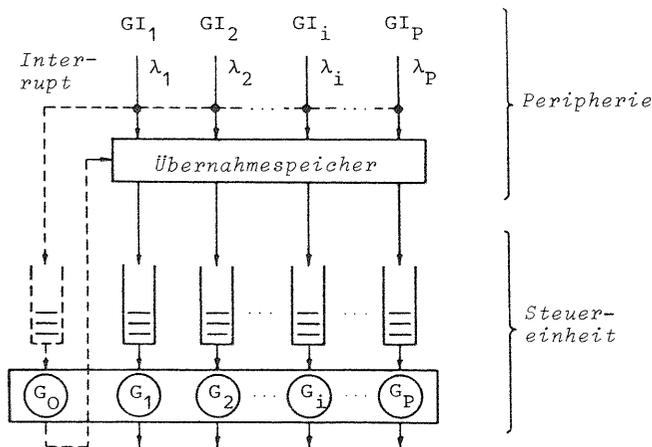


Bild 2.1. Warteschlangenmodell mit interruptgesteuerter Eingabe, P Prioritätsklassen und Overhead.

Ein Modell mit Markoff-Ankunftsprozeß ($GI_i = M, i = 1, 2, \dots, P$), allgemein verteilten Bedienungszeiten für die Anforderungen der verschiedenen Prioritätsklassen (sog. Unterbrechungsdistanzprioritäten, eine Verallgemeinerung der gemischten Prioritäten) und Overhead wurde von Herzog in [8] ausführlich untersucht.

Da der Verwaltungsaufwand proportional mit der Gesamtankunftsrate ansteigt, werden die Wartezeiten in allen Prioritätsklassen sehr ungünstig beeinflusst. Diese interruptgesteuerte Methode wird deshalb nur für dringliche Anforderungen (Alarmer, Fehlermeldungen) oder bei Systemen mit geringer Last angewendet.

2.3.2 Taktgesteuerte Ein-/Ausgabe

Bei der taktgesteuerten E/A (bei Rechnersystemen vergleichbar zur programmgesteuerten Ein-/Ausgabe) geht die gesamte Initiative nur von der Steuereinheit aus. In äquidistanten zeitlichen Abständen führt die Steuereinheit, ausgelöst durch eine Realzeituhr oder durch entsprechende Programmierung, die Ein- und Ausgabe von Anforderungen durch. Der Verwaltungsaufwand ist nahezu lastunabhängig, da das E/A-Programm nur zu den Abfragezeitpunkten aufgerufen wird. Da die Anforderungen von ihrer zufälligen Entstehung bis zum Transferzeitpunkt warten müssen, benötigt man E/A-Puffer zur Zwischenspeicherung. Der zeitliche Abstand zwischen den Abfragezeitpunkten ist meistens konstant, man spricht deshalb von taktgesteuerter E/A.

Ein Warteschlangenmodell zur Beschreibung der taktgesteuerten Eingabe unter Einbeziehung von Prioritätsklassen und Overhead zeigt Bild 2.2. Die in der Peripherie erzeugten Anforderungen werden entsprechend ihrer Dringlichkeit in einer Eingabewarteschlange zwischengespeichert. Die taktmäßige Umspeicherung der Anforderungen wird im Modell durch einen Schalter nachgebildet. Eine zum Taktzeitpunkt in Bedienung befindliche Anforderung wird sofort unterbrochen. Die Bearbeitung wartender Anforderungen erfolgt nach Prioritäten. Durch die Bedienungsphase G_0 wird die Verwaltungszeit für die taktgesteuerte Eingabe berücksichtigt.

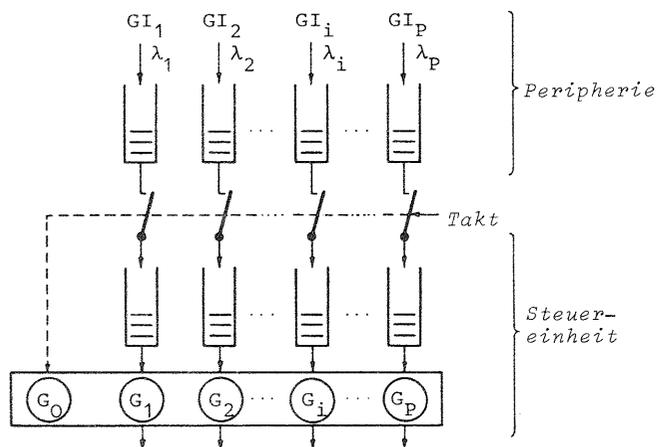


Bild 2.2. Warteschlangenmodell mit taktgesteuerter Eingabe, P Prioritätsklassen und Overhead.

Exakte analytische Untersuchungen für Modelle mit taktgesteuerter Eingabe für eine Klasse von Anforderungen und speziellen Bedienungs-VF findet man in [9,10,11,12]. Nur in [11,12] ist die Overheadphase berücksichtigt, bei den Modellen in [10,11,12] ist die realistischere Annahme einer endlichen Zahl von Warteplätzen in den Arbeitswarteschlangen mit einbezogen. Ein mehrklassiges Modell mit nichtunterbrechenden Prioritäten wurde in [13,14] untersucht. Die Analyse eines Modells mit taktgesteuerter E/A, allgemeinen Ankunfts- und Bedienungsprozessen für die verschiedenen Prioritätsklassen und Overhead ist ein wesentlicher Bestandteil dieser Arbeit und wird in Kap. 4 bzw. Kap. 5 durchgeführt.

Die Methode der taktgesteuerten E/A ist bei größeren Lasten der interruptgesteuerten E/A überlegen und wird deshalb bei vielen modernen programmgesteuerten Vermittlungssystemen oder bei Rechnersystemen angewendet.

2.3.3 Prinzipieller Leistungsvergleich zwischen interruptgesteuerter und taktgesteuerter Eingabe

Ein prinzipieller Leistungsvergleich, bei dem die Durchlaufzeiten der Anforderungen, d. h. die Gesamtzeit vom Eintreffen einer Anforderung bis zum Verlassen des Systems, betrachtet werden, soll die charakteristischen Eigenschaften beider Eingabe-Mechanismen aufzeigen.

a) Eine Klasse von Anforderungen

Für eine Klasse von Anforderungen wurden von Langenbach-Belz in [15] Ergebnisse für Modelle mit interrupt- und taktgesteuerter Eingabe und deterministischer Bedienungszeit miteinander verglichen. Beim Vergleich der numerischen Ergebnisse wurde davon ausgegangen, daß die konstante Verwaltungszeit T_{OH} bei beiden Modellen gleich groß sein soll. Wie das Bild 2.3 zeigt, ist die mittlere Durchlaufzeit im Modell mit taktgesteuerter Eingabe bei kleinen Ankunftsraten wesentlich größer als bei der interruptgesteuerten Eingabe.

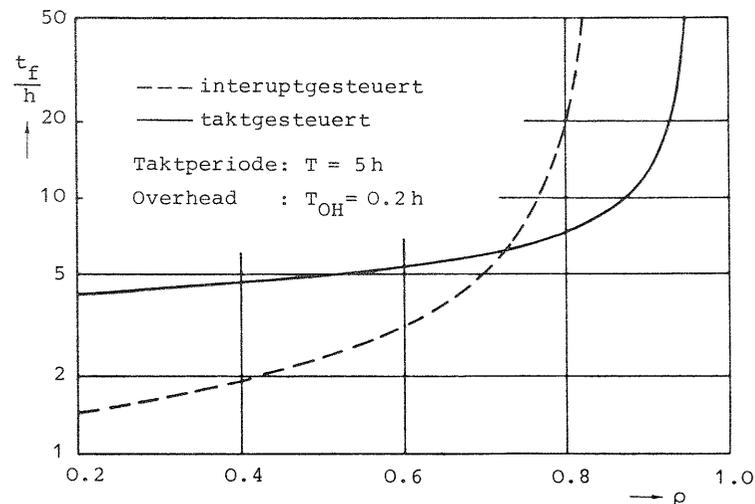


Bild 2.3 Bezogene mittlere Durchlaufzeit t_f/h in Abhängigkeit des Angebots ρ

Dies rührt daher, daß bereits die mittlere Wartezeit in der Peripherie $\geq T/2$ beträgt. Dagegen sind die Verhältnisse bei größeren Ankunftsraten gerade umgekehrt, hier wirkt sich die geringere Anzahl von Verwaltungszeiten pro Zeiteinheit zu Gunsten der taktgesteuerten Eingabe aus.

b) Zwei oder mehr Klassen von Anforderungen

Für Modelle mit zwei oder mehr Klassen von Anforderungen wurden bisher keine Leistungsvergleiche zwischen den beiden Eingabemechanismen durchgeführt, da für das entsprechende taktgesteuerte Modell unter allgemeinen Voraussetzungen keine Ergebnisse bekannt sind. Die Untersuchung eines taktgesteuerten Modells und ein Leistungsvergleich zu dem entsprechenden interruptgesteuerten Modell sind deshalb Hauptbestandteil dieser Arbeit. Es wird gezeigt, daß man für die Durchlaufzeiten aller Klassen im Prinzip die gleichen Kurvenverläufe erhält, wie sie in Bild 2.3 dargestellt sind. Da jedoch die taktgesteuerte Methode bei höheren Lasten der interruptgesteuerten Methode überlegen ist, wird mit ihr insbesondere für die niedrigeren Prioritätsklassen eine erhebliche Reduzierung der Durchlaufzeiten erzielt. Erste Ergebnisse dieser Untersuchungen findet man in [16].

2.4 Anwendung von Modellen in der Vermittlungstechnik

Die in diesem Kapitel kurz dargestellte Methodik der Modellbildung ist fester Bestandteil der Systementwicklung und Planung von Nachrichtenvermittlungs- und Rechnersystemen. Dementsprechend ist die Zahl der Arbeiten auf diesem Gebiet nahezu unüberschaubar groß. Stellvertretend sei deshalb nur auf die Arbeit von Kühn [17] verwiesen, da dort eine umfangreiche Literaturzusammenstellung zu finden ist.

Die mit Hilfe verkehrstheoretischer Modelle gewonnenen Ergebnisse gestatten ein Studium der verschiedenen Parametereinflüsse bereits bei der Entwicklung von Bedienungssystemen. Dadurch lassen sich frühzeitig verschiedene Entwicklungskonzepte miteinander vergleichen und als Konsequenz Fehlplanungen weitgehend vermeiden.

Wie der prinzipielle Leistungsvergleich zwischen der interruptgesteuerten und taktgesteuerten Eingabe zeigt, haben beide Verfahren in verschiedenen Auslastungsbereichen gewisse Vorteile. Für kleinere Nebenstellenanlagen mit geringem Angebot wird deshalb die interruptgesteuerte Methode bevorzugt, während größere Anlagen aufgrund des höheren Angebots die taktgesteuerte Methode verwenden. Neuere Steuerungskonzepte, denen eine modulare Mehrrechner-Konfiguration zugrundeliegt, verwenden wegen der dezentralen Steuerung eine meldungsgekoppelte Übergabe der Anforderungen. Ähnlich wie in Datenkommunikationssystemen sind hierbei alle Steuerungseinheiten an einem gemeinsamen "Übertragungskanal" (Bussystem, Durchschaltenetz) angeschlossen. Der Übertragungsvorgang erfolgt mit Hilfe von Meldungen (messages) nach einem vereinbarten Protokoll, unter Festlegung einer Zuteilungsstrategie des Kanals.

Nach dieser kurzen Darstellung der Modellbildung soll im nachfolgenden Kapitel das Gebiet der Modellanalyse, insbesondere im Hinblick auf die in dieser Arbeit verwendeten Analysemethoden, näher betrachtet werden.

3. METHODEN DER MODELLANALYSE

Bei der Modellanalyse unterscheidet man grundsätzlich zwischen den analytischen und den experimentellen Methoden. Insbesondere die analytischen Methoden sind in großem Umfang in der einschlägigen Fachliteratur beschrieben. Standardwerke der Bedienungstheorie (Verkehrstheorie, Warteschlangentheorie) sind z. B. [18, 19, 20, 21, 22, 23]. In diesem Kapitel sollen beide Analysemethoden näher betrachtet werden.

3.1 Analytische Methoden

Die analytischen Methoden gliedern sich in exakte und approximative Verfahren. Bei einer Vielzahl von Warteschlangenmodellen sind exakte analytische Lösungen entweder nicht bekannt oder mit vertretbarem Aufwand kaum auswertbar. Zur Untersuchung solcher Modelle wird deshalb auf approximative Verfahren zurückgegriffen. Beide Verfahren haben damit ihre Anwendungsbereiche und sollen im folgenden nacheinander betrachtet werden.

3.1.1 Exakte analytische Methoden

Grundlage aller Wahrscheinlichkeitstheoretischen Methoden sind die stochastischen Prozesse, mit deren Hilfe die in den untersuchten Systemen ablaufenden Vorgänge beschrieben werden. Neben den bereits erwähnten Ankunfts- und Bedienungsprozessen, die durch die VF der Ankunftsabstände bzw. Bedienungsdauern beschrieben werden, spielt der Systemzustandsprozeß zur Charakterisierung der zufallsabhängigen Anzahl X(t) von Anforderungen im Bedienungssystem eine wichtige Rolle. Allgemein sind stochastische Prozesse durch die drei Komponenten

- Zustandsraum X(t) (diskret oder kontinuierlich)
- Parameter t (diskret oder kontinuierlich)
- Verteilung F(z,t) bzw. Verteilungsfunktion F(t)

beschreibbar. X(t) = x heißt Zustand des stochastischen Prozesses

zur Zeit t. Ein Klassifizierungsmerkmal stochastischer Prozesse stellt das sog. "Gedächtnis" des Prozesses dar, d. h. die mögliche Abhängigkeit der weiteren Prozeßentwicklung von der Vergangenheit. Hieraus ergibt sich eine grundsätzliche Einteilung in Markoff'sche und Nichtmarkoff'sche Prozesse. Vor der Beschreibung der Analysemethoden beider Prozeßtypen soll die "Gedächtnislosigkeit" Markoff'scher Prozesse erläutert werden.

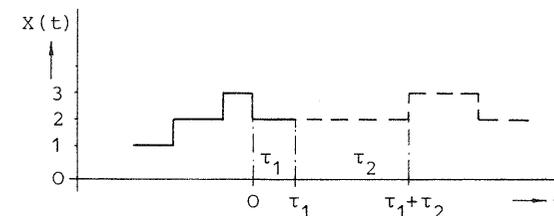
3.1.1.1 Die "Gedächtnislosigkeit" Markoff'scher Prozesse

Ein zufälliger Prozeß X(t) heißt ein Markoff'scher Prozeß, wenn für beliebige n = 0, 1, 2, ... und beliebige Zeitpunkte t_0, t_1, ..., t_n, t_{n+1} (t_0 < t_1 < ... < t_n < t_{n+1}) auf t gilt:

$$P\{X(t_{n+1})=x_{n+1} | X(t_n)=x_n, \dots, X(t_0)=x_0\} = P\{X(t_{n+1})=x_{n+1} | X(t_n)=x_n\} . \tag{3.1}$$

Der wahrscheinliche Zustand des Prozesses im Zeitpunkt t_{n+1} unter Berücksichtigung aller seiner vergangenen Zustände in den Zeitpunkten t_0, t_1, ..., t_n hängt nur vom letzten Zustand x_n im Zeitpunkt t_n ab; die weiter zurückliegende Vergangenheit des Prozesses wirkt nicht mehr nach.

Diese charakteristische Eigenschaft der Markoff'schen Prozesse wird als "Gedächtnislosigkeit" bezeichnet. Sie bildet die Grundlage für die Analyse Markoff'scher Bedienungssysteme. Betrachtet werde z. B. ein Zustandsprozeß zu einem beliebigen Zeitpunkt t, der die Markoff'sche Eigenschaft der Gedächtnislosigkeit in jedem Zeitpunkt besitzt.



Kennt man die Zeitdauer τ_1 seit der letzten Zustandsänderung und fragt nach der VF der Restzeit τ_2 bis zur nächsten Zustandsänderung, so erhält man die gleiche VF wie für die Gesamtzeit $\tau_1 + \tau_2$ zwischen den beiden Systemzuständen. Es gilt die Beziehung:

$$P\{T > \tau_1 + \tau_2 | T > \tau_1\} = P\{T > \tau_2\} . \quad (3.2)$$

Es kann gezeigt werden, daß von allen kontinuierlichen VF nur die Exponentialverteilung diese Eigenschaft besitzt (vgl. z. B. [19]).

$$F(t) = P\{T \leq t\} = 1 - e^{-\alpha t}$$

mit (3.3)

$$E[T] = 1/\alpha .$$

α entspricht der Übergangsrate pro Zeiteinheit und $E[T]$ der mittleren Zeitdauer für das Bestehen eines Zustandes.

3.1.1.2 Analyse Markoff'scher Bedienungssysteme

Zur Beschreibung des Systemzustandsprozesses benötigt man die sogenannte Kolmogoroff'sche Gleichung der Übergangswahrscheinlichkeiten. Ihre Herleitung soll deshalb zunächst angegeben werden.

a) Kolmogoroff'sche Gleichung der Übergangswahrscheinlichkeiten

Betrachtet werde ein Systemzustandsprozeß, der die Markoff'sche Eigenschaft besitzt. Die Übergangswahrscheinlichkeiten sind wie folgt definiert:

$$p_{ij}(u) = P\{X(t+u) = j | X(t) = i\}, \quad u \geq 0 . \quad (3.4)$$

Sie geben die Wahrscheinlichkeit für eine Zustandsänderung vom Zustand i in den Zustand j innerhalb der Zeitspanne u an. Unter der Voraussetzung eines homogenen Prozesses - p_{ij} ist nur von der Zeitspanne u , nicht von der absoluten Zeit t abhängig - und bei Betrachtung der Prozeßentwicklung während zweier aufeinanderfolgender Zeitabschnitte erhält man die als Chapman-Kolmogoroff-

Gleichung bekannte Beziehung [17,18]:

$$p_{ij}(t+u) = \sum_k p_{ik}(t) \cdot p_{kj}(u), \quad t \geq 0 . \quad (3.5)$$

Die Gleichung beschreibt den Übergang von $i \rightarrow j$, ausgedrückt durch die Übergänge $i \rightarrow k$ in der Zeitspanne t und $k \rightarrow j$ in der Zeitspanne u für alle möglichen k . Durch den Grenzübergang $u \rightarrow 0$ erhält man schließlich die Kolmogoroff'sche Vorwärtsgleichung der Übergangswahrscheinlichkeiten

$$\frac{d}{dt} p_{ij}(t) = -q_j p_{ij}(t) + \sum_{k \neq j} q_{kj} p_{ik}(t) \quad (3.6a)$$

mit

$$q_{kj} = \lim_{u \rightarrow 0} \frac{p_{kj}(u)}{u}, \quad q_j = \lim_{u \rightarrow 0} \frac{1 - p_{jj}(u)}{u} . \quad (3.6b)$$

q_{kj} und q_j bezeichnet man als Übergangswahrscheinlichkeitsdichten oder Übergangsraten.

Analog zu Gl. (3.6) läßt sich die sogenannte Kolmogoroff'sche Rückwärtsgleichung herleiten. Diese kann nicht zur Bestimmung der Zustandswahrscheinlichkeiten, sondern zur Berechnung der Wartezeitverteilungsfunktion herangezogen werden. Da sie im Rahmen dieser Arbeit nicht benötigt wird, bleibt sie hier nur erwähnt.

b) Systemzustandsprozeß

Grundlage der Analyse ist die Betrachtung des Systemzustandsprozesses, dessen Systemzustandsfolge $\{X(t), t \geq 0\}$ durch die Zustandswahrscheinlichkeit

$$P_j(t) = P\{X(t) = j\}, \quad j = 0, 1, \dots \quad (3.7)$$

charakterisiert wird. Aus der Anfangsverteilung $P_j(0)$, $j = 0, 1, \dots$ und durch Einführung der (bedingten) Übergangswahrscheinlichkeiten läßt sich die Zustandswahrscheinlichkeit unter Benutzung des Gesetzes der totalen Wahrscheinlichkeit wie folgt ausdrücken:

$$P_j(t) = \sum_i P\{X(0)=i\} \cdot P\{X(t)=j|X(0)=i\} = \sum_i P_i(0) \cdot p_{ij}(t) \quad (3.8)$$

$j = 0, 1, \dots$

Durch Einsetzen der Gl. (3.6) in die Gl. (3.8) erhält man die Kolmogoroff'sche Vorwärtsgleichung der Zustandswahrscheinlichkeiten

$$\frac{d}{dt} P_j(t) = -q_j P_j(t) + \sum_{k \neq j} q_{kj} P_k(t), \quad j = 0, 1, \dots \quad (3.9)$$

Mit der Normierungsbedingung

$$\sum_j P_j(t) = 1 \quad (3.10)$$

und ausgehend von einer vorgegebenen Anfangsverteilung $P_j(0)$, $j = 0, 1, \dots$, können mit Hilfe des Differentialgleichungssystems nach Gl. (3.9) sämtliche zeitabhängigen Zustandswahrscheinlichkeiten und damit das dynamische Systemverhalten berechnet werden.

Bei vielen Anwendungen interessiert das Systemverhalten nur im stationären Fall, d. h. im "eingeschwungenen Zustand".

c) Der stationäre Fall

Für die stationären Zustandswahrscheinlichkeiten

$$P_j = \lim_{t \rightarrow \infty} P_j(t) \quad (3.11)$$

erhält man aus der Gl. (3.9) mit

$$\lim_{t \rightarrow \infty} \frac{d}{dt} P_j(t) = 0$$

ein System von linearen Gleichungen

$$q_j P_j = \sum_{k \neq j} q_{kj} P_k, \quad j = 0, 1, \dots \quad (3.12)$$

Im "stationären Gleichgewichtszustand" und bei Betrachtung des Systems innerhalb der Zeitspanne dt ist also die Wahrscheinlichkeit für das Verlassen des Zustandes j gleich der Wahrscheinlichkeit des Entstehens dieses Zustandes aus allen anderen Zuständen.

Die Auflösung des linearen Gleichungssystems kann mit den Methoden der linearen Algebra durchgeführt werden. Abhängig vom betrachteten System und von den dadurch bedingten Übergangsraten kann die Auflösung rekursiv oder muß iterativ durchgeführt werden.

Einen speziellen Markoff'schen Prozeß stellen die Geburts- und Sterbeprozesse dar, die zur Untersuchung vielfältiger Probleme der Bedienungs-, Lagerhaltungs- und Zuverlässigkeitstheorie herangezogen werden. Sie zeichnen sich dadurch aus, daß nur Übergänge zwischen Nachbarzuständen möglich sind. Dadurch ist eine rekursive Berechnung der Zustandswahrscheinlichkeiten im eindimensionalen Fall möglich.

3.1.1.3 Analyse Nichtmarkoff'scher Prozesse

Die Nichtmarkoff'schen Prozesse sind solche stochastischen Prozesse, die nicht zu jedem Zeitpunkt die für die Analyse günstige Eigenschaft der Gedächtnislosigkeit besitzen. Betrachtet man z. B. ein einfaches Warteschlangenmodell, bei dem schon eine Modellkomponente (Ankunftsprozeß oder Bedienungsprozeß) keine negativ exponentielle Verteilungsfunktion nach Gl. (3.3) aufweist und damit gedächtnisbehaftet ist, so ist der Systemzustandsprozeß ebenfalls gedächtnisbehaftet.

Bei den Analysemethoden Nichtmarkoff'scher Prozesse unterscheidet man zwischen den Methoden, die das "Gedächtnis" in einer Zusatzvariablen festhalten und denjenigen, die den Prozeß nur zu solchen Zeitpunkten beschreiben, welche die Markoff'sche Eigenschaft aufweisen. Da einige dieser Verfahren in dieser Arbeit verwendet werden, folgt eine kurze Beschreibung ihrer Arbeitsweise.

a) Methode der Eingebetteten Markoff-Kette

Diese Methode geht auf Kendall zurück [24,18,19,17] und ist dann anwendbar, wenn der zu untersuchende Systemzustandsprozeß Zeitpunkte aufweist, welche die Markoff'sche Eigenschaft besitzen. In diesen sogenannten "Regenerationszeitpunkten" verliert der Prozeß sein Gedächtnis, d. h. er entwickelt sich von dort aus unabhängig von der Vergangenheit. Zur Analyse des allgemeinen Zustandsprozesses wird in diese Regenerationspunkte eine Markoff-Kette "eingebettet". Die Beschreibung der eingebetteten Markoff-Kette erfolgt durch ein Gleichungssystem, das die Entwicklung des Zustandsprozesses für zwei aufeinanderfolgende Regenerationszeitpunkte i und $i+1$ beschreibt:

$$P\{X(t_{i+1}+0)=x_{i+1}\} = \sum_{x_i} P\{X(t_i+0)=x_i\} \cdot P\{X(t_{i+1}+0)=x_{i+1} | X(t_i+0)=x_i\} . \quad (3.13)$$

Für die (bedingten) Übergangswahrscheinlichkeiten gilt im stationären Gleichgewichtszustand (keine Abhängigkeit vom Zeitindex i)

$$q_{jk} = P\{X(t_{i+1}+0) = k | X(t_i+0) = j\} . \quad (3.14)$$

Mit Hilfe der q_{jk} erhält man im stationären Fall folgendes Gleichungssystem:

$$P_k = \sum_j P_j q_{jk}, \quad k = 0, 1, \dots \quad (3.15a)$$

oder in Matrix-Schreibweise

$$\underline{P} = \underline{Q} \cdot \underline{P} \quad (3.15b)$$

mit \underline{P} als Wahrscheinlichkeitsvektor und \underline{Q} als Übergangswahrscheinlichkeitsmatrix.

In vielen Anwendungsfällen gelingt die Auflösung durch Einführung sogenannter "Erzeugender Funktionen" (vgl. Anhang 1). Gegebenenfalls ist ein iteratives Verfahren zur Bestimmung der Zustandswahrscheinlichkeiten notwendig.

Ausgehend von den Zustandswahrscheinlichkeiten zu den Regenerationszeitpunkten können abschließend die Zustandswahrscheinlichkeiten zu jedem beliebigen Zeitpunkt (outside observer) bestimmt werden.

b) Phasenmethode

Mit Hilfe der Phasenmethode läßt sich eine allgemeine, d. h. Nichtmarkoff'sche VF beliebig genau aus exponentiellen (d. h. Markoff'schen) Teilphasen zusammensetzen. Das Verfahren wurde von Cox [25,18] entwickelt. Ein Beispiel zur Nachbildung einer allgemeinen VF mit Serien-Parallel-Bedienphasen zeigt Bild 3.1.

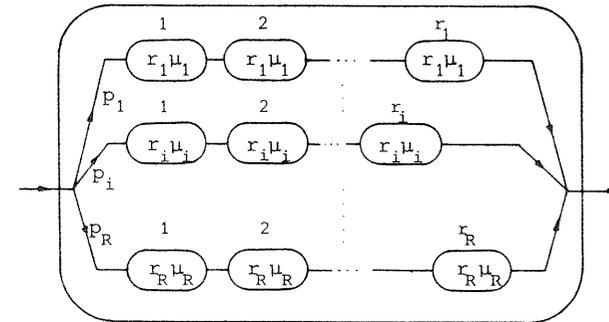


Bild 3.1. Approximation einer allg. VF durch Serien-Parallel-Bedienphasen

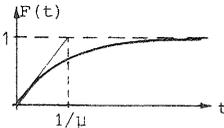
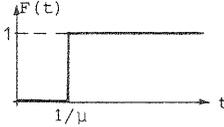
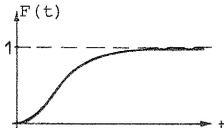
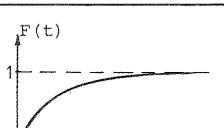
Für die Verteilungsfunktion gilt die Beziehung

$$F(t) = P\{T \leq t\} = 1 - \sum_{i=1}^R p_i e^{-r_i \mu_i t} \sum_{j=0}^{r_i-1} \frac{(r_i \mu_i t)^j}{j!}, \quad t \geq 0 . \quad (3.16)$$

Ersetzt man nun eine vorgegebene allgemeine VF durch die in Bild 3.1 angegebene Anordnung, so ist eine rein Markoff'sche Zustandsbeschreibung des gegebenen Prozesses möglich. Die Zahl der möglichen Zustände, bedingt durch die Teilbedienphasen, nimmt jedoch zu. Das in Bild 3.1 dargestellte Modell läßt sich in ein äquivalentes Phasenmodell, bestehend aus einer Kette von

exponentiellen Teilphasen und Verzweigungen vor und hinter diesen Teilphasen, umwandeln [25,18].

Die folgende Tabelle zeigt die wichtigsten VF der Nachrichtenverkehrstheorie, die als Sonderfälle in der allgemeinen Darstellung nach Bild 3.1 bzw. Gl. (3.16) enthalten sind.

Kurz-bez.	Mathematische Beschreibung $F(t) = P\{T \leq t\}$	Mittelwert $E[T]$ Variationskoeff. c	Graphische Darstellung
M	$F(t) = 1 - e^{-\mu t}$	$E[T] = 1/\mu$ $c = 1$	
D	$F(t) = \begin{cases} 0, & t < 1/\mu \\ 1, & t \geq 1/\mu \end{cases}$	$E[T] = 1/\mu$ $c = 0$	
E_k	$F(t) = 1 - e^{-k\mu t} \cdot \sum_{j=0}^{k-1} \frac{(k\mu t)^j}{j!}$ $k = 1, 2, \dots$	$E[T] = 1/\mu$ $c = 1/\sqrt{k}$	
H_k	$F(t) = 1 - \sum_{i=1}^k p_i e^{-\mu_i t}$ $k = 1, 2, \dots; \sum_{i=1}^k p_i = 1$	$E[T] = \sum_{i=1}^k p_i / \mu_i$ $c = \sqrt{\frac{2 \sum_{i=1}^k p_i / \mu_i^2}{E[T]^2}} - 1$	

c) Methode der Gedächtnisvariablen

Durch die Einführung einer Gedächtnisvariablen (supplementary variable) ist es möglich, einen Nichtmarkoff'schen Prozeß auf einen Markoff'schen Prozeß zurückzuführen. Die Gedächtnisvariable, die abhängig vom betrachteten Prozeß diskret oder kontinuierlich sein kann, berücksichtigt die vorhandene Abhängigkeit der wei-

teren Prozeßentwicklung von der Vergangenheit, z. B. die Zeitdauer zwischen der letzten Zustandsänderung und dem betrachteten Zeitpunkt. Da diese Methode im Rahmen dieser Arbeit nicht benutzt wird, soll sie nicht näher erläutert werden, vgl. [26].

d) Lindley'sche Integralmethode

Dieses Verfahren wurde von Lindley [27] entwickelt und ist für Wartesysteme des Typs GI/G/1 anwendbar. Es nutzt die Tatsache aus, daß unter der Voraussetzung FIFO die Wartezeit w_n für den n-ten Ankommenden identisch ist mit der Restarbeit $U(t)$ im System unmittelbar zum Zeitpunkt der Ankunft. Für die VF der Wartezeiten erhält man unter den genannten Voraussetzungen eine Integralgleichung des Wiener-Hopf-Typs

$$W(t) = \int_{u=-\infty}^t W(t-u) \cdot dC(u), \quad t \geq 0 \quad (3.19)$$

mit

$$C(u) = \int_{t=0}^{\infty} H(u+t) dA(t) .$$

$A(t)$ ist die VF des Ankunftsprozesses, $H(t)$ die VF des Bedienungsprozesses.

Unter bestimmten Voraussetzungen kann diese Gleichung mit Hilfsmitteln der Funktionentheorie gelöst werden [18]. Im Rahmen dieser Arbeit wird für ein Modell des Typs $C^{[x]}/G/1$ eine Lösung durch Zeitdiskretisierung vorgestellt.

e) Momentenmethoden

Ist man nicht am gesamten Zustandsprozeß, sondern nur an den Erwartungswerten charakteristischer Größen interessiert, so kann auf dieses Verfahren zurückgegriffen werden. Durch Betrachtung des Schicksals von sogenannten "Testanforderungen" können mit bekannten Beziehungen der Erneuerungstheorie (Little'sches Gesetz, vgl. Abschnitt 3.3.2) bestimmte Erwartungswerte berechnet werden (z. B. mittlere Wartezeit, mittlere Warteschlangenlänge usw.). Beispiele für die Anwendung des Verfahrens findet man in [28,29].

Die Aufzählung der genannten exakten analytischen Methoden erhebt keinen Anspruch auf Vollständigkeit, weitere Verfahren können der Literatur der Warteschlangentheorie entnommen werden.

3.1.2 Approximative analytische Methoden

Die approximativen analytischen Methoden lassen sich nicht in der gleichen Weise klassifizieren wie die exakten analytischen Methoden. Sie sind stark modellabhängig und entstehen häufig aus intuitiven Ansätzen (Überlegungen) zur approximativen Lösung von komplexen Warteschlangenmodellen oder Warteschlangennetzen. Stellvertretend soll deshalb nur auf die in dieser Arbeit angewendeten Verfahren näher eingegangen werden.

3.1.2.1 Zerlegungsmethoden

Die Zerlegungsmethoden beruhen auf der Aufteilung eines komplexen Gesamtsystems in eine überschaubare Anzahl von Teilsystemen. Die Genauigkeit der Ergebnisse ist von der Wahl geeigneter Schnittstellen zwischen den Teilsystemen abhängig. Gute Ergebnisse erhält man z. B., wenn die Interaktionen zwischen den Teilsystemen im Verhältnis zu Interaktionen innerhalb der Teilsysteme gering sind oder wenn an den Schnittstellen viele Teilverkehre überlagert werden (Grenzwertsatz der Erneuerungstheorie). Sind die Verkehre an den Schnittstellen bekannt oder hinreichend genau berechenbar, so erhält man ebenfalls gute Ergebnisse.

Bei der Analyse eines dreistufigen Modells mit taktgesteuerter Übergabe wird in Kap. 4 und 5 auf dieses Zerlegungsverfahren zurückgegriffen (vgl. auch [30,17,31]).

3.1.2.2 Momentenapproximation

Häufig sind für die Schnittstellenverkehre zwar die ersten zwei oder drei Momente bestimmbar, nicht aber die ganze VF. Bei der

Analyse werden dann die Verkehre durch bekannte VF approximiert. Die so entstehenden Ersatzsysteme sind in der Regel einfacher zu analysieren. Dieses Verfahren, welches z. B. zur Approximation von Wartezeit-VF in [32,33] angewendet wurde, wird bei der später durchgeführten Untersuchung taktgesteuerter Modelle ebenfalls benutzt.

3.1.2.3 Erneuerungsannahmen

Bei der Berechnung von Warteschlangenmodellen greift man häufig auf Methoden der Erneuerungstheorie zurück. Bei Benutzung der Methode der eingebetteten Markoff-Kette ist z. B. die Zahl der Erneuerungen (Ankünfte und/oder Bedienungen) in einem Zeitintervall $(0,t)$ zu bestimmen. Da bei der approximativen Berechnung von getakteten Modellen auf Methoden der Erneuerungstheorie zurückgegriffen werden kann, soll an dieser Stelle auf spätere Kapitel dieser Arbeit verwiesen werden. Begriffe der Erneuerungstheorie, wie z. B. Vorwärts-Rekurrenzzeit, werden dort näher erläutert.

3.2 Experimentelle Methoden

Bei den experimentellen Methoden unterscheidet man zwischen Messungen an realen Systemen und Simulationsmethoden. Durch Messungen an realen Systemen kann die Leistungsfähigkeit durch das Studium der verschiedenen Parametereinflüsse verbessert oder sogar optimiert werden. Bei der Entwicklung neuer Systeme können mit Hilfe von Simulationsmethoden verschiedene Entwicklungskonzepte auf ihre Leistungsfähigkeit hin untersucht werden.

3.2.1 Messung an Systemen

Die Leistungsfähigkeit realer Systeme oder einzelner Systemkomponenten kann meßtechnisch durch Hardware- oder Software-Messungen ermittelt werden. Bei den Hardware-Messungen erfolgt die

Erfassung der Meßgrößen durch das Anschalten von Meßfühlern parallel zu den zu untersuchenden Komponenten. Durch Aufzeichnung der Meßdaten und entsprechende Auswertung ist anschließend eine Leistungsbeurteilung möglich. Bei den Softwaremessungen werden innerhalb der Systemprogramme entsprechende Meßprogramme eingefügt, die bei jedem Aufruf die gewünschten Messungen durchführen und die anfallenden Daten auswerten. Nachteil dieses Verfahrens ist die mögliche ungünstige Beeinflussung des realen Systemsverhaltens. Obwohl Messungen an realen Systemen meistens sehr aufwendig sind, sind sie zur Überprüfung von Modell-Aussagen unerlässlich.

3.2.2 Simulative Methoden

Bei der Synthese neuer Systeme oder zur Überprüfung (Validierung) von approximativ berechneten Ergebnissen greift man auf die Simulationstechnik zurück. Die Simulationstechnik verwendet zwar wie die analytischen Methoden Modelle, jedoch kann hier die Modellierungstiefe, d. h. die Einbeziehung vieler Parameter des realen Systems innerhalb des Modells, theoretisch beliebig weit gehen. Nach der Modellfestlegung wird die Struktur und die Betriebsart des Modells durch ein entsprechendes Programm auf einem Digitalrechner abgebildet. Die erforderlichen stochastischen Prozesse, wie Ankunftsabstände und Bedienungsdauer werden nach Typ, Mittelwert und VF mit Hilfe gleichverteilter Pseudo-Zufallszahlen erzeugt. Bei der i. a. verwendeten zeittreuen, ereignisgesteuerten Simulation ist eine Speicherung des Systemzustandes der einzelnen Betriebsmittel (Bedienungseinheiten, Warteplätze usw.) erforderlich. Die Zustandsänderungen, z. B. Ankunft einer Anforderung, Bedienungsende, Zeitmarke usw., sind durch Ereignisse zu diskreten Zeitpunkten gekennzeichnet. Zur Verwaltung dieser Ereignisse wird ein sogenannter "Ereigniskalender" geführt, in den alle bekannten, zukünftigen Ereignisse eingetragen werden. Der aktuelle Zeitpunkt auf der simulierten Zeitachse wird durch die Zeitvariable "Systemzeit" gekennzeichnet. Abhängig vom Typ des zu bearbeitenden Ereignisses ist neben der

Aktualisierung des Systemzustandes und der Erfassung von statistischen Daten die Ermittlung eines neuen Zukunftsereignisses (z. B. nächster Ankunftszeitpunkt) erforderlich.

Nach Bearbeitung eines Ereignisses wird aus dem Ereigniskalender das zeitlich nächstfolgende Ereignis ermittelt, und auf der Zeitachse erfolgt ein unmittelbarer Sprung zu diesem Ereigniszeitpunkt (event by event simulation). Die Zeit zwischen zwei Ereignissen wird nur simuliert, d. h. das Ablaufgeschehen auf dem Rechner läuft zeitlich gerafft ab.

Die Programmstruktur eines Simulationsprogramms hat einen hierarchischen Aufbau, der in vier Ebenen eingeteilt werden kann. Die erste Ebene bildet das Rahmenprogramm. Hier erfolgt die Initialisierung des Systems (z. B. Aufbau der Datenstruktur entsprechend der Modellbeschreibung, Initialisierung der Ankunftsereignisse usw.) und die statistische Auswertung und Ausgabe der Simulationsergebnisse am Ende des Simulationslaufes. In der zweiten Ebene erfolgt die Steuerung der Ereignisverwaltung. Hier wird die Bestimmung des zeitlich nächstfolgenden Ereignisses, die Verzweigung zur Ereignisablaufsteuerung (Ereignisanalyse) und die statistische Auswertung der Meßdaten nach jedem Teiltest durchgeführt. Die dritte Ebene ist die eigentliche Ereignisablaufsteuerung. Abhängig vom Typ des Ereignisses erfolgt hier die eigentliche Ereignisbearbeitung, die Erfassung statistischer Daten und die Ermittlung eines Zukunftsergebnisses, falls erforderlich. In der vierten und letzten Ebene befindet sich die Prozedurbibliothek. Hier sind die für die oberen Programmebenen erforderlichen Unterprogramme, die während eines Simulationslaufes immer wieder aufgerufen werden, untergebracht.

Aus Gründen der statistischen Aussagesicherheit wird der gesamte Simulationslauf in einen Vorlauf und in N Teiltests unterteilt. Der Vorlauf dient zur Erreichung des stationären Zustands. Nach Beendigung eines jeden Teiltests werden die charakteristischen Meßgrößen berechnet und die Meßzähler wieder auf Null gesetzt. Aus den gewonnenen N Stichproben läßt sich für jede Meßgröße, z. B. Mittelwert und Varianz, und mit Hilfe der Theorie der

Student-t-Verteilung [34] ein Vertrauensintervall bestimmen. Die Zahl der Teiltests und die Zahl der Ereignisse innerhalb eines Teiltests sind von der gewünschten Genauigkeit der Meßgrößen abhängig (Richtwert 10 Teiltests mit je 10 000 Ereignissen). Das Kriterium für das Teiltestende ist i. a. eine vorgebbare Anzahl von eintreffenden Anforderungen im System oder eine vorgebbare Teiltestdauer.

Zur Validierung der in Kap. 4 und 5 entwickelten approximativen, analytischen Methoden zur Untersuchung mehrstufiger Wartemodelle mit taktgesteuerter Ein-/Ausgabe, Verwaltungszeit und Prioritäten wurde ein Simulationsprogramm in der Programmiersprache FORTRAN IV entwickelt [35].

3.3 Charakteristische Größen

Ziel aller verkehrstheoretischen Untersuchungen ist die Bestimmung "charakteristischer Größen" für das zu untersuchende Bedienungssystem. Bei den analytischen Methoden erfolgt dies z. B. aus den Zustandswahrscheinlichkeiten, bei den experimentellen Methoden durch statistische Auswertung der Meßgrößen.

3.3.1 Kenngrößen für Systemeigenschaften

Im folgenden sind die wichtigsten Kenngrößen für Systemeigenschaften zusammengefaßt, die im Rahmen dieser Arbeit verwendet werden:

Mittlere Durchlaufzeit t_f : Die mittlere Durchlaufzeit einer bestimmten Anforderung ist gleich der Summe aller Warte-, Bedienungs- und Transportzeiten. Man benutzt häufig auch die Begriffe Antwort- oder Reaktionszeit.

Wartewahrscheinlichkeit W : Wahrscheinlichkeit, daß eine eintreffende Anforderung überhaupt warten muß.

Mittlere Wartezeit w : Mittlere Wartezeit einer eintreffenden Anforderung bezogen auf alle Anforderungen, d. h. wartende und

nichtwartende Anforderungen.

Mittlere Wartezeit t_w : Mittlere Wartezeit einer eintreffenden Anforderung bezogen auf alle wartenden Anforderungen. Es gilt der Zusammenhang

$$t_w = \frac{w}{W} .$$

Wartebelastung Ω : Die Wartebelastung, häufig auch mittlere Warteschlangenlänge genannt, ist gleich der mittleren Anzahl gleichzeitig auf Bedienung wartender Anforderungen.

Verteilungsfunktion VF : Charakteristische Mittelwerte und Wahrscheinlichkeiten reichen zur Beschreibung des Verkehrsverhaltens eines Systems nicht immer aus. In vielen Fällen benötigt man die VF zur vollständigen Beschreibung der Schwankungen der Größen um den Mittelwert, z. B. Wartezeitverteilungsfunktionen.

Neben diesen Größen gibt es eine Reihe weiterer Kenngrößen, die zur Beschreibung der Systemeigenschaften eines Bedienungssystems herangezogen werden können. Einige seien hier nur aufgezählt:

- Verlustwahrscheinlichkeit
- Blockierungswahrscheinlichkeit
- Momente der VF
- Durchsatz
- Belastung
- Unterbrechungswahrscheinlichkeit

Abhängig von der Struktur und Betriebsart können zur Charakterisierung des betrachteten Modells alle oder nur ein gewisser Teil dieser Kenngrößen sinnvoll angegeben werden.

Zwischen einigen Größen gibt es eine mathematische Beziehung, die als Little'sches Theorem bekannt ist. Dieses Theorem hat in der Verkehrstheorie grundlegende Bedeutung und soll deshalb kurz diskutiert werden.

3.3.2 Das Little'sche Theorem

Betrachtet werde ein beliebiges System, z. B. ein Bedienungssystem oder nur ein Teilsystem, dem ein beliebiger Ankunftsprozeß mit der Ankunftsrate λ angeboten wird. Das System befinde sich im stationären Gleichgewichtszustand. Bezeichnet man mit $E[X]$ die mittlere Anzahl von Anforderungen und mit $E[T_F]$ die mittlere Durchlaufzeit im betrachteten System, so gilt folgende Gesetzmäßigkeit zwischen diesen Größen

$$\lambda \cdot E[T_F] = E[X] . \tag{3.20}$$

Diese Beziehung wird als Little'sches Theorem bezeichnet [36]. Ein prinzipieller Beweis dieses Theorems ist leicht möglich. Betrachtet man ein System mit den Systemzuständen $X(t)$ über einen langen Zeitraum t_0 ($t_0 \rightarrow \infty$) und treffen in dieser Zeit genau c_A Anforderungen ein, so gilt für die mittlere Durchlaufzeit und für die mittlere Anzahl im System:

$$E[T_F] = \frac{1}{c_A} \sum_{i=1}^{c_A} T_{F_i} \approx \frac{1}{c_A} \int_0^{t_0} X(t) dt$$

$$E[X] = \frac{1}{t_0} \int_0^{t_0} X(t) dt .$$

Aus den beiden Gleichungen ergibt sich unmittelbar

$$E[X] \approx \frac{c_A}{t_0} \cdot E[T_F] .$$

Im Grenzübergang $t_0 \rightarrow \infty$ ist $\lambda = c_A/t_0$ und man erhält die Gl. (3.20).

Ein Zusammenhang zwischen der mittleren Wartezeit w bzw. t_w und der Wartebelastung Ω kann ebenfalls mit Hilfe des Little'schen Theorems direkt angegeben werden. Es gilt die Beziehung

$$t_w = \frac{\Omega}{\lambda \cdot W} , \quad w = \frac{\Omega}{\lambda} . \tag{3.21}$$

$\lambda \cdot W$ entspricht der mittleren Zahl von Anforderungen die pro Zeiteinheit eintreffen und warten müssen.

Das Little'sche Theorem wird häufig auch für Kontrollrechnungen verwendet, wenn die Berechnung der charakteristischen Größen getrennt durchgeführt wurde, z. B. über die Analyse der Wartezeitverteilung oder bei Simulationen.

3.3.3 Erhaltungssätze für Bedienungssysteme

Mit Hilfe der sog. "Erhaltungssätze" lassen sich bei Kenntnis gewisser charakteristischer Größen weitere Kenngrößen eines Bedienungssystems bestimmen. So wird z. B. im Kap. 5 ein Erhaltungssatz für Prioritätssysteme mit unterbrechenden Prioritäten abgeleitet, der von der Überlegung ausgeht, daß die in einem System zu einem beliebigen Zeitpunkt anstehende Arbeit nicht von der Bedienungsreihenfolge abhängt, solange das System "arbeitserhaltend" (im engl.: "work conservative") ist (vgl. [18]). Für die weitere Diskussion wird auf den Abschnitt 5.2.2.1 verwiesen.

3.4 Zusammenfassung

Der in diesem Kapitel gegebene Überblick über die wesentlichen Methoden der Modellanalyse erhebt keinen Anspruch auf Vollständigkeit, da hier nur solche Methoden diskutiert wurden, die im Rahmen dieser Arbeit angewendet werden. Zur weitergehenden Information kann deshalb nur wiederholt auf die einschlägige Fachliteratur verwiesen werden; einige Angaben hierzu wurden am Anfang dieses Kapitels und bei der Beschreibung der verschiedenen Analysemethoden vorgenommen.

In den folgenden Kapiteln werden spezielle Modelle untersucht, für die bisher entweder keine oder nur unter gewissen einschränkenden Voraussetzungen Ergebnisse vorliegen.

4. ANALYSE DES DREISTUFIGEN WARTEMODELLS MIT TAKTGESTEUERTER EIN-/AUSGABE, VERWALTUNGSZEIT UND EINER ANFORDERUNGSKLASSE

4.1 Struktur und Betriebsweise

Das Modell in Bild 4.1 bildet die Abläufe zwischen einer Steuerungseinheit und der Peripherie in vereinfachter Form nach, wenn die Übertragung von und zur Peripherie nur zu äquidistanten, d. h. taktgesteuerten Zeitpunkten stattfindet. Die Modellstruktur berücksichtigt nur eine Anforderungsklasse, ein Modell mit mehreren Prioritätsklassen wird im Kap. 5 untersucht. Die Steuerungseinheit entspricht z. B. bei einem zentralgesteuerten Vermittlungssystem dem Zentralrechner, bei einem dezentral- oder teilzentralgesteuerten Vermittlungssystem einem peripheren Rechner. Die Modellstruktur ist auch anwendbar für einen Prozessor mit Peripherie, wenn die Ein-/Ausgabe programmgesteuert abläuft, d. h. vom Rechner aktiviert und durchgeführt wird.

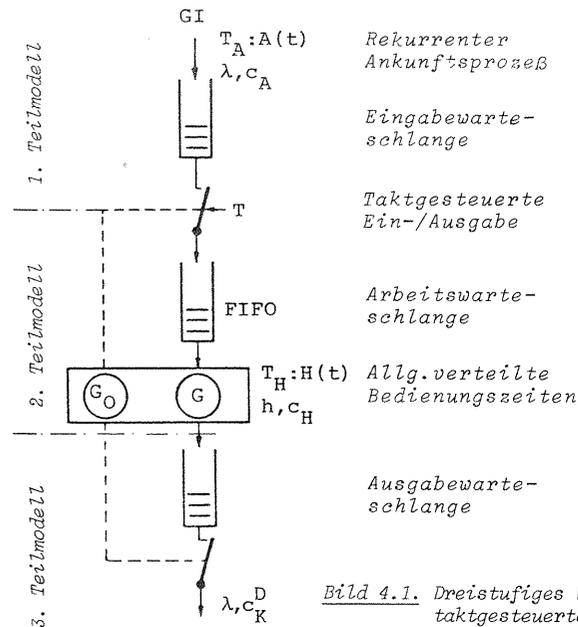


Bild 4.1. Dreistufiges Wartemodell mit taktgesteuerter Ein-/Ausgabe, Verwaltungszeit und einer Anforderungsklasse

Der Ankunftsprozeß wird durch allgemein verteilte Ankunftsabstände (GI) mit der mittleren Ankunftsrate λ beschrieben. Diese in der Peripherie erzeugten Anforderungen werden in der Eingabewarteschlange zwischengespeichert.

Zu jeweils äquidistanten Taktzeitpunkten mit der Periode T werden alle wartenden Anforderungen von der Eingabewarteschlange in die Arbeitswarteschlange (work queue) übertragen. Die z. B. von einer Realzeituhr erzeugte Unterbrechungsanforderung (interrupt request) legt die Taktzeitpunkte fest und aktiviert die Steuerungseinheit zur Durchführung des Transfers der Anforderungen. Die zur Übertragung erforderliche Bedienungsphase G_O wird als konstant angenommen, die zugehörige Zeitdauer T_{OH} bezeichnet man als Verwaltungszeit (overhead time). Die Verwaltungszeit setzt sich im wesentlichen aus zwei Anteilen zusammen, aus der Initialisierungszeit (Aufruf des Unterbrechungsprogramms) und der eigentlichen Übertragungszeit. Die Übertragungszeit schwankt zwar abhängig von der Zahl der zu übertragenden Anforderungen, jedoch ist sie bei realen Systemen vernachlässigbar klein verglichen mit der Initialisierungszeit. Da zur Initialisierung stets das gleiche Unterbrechungsprogramm aufgerufen wird, stellt die Annahme einer konstanten Verwaltungszeit eine gute Approximation dar.

Eine zum Taktzeitpunkt in Bedienung befindliche Anforderung wird sofort unterbrochen. Die Fortsetzung der Bedienung erfolgt erst nach Ablauf der Overhead-Phase, weitere wartende Anforderungen werden nach der Strategie FIFO abgearbeitet.

Der Bedienungsprozeß wird durch allgemein verteilte Bedienungszeiten (G) mit der mittleren Bedienungszeit h charakterisiert.

Die bei der Bearbeitung der Anforderungen erzeugten Rückmeldungen (Befehle) für die Peripherie werden in der Ausgabewarteschlange (output queue) zwischengespeichert. Ebenfalls zu den Taktzeitpunkten findet der Transfer zwischen der Ausgabewarteschlange und der Peripherie statt. Repräsentiert das Modell nach Bild 4.1 den Knoten eines Warteschlangennetzes, so entspricht der Transfer der im Abschnitt 2.2.4 besprochenen Wegelenkung (routing).

Zur Untersuchung des Modells und zur Berechnung der charakteristischen Verkehrsgrößen wird auf die in Abschnitt 3.1.2.1 angegebenen Zerlegungsmethoden zurückgegriffen. Dabei wird das Gesamtmodell in drei Teilmodelle zerlegt. Das erste Teilmodell umfaßt den Ankunftsprozeß, die Eingabewarteschlange und die taktmäßige Umspeicherung von Anforderungen. Da diese Umspeicherung wie eine Bedienung interpretiert werden kann, ergibt sich für dieses Teilmodell ein Warteschlangenmodell des Typs $GI/C/\infty$, d. h. allgemeiner Ankunftsprozeß und taktmäßige Bedienung aller wartenden Anforderungen.

Das zweite Teilmodell berücksichtigt den taktmäßigen Gruppenankunftsprozeß, die Arbeitswarteschlange und die eigentliche Bedienungseinheit. Die Kurznotation hierfür ist $C^{[x]}/G/1$, d. h. taktmäßiger Gruppenankunftsprozeß und eine Bedienungseinheit mit allgemeiner Bedienungs-VF $H(t)$. Das Modell berücksichtigt zusätzlich die Overhead-Phase G_0 mit der konstanten Bearbeitungsdauer T_{OH} .

Das dritte Teilmodell entspricht im Prinzip dem ersten Teilmodell, jedoch wird hier der Eingangsprozeß durch den Bedienungsprozeß des zweiten Teilmodells gebildet.

Wie bereits im Abschnitt 3.1.2.1 beschrieben, gehören die Zerlegungsmethoden zu den approximativen analytischen Methoden. Die Ergebnisse sind um so besser, je genauer die Schnittstellenverkehre beschreibbar sind oder je weniger die Teilmodelle voneinander abhängen.

Betrachtet werde zunächst die Schnittstelle zwischen dem ersten und dem zweiten Teilmodell. Unter der Voraussetzung, daß die Arbeitswarteschlange unendlich groß sein soll, können zum Taktzeitpunkt immer alle Anforderungen übertragen werden, d. h. die Eingabewarteschlange wird zu jedem Taktzeitpunkt völlig entleert. Das bedeutet aber, daß in keiner Weise ein Zustand des ersten Teilmodells durch irgendeinen Zustand des zweiten Teilmodells beeinflusst wird. Deshalb können beide Teilmodelle unabhängig voneinander untersucht werden. Ist die Arbeitswarteschlange jedoch begrenzt, dann können nicht immer alle Anforderungen übertragen

werden. Aufgrund dieser Rückwirkung können die beiden Teilmodelle nicht unabhängig voneinander untersucht werden. Eine derartige Modellstruktur, die beide Teilmodelle umfaßt und die dadurch bedingten gegenseitigen Abhängigkeiten mit einbezieht, wurde in [37] ausführlich untersucht.

Zur Untersuchung des zweiten Teilmodells muß der Gruppenausgangsprozeß des ersten Teilmodells genauer betrachtet werden. Die Gruppengrößen-Verteilung ist identisch mit der Anzahl wartender Anforderungen in der Eingabewarteschlange kurz vor dem Taktzeitpunkt. Da stets alle wartenden Anforderungen zwischen der Eingabe- und Arbeitswarteschlange transferiert werden, ist die Gruppengrößen-Verteilung gleich der Verteilung der Anzahl von eintreffenden Anforderungen in einem Taktintervall. Die Berechnung dieser Verteilung mit Methoden der Erneuerungstheorie wird im Abschnitt 4.2.1 durchgeführt.

Die zweite Schnittstelle liegt zwischen der Bedienungseinheit und der Ausgabewarteschlange. Der Schnittstellenverkehr kann unter Erneuerungsannahmen zu den Taktzeitpunkten ebenfalls approximativ berechnet werden. Ausgehend von den Zustandswahrscheinlichkeiten kurz nach den Taktzeitpunkten im zweiten Teilmodell kann mit Hilfe der Sterbewahrscheinlichkeiten die VF der Zahl von Anforderungen kurz vor dem Takt in der Ausgabewarteschlange und damit die Gruppengrößen-Verteilung des Ausgangsprozesses bestimmt werden.

Im folgenden werden die drei Teilmodelle nacheinander untersucht und die zugehörigen charakteristischen Verkehrsgrößen bestimmt. Anschließend werden die berechneten Ergebnisse diskutiert. Die aufgrund von approximativen Analysemethoden ermittelten Ergebnisse werden hierbei grundsätzlich mit entsprechenden Simulationsergebnissen verglichen.

4.2 Das erste Teilmodell GI/C/∞

Das erste Teilmodell, welches den Ankunftsprozeß, die Eingabewarteschlange und die taktmäßige Bedienung aller wartenden Anforderungen zum Taktzeitpunkt berücksichtigt, kann aufgrund der Rückwirkungsfreiheit durch ein Warteschlangenmodell des Typs GI/C/∞ nachgebildet werden. Bild 4.2 zeigt dieses Modell mit allgemeinem Ankunftsprozeß GI, der Eingabewarteschlange und dem Schalter, der die taktmäßige Bedienung nachbildet.

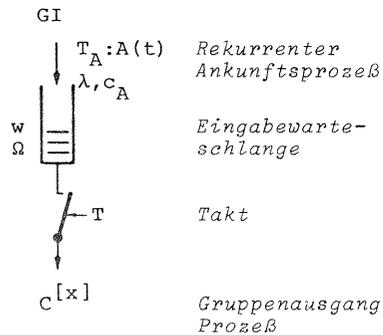


Bild 4.2. Das erste Teilmodell GI/C/∞

Analytische Untersuchungen dieses Modelltyps für spezielle Ankunftsprozesse wurden in [9, 38, 39, 40, 41] durchgeführt. Die Modelle M/C/n in [9], M/C/1 - s in [38], M/C/n - s in [39, 41] und E_k/C/n - s in [40] berücksichtigen eine Begrenzung der zum Taktzeitpunkt transferierten Anforderungen auf 1 bzw. n. Bei den meisten Untersuchungen wird außerdem eine endliche Eingabewarteschlange (s) mit einbezogen. Die Analyse erfolgt einheitlich mit Hilfe einer eingebetteten Markoff-Kette.

Die zu berechnenden charakteristischen Verkehrsgrößen sind die mittlere Wartezeit w und die mittlere Warteschlangenlänge Ω. Hinzu kommt der taktmäßige Gruppenausgangsprozeß C^[x], der den Eingangsprozeß für das zweite Teilmodell darstellt.

Für die Bestimmung der mittleren Wartezeit wird eine "Testanforderung" betrachtet, die im Intervall (t_v, t_v - dt_v) eintrifft und

bereits x Anforderungen in der Zeitspanne (T - t_v) mit der Wahrscheinlichkeit g_x(T - t_v) eingetroffen sind (vgl. Bild 4.3).

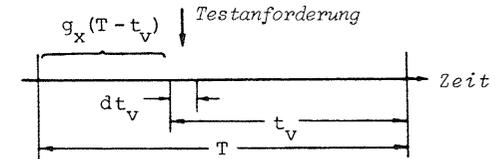


Bild 4.3. Testanforderung zur Bestimmung der mittleren Wartezeit

Die Wahrscheinlichkeit für das Eintreffen der Testanforderung im Zeitintervall dt_v ist gleichverteilt über das Intervall T und beträgt dt_v/T. Da die beiden Ereignisse

- Eintreffen der Testanforderung in dt_v
- bereits x Anforderungen im System

unabhängig voneinander sind, ergibt sich für die Wahrscheinlichkeit P{t_v - dt_v, t_v}, daß die Testanforderung zwischen t_v - dt_v und t_v warten muß

$$P\{t_v - dt_v, t_v\} = g_x(T - t_v) \cdot \frac{dt_v}{T}, \quad 0 \leq t_v \leq T. \quad (4.1)$$

Durch Multiplikation mit t_v, Addition über alle möglichen Systemzustände und Mittelung über das Taktintervall T erhält man für die mittlere Wartezeit

$$w = \frac{1}{T} \int_0^T \sum_{x=0}^{\infty} g_x(T - t_v) t_v dt_v.$$

Da die Summation über alle Systemzustandswahrscheinlichkeiten zu jedem Zeitpunkt identisch "1" ist, erhält man für die mittlere Wartezeit die einfache Beziehung

$$w = \frac{T}{2}. \quad (4.2)$$

Das Ergebnis ist unmittelbar interpretierbar. Eine beliebig herausgegriffene Anforderung, die gleichwahrscheinlich zu jedem Zeitpunkt innerhalb eines Taktintervalls eintrifft, muß im Mittel die halbe Taktperiode warten.

Da alle Anforderungen warten müssen ($w=1$), gilt $t_w = w$ und für die mittlere Warteschlangenlänge Ω unter Verwendung des Little'schen Theorems nach Gl. (3.21)

$$\Omega = \lambda w \quad (4.3)$$

Unter der Voraussetzung eines allgemeinen Ankunftsprozesses sind die Taktzeitpunkte keine Regenerationspunkte des Prozesses, da die zufällige Zeitdauer T_1 bis zum ersten Ankunftszeitpunkt gedächtnisbehaftet ist. Nur bei exponentiell verteilten Ankunftsabständen ($GI=M$) ist jeder Zeitpunkt und damit auch der Taktzeitpunkt ein Regenerationszeitpunkt des Prozesses (vgl. Kap. 3.1.1.1).

Die Berechnung des Gruppenausgangsprozesses erfolgt deshalb mit Methoden der Erneuerungstheorie. Da zu den Taktzeitpunkten stets alle wartenden Anforderungen übertragen werden, ist die Gruppengrößen-Verteilung des Ausgangsprozesses identisch der Verteilung der Anzahl von Ankünften in einem Taktintervall T . Die Berechnung der Gruppengrößen-Verteilung wird im folgenden für allgemeine Ankunftsprozesse unter Erneuerungsannahmen zu den Taktzeitpunkten durchgeführt.

4.2.1 Berechnung der Gruppengrößen-Verteilung

Entsprechend Bild 4.4 beobachtet man einen rekurrenten Prozeß mit allgemein verteilten Ankunftsabständen und fragt nach der Zahl der Anforderungen (Erneuerungen des Prozesses), die in ein Intervall $(0, t)$ hineinfallen. Die Zwischenankunftszeiten werden durch die stetigen ZV T_i , $i = 1, 2, \dots$, gekennzeichnet. T_1 habe die Verteilungsdichtefunktion (VDF) $f_1(t)$, alle anderen T_i , $i > 1$, haben die gleiche VDF $f(t)$ (rekurrenter Prozeß).

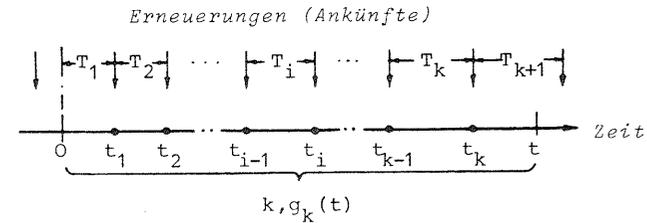


Bild 4.4. Bestimmung der Anzahl von Erneuerungen eines rekurrenten Prozesses

Abhängig von der VDF $f_1(t)$ unterscheidet man zwischen drei verschiedenen Erneuerungsprozeßtypen:

1. Gewöhnlicher oder normaler Erneuerungsprozeß
 - $f_1(t) = f(t)$, alle ZV T_i , $i = 1, 2, \dots$, sind gleich verteilt
2. Modifizierter Erneuerungsprozeß
 - $f_1(t) \neq f(t)$, alle ZV T_i , $i > 1$, sind gleich verteilt
3. Stationärer Erneuerungsprozeß
 - T_1 entspricht der "Vorwärts-Rekurrenzzeit", d. h. $f_1(t)$ ist die VDF der Vorwärts-Rekurrenzzeit nach Gl. (4.4a). Alle anderen ZV T_i , $i > 1$, sind gleich verteilt

Zur Berechnung der Gruppengrößen-Verteilung wird hier auf den stationären Erneuerungsprozeß zurückgegriffen. Die physikalische Interpretation eines stationären Erneuerungsprozesses ist folgende: Ein gewöhnlicher Erneuerungsprozeß wird weit vor dem zufälligen Beobachtungszeitpunkt gestartet. Dann bezeichnet man die Zeitdauer T_1 zwischen Beobachtungszeitpunkt und der 1. Erneuerung als "Vorwärts-Rekurrenzzeit" des Prozesses. Die Vorwärts-Rekurrenzzeit hat die VDF

$$f_1(t) = \lambda [1 - F(t)] \quad (4.4a)$$

mit dem Mittelwert

$$E[T_1] = \frac{\lambda}{2} E[T_i^2], \quad i > 1 \quad (4.4b)$$

und der Laplace-Transformierten (LT)

$$\Phi_1(s) = \lambda \frac{1 - \Phi(s)}{s} . \quad (4.4c)$$

$F(t)$ ist die VF der regulären Erneuerungen ($T_i, i > 1$) mit dem Mittelwert $1/\lambda$, $\Phi(s)$ ist die Laplace-Stieltjes-Transformierte (LST) von $F(t)$. Für die Herleitung der Gl. (4.4) wird z. B. auf [42,18] verwiesen.

Wie im folgenden gezeigt wird, besteht zwischen dem stationären und dem gewöhnlichen Erneuerungsprozeß ein fester Zusammenhang. Beide Prozesse werden deshalb näher betrachtet. Zur Unterscheidung werden die Größen des gewöhnlichen Erneuerungsprozesses mit $\bar{}$, die des stationären Erneuerungsprozeß mit $\tilde{}$ gekennzeichnet.

Vor der Berechnung der Wahrscheinlichkeiten g_k des Gruppenausgangsprozesses werden einige grundsätzliche Beziehungen der Erneuerungsprozesse abgeleitet. Zur Verallgemeinerung wird zunächst ein beliebiges Intervall $(0,t)$ betrachtet und die Wahrscheinlichkeit $g_k(t)$ für $K=k$ Erneuerungen innerhalb dieses Intervalls berechnet (vgl. auch [42]).

Ist die ZV S_k die Zeit bis zur k -ten Erneuerung, $F_k(t)$ die VF von S_k mit $F_0(t) = 1$, dann gilt

$$P\{K < k\} = P\{S_k > t\} = 1 - F_k(t)$$

und damit

$$g_k(t) = P\{K = k\} = F_k(t) - F_{k+1}(t) . \quad (4.5)$$

Ist $G(z,t)$ die erzeugende Funktion für die Wahrscheinlichkeiten $g_k(t)$,

$$G(z,t) = \sum_{k=0}^{\infty} g_k(t) z^k , \quad (4.6)$$

dann erhält man durch Einsetzen der Gl. (4.5) in die Gl. (4.6)

$$G(z,t) = 1 + \sum_{k=1}^{\infty} F_k(t) (z-1) z^{k-1} . \quad (4.7a)$$

Bezeichnet man mit $\Phi_k(s)$ die LT von $f_k(t)$, so ist $\Phi_k(s)/s$ die LST von $F_k(t)$. Damit gilt für die LT $\Phi_G(z,s)$ von $G(z,t)$:

$$\Phi_G(z,s) = \frac{1}{s} + \frac{1}{s} \sum_{k=1}^{\infty} \Phi_k(s) (z-1) z^{k-1} . \quad (4.7b)$$

Da die ZV S_k gleich der Summe der ZV $T_i, i = 1, 2, \dots, k$, ist, ergibt sich deren LT $\Phi_k(s)$ aus dem Produkt der LT $\Phi_i(s)$ der VDF $f_i, i = 1, 2, \dots, k$. Da die VDF $f_1(t)$ für die beiden betrachteten Erneuerungsprozesse verschieden ist, erhält man unterschiedliche Beziehungen für $\Phi_k(s)$.

1. Gewöhnlicher Erneuerungsprozeß

Mit

$$\bar{\Phi}_k(s) = \Phi_1(s) \Phi^{k-1}(s) = \Phi^k(s) \quad (4.8)$$

gilt für die Gl. (4.7b)

$$\bar{\Phi}_G(z,s) = \frac{1 - \Phi(s)}{s[1 - z\Phi(s)]} . \quad (4.9)$$

2. Stationärer Erneuerungsprozeß

Mit

$$\tilde{\Phi}_k(s) = \Phi_1(s) \Phi^{k-1}(s) = \lambda \frac{1 - \Phi(s)}{s} \Phi^{k-1}(s) \quad (4.10)$$

gilt für die Gl. (4.7b)

$$\begin{aligned} \tilde{\Phi}_G(z,s) &= \frac{1}{s} + \lambda \frac{(z-1)[1 - \Phi(s)]}{s^2[1 - z\Phi(s)]} \\ &= \frac{1}{s} + \lambda \frac{z-1}{s} \bar{\Phi}_G(z,s) . \end{aligned} \quad (4.11a)$$

Die Rücktransformation dieser Gleichung liefert

$$\tilde{G}(z,t) = 1 + \lambda(z-1) \int_0^t \bar{G}(z,u) du . \quad (4.11b)$$

Durch Koeffizientenvergleich von z^k erhält man für die Wahrscheinlichkeiten $\tilde{g}_k(t)$, ausgedrückt durch die einfacher zu berechnenden Wahrscheinlichkeiten $\bar{g}_k(t)$:

$$\tilde{g}_k(t) = \lambda \int_0^t [\bar{g}_{k-1}(u) - \bar{g}_k(u)] du, \quad k=1,2,\dots \quad (4.12a)$$

und

$$\tilde{g}_0(t) = 1 - \lambda \int_0^t \bar{g}_0(u) du. \quad (4.12b)$$

Die gesuchten Gruppengrößenwahrscheinlichkeiten g_k sind nun gleich den Wahrscheinlichkeiten $\tilde{g}_k(T)$, d. h. die obere Integrationsgrenze ist gleich T zu setzen. λ ist die mittlere Ankunftsrate des rekurrenten Ankunftsprozesses.

Ist die numerische Berechnung der Wahrscheinlichkeiten $\bar{g}_k(t)$ für einen vorgegebenen Ankunftsprozeß mit vertretbarem Aufwand möglich, dann stellt die Berechnung der Wahrscheinlichkeiten $\tilde{g}_k(t)$ und damit die Bestimmung der Gruppengrößenwahrscheinlichkeiten g_k keine Schwierigkeit dar. Speziell für einen Markoff-Ankunftsprozeß ist die Berechnung besonders einfach und man erhält für die Gruppengrößen-Verteilung eine Poisson-Verteilung.

Für einen Erlang-verteilten Ankunftsprozeß der Ordnung k erhält man die Beziehung

$$\bar{g}_r(t) = \sum_{m=rk}^{(r+1)k-1} \frac{(k\lambda t)^m}{m!} e^{-k\lambda t}, \quad r=0,1,\dots \quad (4.13)$$

Die Gl. (4.13) folgt unmittelbar aus der Überlegung, daß die zufällige Zahl R der Erneuerungen genau dann den Wert r annimmt, wenn die Anzahl der durchlaufenen Stufen des zugrunde liegenden Markoff-Prozesses einen der Werte $rk, rk+1, \dots, (r+1)k-1$ annimmt. Hierbei entspricht die Anzahl der durchlaufenen Stufen in der Zeit t einer Poisson-Verteilung mit dem Mittelwert $k\lambda t$. Setzt man die Gl. (4.13) in die Gl. (4.12) ein, dann ergibt sich für die Gruppengrößen-Verteilung des Ausgangsprozesses:

$$g_0 = (1 - \lambda T) \bar{g}_0(T) + \lambda T f_{k-1}, \quad (4.14a)$$

$$g_r = (r+1 - \lambda T) \bar{g}_r(T) - (r-1 - \lambda T) \bar{g}_{r-1}(T) + \lambda T (f_{(r-1)k-1} - 2f_{rk-1} + f_{(r+1)k-1}), \quad r=1,2,\dots \quad (4.14b)$$

mit

$$f_j = \frac{(k\lambda T)^j}{j!} e^{-k\lambda T}, \quad f_j = 0 \quad \text{für } j < 0.$$

Der Indextausch k gegen r für die Wahrscheinlichkeiten g_r wurde durchgeführt, da k für die E_k -VF benutzt wird.

Die numerische Berechnung der Gruppengrößen-Verteilung für einen hyperexponentiellen Ankunftsprozeß (z. B. H_2) läßt sich mit vertretbarem Rechenaufwand nicht mehr durchführen. Lediglich die Berechnung der ausgezeichneten Größe g_0 , d. h. die Wahrscheinlichkeit, daß keine Anforderung innerhalb eines Taktintervalls eintrifft, ist leicht möglich. Man erhält für eine H_2 -Verteilung bei Vorgabe des Mittelwertes $t_A = 1/\lambda$ und des Variationskoeffizienten C_A unter Einbeziehung der üblicherweise verwendeten Relation $P_2 \lambda_1 = P_1 \lambda_2$:

$$g_0 = e^{-\lambda T} \cdot \cosh\left(\sqrt{\frac{C_A^2 - 1}{C_A^2 + 1}} \lambda T\right) \quad (4.15)$$

Die Größe g_0 wird für die Berechnung der Wartewahrscheinlichkeit bei der Untersuchung des zweiten Teilmodells vorausgesetzt.

Da die numerische Berechnung der gesamten Gruppengrößen-Verteilung nur für spezielle Ankunftsprozesse sinnvoll durchführbar ist, wird für die Untersuchung des zweiten Teilmodells eine approximative Berechnungsmethode vorgestellt, die nur den Mittelwert $E[K]$ und die Varianz $\text{Var}[K]$ der Gruppengrößen-Verteilung benötigt. Die Bestimmung dieser Größen aus der erzeugenden Funktion $G(z,s)$ wird im folgenden kurz diskutiert; abschließend werden Ergebnisse für die Ankunftsprozeßtypen $GI = E_k$ und $GI = H_2$ angegeben.

4.2.2 Mittelwert und Varianz der Gruppengrößen-Verteilung

Der Mittelwert $E[K]$ bzw. die Varianz $\text{Var}[K]$ ist identisch dem Mittelwert bzw. der Varianz der Erneuerungen im Intervall $(0, T)$. Diese Größen sind für beide Erneuerungsprozessypen direkt aus der erzeugenden Funktion $\Phi_G(z, s)$ berechenbar. Die LT sind:

$$\bar{\Phi}_H(s) \triangleq \left. \frac{d\bar{\Phi}_G(z, s)}{dz} \right|_{z=1} = \frac{\Phi(s)}{s[1 - \Phi(s)]} , \quad (4.16a)$$

$$\begin{aligned} \bar{\Phi}_V(s) &\triangleq \left. \frac{d^2\bar{\Phi}_G(z, s)}{dz^2} \right|_{z=1} + \bar{\Phi}_H(s) - \bar{\Phi}_H^2(s) \\ &= \bar{\Phi}_H(s) [1 + \bar{\Phi}_H(s) (2s - 1)] \end{aligned} \quad (4.16b)$$

und

$$\tilde{\Phi}_H(s) \triangleq \left. \frac{d\tilde{\Phi}_G(z, s)}{dz} \right|_{z=1} = \frac{\lambda}{s} , \quad (4.17a)$$

$$\begin{aligned} \tilde{\Phi}_V(s) &\triangleq \left. \frac{d^2\tilde{\Phi}_G(z, s)}{dz^2} \right|_{z=1} + \tilde{\Phi}_H(s) - \tilde{\Phi}_H^2(s) \\ &= \frac{2\lambda}{s} \bar{\Phi}_H(s) + \frac{\lambda}{s} - \frac{\lambda^2}{s^2} . \end{aligned} \quad (4.17b)$$

Die Rücktransformation der LT nach Gl. (4.17) liefern die gesuchten Größen $E[K]$ und $\text{Var}[K]$, wenn die obere Integrationsgrenze gleich T gesetzt wird:

$$E[K] = \lambda \int_0^T du = \lambda T \quad (4.18)$$

und

$$\text{Var}[K] = 2\lambda \int_0^T (\bar{H}(u) - \lambda u + \frac{1}{2}) du . \quad (4.19)$$

Die noch unbekannte Größe $\bar{H}(t)$ ist die sog. "Erneuerungsfunktion". Sie ergibt sich aus der Rücktransformation der Gl. (4.16a) und ist nur abhängig von $\Phi(s)$. $\Phi(s)$ ist die LT der VDF $f(t)$ der re-

gulären Erneuerungen. Vor der Betrachtung spezieller Ankunftsprozesse wird die LST $\bar{\Phi}_H(s)$ der Erneuerungsfunktion $\bar{H}(t)$ etwas genauer untersucht.

Ist $\Phi(s)$ eine rationale Funktion in s , dann kann $\bar{\Phi}_H(s)$ in Partialbrüche entwickelt werden. Die Pole der Funktion liegen bei $s=0$ und bei den nichtverschwindenden Wurzeln s_i , $i=1, 2, \dots$, der Gleichung

$$\Phi(s) = 1 .$$

Da diese Gleichung wegen $\Phi'(0) = -1/\lambda < 0$ ebenfalls eine einfache Nullstelle bei $s=0$ aufweist, enthält die Partialbruchzerlegung von $\bar{\Phi}_H(s)$ immer Glieder mit $1/s^2$ und $1/s$ und man erhält:

$$\bar{\Phi}_H(s) = \frac{\lambda}{s^2} + \frac{1}{2s}(c^2 - 1) + R(s) . \quad (4.20)$$

$R(s)$ ist eine rationale Funktion in s mit den Polen s_i , $i=1, 2, \dots$. Die nichtverschwindenden Wurzeln s_i treten entweder alle konjugiert komplex auf oder eine Wurzel ist negativ reell. $1/\lambda$ ist der Mittelwert, c ist der Variationskoeffizient der VF $F(t)$ der regulären Erneuerungen.

Für jede einfache, nichtverschwindende Wurzel s_i ergibt sich der folgende Ausdruck für die Funktion $R(s)$ in der Gl. (4.20)

$$- \frac{1}{s_i \Phi'(s_i) (s - s_i)} . \quad (4.21)$$

Für die speziellen Ankunftsprozesse $GI = E_k$ und $GI = H_2$ sollen nun diese Partialbruchentwicklung durchgeführt und nach der Rücktransformation die gesuchten Größen $E[K]$ und $\text{Var}[K]$ berechnet werden.

a) E_k -Ankunftsprozeß

Ist der Ankunftsprozeß E_k -verteilt mit der Ankunftsrate λ , dann haben die einzelnen M-Phasen die Rate $\epsilon = k \cdot \lambda$. Damit ergibt sich für die LT $\Phi(s)$ die Beziehung

$$\Phi(s) = \left(\frac{\epsilon}{s + \epsilon} \right)^k . \quad (4.22)$$

Durch Einsetzen von $\Phi(s)$ in Gl. (4.16a) erhält man:

$$\bar{\Phi}_H(s) = \frac{\epsilon^k}{s[(s+\epsilon)^k - \epsilon^k]} \quad (4.23)$$

Die nichtverschwindenden Wurzeln der Gleichung

$$(s+\epsilon)^k = \epsilon^k \quad (4.24a)$$

sind

$$s_i = \epsilon(e^{j\frac{2\pi}{k}i} - 1); \quad i=1,2,\dots,k-1 \quad (4.24b)$$

Ist k eine gerade Zahl, dann erhält man für $i=k/2$ die negativ reelle Wurzel $s_{k/2} = -2\epsilon$. Alle anderen Wurzeln treten - wie oben erwähnt - konjugiert komplex auf. Die Ableitung $\Phi'(s_i)$ liefert unter Benutzung der Gl. (4.24a)

$$\Phi'(s_i) = -\frac{k}{s_i + \epsilon} \quad (4.25)$$

Durch Einsetzen der Gl. (4.25) und der Gl. (4.24b) in Gl. (4.21) erhält man $R(s)$ und damit für die Entwicklung von $\bar{\Phi}_H(s)$ nach Gl. (4.20) mit $c=c_A$:

$$\bar{\Phi}_H(s) = \frac{\lambda}{s^2} + \frac{1}{2s} (c_A^2 - 1) + \frac{1}{k} \sum_{i=1}^{k-1} \frac{s_i + \epsilon}{s_i(s - s_i)} \quad (4.26)$$

Die konjugiert komplexen Wurzeln $s_i = \sigma_i + j\omega_i$ und $s_i^* = \sigma_i - j\omega_i$ können unter der Summe zusammengefaßt werden. Durch Einsetzen der Wurzeln und $\epsilon = k\lambda$ ergibt sich nach einigen Umformungen

$$\begin{aligned} \bar{\Phi}_H(s) &= \frac{\lambda}{s^2} + \frac{1}{2s} (c_A^2 - 1) \\ &+ \frac{1}{k} \sum_{i=1}^j \left[\frac{s - \sigma_i}{(s - \sigma_i)^2 + \omega_i^2} + \frac{\sigma_i + 2\epsilon}{\omega_i} \cdot \frac{\omega_i}{(s - \sigma_i)^2 + \omega_i^2} \right] \\ &+ \frac{1}{2k} \frac{1}{s + 2\epsilon} \Big|_{\text{nur für gerade } k} \end{aligned} \quad (4.27a)$$

und durch Rücktransformation

$$\begin{aligned} \bar{H}(t) &= \lambda t + \frac{1}{2}(c_A^2 - 1) \\ &+ \frac{1}{k} \sum_{i=1}^j e^{\sigma_i t} \left[\cos(\omega_i t) + \frac{\sigma_i + 2\epsilon}{\omega_i} \sin(\omega_i t) \right] \\ &+ \frac{1}{2k} e^{-2\epsilon t} \Big|_{\text{nur für gerade } k} \end{aligned} \quad (4.27b)$$

mit

$$j = \left[\frac{k-1}{2} \right] = \text{das Ganze der Division } \frac{k-1}{2}$$

und

$$\sum_{i=1}^j (\cdot) = 0 \quad \text{für } j < i.$$

Setzt man die Gl. (4.27b) in die Gl. (4.19) ein, so erhält man schließlich für die Varianz der Gruppengrößen-Verteilung des Ausgangsprozesses:

$$\begin{aligned} \text{Var}[K] &= \frac{\lambda T}{k} + \frac{2}{k^2} \sum_{i=1}^j \frac{1}{1 - \cos\varphi_i} [1 - e^{\sigma_i T} \cos(\omega_i T)] \\ &+ \frac{1}{2k^2} (1 - e^{-2k\lambda T}) \Big|_{\text{nur für gerade } k} \end{aligned} \quad (4.28)$$

mit

$$\varphi_i = \frac{2\pi}{k} i, \quad \sigma_i = k\lambda(\cos\varphi_i - 1), \quad \omega_i = k\lambda\sin\varphi_i.$$

Für den speziellen Fall $k=1$ (Markoff-Ankunftsprozeß) ist $c_A^2=1$, k ungerade, $j < i$ und man erhält:

$$\text{Var}[K] = \lambda T.$$

Da der Mittelwert für den stationären Erneuerungsprozeß immer gleich λT ist, erhält man wegen $E[K] = \text{Var}[K]$ erwartungsgemäß eine Poisson-Verteilung für die Gruppengrößen-Verteilung.

b) H_2 -Ankunftsprozeß

Der Ankunftsprozeß ist H_2 -verteilt mit der Ankunftsrate λ . Die Verzweigungswahrscheinlichkeiten sind p_1 bzw. p_2 ($p_1 + p_2 = 1$) und die Raten der beiden parallelen M-Phasen λ_1 bzw. λ_2 . Damit ergibt sich für die LT $\Phi(s)$ die Beziehung

$$\Phi(s) = \frac{p_1 \lambda_1}{s + \lambda_1} + \frac{p_2 \lambda_2}{s + \lambda_2} \quad (4.29)$$

Mit der nichtverschwindenden Wurzel

$$s_1 = - (p_2 \lambda_1 + p_1 \lambda_2) = - \frac{\lambda_1 \lambda_2}{\lambda} \quad (4.30)$$

der Gleichung $\Phi(s) = 1$ erhält man für die Entwicklung von $\bar{\Phi}_H(s)$ nach Gl. (4.20) mit $c = c_A$

$$\bar{\Phi}_H(s) = \frac{\lambda}{s^2} + \frac{1}{2s} (c_A^2 - 1) - \frac{1}{2} (c_A^2 - 1) \frac{1}{s - s_1} \quad (4.31a)$$

und durch Rücktransformation

$$\bar{H}(t) = \lambda t + \frac{1}{2} (c_A^2 - 1) (1 - e^{s_1 t}) \quad (4.31b)$$

Setzt man die Gl. (4.31b) in die Gl. (4.19) ein, so ergibt sich für die Varianz der Gruppengrößen-Verteilung des Ausgangsprozesses nach einigen Umformungen

$$\text{Var}[K] = \lambda T c_A^2 - \frac{\lambda^2}{\lambda_1 \lambda_2} (c_A^2 - 1) (1 - e^{-\lambda_1 \lambda_2 T / \lambda}) \quad (4.32a)$$

Unter Verwendung der üblichen Zusatzbedingung $p_1 \lambda_2 = p_2 \lambda_1$ kann λ_1 bzw. λ_2 als Funktion von c_A ausgedrückt werden und man erhält die einfachere Beziehung

$$\text{Var}[K] = \lambda T c_A^2 - \frac{1}{2} (c_A^4 - 1) (1 - e^{-2\lambda T / (c_A^2 + 1)}) \quad (4.33)$$

Für den speziellen Fall $c_A = 1$ (Markoff-Ankunftsprozeß) ergibt

sich mit $\text{Var}[K] = \lambda T$ wiederum die Poisson-Verteilung für die Gruppengrößen-Verteilung.

Aus dem Mittelwert $E[K]$ und der Varianz $\text{Var}[K]$ kann der Variationskoeffizient c_K der Gruppengrößen-Verteilung nach der Definition

$$c_K = \frac{\sqrt{\text{Var}[K]}}{E[K]} \quad (4.34)$$

berechnet werden. Bild 4.5 zeigt die Abhängigkeit des Variationskoeffizienten vom Mittelwert $E[K] = \lambda T$ für die Ankunftsprozeßtypen M, E_k und H_2 .

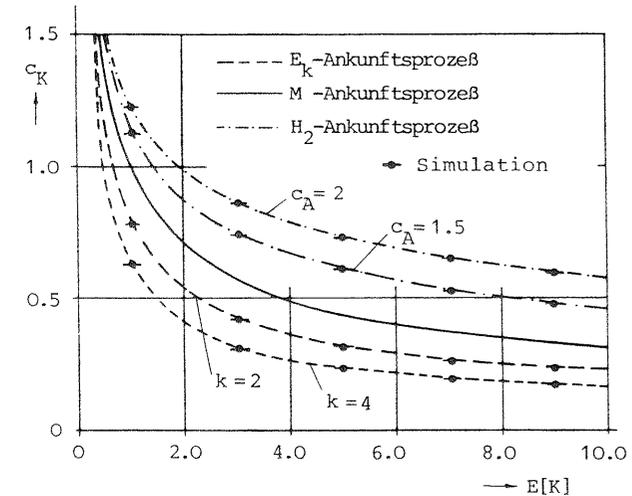


Bild 4.5. Variationskoeffizient c_K des Gruppenausgangsprozesses in Abhängigkeit von der mittleren Gruppengröße $E[K]$ für verschiedene Ankunftsprozeßtypen

Abhängig von der Größe k für die E_k -VF und c_A für die H_2 -VF ergeben sich sehr unterschiedliche Werte für c_K , verglichen mit der Poisson-Verteilung für die M-VF. Für große Werte von $E[K] = \lambda T$ erhält man unter Einbeziehung der Gl. (4.28) bzw. (4.33) für alle drei Ankunftsprozeßtypen $c_K = c_A / \sqrt{\lambda T}$.

Mit der Berechnung des Variationskoeffizienten der Gruppengrößen-Verteilung des Ausgangsprozesses wird die Untersuchung des ersten Teilmodells abgeschlossen.

4.3 Das zweite Teilmodell $C^{[x]}/G/1$

Das zweite Teilmodell, welches den taktmäßigen Gruppenankunftsprozeß, die Arbeitswarteschlange und die Bedienungseinheit berücksichtigt, wird durch ein Warteschlangenmodell des Typs $C^{[x]}/G/1$ nachgebildet. Bild 4.6 zeigt die zugehörige Modellstruktur.

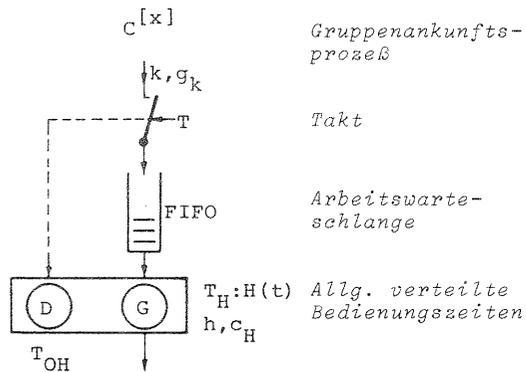


Bild 4.6. Das zweite Teilmodell $C^{[x]}/G/1$

Der Ankunftsprozeß entspricht dem Ausgangsprozesse des ersten Teilmodells. Jeweils zu den Taktzeitpunkten trifft eine Gruppe mit einer zufälligen Anzahl von $k \geq 0$ Anforderungen mit der Wahrscheinlichkeit g_k ein. Alle eintreffenden Anforderungen werden in der als unbegrenzt angenommenen Arbeitswarteschlange abgelegt. Die wartenden Anforderungen werden nach der Abfertigungsstrategie FIFO bedient. Für die Übernahme einer Gruppe wird im Modell eine konstante Verwaltungszeit T_{OH} berücksichtigt, die in realen Systemen der Bearbeitungszeit der Unterbrechungsroutine entspricht. Eine zum Taktzeitpunkt in Bedienung befindliche Anforderung

wird sofort unterbrochen und erst nach Ablauf der Verwaltungszeit T_{OH} weiterbedient.

Analytische Untersuchungen dieses Modells für spezielle Bedienungs-VF wurden in [9, 10, 12] durchgeführt. Die Untersuchung des Modells $C^{[x]}/D/1$ in [9] und $C^{[x]}/M/1$ -s in [10] berücksichtigt keine Verwaltungszeit. In [12] wird das Modell $C^{[x]}/E_k/1$ -s unter Einbeziehung einer Verwaltungszeit und endlicher Arbeitswarteschlangenlänge auf ein Ersatzmodell mit exponentiell verteilten Bedienungszeiten zurückgeführt. Für die Gruppengrößen-Verteilung des Ankunftsprozesses wird bei den genannten Untersuchungen eine Poisson-Verteilung vorausgesetzt. Aufgrund dieser Poisson-Verteilung und den speziellen Bedienungs-VF sind die Taktzeitpunkte stets Regenerationszeitpunkte des Systemzustandsprozesses. Allen Untersuchungen liegt deshalb die Methode der eingebetteten Markoff-Kette zugrunde (vgl. Abschnitt 3.1.1.3a), und man erhält exakte Ergebnisse für die charakteristischen Verkehrsgrößen.

Die hier durchgeführten Untersuchungen berücksichtigen allgemeine Gruppengrößen-Verteilungen und allgemein verteilte Bedienungsdauern. Hierbei ist die für eine exakte analytische Berechnung notwendige Bedingung der Gedächtnislosigkeit zu den Taktzeitpunkten nicht mehr gegeben. Die durchgeführten approximativen analytischen Untersuchungen stützen sich deshalb auf die Methoden der Erneuerungstheorie, d. h. bei Betrachtung des Systems zu den Taktzeitpunkten und unter stationären Voraussetzungen wird für die Zeitdauer bis zur nächsten Systemzustandsänderung (Bedienungs-ende) die Vorwärts-Rekurrenzzeit eines stationären Erneuerungsprozesses angenommen (vgl. Abschnitt 4.2.1). Modelle mit den oben genannten speziellen Voraussetzungen für den Ankunfts- und Bedienungsprozeß sind als Sonderfälle in den Untersuchungen enthalten. Die Überprüfung der approximativ berechneten Ergebnisse erfolgt stets mit Hilfe der Simulation.

Im folgenden werden drei verschiedene approximative Analysemethoden beschrieben und anschließend bezüglich Aufwand und Genauigkeit der Ergebnisse miteinander verglichen.

4.3.1 Analyse mit Hilfe einer eingebetteten Markoff-Kette kurz vor den Taktzeitpunkten

Entsprechend Bild 4.7 wird das System zu zwei aufeinanderfolgenden Taktzeitpunkten i und $i+1$ beobachtet. Die Wahrscheinlichkeit, daß sich das System kurz vor dem Taktzeitpunkt $i+1$ im Zustand $\{x\}$ befindet, wird als Zustandswahrscheinlichkeit $p_{i+1}(x, T)$ bezeichnet.

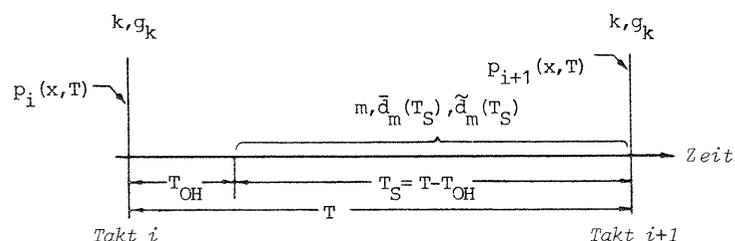


Bild 4.7. Beobachtungszeitpunkte kurz vor dem Takt

Der Gruppenankunftsprozeß, der dem Ausgangsprozeß des ersten Teilmodells entspricht, wurde im Abschnitt 4.2.1 ausführlich untersucht. Dabei wurden die Wahrscheinlichkeiten g_k , d. h. die ankommende Gruppe besteht aus genau k Anforderungen, mit Hilfe von Erneuerungsannahmen zu den Taktzeitpunkten berechnet.

Zur Bedienung der wartenden Anforderungen steht die Zeitspanne $T_S = T - T_{OH}$ zur Verfügung. Da das System nur zu den aufeinanderfolgenden Taktzeitpunkten beobachtet wird, kann in dieser Zeitspanne mit der Wahrscheinlichkeit $\bar{d}_m(T_S)$ bzw. $\tilde{d}_m(T_S)$ eine Gruppe von m Anforderungen bedient werden. $\tilde{d}_m(T_S)$ berücksichtigt die Restbedienzeit T_R einer unterbrochenen Anforderung; bei der Bestimmung von $\bar{d}_m(T_S)$ geht man von einem leeren System kurz vor dem Takt aus. Die Berechnung dieser Wahrscheinlichkeiten, die einen Gruppenbedienungsprozeß charakterisieren, kann analog zur Berechnung des Gruppenausgangsprozesses nach Abschnitt 4.2.1

durchgeführt werden. Wegen der "Gedächtnislosigkeit" Markoff'scher Prozesse sind die Wahrscheinlichkeiten $\bar{d}_m(t)$ und $\tilde{d}_m(t)$ für exponentiell verteilte Bedienungsdauern identisch und entsprechen einer Poisson-Verteilung. Für Erlang- k -verteilte Bedienungsdauern kann $\bar{d}_m(t)$ nach Gl. (4.13) und $\tilde{d}_m(t)$ nach Gl. (4.14) berechnet werden, wenn dort die Ankunftsrate λ durch die Bedienrate μ und T durch t ersetzt wird.

Aufgrund der Erneuerungsannahmen, sowohl für die Berechnung der Gruppengrößen-Verteilung des Ankunfts- und Bedienungsprozesses, können die Taktzeitpunkte quasi als Regenerationszeitpunkte des Systemzustandsprozesses betrachtet werden. Durch diese Approximation ist nun die Berechnung der Zustandswahrscheinlichkeiten $p(x, T)$ kurz vor den Taktzeitpunkten mit Hilfe einer eingebetteten Markoff-Kette möglich.

4.3.1.1 Bestimmung der Zustandswahrscheinlichkeiten

Im stationären Gleichgewichtszustand erhält man für die Zustandswahrscheinlichkeiten kurz vor dem Takt

$$p(x, T) = \sum_{j=0}^N p(j, T) q_{jx}(T), \quad x = 0, 1, \dots, N. \quad (4.35)$$

Eine Begrenzung des an sich unbegrenzten Zustandsraumes (maximal N Anforderungen im System) ist wegen der iterativen Bestimmung der Zustandswahrscheinlichkeiten zwingend notwendig. Wählt man N genügend groß, dann ist der durch die Begrenzung der Arbeitswarteschlange auf $s = N - 1$ Warteplätze bedingte Fehler vernachlässigbar klein. Die Übergangswahrscheinlichkeiten $q_{jx}(T)$, die nur vom Gruppenankunfts- und -bedienungsprozeß abhängig sind, können aus kombinatorischen Überlegungen direkt bestimmt werden und sind im Anhang 2 angegeben.

Durch Einsetzen dieser Übergangswahrscheinlichkeiten in die Gl. (4.35) und unter Einbeziehung der Normierungsbedingung

$$\sum_{x=0}^N p(x,T) = 1 \quad (4.36)$$

erhält man ein lineares Gleichungssystem für die $N+1$ Zustandswahrscheinlichkeiten $p(x,T)$. Zur Auflösung dieses Gleichungssystems wurde das Relaxationsverfahren [43] mit Überrelaxation verwendet.

Die Zustandswahrscheinlichkeiten $p(x,0)$ kurz nach den Taktzeitpunkten erhält man aus der Faltung der Zustandswahrscheinlichkeiten $p(x,T)$ kurz vor dem Takt und den Gruppengrößenwahrscheinlichkeiten g_k .

$$p(x,0) = \sum_{j=0}^x p(j,T) g_{x-j}, \quad x = 0, 1, \dots, N-1, \quad (4.37a)$$

$$p(N,0) = \sum_{j=0}^N p(j,T) g_{\geq(N-j)}, \quad x = N. \quad (4.37b)$$

Zur Bestimmung der Zustandswahrscheinlichkeiten $p(x,t)$ zu einem beliebigen Zeitpunkt t nach dem Takt muß zwischen den beiden Zeitabschnitten

1. $0 < t < T_{OH}$
2. $T_{OH} < t < T$

unterschieden werden (vgl. Bild 4.8).

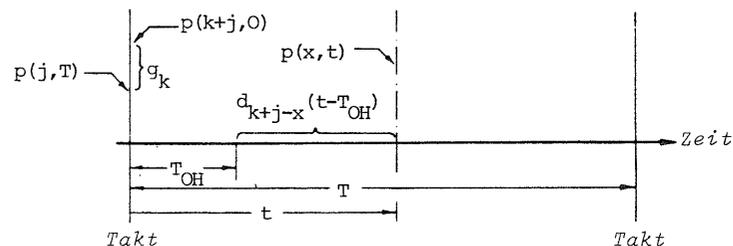


Bild 4.8. Zufälliger Beobachtungszeitpunkt

1. $0 < t < T_{OH}$

Da während der Overheadphase keine Anforderungen bedient werden, sind die Zustandswahrscheinlichkeiten innerhalb dieser Zeitspanne identisch den Zustandswahrscheinlichkeiten kurz nach dem Takt entsprechend Gl. (4.37).

$$p(x,t) = p(x,0), \quad x = 0, 1, \dots, N. \quad (4.38a)$$

2. $T_{OH} < t < T$

Zur Berechnung der Zustandswahrscheinlichkeiten $p(x,t)$ kann nicht von den Zustandswahrscheinlichkeiten $p(x,0)$ kurz nach dem Takt ausgegangen werden, da dann keine Unterscheidung mehr möglich ist, ob das System kurz vor dem Takt leer war oder nicht. Da aber die Fortsetzung der Bedienung nach Ablauf der Verwaltungszeit vom Systemzustand vor dem Takt abhängig ist, müssen zur Bestimmung der $p(x,t)$ die Zustandswahrscheinlichkeiten $p(j,T)$ herangezogen werden. Die Wahrscheinlichkeiten $p(x,t)$ sind deshalb direkt aus der Gl. (4.35) berechenbar, wenn für die Übergangswahrscheinlichkeiten nach Anhang 2 $T = t$ gesetzt wird.

$$p(x,t) = \sum_{j=0}^N p(j,T) q_{jx}(t), \quad x = 0, 1, \dots, N. \quad (4.38b)$$

Aus den Zustandswahrscheinlichkeiten $p(x,t)$ werden im folgenden die charakteristischen Verkehrsgrößen

- mittlere Warteschlangenlänge Ω
- mittlere Wartezeit w bis zum Bedienungsbeginn
- mittlere Gesamtwartezeit w_G
- mittlere Durchlaufzeit t_f
- Wartewahrscheinlichkeit W bzw. W_0

abgeleitet.

4.3.1.2 Berechnung der charakteristischen Größen

a) Mittlere Warteschlangenlänge Ω

Zur Bestimmung der mittleren Warteschlangenlänge Ω wird zunächst die Wartebelastung $\Omega(t)$ zu einem beliebigen Zeitpunkt t berechnet. $\Omega(t)$ ist direkt aus den Zustandswahrscheinlichkeiten $p(x,t)$ für die beiden Intervalle $0 < t < T_{OH}$ und $T_{OH} < t < T$ berechenbar.

1. $0 < t < T_{OH}$

Da während dieser Zeitspanne keine Anforderungen bedient werden, ist $\Omega(t)$ gleich der Wartebelastung $\Omega(0)$ kurz nach dem Takt. Unter der Bedingung, daß die Verlustwahrscheinlichkeit durch geeignete Wahl von $N = s + 1$ vernachlässigbar klein bleibt, erhält man aus den Zustandswahrscheinlichkeiten $p(x,0)$ nach Gl. (4.37) explizit

$$\Omega(0) = \sum_{x=2}^N (x-1)p(x,0) \approx E[X,T] + E[K] - 1 + p(0,T)g_0 \quad (4.39a)$$

2. $T_{OH} < t < T$

In diesem Bereich gilt für die Wartebelastung

$$\Omega(t) = \sum_{x=2}^N (x-1)p(x,t) \quad (4.39b)$$

Durch Mittelwertbildung über das gesamte Taktintervall T ergibt sich für die mittlere Warteschlangenlänge

$$\Omega = \frac{1}{T} \int_0^T \Omega(t) dt \quad (4.40)$$

Nach Einsetzen der Zustandswahrscheinlichkeiten $p(x,t)$ in die Gl. (4.39b) und Integration über die zwei Zeitabschnitte folgt schließlich

$$\begin{aligned} \Omega &= \Omega(0) \frac{T_{OH}}{T} + \frac{1}{T} \sum_{j=0}^N p(j,T) \sum_{x=2}^N (x-1) \int_{T_{OH}}^T q_{jx}(t) dt \\ &= \Omega(0) \frac{T_{OH}}{T} + \frac{1}{T} \sum_{j=0}^N p(j,T) \cdot c_j \end{aligned} \quad (4.41)$$

mit

$$\begin{aligned} c_0 &= \sum_{x=2}^N (x-1) \left[\sum_{k=x}^{N-1} g_k \bar{d}_{k-x}^I(T_S) + g_{\geq N} \bar{d}_{N-x}^I(T_S) \right], \\ c_j &= \sum_{x=2}^N (x-1) \left[\sum_{k=\max(0, x-j)}^{N-j-1} g_k \bar{d}_{k+j-x}^I(T_S) + g_{\geq (N-j)} \bar{d}_{N-x}^I(T_S) \right], \quad j > 0 \end{aligned}$$

und

$$\sum_1^j (\cdot) = 0 \quad \text{für } j < i.$$

Die Bestimmung der durch I gekennzeichneten Größen, die sich nach Einsetzen der Übergangswahrscheinlichkeiten durch die Integration ergeben, ist im Anhang 3 durchgeführt.

Für exponentiell verteilte Bedienungszeiten besitzt der Bedienungsprozeß zu jedem Zeitpunkt die Markoff-Eigenschaft der Gedächtnislosigkeit. Unter dieser Voraussetzung kann zur Berechnung der Zustandswahrscheinlichkeiten $p(x,t)$ direkt von den Zustandswahrscheinlichkeiten $p(x,0)$ kurz nach den Taktzeitpunkten ausgegangen werden und man erhält den einfacheren Zusammenhang

$$p(x,t) = \sum_{j=x}^N p(j,0) \bar{d}_{j-x}(t), \quad x = 0, 1, \dots, N.$$

Setzt man diese $p(x,t)$ in die Gl. (4.39b) ein, dann liefert die Integration der Gl. (4.40) für die Warteschlangenlänge

$$\Omega = \Omega(0) \frac{T_{OH}}{T} + \frac{h}{T} \sum_{j=2}^N p(j,0) \sum_{x=2}^j (x-1) \bar{d}_{\geq(j-x+1)}(T_S) \quad (4.42a)$$

Die Integration über die Poisson-Therme ist im Anhang 3 durchgeführt.

Der erste Summand dieser Gleichung berücksichtigt den Einfluß der Verwaltungszeit, der zweite Summand entspricht der mittleren Warteschlangenlänge Ω_0 für ein System mit der reduzierten Taktdauer $T_S = T - T_{OH}$ und unveränderter Gruppengrößen-Verteilung. Mit Ω_0 für den zweiten Summanden läßt sich Ω einfacher darstellen,

$$\Omega = \Omega(0) \frac{T_{OH}}{T} + \Omega_0 . \quad (4.42b)$$

b) Mittlere Wartezeit w bis zum Bedienungsbeginn

Mit dem Little'schen Theorem kann die mittlere Wartezeit w direkt aus der mittleren Warteschlangenlänge Ω berechnet werden. Mit $\lambda T = E[K]$ und der Festlegung $w_0 = \Omega_0 / \lambda$ erhält man

$$w = \frac{\Omega(0)}{E[K]} T_{OH} + w_0 . \quad (4.43)$$

Diese Wartezeit entspricht der mittleren Wartezeit bis zum Bedienungsbeginn, bezogen auf alle ankommenden Anforderungen.

c) Mittlere Gesamtwartezeit w_G

Ist die Verwaltungszeit T_{OH} größer Null, dann ergibt sich durch die taktmäßige Unterbrechung der Bedienungen ein weiterer Wartezeitanteil, der entweder der mittleren Bedienungsdauer h oder der mittleren Wartezeit w zugeschlagen werden kann. Addiert man diesen Anteil zur Wartezeit w , dann ergibt sich die "effektive" Gesamtwartezeit w_G , die alle Warteanteile einer Anforderung enthält. Berücksichtigt man diesen Anteil bereits bei der Berechnung einer mittleren Gesamtwartebelastung Ω_G , dann ist $\Omega(0)$ in Gl. (4.42b) durch $E[X,0]$, d. h. durch die mittlere Anzahl von wartenden Anforderungen nach dem Takt zu ersetzen.

Mit $E[X,0] = E[X,T] + E[K]$ erhält man analog zur Gl. (4.43)

$$w_G = \left\{ \frac{E[X,T]}{E[K]} + 1 \right\} T_{OH} + w_0 .$$

Die mittlere Wartezeit $w_0 = \Omega_0 / \lambda$, die der Wartezeit in einem System mit der Taktperiode $T_S = T - T_{OH}$ bei unveränderter Gruppengrößen-Verteilung entspricht, kann in zwei Anteile zerlegt werden:

- mittlere Wartezeit w_{01} des Gruppenersten
- mittlere Wartezeit w_{02} der restlichen Anforderungen der Gruppe nach Bedienungsbeginn des Gruppenersten

$$w_0 = w_{01} + w_{02} . \quad (4.44)$$

Die Wartezeit w_{01} entspricht der Bedienungszeit aller zum Taktzeitpunkt im System angetroffenen Anforderungen. Für allgemein verteilte Bedienungszeiten muß jedoch aufgrund der Erneuerungsannahme zum Taktzeitpunkt eine Restbedienungszeit berücksichtigt werden. Lediglich für deterministische Bedienungs-VF ($G=D$) mit der Synchronisationsbedienungs

$$j = \left[\frac{T - T_{OH}}{h} \right] , \quad j \text{ ganzzahlig}$$

(diese Bedingung wird bei der Untersuchung in [9] vorausgesetzt) und für exponentiell verteilte Bedienungszeiten ($G=M$, hier ist die mittlere Restbedienungszeit h_R gleich der mittleren Bedienungszeit h) läßt sich w_{01} durch die folgende Gleichung exakt angeben

$$w_{01} = E[X,T] h . \quad (4.45)$$

Für die mittlere Wartezeit w_{02} gilt nach [9,44] für beliebige Gruppengrößen-Verteilungen und Bedienungs-VF

$$w_{02} = \frac{h}{2} \left\{ \frac{\text{Var}[K]}{E[K]} + E[K] - 1 \right\} . \quad (4.46)$$

Setzt man w_{01} und w_{02} in die Gl. (4.44) ein, so erhält man schließlich für die mittlere Gesamtwartezeit

$$w_G = \left\{ \frac{E[X,T]}{E[K]} + 1 \right\} T_{OH} + h E[X,T] + \frac{h}{2} \left\{ \frac{\text{Var}[K]}{E[K]} + E[K] - 1 \right\} . \quad (4.47)$$

Wie oben erwähnt, liefert diese Gleichung nur für spezielle Bedienungs-VF exakte Ergebnisse. Der entscheidende Vorteil dieser Gleichung gegenüber der Gl. (4.43) besteht nun darin, daß man zur Berechnung der mittleren Gesamtwartezeit nicht die aufwendige Be-

rechnung der mittleren Warteschlangenlänge Ω_0 durchführen muß, sondern nur die mittlere Anzahl von Anforderungen $E[X,T]$ kurz vor dem Takt, den Mittelwert $E[K]$ und die Varianz $\text{Var}[K]$ der Gruppengrößen-Verteilung benötigt. Die Berechnung von $E[K]$ und $\text{Var}[K]$ des Gruppenankunftsprozesses wurde ausführlich im Abschnitt 4.2.2 diskutiert, wenn die Ankunftsabstände im ersten Teilmodell durch eine M -, E_k - oder H_2 -VF beschreibbar sind.

d) Mittlere Durchlaufzeit t_f

Da die mittlere Gesamtwarezeit w_G bereits alle durch die Verwaltungszeit bedingten Wartezeiten einer Anforderung berücksichtigt, ergibt sich für die Durchlaufzeit (flow time) t_f die einfache Beziehung

$$t_f = w_G + h \quad (4.48)$$

e) Wartewahrscheinlichkeit W bzw. W_0

Ist die Verwaltungszeit $T_{OH} > 0$, dann muß eine eintreffende Anforderung mit der Wahrscheinlichkeit $W=1$ warten, da die Bedienungseinheit durch die Bearbeitung des Unterbrechungsprogramms belegt wird. Da W somit keine weitere Information über das System liefert, wird hier die Wahrscheinlichkeit W_0 berechnet, mit der eine zum Taktzeitpunkt eintreffende Anforderung wenigstens eine Anforderung im System antrifft, und damit nach Ablauf der Verwaltungszeit noch warten muß. Wegen des Gruppenankunftsprozesses muß die erste Anforderung der Gruppe betrachtet werden, alle anderen Anforderungen der gleichen Gruppe müssen mit Wahrscheinlichkeit "1" warten. Für die Wahrscheinlichkeit, daß eine beliebige Anforderung einer Gruppe mit $k > 0$ Anforderungen die erste Anforderung ist, gilt

$$q = \frac{1}{E[K|K>0]} = \frac{1 - g_0}{E[K]}$$

mit g_0 nach Gl. (4.14b) bzw. Gl. (4.15).

Die erste Anforderung einer Gruppe muß nach Ablauf der Verwaltungszeit nicht warten, wenn kurz vor dem Taktzeitpunkt keine An-

forderung mehr im System ist. Für diese komplementäre Wartewahrscheinlichkeit $1 - W_0$ gilt

$$1 - W_0 = p(O,T)q = p(O,T) \frac{1 - g_0}{E[K]}$$

und damit

$$W_0 = 1 - p(O,T) \frac{1 - g_0}{E[K]} \quad (4.49)$$

Voraussetzung ist wiederum, daß die Verlustwahrscheinlichkeit vernachlässigbar klein ist.

Mit der Berechnung der Wartewahrscheinlichkeit W_0 wird die Analyse mit Hilfe der eingebetteten Markoff-Kette abgeschlossen. Voraussetzung für diese Analyse ist die Kenntnis der Gruppengrößen-Verteilung des Ankunftsprozesses und die Gruppenbedienungs-Verteilung für ein Taktintervall. Wie dargestellt, ist die Berechnung dieser Größen nur für spezielle Prozeßtypen (D, M, E_k) mit vertretbarem numerischem Aufwand möglich. Insbesondere für hyperexponentielle Bedienungs-VF, die bei der im Kap. 5 durchgeführten approximativen Analyse mehrklassiger Modelle eine wichtige Rolle spielen, ist dieses Verfahren mit vertretbarem Aufwand nicht mehr anwendbar. Zusätzlich ergeben sich für die iterative Auflösung des linearen Gleichungssystems zur Bestimmung der Zustandswahrscheinlichkeiten $p(x,T)$ kurz vor den Taktzeitpunkten erhebliche Rechenzeiten.

Zur Untersuchung des Modells unter allgemeinen Voraussetzungen wurde deshalb das im folgenden beschriebene Analyseverfahren entwickelt. Grundlage dieses Verfahrens ist die Lindley'sche Integralmethode. Für allgemein verteilte Bedienungszeiten und Gruppengrößen-Verteilungen des Ankunftsprozesses ist die mittlere Anzahl im System $E[X,T]$ kurz vor dem Takt berechenbar. Durch Einsetzen von $E[X,T]$ in die Gl. (4.47) ergibt sich daraus die mittlere Gesamtwarezeit w_G , aus der sofort die mittlere Gesamtwarebelastung Ω_G und die Durchlaufzeit t_f berechenbar sind. Das Verfahren läßt auch die Bestimmung der Wartewahrscheinlichkeit W_0 zu.

4.3.2 Analyse mit Hilfe der Lindley'schen Integralmethode

Die Lindley'sche Integralmethode, die im Abschnitt 3.1.1.3d kurz dargestellt wurde, gehört zu den Analysemethoden Nichtmarkoff'scher Prozesse. Die Methode wurde von Lindley [27] entwickelt und ist für Wartesysteme des Typs GI/G/1 anwendbar. Eine direkte Anwendung auf Modelle mit Gruppenankünften, wie das zu untersuchende Modell des Typs $C^{[x]}/G/1$, ist nicht möglich. Betrachtet man jedoch die gesamte ankommende Gruppe wie eine zusammengesetzte Einzelanforderung (super request), so kann das Modell auf ein Ersatzmodell des Typs $C/G_x/1$ zurückgeführt werden. Die VF des Ersatzbedienungsprozesses G_x berücksichtigt hierbei die Bedienung der gesamten zum Taktzeitpunkt eintreffenden Gruppe mit $k \geq 0$ Anforderungen. Aus der Sicht der Bedienungseinheit bringt jede Anforderung einen gewissen Arbeitsumfang in das System. Grundlage des Analyseverfahrens ist deshalb die Bestimmung der angetroffenen Restarbeit im System kurz vor den Taktzeitpunkten.

4.3.2.1 Bestimmung der Restarbeit im System kurz vor den Taktzeitpunkten

Analog zur Untersuchung im Abschnitt 4.3.1 wird das System zu zwei aufeinanderfolgenden Taktzeitpunkten i und $i+1$ betrachtet. Hier wird jedoch die Restarbeit im System kurz vor den Taktzeitpunkten, die durch die kontinuierliche ZV T_U charakterisiert wird, genauer untersucht. Diese Restarbeit entspricht genau der Gesamtbedienungs-dauer aller zum Taktzeitpunkt im System befindlichen Anforderungen und somit genau der Wartezeit, welche die erste Anforderung der ankommenden Gruppe erfährt. Die Begriffe Systemarbeit und zugehörige Bedienungszeit sind quasi austauschbar. Bild 4.9 zeigt den zeitlichen Verlauf der Systemarbeit im betrachteten Taktintervall. Die ZV T_B beschreibt die gesamte zum Taktzeitpunkt eintreffende Arbeit, die der zufälligen Bedienungsdauer aller Anforderungen der ankommenden Gruppe entspricht.

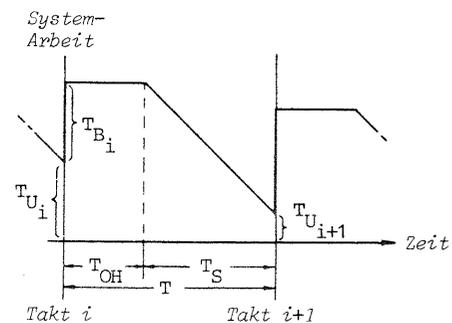


Bild 4.9. Zeitlicher Verlauf der Systemarbeit

Zur Berechnung werden die folgenden VF definiert:

$$U_i(t) = P\{T_{U_i} \leq t\} \\ = P\{\text{Gesamtbedienungs-dauer aller zum Taktzeitpunkt im System befindlichen Anforderungen} \leq t\}$$

$$B_i(t) = P\{T_{B_i} \leq t\} \\ = P\{\text{Gesamtbedienungs-dauer aller zum Taktzeitpunkt eintreffenden Anforderungen} \leq t\}$$

Betrachtet man die Restarbeit kurz vor den beiden Taktzeitpunkten, so gilt die Beziehung

$$T_{U_{i+1}} = \{T_{U_i} + T_{B_i} - T_S\}^+ \quad (4.50)$$

mit

$$\{T\}^+ = \text{Max}\{T, 0\} .$$

Die Restarbeit $T_{U_{i+1}}$ zum Taktzeitpunkt $i+1$ setzt sich additiv zusammen aus der Restarbeit T_{U_i} und der eintreffenden Arbeit T_{B_i} zum Taktzeitpunkt i , reduziert um die konstante, effektiv vom System bewältigbare Arbeit $T_S = T - T_{OH}$. Ist die Summe T_{U_i} plus T_{B_i} kleiner als T_S , dann ist $T_{U_{i+1}}$ gleich Null, d. h. die Restarbeit

kurz vor den Taktzeitpunkten kann nicht negativ werden.

Durch Einführung einer neuen ZV T_{C_i} , die definiert ist als

$$T_{C_i} = T_{B_i} - T_S \quad (4.51a)$$

mit der VF

$$C_i(t) = P\{T_{C_i} \leq t\}, \quad -T_S \leq t < \infty \quad (4.51b)$$

und bei Betrachtung des Systems im stationären Gleichgewichtszustand, d. h.

$$U_i(t) = U_{i+1}(t) = U(t)$$

erhält man für die VF der Restarbeit kurz vor den Taktzeitpunkten die folgende Integralgleichung:

$$U(t) = \begin{cases} \int_{-\infty}^t U(t-\tau) dC(\tau) = \int_{-\infty}^t U(t-\tau) c(\tau) d\tau, & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (4.52)$$

mit $c(t)$ als VDF der ZV T_C .

Eine detaillierte Herleitung dieser Integralgleichung, die der Form nach einer Wiener-Hopf Integralgleichung entspricht, findet man z. B. in [18]. Sie beschreibt prinzipiell die VF der Wartezeit für das Modell des Typs GI/G/1. Für spezielle Ankunfts- und Bedienungsprozesse kann die Gleichung mit Methoden der Funktionentheorie gelöst werden. Diese Lösungsmethode ist für die Warteschlangensysteme M/M/1 und M/G/1 in [18] ausführlich dargestellt. Für das zu untersuchende Ersatzmodell des Typs C/G_x/1 ist keine Lösung bekannt. Die Grundidee des hier verwendeten Lösungsansatzes ist die Zeitdiskretisierung der kontinuierlichen Verteilungsfunktion.

4.3.2.2 Lösungsansatz durch Zeitdiskretisierung

Für die weiteren Betrachtungen wird die kontinuierliche Zeitachse t in äquidistante Abschnitte der Länge Δt eingeteilt. Alle kontinuierlichen VF werden somit durch diskrete Verteilungen approximiert, daraus ergibt sich für die Dichte eine Folge von gewichteten Dirac-Funktionen im Abstand Δt . Durch Verkleinerung von Δt reduziert sich der Diskretisierungsfehler, im Grenzfall $\Delta t \rightarrow 0$ geht der Fehler ebenfalls gegen Null. Bild 4.10 zeigt die Dichte der Verteilung, die durch die Diskretisierung der Bedienungs-VF $B(t)$ entsteht.

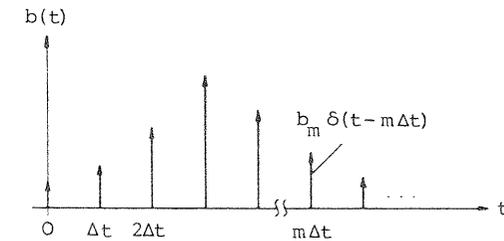


Bild 4.10. Dichte $b(t)$ der diskretisierten VF $B(t)$

Die Dichte $b(t)$ der Verteilung ist darstellbar durch

$$b(t) = B'(t) \triangleq \sum_{m=0}^{\infty} b_m \delta(t - m \Delta t) \quad (4.53)$$

Die Zeitdiskretisierung verschiedener Bedienungs-VF wird im Abschnitt 4.3.2.4 genauer untersucht.

Die Dichte $c(t)$ der diskretisierten VF $C(t)$ ergibt sich nach Gl. (4.51) aus der Dichte $b(t)$ durch Verschiebung um die konstante Zeitspanne T_S nach links (vgl. Bild 4.11).

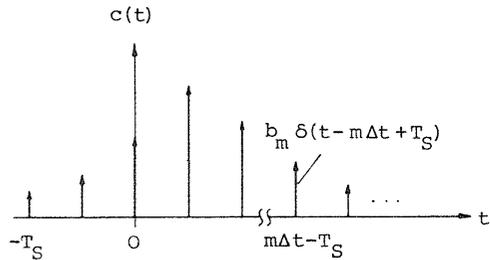


Bild 4.11. Dichte $c(t)$ der diskretisierten VF $C(t)$

Wählt man bei der Diskretisierung Δt so, daß T_S ein ganzzahliges Vielfaches von Δt darstellt,

$$T_S = \ell \cdot \Delta t, \quad \ell \text{ ganzzahlig}, \quad (4.54)$$

dann gilt für die Dichte $c(t)$:

$$c(t) = C'(t) \triangleq \sum_{m=0}^{\infty} b_m \delta(t - m\Delta t + \ell \cdot \Delta t). \quad (4.55)$$

Durch Einsetzen von $c(t)$ in die Gl. (4.52) erhält man

$$U(t) = \begin{cases} \int_{-\infty}^t U(t-\tau) \sum_{m=0}^{\infty} b_m \delta(\tau - m\Delta t + \ell \Delta t) d\tau & t \geq 0 \\ 0 & t < 0. \end{cases}$$

Ersetzt man t aufgrund der Diskretisierung durch $i \Delta t$, vertauscht die Integration und Summation und berücksichtigt die Ausblendeneigenschaft der δ -Funktion, dann gilt

$$U(i \Delta t) = \begin{cases} \sum_{m=0}^{\infty} b_m U(i \Delta t - m \Delta t + \ell \Delta t) & i = 0, 1, \dots \\ 0 & i < 0 \end{cases}$$

Da die Verteilungsfunktion $U(t)$ für negative Argumente identisch Null ist, ergibt sich für die obere Grenze der Summation

$$m_{\max} = i + \ell$$

und damit für die Diskretisierung der VF $U(t)$:

$$U(i) = \sum_{m=0}^{i+\ell} b_m U(i - m + \ell), \quad i = 0, 1, \dots \quad (4.56)$$

$U(i)$ gibt die Wahrscheinlichkeit an, daß die Restarbeit kurz vor den Taktzeitpunkten höchstens den Betrag $i \Delta t$ aufweist. Da diese Restarbeit in Δt -Anteile gequantelt ist, stellt sie eine diskrete Verteilung dar und kann als "Zustand" des Systems interpretiert werden. Bezeichnet man mit $u(i, T)$ die "Zustandswahrscheinlichkeit", daß sich das System kurz vor dem Takt im Zustand $\{i\}$ befindet, d. h. die Restarbeit genau $i \Delta t$ beträgt, dann gilt der Zusammenhang

$$u(i, T) = \sum_{m=0}^{i+\ell} b_m u(i - m + \ell, T), \quad i = 0, 1, \dots \quad (4.57)$$

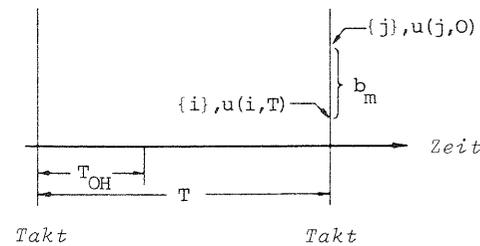


Bild 4.12. Beobachtungszeitpunkt zum Takt

Entsprechend Bild 4.12 erhält man aus der Faltung von $u(i, T)$ und b_m für die Zustandswahrscheinlichkeit

$$u(j, 0) = \sum_{i=0}^j u(i, T) b_{j-i}, \quad j = 0, 1, \dots \quad (4.58)$$

Zur weiteren Betrachtung werden die folgenden erzeugenden Funk-

tionen für die einzelnen Wahrscheinlichkeiten der diskreten Zufallsvariablen definiert:

$$G_i(z, T) = \sum_{i=0}^{\infty} u(i, T) z^i \quad (4.59a)$$

$$G_i(z, 0) = \sum_{i=0}^{\infty} u(i, 0) z^i \quad (4.59b)$$

$$G_m(z) = \sum_{m=0}^{\infty} b_m z^m \quad (4.59c)$$

Durch Einsetzen der erzeugenden Funktionen in die Gl. (4.58) folgt direkt

$$G_i(z, 0) = G_i(z, T) \cdot G_m(z) \quad (4.60)$$

Für die Berechnung der erzeugenden Funktion $G_i(z, T)$, aus der man durch Differentiation an der Stelle $z=1$ sofort die gesuchte mittlere Restarbeit $E[I, T]$ kurz vor dem Takt erhält, ergeben sich einfachere Beziehungen, wenn zur Herleitung vom Systemzustand kurz nach dem Takt ausgegangen wird. Berücksichtigt man, daß mit der Wahrscheinlichkeit "1" genau ℓ Bedienungsphasen der Länge Δt pro Taktintervall zur Verfügung stehen, so erhält man für den stationären Fall folgenden Zusammenhang für die Arbeit im System kurz nach dem Takt (vgl. Bild 4.13):

$$u(0, 0) = \sum_{i=0}^{\ell} u(i, 0) b_0 \quad (4.61a)$$

$$u(j, 0) = \sum_{i=0}^{\ell} u(i, 0) b_j + \sum_{i=\ell+1}^{\ell+j} u(i, 0) b_{\ell+j-i}, \quad j = 1, 2, \dots \quad (4.61b)$$

Die Gl. (4.61) für die Restarbeit entsprechen den Gl. des speziellen Modells $C^{[x]}/D/1$ ohne Overhead mit einer Taktdauer als ganzzahligem Vielfachen der konstanten Bedienungszeit, welche in [9] abgeleitet wurden. Die vorstehende Ableitung zeigt, daß der in [9] durchgeführte Ansatz als Sonderfall des allgemeinen mit zeitdiskretisierter, allgemein verteilter Bedienungszeit aufgefaßt werden kann.

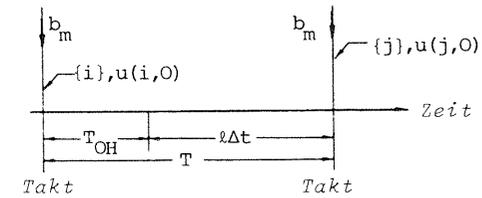


Bild 4.13. Beobachtungszeitpunkte kurz nach dem Takt

Durch Einsetzen der in Gl. (4.59) definierten erzeugenden Funktionen erhält man nach einigen Umformungen:

$$G_i(z, 0) = \frac{\sum_{i=0}^{\ell-1} u(i, 0) (z^\ell - z^i)}{z^\ell - G_m(z)} G_m(z) \quad (4.62)$$

Diese Gleichung enthält noch die unbekannt Wahrscheinlichkeiten $p(0,0), p(1,0), \dots, p(\ell-1,0)$. Mit dem Satz von Rouché wird in [9] gezeigt, daß der Nenner der Gl. (4.62) genau ℓ Nullstellen innerhalb und auf dem Einheitskreis, d. h. $|z| \leq 1$, besitzt. Da die erzeugende Funktion $G_i(z, 0)$ im Bereich $|z| \leq 1$ endlich ist, muß der Zähler der Gl. (4.62) die gleichen Nullstellen besitzen. Bezeichnet man die Nullstellen mit $z_0, z_1, \dots, z_{\ell-1}$ und berücksichtigt man, daß $z_0 = 1$ eine Nullstelle des Nenners ist, dann kann der Zähler in Produktform dargestellt werden und man erhält:

$$G_i(z, 0) = \frac{K(z-1)(z-z_1)\dots(z-z_{\ell-1})}{z^\ell - G_m(z)} G_m(z) \quad (4.63)$$

Die unbekannte Konstante K kann aus der Bedingung $G_i(1, 0) = 1$ unter Anwendung der Regel von l'Hospital bestimmt werden. Es ergibt sich

$$K = \frac{\ell - G_m'(1)}{\prod_{v=1}^{\ell-1} (1 - z_v)} \quad (4.64)$$

Da die erste Ableitung $G'_m(1)$ dem Mittelwert $E[M]$ entspricht, erhält man durch Einsetzen der Gl. (4.64) in die Gl. (4.63) und unter Verwendung der Gl. (4.60) die gesuchte erzeugende Funktion

$$G_i(z, T) = \frac{\ell - E[M]}{\prod_{v=1}^{\ell-1} (1 - z_v)} \cdot \frac{\prod_{v=0}^{\ell-1} (z - z_v)}{z^\ell - G_m(z)} \quad (4.65)$$

mit

ℓ nach Gl. (4.54),

$$E[M] = \sum_{m=0}^{\infty} m b_m ,$$

und

z_v Wurzeln der Gleichung $z^\ell - G_m(z) = 0$.

$E[M]$ ist die mittlere Anzahl von Δt -Arbeitsintervallen, die eine zum Taktzeitpunkt ankommende Gruppe in das System hineinbringt. Da pro Taktintervall maximal ℓ Bedienungsintervalle der Länge Δt zur Verfügung stehen, ist die Stationarität des Systems durch folgende Bedingung gekennzeichnet:

$$E[M] < \ell . \quad (4.66)$$

Vor der Bestimmung der charakteristischen Verkehrsgrößen, die mit Hilfe der erzeugenden Funktion $G_i(z, T)$ ermittelt werden können, wird die erzeugende Funktion $G_m(z)$ genauer untersucht.

4.3.2.3 Berechnung der erzeugenden Funktion $G_m(z)$

Wie oben beschrieben, wurde das gegebene Modell des Typs $C^{[x]}/G/1$ zunächst auf das Ersatzmodell des Typs $C/G_x/1$ zurückgeführt. Der Bedienungsprozeß G_x mit der Bedienungs-VF $B(t)$ berücksichtigt die Bedienung der gesamten zum Taktzeitpunkt ankommenden Gruppe. Durch die anschließende Zeitdiskretisierung wurde diese zusammengesetzte, kontinuierliche VF durch eine diskrete Verteilung approximiert. Für die Berechnung der so entstandenen zusammengesetzten

Verteilung und deren erzeugenden Funktion $G_m(z)$ benötigt man:

1. Die Verteilung, die durch die Diskretisierung der Bedienungs-VF $H(t)$ des gegebenen Modells entsteht (innere Verteilung).
2. Die Verteilung der ankommenden Gruppe (äußere Verteilung).

Die innere Verteilung wird wie folgt definiert:

$$a_n = P\{N=n\} = P \left\{ \begin{array}{l} \text{die Bedienung einer Anforderung benötigt} \\ \text{genau } N=n \text{ Intervalle der Länge } \Delta t \end{array} \right\}$$

mit der erzeugenden Funktion

$$G_n(z) = \sum_{n=0}^{\infty} a_n z^n . \quad (4.67)$$

Die Berechnung dieser erzeugenden Funktion wird für verschiedene Bedienungs-VF im nächsten Abschnitt durchgeführt.

Die äußere Verteilung entspricht der Gruppengrößen-Verteilung mit

$$g_k = P\{K=k\} = P \left\{ \begin{array}{l} \text{die ankommende Gruppe besteht aus} \\ \text{genau } K=k \text{ Anforderungen} \end{array} \right\}$$

mit der erzeugenden Funktion

$$G_k(z) = \sum_{k=0}^{\infty} g_k z^k . \quad (4.68)$$

Besteht die zufällig eintreffende Gruppe aus K Anforderungen, dann benötigt man zur Bedienung aller Anforderungen dieser Gruppe die zufällige Anzahl von

$$M = n_1 + n_2 + \dots + n_K \quad (4.69)$$

Intervallen der Länge Δt ; hierin bedeutet n_i die zufällige Anzahl von Δt -Bedienungsintervallen der i -ten Anforderung.

Mit

$$G_m(z) = \sum_{m=0}^{\infty} P\{M=m\} z^m \quad (4.70)$$

gilt für die bedingte erzeugende Funktion $G_m(z|k)$:

$$G_m(z|k) = \sum_{m=0}^{\infty} P\{M=m|K=k\} z^m \quad (4.71)$$

Durch Multiplikation mit g_k und Summation über alle k folgt

$$\begin{aligned} G_m(z) &= \sum_{m=0}^{\infty} \left(\sum_{k=0}^{\infty} g_k P\{M=m|K=k\} \right) z^m \\ &= \sum_{k=0}^{\infty} g_k \sum_{m=0}^{\infty} P\{M=m|K=k\} z^m = \sum_{k=0}^{\infty} g_k G_m(z|k) \end{aligned} \quad (4.72)$$

Da die Wahrscheinlichkeit $P\{M=m|K=k\}$ der zusammengesetzten Verteilung aus der k -fachen Faltung der Wahrscheinlichkeit $P\{N=n\}$ berechenbar ist (vgl. Gl. (4.69)), ergibt sich $G_m(z|k)$ aus dem k -fachen Produkt der erzeugenden Funktion $G_n(z)$ und man erhält

$$G_m(z) = \sum_{k=0}^{\infty} g_k G_n^k(z) \quad (4.73)$$

Ersetzt man in der erzeugenden Funktion $G_k(z)$ der Gl. (4.68) die Variable z durch $G_n(z)$, so gilt schließlich für die erzeugende Funktion $G_m(z)$ die Beziehung

$$G_m(z) = G_k[G_n(z)] \quad (4.74)$$

Die für die späteren Berechnungen benötigten Größen $E[M]$ und $\text{Var}[M]$ können direkt aus der erzeugenden Funktion bestimmt werden. Mit

$$E[M] = \left. \frac{dG_m(z)}{dz} \right|_{z=1} = \left. \frac{dG_k}{dG_n} \cdot \frac{dG_n}{dz} \right|_{z=1} = G_k'(1) \cdot G_n'(1) ,$$

$G_k'(1) = E[K]$ und $G_n'(1) = E[N]$ erhält man

$$E[M] = E[K] \cdot E[N] \quad (4.75)$$

Das Ergebnis ist direkt interpretierbar: Ist die mittlere Gruppengröße gleich $E[K]$ und benötigt man zur Bedienung einer Anforderung im Mittel $E[N]$ Intervalle der Länge Δt , dann benötigt man zur Bedienung der gesamten Gruppe im Mittel $E[M] = E[K] \cdot E[N]$ Intervalle der Länge Δt .

Zur Berechnung der Varianz der zusammengesetzten Verteilung gilt mit

$$\begin{aligned} G_m''(1) &= \left. \frac{d^2 G_m(z)}{dz^2} \right|_{z=1} = \left. \frac{dG_k}{dG_n} \cdot \frac{d^2 G_n}{dz^2} \right|_{z=1} + \left. \frac{dG_n}{dz} \cdot \frac{d^2 G_k}{dG_n^2} \cdot \frac{dG_n}{dz} \right|_{z=1} \\ &= G_k'(1) \cdot G_n''(1) + G_k''(1) \cdot [G_n'(1)]^2 \end{aligned}$$

und der im Anhang 1 abgeleiteten Beziehung

$$\begin{aligned} \text{Var}[M] &= G_m''(1) + G_m'(1) - [G_m'(1)]^2 \\ \text{Var}[M] &= E[K] \cdot \text{Var}[N] + E[N]^2 \cdot \text{Var}[K] \end{aligned} \quad (4.76)$$

Die noch unbekanntenen Größen $E[N]$ und $\text{Var}[N]$ der diskretisierten Bedienungs-VF $H(t)$ werden im folgenden für verschiedene Bedienungs-VF berechnet.

4.3.2.4 Zeitdiskretisierung verschiedener Bedienungs-VF

Durch die Zeitdiskretisierung wird eine kontinuierliche Bedienungs-VF durch eine diskrete Verteilung approximiert. Bild 4.14 zeigt die Vorgehensweise am Beispiel einer beliebigen Bedienungs-VF $H(t)$.

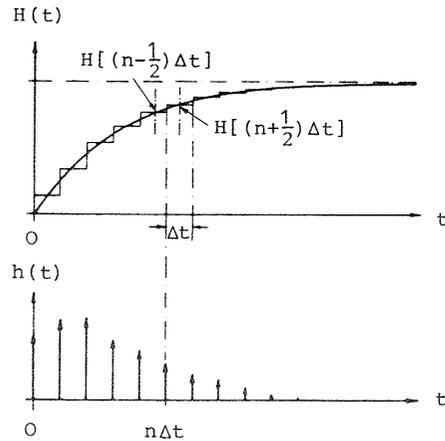


Bild 4.14. Diskretisierung der kontinuierlichen VF $H(t)$

Die kontinuierliche Zeitachse t wird in äquidistante Intervalle der Länge Δt unterteilt und die VF $H(t)$ durch eine "Treppenfunktion" mit der Stufenbreite Δt approximiert. Damit kann $H(t)$ durch eine Folge von Sprungfunktionen unterschiedlicher Höhe, die um ganzzahlige Vielfache von Δt verschoben sind, mathematisch dargestellt werden. Die zugehörige Dichtefunktion $h(t)$ dieser Verteilung ist dann durch eine gewichtete Folge von Dirac-Funktionen mit gleicher Zeitabhängigkeit darstellbar, für die gilt

$$h(t) = \sum_{n=0}^{\infty} a_n \delta(t - n\Delta t) \quad (4.77)$$

Da die Schnittpunkte zwischen der Treppenkurve und der VF $H(t)$ an jeder beliebigen Stelle innerhalb der Stufenbreite Δt liegen können, gibt es theoretisch unendlich viele Möglichkeiten der Annäherung. Die im Bild 4.14 gewählte Realisierung, bei der die Schnittpunkte in der Stufenmitte liegen, ergibt eine gute Übereinstimmung für die ersten beiden Momente bereits bei größer gewählten Δt -Intervallen. Bei dieser Wahl erhält man für die Ampli-

tuden der Dirac-Impulse

$$a_0 = H\left(\frac{\Delta t}{2}\right) \quad (4.78a)$$

$$a_n = H\left[\left(n + \frac{1}{2}\right)\Delta t\right] - H\left[\left(n - \frac{1}{2}\right)\Delta t\right], \quad n = 1, 2, \dots \quad (4.78b)$$

Für verschiedene Bedienungs-VF ($D, M, E_k, E_k^* M, H_k$), die beliebige Variationskoeffizienten $0 \leq c_H < \infty$ zulassen, werden im folgenden die Wahrscheinlichkeiten a_n , daraus die erzeugende Funktion $G_n(z)$ nach Gl. (4.67) und schließlich die Größen $E[N]$ und $\text{Var}[N]$ berechnet.

a) Zeitdiskretisierung der D-Bedienungs-VF

Für die VDF gilt

$$h(t) = \delta(t - h)$$

mit

$$h = \text{mittlere Bedienungsdauer.}$$

Definiert man

$$n_0 = \left[\frac{h}{\Delta t} \right] = \text{das Ganze der Division } \frac{h}{\Delta t}$$

und setzt $H(t)$ in die Gl. (4.78) ein, dann folgt mit

$$a_{n_0} = 1, \quad a_n = 0 \quad \text{für } n \neq n_0$$

für die erzeugende Funktion nach Gl. (4.67)

$$G_n(z) = z^{n_0} \quad (4.79a)$$

Für den Mittelwert $E[N]$ und die Varianz $\text{Var}[N]$ erhält man durch entsprechende Differentiation

$$E[N] = n_0, \quad (4.79b)$$

$$\text{Var}[N] = 0. \quad (4.79c)$$

Die Berechnung der einzelnen Größen für die weiteren Bedienungs-VF verläuft analog. Deshalb werden hier nur die Ergebnisse angegeben.

b) Zeitdiskretisierung der M-Bedienungs-VF

$$H(t) = 1 - e^{-\mu t} .$$

$h = 1/\mu =$ mittlere Bedienungsdauer.

$$G_n(z) = (1-x) \frac{1+zx}{1-zx^2} \quad (4.80a)$$

mit

$$x = e^{-\mu \Delta t / 2} .$$

$$E[N] = \frac{x}{1-x^2} . \quad (4.80b)$$

$$\text{Var}[N] = x \frac{1-x+x^2}{(1-x^2)^2} . \quad (4.80c)$$

c) Zeitdiskretisierung der E_k -Bedienungs-VF

Da die E_k -VF aus k hintereinanderliegenden exponentiell verteilten Phasen mit dem Mittelwert $1/\mu k$ besteht, ergibt sich die erzeugende Funktion aus dem k -fachen Produkt der erzeugenden Funktion für den M-Bedienungsprozess:

$$G_n(z) = (1-x)^k \left(\frac{1+zx}{1-zx^2} \right)^k \quad (4.81a)$$

mit

$$x = e^{-\mu k \Delta t / 2} .$$

$$E[N] = k \frac{x}{1-x^2} . \quad (4.81b)$$

$$\text{Var}[N] = kx \frac{1-x+x^2}{(1-x^2)^2} . \quad (4.81c)$$

Die E_k -VF läßt nur die speziellen Werte des Variationskoeffizienten $c_H = \frac{1}{\sqrt{k}}$ zu. Zur Realisierung beliebiger Variationskoeffizienten im Bereich $0 < c_H < 1$ wird deshalb auf eine modifizierte E_k -VF zurückgegriffen. Die zugehörige Kurznotation wird mit E_k^*M bezeichnet.

d) Zeitdiskretisierung der E_k^*M -Bedienungs-VF

Diese VF besteht aus k hintereinanderliegenden exponentiell verteilten Phasen mit dem gleichen Mittelwert h_1 (E_k -VF) und eine zusätzliche M-Phase mit dem Mittelwert h_2 (vgl. Bild 4.15).

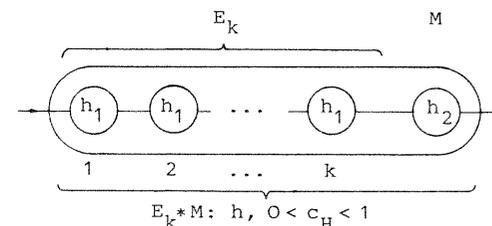


Bild 4.15. Zeitdiskretisierung der E_k^*M -VF

Gibt man den Gesamtmittelwert

$$h = kh_1 + h_2$$

und den Variationskoeffizienten

$$c_H = \frac{\sqrt{kh_1^2 + h_2^2}}{kh_1 + h_2}$$

der VF vor, dann erhält man für die Variablen k , h_1 und h_2 :

$$k = \left[1/c_H^2 \right] = \text{das Ganze der Division } 1/c_H^2$$

$$h_1 = \frac{h}{k+1} \left[1 + \sqrt{c_H^2 + \frac{c_H^2}{k} - \frac{1}{k}} \right] ,$$

$$h_2 = h - kh_1 .$$

Analog zu den Überlegungen für die E_k -VF gilt für die erzeugende Funktion

$$G_n(z) = (1-x_1)^k \left(\frac{1+zx_1}{1-zx_1^2} \right)^k (1-x_2) \frac{1+zx_2}{1-zx_2^2} \quad (4.82a)$$

mit

$$x_1 = e^{-\Delta t/2h_1}, \quad x_2 = e^{-\Delta t/2h_2}.$$

$$E[N] = k \frac{x_1}{1-x_1^2} + \frac{x_2}{1-x_2^2}. \quad (4.82b)$$

$$\text{Var}[N] = kx_1 \frac{1-x_1+x_1^2}{(1-x_1^2)^2} + x_2 \frac{1-x_2+x_2^2}{(1-x_2^2)^2}. \quad (4.82c)$$

e) Zeitdiskretisierung der H_k -Bedienungs-VF

Da die H_k -VF aus k parallelen M -Phasen mit den Verzweigungswahrscheinlichkeiten p_i und den Mittelwerten h_i , $i=1,2,\dots,k$ besteht, ergibt sich unter Verwendung der Gl. (4.80a)

$$G_n(z) = \sum_{i=1}^k p_i (1-x_i) \frac{1+zx_i}{1-zx_i^2} \quad (4.83a)$$

mit

$$x_i = e^{-\Delta t/2h_i}.$$

$$E[N] = \sum_{i=1}^k p_i \frac{x_i}{1-x_i^2}. \quad (4.83b)$$

$$\text{Var}[N] = 2 \sum_{i=1}^k \frac{p_i x_i^3}{(1-x_i^2)^2} + E[N] - E[N]^2. \quad (4.83c)$$

Die für die verschiedenen Bedienungs-VF berechneten Größen $E[N]$ und $\text{Var}[N]$ können nun in die Gl. (4.75) bzw. (4.76) zur Bestimmung des Mittelwertes $E[M]$ und der Varianz $\text{Var}[M]$ eingesetzt werden. Für die weiteren Betrachtungen benötigt man noch die Wurzeln der charakteristischen Gleichung $z^\ell - G_m(z) = 0$.

4.3.2.5 Lösung der charakteristischen Gleichung $z^\ell - G_m(z) = 0$

Bei der Bestimmung der Wurzeln z_ν der charakteristischen Gleichung spielt die Gruppengrößen-Verteilung des Ankunftsprozesses eine entscheidende Rolle. Es wird deshalb unterschieden zwischen

- Poisson-verteilten Gruppengrößen
- Allgemein verteilten Gruppengrößen, die durch den Mittelwert $E[K]$ und die Varianz $\text{Var}[K]$ beschrieben werden (Zwei-Momenten-Approximation).

a) Poisson-verteilte Gruppengrößen

Die erzeugende Funktion einer Poisson-Verteilung mit dem Mittelwert $E[K] = \lambda T$ lautet

$$G_k(z) = e^{-\lambda T(1-z)}. \quad (4.84)$$

Ersetzt man z durch $G_n(z)$, so erhält man nach Gl. (4.74) die erzeugende Funktion $G_m(z)$ und damit die charakteristische Gleichung

$$z^\ell - e^{-\lambda T[1-G_n(z)]} = 0. \quad (4.85)$$

Die ℓ Wurzeln $z_\nu = x_\nu + jy_\nu$ dieser Gleichung im Bereich $|z| \leq 1$ können nur iterativ bestimmt werden. Die zu iterierenden Gleichungen für den Realteil x_ν und Imaginärteil y_ν lauten

$$x_\nu = e^{A \cdot \text{Re}\{G_n(z_\nu)-1\}} \cdot \cos[A \cdot \text{Im}\{G_n(z_\nu)\} + \frac{2\nu\pi}{\ell}], \quad (4.86a)$$

$$y_\nu = e^{A \cdot \text{Re}\{G_n(z_\nu)-1\}} \cdot \sin[A \cdot \text{Im}\{G_n(z_\nu)\} + \frac{2\nu\pi}{\ell}], \quad (4.86b)$$

$$\nu = 0, 1, \dots, \ell - 1$$

mit

$$A = \frac{\lambda T}{\ell}.$$

Die Aufspaltung der erzeugenden Funktion $G_n(z)$ in Real- und Imaginärteil für die verschiedenen Bedienungs-VF ist hier nicht angegeben, sie stellt jedoch keine grundsätzliche Schwierigkeit dar.

b) Allgemein verteilte Gruppengrößen

Für allgemein verteilte Gruppengrößen wird eine Zwei-Momenten-Approximation vorgenommen. Die hierfür notwendigen Angaben sind der Mittelwert $E[K]$ und die Varianz $\text{Var}[K]$ der Gruppengrößen-Verteilung. Die Berechnung dieser Größen wurde im Abschnitt 4.2.2 für verschiedene Ankunftsprozesse des ersten Teilmodells (M, E_k, H_2) durchgeführt. Wird das zweite Teilmodell für sich betrachtet, d. h. ist der Ankunftsprozeß des zweiten Teilmodells nicht der Ausgangsprozeß des ersten Teilmodells, dann können beliebige Werte für $E[K]$ und $\text{Var}[K]$ vorgegeben werden.

Aus $E[M]$ nach Gl. (4.75) und $\text{Var}[M]$ nach Gl. (4.76) erhält man für den Variationskoeffizienten c_M der Bedienungs-VF $B(t)$ des Ersatzmodells $C/G_x/1$

$$c_M = \sqrt{c_K^2 + \frac{c_N^2}{E[K]}} \approx \sqrt{c_K^2 + \frac{c_H^2}{E[K]}} \quad (4.87)$$

c_K ist direkt aus $E[K]$ und $\text{Var}[K]$ bestimmbar, c_H ist der Variationskoeffizient der Bedienungs-VF $H(t)$ des gegebenen Modells $C^{[x]}/G/1$. Aus der mittleren Bedienungsdauer h der VF $H(t)$ und der mittleren Gruppengröße $E[K]$ erhält man für die mittlere Bedienungsdauer der VF $B(t)$ des Ersatzbedienungsprozesses

$$h_M = h E[K] \quad .$$

Durch diesen Mittelwert h_M und den Variationskoeffizienten c_M nach Gl. (4.87) kann nun die zusammengesetzte Bedienungs-VF $B(t)$ des Ersatzmodells durch zwei Momente approximiert werden. Abhängig von der Größe des Variationskoeffizienten c_M ($c_M < 1$, $c_M = 1$, $c_M > 1$) ergibt sich für die Approximation ein spezieller VF-Typ ($E_k * M, M, H_2$), der anschließend wie im Abschnitt 4.3.2.4 dargestellt zeitdiskretisiert wird. Die gesuchte erzeugende Funktion $G_m(z)$ der diskretisierten Verteilung ist dann identisch der jeweiligen erzeugenden Funktion $G_n(z)$, wenn bei den verschiedenen VF-Typen für die mittlere Bedienstzeit $h = h_M$ eingesetzt wird. Bei der Approximation durch eine H_2 -VF ($c_M > 1$) wird hierbei auf die übliche Zusatzbedingung $p_1 h_1 = p_2 h_2$ zurückgegriffen, d. h. bei der Berech-

nung von x_i , $i = 1, 2$, nach Gl. (4.83a) ist für die Mittelwerte h_1 und h_2 einzusetzen:

$$h_1 = \frac{h}{2p} \quad , \quad (4.88)$$

$$h_2 = \frac{h}{2(1-p)} \quad ,$$

mit

$$p = \frac{1}{2} \left(1 - \sqrt{1 - \frac{2}{c_M^2 + 1}} \right) \quad .$$

Damit ist die erzeugende Funktion $G_m(z)$ bekannt und die Wurzeln der charakteristischen Gleichung

$$z^\ell - G_m(z) = 0$$

können berechnet werden. Die Bestimmung dieser Wurzeln $z_\nu = d_\nu e^{j\varphi_\nu}$, die nach Betrag d_ν und Phase φ_ν vorgenommen wurde, ist durch Iteration der Gl. (4.89a) und Gl. (4.89b) möglich.

$$d_\nu = \left[\text{Re}\{G_m(z_\nu)\}^2 + \text{Im}\{G_m(z_\nu)\}^2 \right]^{\frac{1}{2\ell}} \quad (4.89a)$$

$$\varphi_\nu = \frac{2\nu\pi}{\ell} + \frac{1}{\ell} \arctan \frac{\text{Im}\{G_m(z_\nu)\}}{\text{Re}\{G_m(z_\nu)\}} \quad (4.89b)$$

$$\nu = 0, 1, 2, \dots, \ell - 1 \quad .$$

Nach der Bestimmung der Wurzeln sind alle Werte bekannt, die für die Berechnung der charakteristischen Verkehrsgrößen benötigt werden.

4.3.2.6 Berechnung der charakteristischen Größen

Die mittlere Anzahl im System $E[X, T]$ ist direkt aus der mittleren Anzahl $E[I, T]$ von Δt -Intervallen kurz vor dem Taktzeitpunkt berechenbar. Da zur Bedienung einer Anforderung im Mittel $E[N]$ In-

tervalle der Länge Δt erforderlich sind, gilt der Zusammenhang

$$E[X,T] = \frac{E[I,T]}{E[N]} \quad (4.90)$$

Für die mittlere Anzahl $E[I,T]$ von Δt -Intervallen kurz vor dem Takt gilt

$$E[I,T] = \left. \frac{dG_{\perp}(z,T)}{dz} \right|_{z=1}$$

und man erhält nach einigen Umformungen:

$$E[I,T] = \sum_{\nu=1}^{\ell-1} \frac{1}{1-z_{\nu}} + \frac{1}{2} \left\{ 1 - E[M] - \ell - \frac{\text{Var}[M]}{E[M] - \ell} \right\} \quad (4.91a)$$

mit

$$\begin{aligned} \ell & \text{ nach Gl. (4.54),} \\ E[M] & \text{ nach Gl. (4.75),} \\ \text{Var}[M] & \text{ nach Gl. (4.76),} \\ z_{\nu} & \text{ nach Abschnitt 4.3.2.5 .} \end{aligned}$$

Die Wurzeln $z_{\nu} = x_{\nu} + jy_{\nu}$ sind entweder alle konjugiert komplex oder eine Wurzel ist negativ reell (für gerade ℓ). Faßt man die konjugiert komplexen Wurzeln unter der Summe zusammen, dann gilt

$$\begin{aligned} E[I,T] &= 2 \sum_{\nu=1}^{\frac{\mu}{2}} \frac{1-x_{\nu}}{(1-x_{\nu})^2 + y_{\nu}^2} + \frac{1}{2} \left\{ 1 - E[M] - \ell - \frac{\text{Var}[M]}{E[M] - \ell} \right\} \\ &+ \frac{1}{1-x_{\mu+1}} \Big|_{\text{nur für gerade } \ell} \end{aligned} \quad (4.91b)$$

mit

$$\mu = \left[\frac{\ell-1}{2} \right]$$

Durch Einsetzen von $E[I,T]$ in die Gl. (4.90) erhält man schließlich die für die Berechnung der charakteristischen Größen notwendige mittlere Anzahl im System $E[X,T]$ kurz vor dem Takt.

a) Mittlere Gesamtwarezeit w_G

Zur Berechnung der mittleren Gesamtwarezeit wird $E[X,T]$ in die Gl. (4.47) eingesetzt, d. h.

$$w_G = \left\{ \frac{E[X,T]}{E[K]} + 1 \right\} T_{OH} + h E[X,T] + \frac{h}{2} \left\{ \frac{\text{Var}[K]}{E[K]} + E[K] - 1 \right\}$$

T_{OH} , h , $E[K]$ und $\text{Var}[K]$ sind gegebene Größen des Modells $C^{[X]}/G/1$.

b) Mittlere Warteschlangenlänge Ω

Unter Verwendung des Little'schen Theorems erhält man aus der mittleren Gesamtwarezeit die mittlere Gesamtwarebelastung

$$\Omega_G = w_G \lambda$$

Wie im Abschnitt 4.3.1.2c gezeigt, ist Ω_G auch aus der folgenden Gleichung berechenbar:

$$\Omega_G = E[X,O] \frac{T_{OH}}{T} + \Omega_O$$

Für die Warteschlangenlänge Ω gilt nach Gl. (4.42b):

$$\Omega = \Omega(O) \frac{T_{OH}}{T} + \Omega_O$$

Aus der Differenz der beiden Gleichungen und mit der Beziehung

$$E[X,O] - \Omega(O) = 1 - p(O,T) g_O$$

nach Gl. (4.39a) erhält man schließlich

$$\Omega = \Omega_G - [1 - p(O,T) g_O] \frac{T_{OH}}{T} \quad (4.92)$$

Die Bestimmung der Wahrscheinlichkeit $p(O,T)$ erfolgt im Punkt e) dieses Abschnitts.

c) Mittlere Wartezeit w bis zum Bedienungsbeginn

Ist die mittlere Warteschlangenlänge Ω bekannt, dann gilt für die mittlere Wartezeit

$$w = \frac{\Omega}{\lambda}$$

d) Mittlere Durchlaufzeit t_f

Die mittlere Durchlaufzeit setzt sich nach Gl. (4.48) additiv aus der mittleren Gesamtwartezeit w_G und der mittleren Bedienungszeit h zusammen

$$t_f = w_G + h .$$

e) Wartewahrscheinlichkeit W_0

Analog zur Gl. (4.49) kann hier die Wartewahrscheinlichkeit W_0 aus den Größen $u(O,T)$, b_0 und $E[K]$ berechnet werden. $u(O,T)$ ist die Wahrscheinlichkeit, daß die Restarbeit kurz vor dem Takt Null ist. Damit gilt der Zusammenhang $u(O,T) \approx p(O,T)$. b_0 entspricht der Wahrscheinlichkeit, daß die zum Taktzeitpunkt eintreffende Arbeit Null ist. Beide Größen sind direkt aus der zugehörigen erzeugenden Funktion für $z=0$ berechenbar:

$$u(O,T) = G_i(O,T) = \frac{\ell - E[M]}{b_0} \prod_{v=0}^{\ell-1} \frac{z_v}{z_v - 1} . \quad (4.93a)$$

Durch Zusammenfassung der konjugiert komplexen Wurzeln erhält man die Beziehung

$$u(O,T) = \frac{\ell - E[M]}{b_0} \left[\prod_{v=1}^{\mu} \frac{|z_v|^2}{(x_v - 1)^2 + y_v^2} + \frac{x_{\mu+1}}{x_{\mu+1} - 1} \Big|_{\text{nur für gerade } \ell} \right] \quad (4.93b)$$

mit

$$\mu = \left[\frac{\ell - 1}{2} \right] = \text{das Ganze der Division } \frac{\ell - 1}{2}$$

Aus der Gl. (4.74) folgt

$$b_0 = G_m(0) = G_k[G_n(0)] . \quad (4.94)$$

Setzt man in die erzeugenden Funktionen $G_n(z)$ (Abschnitt 4.3.2.4) $z=0$ ein und wählt Δt voraussetzungsgemäß genügend klein ($\Delta t \ll h$), dann gilt für alle betrachteten Bedienungs-VF $G_n(0) \approx 0$, in der Grenze

$$\lim_{\Delta t \rightarrow 0} G_n(0) = 0 .$$

Daraus ergibt sich für die Wahrscheinlichkeit

$$b_0 = G_k(0) \approx g_0 \quad (4.95)$$

und schließlich für die Wartewahrscheinlichkeit

$$W_0 = 1 - u(O,T) \frac{1 - b_0}{E[K]} .$$

Mit der Berechnung der Wartewahrscheinlichkeit W_0 wird die Analyse mit Hilfe der Lindley'schen Integralmethode abgeschlossen. Im Gegensatz zur Methode der eingebetteten Markoff-Kette, die nur für spezielle Gruppengrößen-Verteilungen und Bedienungs-VF mit vertretbarem numerischem Aufwand anwendbar ist, kann hier ein Modell des Typs $c^{[x]}/G/1$ mit allgemeiner Gruppengrößen-Verteilung und allgemein verteilten Bedienungszeiten approximativ berechnet werden. Wählt man bei der Zeitdiskretisierung Δt genügend klein, so zeigen die erzielten Ergebnisse sehr gute Übereinstimmung mit den Simulationsergebnissen.

Im folgenden wird kurz eine weitere Analyseverfahren für das zweite Teilmodell diskutiert, bei der die von Krämer und Langenbach-Belz in [45] angegebenen approximativen Beziehungen für die Wartezeit und Wartewahrscheinlichkeit für das Warteschlangenmodell des Typs GI/G/1 verwendet werden.

4.3.3 Approximative Berechnung durch Rückführung auf das Modell des Typs GI/G/1

Während das gegebene Modell $c^{[x]}/G/1$ bei der Analyse mit Hilfe der Lindley'schen Integralgleichung auf ein Ersatzmodell des Typs $C/G_x/1$ zurückgeführt wurde, erfolgt hier die Berechnung durch Rückführung auf ein Ersatzmodell des Typs $GI/G^*/1$. Zur Berechnung der mittleren Wartezeit und der Wartewahrscheinlichkeit werden die von Krämer und Langenbach-Belz in [45] angegebenen approximativen Formeln für das Warteschlangenmodell GI/G/1 verwendet. Die dort angegebenen Gleichungen berücksichtigen jedoch keine Verwaltungszeit.

Ausgangspunkt der weiteren Überlegungen ist deshalb die Gl. (4.47), welche die Gesamt-wartezeit w_G einer beliebigen Anforderung in zwei Anteile zerlegt. Der erste Anteil berücksichtigt den Einfluß der Verwaltungszeit, der zweite Anteil w_O entspricht der Wartezeit für ein Modell ohne Verwaltungszeit, jedoch mit der verkürzten Taktperiode $T_S = T - T_{OH}$. Die Gruppengrößen-Verteilung bleibt dabei unverändert. Nach Gl. (4.44) gilt für die Wartezeit w_O der Zusammenhang

$$w_O = w_{O1} + \frac{h}{2} \left\{ \frac{\text{Var}[K]}{E[K]} + E[K] - 1 \right\} . \quad (4.96)$$

w_{O1} ist die Wartezeit des Gruppenersten im Modell $C^{[x]}/G/1$ mit der Taktperiode $T_S = T - T_{OH}$ bei gleicher Gruppengrößen-Verteilung. Die gleiche Wartezeit würde sich ergeben, wenn die gesamte Gruppe ($k > 0$) wie eine zusammengefaßte Einzelanforderung (super-request) betrachtet und für den Bedienungsprozeß eine neue Bedienungs-VF $B^*(t)$, welche die Bedienung der gesamten Gruppe berücksichtigt, definiert wird (vgl. Abschnitt 4.3.2). Da zum Taktzeitpunkt auch Gruppen ankommen können, die keine Anforderung enthalten, ist ein neuer Ankunftsprozeß GI zu bestimmen. Für das so erhaltene Ersatzmodell $GI/G^*/1$ soll im folgenden der Ersatzankunftsprozeß GI und der Ersatzbedienungsprozeß G^* berechnet werden, wobei im wesentlichen auf die in [45] angegebenen Beziehungen zurückgegriffen wird. Beide Prozesse werden durch die ersten zwei Momente approximiert. Die Größen des Ersatzmodells werden einheitlich mit einem * gekennzeichnet.

4.3.3.1 Der Ersatzankunftsprozeß GI

Der mittlere Ankunftsabstand t_A^* des Ersatzankunftsprozesses ergibt sich aus der um T_{OH} verkleinerten Taktperiode unter der Bedingung, daß zum Taktzeitpunkt mindestens eine Anforderung ankommt. Man erhält unmittelbar

$$t_A^* = \frac{T - T_{OH}}{1 - g_O} . \quad (4.97)$$

$1 - g_O$ ist die Wahrscheinlichkeit, daß die ankommende Gruppe mindestens eine Anforderung enthält. Für den Variationskoeffizienten c_A^* gilt nach [45]

$$c_A^* = \sqrt{g_O} . \quad (4.98)$$

4.3.3.2 Der Ersatzbedienungsprozeß G^*

Der Ersatzbedienungsprozeß berücksichtigt die Bedienung der gesamten Gruppe unter der Bedingung, daß die Gruppe mindestens eine Anforderung enthält. Die Ersatzbedienungs-VF $B^*(t)$ hat damit den Mittelwert

$$h^* = E[K|K > 0] \cdot h = \frac{E[K]}{1 - g_O} h \quad (4.99)$$

und die Varianz (vgl. Gl. (4.76))

$$\text{Var}[T_H^*] = E[T_H]^2 \text{Var}[K|K > 0] + E[K|K > 0] \cdot \text{Var}[T_H] . \quad (4.100)$$

Nach dem Einsetzen der einzelnen Größen ergibt sich daraus für den Variationskoeffizienten

$$c_H^* = \frac{1 - g_O}{E[K]} \left\{ \frac{\text{Var}[K]}{E[K]} + c_H^2 \right\} - g_O . \quad (4.101)$$

h und c_H sind gegebene Größen der Bedienungs-VF $H(t)$ des zu untersuchenden Modells $C^{[x]}/G/1$.

4.3.3.3 Berechnung des Angebotes ρ^*

Nach Gl. (2.5) gilt für das Angebot $\rho^* = h^*/t_A^*$. Mit $E[K] = \lambda T$ für die mittlere Gruppengröße und $\rho = h \cdot \lambda$ erhält man für das Angebot des Ersatzmodells

$$\rho^* = h^* \cdot \frac{1}{t_A^*} = \frac{E[K] h}{1 - g_O} \cdot \frac{1 - g_O}{T - T_{OH}} = \rho \frac{T}{T - T_{OH}} . \quad (4.102)$$

Damit sind alle Größen für das Ersatzmodell GI/G*/1 bekannt, aus denen die charakteristischen Größen approximativ berechnet werden können.

4.3.3.4 Berechnung der charakteristischen Größen

Ausgangspunkt für die Berechnung der charakteristischen Größen sind die in [45] angegebenen Formeln für die Wartezeit w_{O1} und die Wartewahrscheinlichkeit W_1 des Warteschlangenmodells GI/G*/1.

$$w_{O1} = \frac{\rho_G^* \cdot h^*}{1 - \rho_G^*} (c_A^{*2} + c_H^{*2}) e^{-\frac{2(1-\rho_G^*)}{(1,5+3\rho_G^*)\rho_G^*} \cdot \frac{(1-c_A^{*2})^2}{c_A^{*2} + c_H^{*2}}}, \quad (4.103a)$$

$$W_1 = \rho_G^* + (c_A^{*2} - 1)\rho_G^*(1 - \rho_G^*) \cdot \frac{1 + c_A^{*2} + \rho_G^* c_H^{*2}}{1 + \rho_G^*(c_A^{*2} - 1) + \rho_G^{*2}(4c_A^{*2} + c_H^{*2})}. \quad (4.103b)$$

Die Gl. (4.103a) ist bis auf eine geringfügige Änderung im Exponenten identisch mit der in [45] angegebenen Beziehung. Hierdurch wurde eine bessere Näherung für das gegebene Modell $C^{[x]}/G/1$ erreicht. Zu der grundsätzlichen Herleitung der Gleichungen wird auf [45] verwiesen.

Setzt man w_{O1} in die Gl. (4.96) ein, dann ist die Wartezeit w_O bekannt. Zur Berechnung der Gesamtwartezeit nach Gl. (4.47) benötigt man noch die mittlere Anzahl im System $E[X,T]$ kurz vor den Taktzeitpunkten. Da die Wartezeit w_{O1} der Bedienungszeit aller zum Taktzeitpunkt im System befindlichen Anforderungen entspricht, gilt näherungsweise

$$E[X,T] = \frac{w_{O1}}{h}. \quad (4.104)$$

Für eine exakte Berechnung müßte in dieser Gleichung die Restbedienungszeit einer zum Taktzeitpunkt unterbrochenen Anforderung berücksichtigt werden. Nur für exponentiell verteilte Bedienungszeiten ist diese Beziehung exakt.

Unter Verwendung der Beziehung $p(O,T) = 1 - W_1$ können nun entsprechend Abschnitt 4.3.2.6 alle charakteristischen Größen berechnet werden. Damit wird die approximative Berechnung des Modells $C^{[x]}/G/1$ durch Rückführung auf ein Ersatzmodell des Typs GI/G/1 abgeschlossen. Im Abschnitt 4.5 werden alle drei Analyseverfahren bezüglich numerischem Aufwand und erreichbarer Genauigkeit miteinander verglichen.

4.4 Das dritte Teilmodell GO/C/∞

Das dritte Teilmodell berücksichtigt die Ausgabewarteschlange und die taktmäßige Bedienung. Es entspricht im Prinzip dem ersten Teilmodell, jedoch wird hier der Eingangsprozeß durch den Ausgangsprozeß (GO: general output process) des zweiten Teilmodells gebildet. Bild 4.16 zeigt die Modellstruktur mit allgemeinem Ankunftsprozeß GO, der Ausgabewarteschlange und den Schalter, der die taktmäßige Bedienung nachbildet (vgl. Bild 4.2).

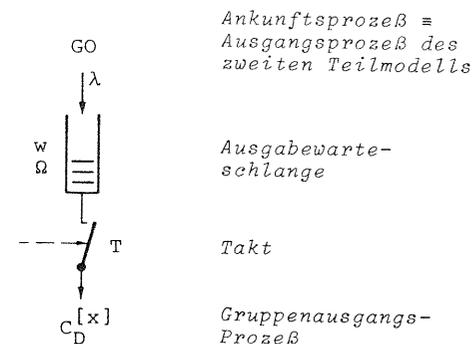


Bild 4.16. Das dritte Teilmodell GO/C/∞

Die zu berechnenden charakteristischen Verkehrsgrößen sind analog zum Abschnitt 4.2 die mittlere Wartezeit w und die Warteschlangenlänge Ω . Hinzu kommt der taktmäßige Gruppenausgangsprozeß $C_D^{[x]}$.

Bedingt durch die drei verschiedenen Analyseverfahren zur Untersuchung des zweiten Teilmodells muß die Analyse des dritten Teilmodells ebenfalls unterschiedlich durchgeführt werden. Wurde das zweite Teilmodell mit Hilfe einer eingebetteten Markoff-Kette (Abschnitt 4.2.1) untersucht, dann können die Zustandswahrscheinlichkeiten des dritten Teilmodells aus den Zustandswahrscheinlichkeiten des zweiten Teilmodells bestimmt und daraus die charakteristischen Größen und die Gruppengrößen-Verteilung des Ausgangsprozesses berechnet werden. Dagegen kann bei Verwendung der beiden anderen Analyseverfahren der Ausgangsprozess des zweiten Teilmodells nicht explizit berechnet werden, und das dritte Teilmodell läßt sich nur approximativ berechnen.

4.4.1 Bestimmung der Zustandswahrscheinlichkeiten zu einem beliebigen Zeitpunkt bei Verwendung einer eingebetteten Markoff-Kette

Entsprechend Bild 4.17 wird das System über eine Taktperiode T beobachtet.

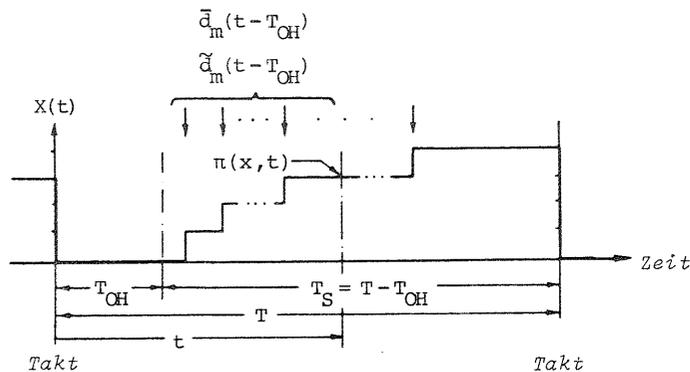


Bild 4.17. Zeitliches Systemverhalten

$\pi(x, t)$ ist die Wahrscheinlichkeit, daß zum Zeitpunkt t nach dem Takt genau x Anforderungen in der Ausgabewarteschlange zwischengespeichert sind. Da aus der Ausgabewarteschlange zum Taktzeitpunkt stets alle wartenden Anforderungen entnommen werden, ist die Wahrscheinlichkeit $\pi(x, t)$ gleich der Zahl der abgeschlossenen Bedienungen in der Zeitspanne $t - T_{OH}$. Während der Verwaltungszeit T_{OH} werden keine Anforderungen bedient, in dieser Zeitspanne ist also die Ausgabewarteschlange leer.

Im Bereich $0 < t < T_{OH}$ gilt deshalb

$$\Omega(t) = 0 .$$

Die Zahl der maximal möglichen Bedienungen in der Zeitdauer $t - T_{OH}$ ist vom Systemzustand des zweiten Teilmodells zum Taktzeitpunkt abhängig. Da nach Ablauf der Verwaltungszeit die Fortsetzung der Bedienung für allgemein verteilte Bedienungszeiten direkt abhängig ist vom Systemzustand des zweiten Teilmodells kurz vor dem Takt (entweder muß eine unterbrochene Anforderung weiterbedient oder bei leerem System beginnt die Bedienung einer neu eingetroffenen Anforderung), müssen für die Berechnung von $\pi(x, t)$ die Zustandswahrscheinlichkeiten $p(j, T)$ des zweiten Teilmodells benutzt werden. Die Zustandswahrscheinlichkeiten $\pi(x, t)$ sind wie folgt von der Gesamtanzahl der Anforderungen im zweiten Teilmodell kurz nach dem Takt (Summe aus den Anforderungen vor dem Takt plus eintreffende Anforderungen zum Taktzeitpunkt) abhängig:

- Ist die Summe genau x , dann müssen in der Zeit $t - T_{OH}$ mindestens x Anforderungen bedient werden ($\bar{d}_{\geq x}(t - T_{OH})$ bzw. $\tilde{d}_{\geq x}(t - T_{OH})$).
- Ist die Summe größer als x , dann müssen in der Zeit $t - T_{OH}$ genau x Anforderungen bedient werden ($\bar{d}_x(t - T_{OH})$ bzw. $\tilde{d}_x(t - T_{OH})$).

Mit diesen Überlegungen gilt für die Zustandswahrscheinlichkeiten $\pi(x, t)$ im Bereich $T_{OH} < t < T$:

$$\begin{aligned} \pi(N,t) &= p(0,T) g_{\geq N} \bar{d}_{\geq N}(t-T_{OH}) + \sum_{j=1}^N p(j,T) g_{\geq (N-j)} \tilde{d}_{\geq N}(t-T_{OH}) , \\ \pi(x,t) &= p(0,T) \left[g_x \bar{d}_{\geq x}(t-T_{OH}) + g_{\geq (x+1)} \bar{d}_x(t-T_{OH}) \right] \\ &+ \sum_{j=1}^N p(j,T) \left[g_{x-j} \tilde{d}_{\geq x}(t-T_{OH}) + g_{\geq (x-j+1)} \tilde{d}_x(t-T_{OH}) \right] , \end{aligned} \quad (4.105a)$$

$$x = 0, 1, \dots, N-1$$

mit

$$g_k = 0 \text{ für } k < 0 \text{ und } g_{\geq k} = 1 \text{ für } k \leq 0 .$$

Für exponentiell verteilte Bedienungszeiten ist der Bedienungsprozeß zum Taktzeitpunkt nicht gedächtnisbehaftet, und für die Berechnung von $\pi(x,t)$ kann direkt von den Zustandswahrscheinlichkeiten $p(j,0)$ nach dem Taktzeitpunkt ausgegangen werden. Daraus ergibt sich die wesentlich einfachere Beziehung

$$\pi(x,t) = p(x,0) \bar{d}_{\geq x}(t-T_{OH}) + \sum_{j=x+1}^N p(j,0) \bar{d}_x(t-T_{OH}) \quad (4.105b)$$

$$x = 0, 1, \dots, N$$

mit

$$\sum_{i=1}^j (\cdot) = 0 \text{ für } j < i .$$

Die Sterbewahrscheinlichkeiten $\bar{d}_x(t-T_{OH})$ sind Poisson verteilt und $p(x,0)$ nach Gl. (4.37) einzusetzen.

Aus diesen Zustandswahrscheinlichkeiten können durch Mittelwertbildung die charakteristischen Verkehrsgrößen berechnet werden. Die Gruppengrößen-Verteilung des Ausgangsprozesses kann direkt aus den Zustandswahrscheinlichkeiten $\pi(x,t)$ bestimmt werden, wenn $t = T$ gesetzt wird.

4.4.2 Berechnung der charakteristischen Größen

a) Mittlere Warteschlangenlänge

Die im Abschnitt 4.3.1.2a dargestellten Überlegungen können hier unmittelbar übernommen werden. Durch Mittelwertbildung über das gesamte Taktintervall T erhält man für die Warteschlangenlänge Ω (vgl. Gl. (4.40))

$$\Omega = \frac{1}{T} \int_0^T \Omega(t) dt$$

mit

$$\Omega(t) = \sum_{x=1}^N x \pi(x,t) \text{ für } T_{OH} < t < T .$$

Für den allgemeinen Fall erhält man durch Einsetzen der Gl. (4.105a) und nach kurzer Umformung

$$\Omega = \frac{1}{T} \sum_{j=0}^N p(j,T) \cdot c_j \quad (4.106a)$$

mit

$$c_0 = \sum_{x=1}^{N-1} x \left[g_x \bar{d}_{\geq x}^I(T_S) + g_{\geq (x+1)} \bar{d}_x^I(T_S) \right] + N g_{\geq N} \bar{d}_{\geq N}^I(T_S) ,$$

$$c_j = \sum_{x=1}^{j-1} x \tilde{d}_x^I(T_S) + \sum_{x=j}^{N-1} x \left[g_{x-j} \tilde{d}_{\geq x}^I(T_S) + g_{\geq (x-j+1)} \tilde{d}_x^I(T_S) \right] + N g_{\geq (N-j)} \tilde{d}_{\geq N}^I(T_S) ,$$

$$j = 1, 2, \dots, N$$

und

$$\sum_{i=1}^j (\cdot) = 0 \text{ für } j < i .$$

Die Bestimmung der durch I gekennzeichneten Größen, die sich aus der Integration über die Sterbewahrscheinlichkeiten ergeben, ist im Anhang 3 durchgeführt.

Für exponentiell verteilte Bedienungszeiten erhält man aufgrund der einfacheren Beziehung für $\pi(x,t)$ nach Gl. (4.105b) für die mittlere Warteschlangenlänge nach kurzer Umformung

$$\Omega = E[X, 0] \frac{T - T_{OH}}{T} - \frac{h}{T} \sum_{x=1}^N x p(x, 0) \sum_{m=0}^{x-1} d_{\geq(m+1)}(T_S) + \frac{h}{T} \sum_{x=1}^{N-1} x d_{\geq(x+1)}(T_S) \sum_{j=x+1}^N p(j, T) . \quad (4.106b)$$

b) Mittlere Wartezeit w

Für die mittlere Wartezeit erhält man unter Verwendung des Little'schen Theorems

$$w = \frac{\Omega}{\lambda} . \quad (4.107)$$

4.4.3 Berechnung des Ausgangsprozesses $C_D^{[x]}$

Der taktmäßige Gruppenausgangsprozeß $C_D^{[x]}$ kann direkt aus den Zustandswahrscheinlichkeiten $\pi(x, t)$ berechnet werden, wenn $t = T$ gesetzt wird. Ist K_D die ZV für die Gruppengröße, dann gilt für die Wahrscheinlichkeiten g_{k_D} der Gruppengrößen-Verteilung

$$g_{k_D} = P\{K_D = k\} = \pi(k, T) , \quad k = 0, 1, \dots, N , \quad (4.108)$$

mit $\pi(k, T)$ nach Gl. (4.105a) bzw. Gl. (4.105b).

Durch Mittelwertbildung erhält man hieraus

$$E[K_D] = \sum_{k=0}^N k g_{k_D} . \quad (4.109)$$

Ist N voraussetzungsgemäß groß, dann ist die mittlere Gruppengröße $E[K_D]$ des Ausgangsprozesses gleich der mittleren Gruppengröße $E[K] = \lambda T$ des Ankunftsprozesses des zweiten Teilmodells.

Für die Varianz $Var[K_D]$ ergibt sich aus der Gl. (4.108)

$$Var[K_D] = \sum_{k=1}^N k^2 g_{k_D} - E[K_D]^2 . \quad (4.110)$$

Schließlich erhält man für den Variationskoeffizienten der Gruppengrößen-Verteilung

$$c_K^D = \frac{\sqrt{Var[K_D]}}{E[K_D]} . \quad (4.111)$$

4.5 Ergebnisdiskussion und Simulationsvergleich

Die im folgenden berechneten und dargestellten Ergebnisse dienen hauptsächlich zur Überprüfung der verschiedenen Analyseverfahren. Der Genauigkeitsnachweis der approximativen Analyseverfahren erfolgt durch Vergleich mit exakt berechneten (Eingebettete Markoff-Kette) oder simulativ bestimmten Ergebnissen. Ein hierzu erstelltes Simulationsprogramm [35], das unter anderem das Ablaufgeschehen für das dreistufige Wartemodell (Bild 4.1) unter Berücksichtigung aller gegenseitigen Abhängigkeiten exakt nachbildet, läßt die Bestimmung der charakteristischen Verkehrsgrößen mit den zugehörigen 95%-Vertrauensintervallen für die drei Teilmodelle bei Variation aller Systemparameter zu. Eine genauere Diskussion der Ergebnisse, insbesondere im Hinblick auf Dimensionierungsfragen, wird im Abschnitt 5.4 durchgeführt. Die dort angegebenen Ergebnisse für die erste Prioritätsklasse gelten, wegen der Voraussetzung unterbrechender Prioritäten, auch für das hier betrachtete einklassige Modell.

Entsprechend der Zerlegung des Gesamtmodells in drei Teilmodelle erfolgt die Validierung der numerisch berechneten Ergebnisse.

4.5.1 Numerische Ergebnisse für das erste Teilmodell GI/C/∞

Wie im Abschnitt 4.2 abgeleitet, ist die mittlere Wartezeit in der Eingabewarteschlange für allgemein verteilte Ankunftsabstände (rekurrenter Ankunftsprozeß) stets gleich der halben Taktperiode. Abweichungen hiervon ergeben sich nur für spezielle, deterministische Ankunftsprozesse, bei denen die Ankunfts- und Taktzeitpunkte eine feste zeitliche Zuordnung zueinander aufweisen, z. B.

wenn eine Anforderung stets zum gleichen Zeitpunkt innerhalb eines Taktintervalls eintrifft. Die mittlere Wartezeit ist hierbei wegen der festen zeitlichen Zuordnung sofort angebar.

Für das erste Teilmodell verbleibt damit die Validierung des taktmäßigen Gruppenausgangsprozesses, d. h. die unter Erneuerungsannahmen zu den Taktzeitpunkten berechnete Gruppengrößen-Verteilung nach Abschnitt 4.2.1 bzw. deren Mittelwert und Varianz nach Abschnitt 4.2.2. Mit der Annahme exponentiell verteilter Ankunftsabstände konnte das Simulationsprogramm überprüft werden, da sich hierbei für die Gruppengrößen-Verteilung eine Poisson-Verteilung mit dem Mittelwert $E[K] = \lambda T$ und der Varianz $Var[K] = \lambda T$ ergeben muß. Für Erlang-k- bzw. H_2 -verteilte Ankunftsabstände lagen die berechneten Werte für

- die Wahrscheinlichkeit g_0 nach Gl. (4.14a) bzw. Gl. (4.15),
- den Mittelwert $E[K]$ nach Gl. (4.18),
- die Varianz $Var[K]$ nach Gl. (4.28) bzw. Gl. (4.33) und
- den Variationskoeffizienten c_K nach Gl. (4.34)

bei Variation der Ordnung k bzw. des Variationskoeffizienten c_A des Ankunftsprozesses sowie λ und T stets innerhalb der sehr kleinen Vertrauensintervalle. Damit stellt die Erneuerungsannahme zu den Taktzeitpunkten zur Berechnung der Gruppengrößen-Verteilung für hypo- und hyperexponentielle VF (E_k und H_2) eine sehr gute Approximation dar.

Da die Abhängigkeit des Variationskoeffizienten c_K vom Mittelwert $E[K] = \lambda T$ für die Ankunftsprozeßtypen E_k , M und H_2 bereits ausführlich am Ende des Abschnitts 4.2 diskutiert wurde (siehe Bild 4.5), kann hier auf eine weitere Betrachtung verzichtet werden.

4.5.2 Numerische Ergebnisse für das zweite Teilmodell $C^{[x]}/G/1$

a) Gegenüberstellung der verschiedenen Analyseverfahren

Zur Untersuchung des zweiten Teilmodells wurden im Abschnitt 4.3 drei verschiedene Analyseverfahren entwickelt.

1. Analyse mit Hilfe einer eingebetteten Markoff-Kette
2. Analyse mit Hilfe der Lindley'schen Integralmethode
3. Approximative Berechnung durch Rückführung auf das Modell GI/G/1.

Die einzelnen Analyseverfahren sind grundsätzlich verschieden. Insbesondere unterscheiden sie sich hinsichtlich des numerischen Aufwandes und des erforderlichen Speicherplatz- und Rechenzeitbedarfs. Die Methode der eingebetteten Markoff-Kette benötigt wegen der Bestimmung der Übergangswahrscheinlichkeiten und der iterativen Berechnung der Zustandswahrscheinlichkeiten einen erheblichen numerischen Aufwand. Da sie jedoch für Poisson-verteilte Gruppengrößen und exponentiell verteilte Bedienungszeiten exakte Ergebnisse liefert (Voraussetzung ist wegen der notwendigen Abschneidung des Zustandsraumes eine vernachlässigbare Verlustwahrscheinlichkeit), kann sie zur Überprüfung der approximativen Verfahren und insbesondere des Simulationsprogramms herangezogen werden. Die analog zur Berechnung der Gruppengrößen-Verteilung des Ankunftsprozesses (Ausgangsprozeß des ersten Teilmodells) gewählte Erneuerungsannahme zur Bestimmung der Sterbewahrscheinlichkeiten innerhalb eines Taktintervalls, die für E_k -verteilte Bedienungszeiten explizit durchgeführt wurde (vgl. Abschnitt 4.3.1.1) liefert sehr gute Ergebnisse verglichen mit der Simulation. Da das Verfahren der eingebetteten Markoff-Kette jedoch sehr aufwendig und für allgemeine Gruppengrößen-Verteilungen und/oder Bedienungs-VF nicht mehr mit vertretbarem numerischem Aufwand auswertbar ist, wurden die beiden approximativen Analyseverfahren entwickelt.

Bei der Analyse mit Hilfe der Lindley'schen Integralmethode liegt die Approximation für Poisson-verteilte Gruppengrößen nur in der Zeitdiskretisierung der allgemein vorgegebenen, kontinuierlichen Bedienungs-VF. Wählt man Δt bei der Diskretisierung genügend klein, so stimmen die Ergebnisse bei exponentiell verteilten Bedienungszeiten sehr genau mit den exakten Ergebnissen der eingebetteten Markoff-Kette überein. Dies gilt auch für einen deterministischen Bedienungsprozeß, der die auf der Seite 72 angegebene Synchronisationsbedingung erfüllt [9]. Für andere Be-

dienungs-VF (E_k, H_2) zeigt der Simulationsvergleich ebenfalls eine gute Übereinstimmung. Für allgemein verteilte Gruppengrößen wird sowohl die Gruppengrößen-Verteilung als auch die Bedienungs-VF durch die ersten beiden Momente approximiert und die hieraus entstehende zusammengesetzte Verteilung anschließend zeitdiskretisiert (vgl. Abschnitt 4.3.2.5). Die erzielten Ergebnisse zeigen die Tabelle 1 und die Bilder 4.18 und 4.19. Der wesentliche numerische Aufwand dieses Verfahrens liegt in der iterativen Bestimmung der Wurzeln der charakteristischen Gleichung $z^\lambda - G_m(z) = 0$. λ ist ein Maß für die Zeitdiskretisierung und nach Gl. (4.54) von der effektiv zur Verfügung stehenden Bedienungszeit $T_S = T - T_{OH}$ und dem gewählten Δt -Intervall abhängig. Für die Durchlaufzeit ergeben sich bereits gute Ergebnisse, wenn für $\Delta t \approx h/10$ ($h =$ mittlere Bedienungsdauer) angenommen wird. Dagegen muß zur genaueren Bestimmung der Wartewahrscheinlichkeit $\Delta t \leq h/100$ vorausgesetzt werden. Trotzdem sind selbst bei dieser Wahl von Δt die erforderlichen Rechenzeiten zur Bestimmung der charakteristischen Größen wesentlich geringer, verglichen zu denjenigen bei Verwendung einer eingebetteten Markoff-Kette.

Das dritte Analyseverfahren verwendet die von Krämer und Langenbach-Belz in [45] angegebenen approximativen Beziehungen für ein Modell des Typs GI/G/1. Die Gleichung zur Bestimmung der Wartezeit des Gruppenersten wurde hierbei jedoch geringfügig geändert und zusätzlich der Einfluß der Verwaltungszeit durch eine Erhöhung des Angebots berücksichtigt. Aufgrund der einfachen Beziehungen stellt dieses Verfahren zwar den geringsten numerischen Aufwand dar, jedoch zeigen die Ergebnisse teilweise größere Abweichungen von den exakten oder simulativen Werten, abhängig von der Wahl der einzelnen Systemparameter.

Es sei besonders darauf hingewiesen, daß bei der GI/G/1-Approximation für die Gruppengrößen-Verteilung neben Mittelwert $E[K]$ und Varianz $Var[K]$ auch die Wahrscheinlichkeit g_0 (Wahrscheinlichkeit für eine leere Gruppe) bekannt sein muß, da sie zur Bestimmung des Ersatzankunftsprozesses erforderlich ist. Die Größe g_0 ist jedoch häufig nicht vorgegeben oder nur schwierig, d. h. aufwendig

berechenbar. Bei der Lindley'schen Integralmethode wird die Größe g_0 nicht benötigt, die approximative Beschreibung der Gruppengrößen-Verteilung erfolgt ausschließlich durch die ersten beiden Momente ($E[K], Var[K]$).

Zur Validierung der Ergebnisse aller drei Analyseverfahren wird nacheinander der Einfluß der wichtigsten Systemparameter auf die mittlere Durchlaufzeit betrachtet. Der Ergebnisvergleich erfolgt in Form von Wertetabellen und graphischen Darstellungen am Ende dieses Abschnittes. Für die Ergebnisdarstellung gelten einheitlich folgende Vereinbarungen;

- ▲ Eingebettete Markoff-Kette (EMK)
- Lindley'sche Integralmethode (LIM)
- GI/G/1-Approximation (GI/G/1-Appr.)
- ‡ Simulation mit 95%-Vertrauensintervall und 10 Teiltests mit je 20 000 Ankunftsereignissen.

Alle Zeitgrößen sind normiert auf die mittlere Bedienungsdauer $h = 1$.

b) Einfluß der Gruppengrößen-Verteilung

Bild 4.18 zeigt den Einfluß der Gruppengrößen-Verteilung, die sich aus dem Ankunftsprozeß des ersten Teilmodells D, E_k, M und H_2 ergibt, auf die mittlere Durchlaufzeit als Funktion des Angebots. Die Bedienungszeiten sind exponentiell verteilt, die Verwaltungszeit $T_{OH} = 0$. Erwartungsgemäß ergeben sich für konstante Gruppengrößen (D -Ankunftsprozeß im ersten Teilmodell) die geringsten Durchlaufzeiten. Wie die graphische Darstellung zeigt, stehen die approximativ berechneten Ergebnisse in guter Übereinstimmung zu denjenigen der eingebetteten Markoff-Kette bzw. der Simulation. Die Tabelle 1 zeigt, daß die unter Erneuerungsannahmen für den E_2 -Ankunftsprozeß des ersten Teilmodells berechnete Gruppengrößen-Verteilung zu guten Ergebnissen führt. Bei der Lindley'schen Integralmethode und der GI/G/1-Appr. wird hierbei nur auf die ersten zwei Momente zurückgegriffen.

Die hier gemachten Aussagen gelten analog für die Ergebnisse nach Bild 4.19. Bis auf $T_{OH} = 1$ gelten die genannten Voraussetzungen.

Angebot	Analyseverfahren			
	EMK	LIM ($\Delta t = h/10$)	GI/G/1-Appr.	Simulation
ρ				
0.1	1.3116	1.4171	1.3124	1.3112 ± 0.011
0.3	2.2835	2.2911	2.2786	2.2850 ± 0.015
0.5	3.3829	3.4060	3.3483	3.3921 ± 0.028
0.7	4.9629	5.0069	4.8495	4.9154 ± 0.110
0.9	10.4842	10.5416	10.2309	10.2321 ± 0.384

Tabelle 1 Mittlere Durchlaufzeit des zweiten Teilmodells für E_2 -verteilte Ankunftsabstände im ersten Teilmodell und exponentiell verteilte Bedienungszeiten ($T=10, T_{OH}=0$)

c) Einfluß des Bedienungsprozeßtyps

Die Bilder 4.20 bis 4.22 zeigen den Einfluß des Bedienungsprozeßtyps (D, E_2, M, H_2) auf die mittlere Durchlaufzeit als Funktion des Angebots. Der taktmäßige Ankunftsprozeß hat Poisson-verteilte Gruppengrößen. Wegen dieser Voraussetzung liefert die Lindley'sche Integralmethode für alle Bedienungs-VF sehr gute Ergebnisse. Bild 4.22 zeigt, daß die GI/G/1-Appr. dagegen für kleine Taktdauern ($T=3$) und bei Berücksichtigung der Verwaltungszeit ($T_{OH}=1$), insbesondere im unteren Angebotsbereich, größere Fehler aufweist. Für die E_2 -Bedienungs-VF liegen die Ergebnisse bei Verwendung der eingebetteten Markoff-Kette und unter Erneuerungsannahmen zu den Taktzeitpunkten stets innerhalb der Vertrauensintervalle.

d) Einfluß der Verwaltungszeit

Im Bild 4.23 ist der Einfluß der Verwaltungszeit auf die Durchlaufzeit als Funktion der Taktdauer dargestellt. Das Angebot ist konstant ($\rho=0.3$ bzw. $\rho=0.7$), die Gruppengrößen sind Poisson-verteilt und die Bedienungszeiten negativ exponentiell. Da der Fehler bei der Lindley'schen Integralmethode hierbei nur in der Zeitdiskretisierung liegt, stimmen die Ergebnisse sehr gut mit den eingezeichneten exakten Ergebnissen der eingebetteten Markoff-Kette überein. Die GI/G/1-Appr. zeigt gewisse Abweichungen, insbesondere im Bereich der minimalen Durchlaufzeit und für größere Werte von T und ρ . Einen Ergebnisvergleich zeigt die Tabelle 2.

Takt-dauer T	Angebot ρ	Analyseverfahren		
		EMK	LIM ($\Delta t = h/10$)	GI/G/1-Appr.
2	0.3	5.6271	5.6271	6.0871
3	0.3	3.5044	3.5044	3.8237
5	0.3	3.0966	3.0966	3.2272
10	0.3	3.5668	3.5668	3.5682
6	0.7	9.2154	9.2152	9.4430
15	0.7	8.7463	8.7463	8.5421
20	0.7	10.0777	10.0775	9.7797

Tabelle 2 Mittlere Durchlaufzeit des zweiten Teilmodells bei Verwendung der verschiedenen Analyseverfahren (Poisson-verteilte Gruppengrößen, exponentiell verteilte Bedienungszeiten, $T_{OH}=0$)

Für $T_{OH}=0$ und $T \rightarrow 0$ ergibt sich erwartungsgemäß die Durchlaufzeit für ein Warteschlangenmodell des Typs M/M/1. Für $T_{OH} > 0$ erhält man eine minimale Taktdauer T_{min} , die unter stationären Voraussetzungen nicht unterschritten werden darf. Sie ergibt sich aus der Bedingung, daß die mittlere Anzahl von eintreffenden Anforderungen höchstens gleich der mittleren Anzahl von Bedienungen innerhalb eines Taktintervalls sein darf. Da die effektiv zur Verfügung stehende Bedienungszeit $T_S = T - T_{OH}$ beträgt, gilt die Beziehung

$$\lambda T \leq \mu(T - T_{OH})$$

Hieraus ergibt sich in der Grenze

$$T_{min} = \frac{T_{OH}}{1 - \rho} \quad (4.112)$$

In [37] wurde gezeigt, daß unter der Annahme einer endlichen Arbeitswarteschlange und bei Berücksichtigung der gegenseitigen Abhängigkeiten zwischen dem ersten und zweiten Teilmodell neben dieser minimalen Taktdauer T_{min} auch eine maximale Taktdauer T_{max} existiert. Das System zeigt dabei unter stationären Voraussetzungen nur im Bereich $T_{min} < T < T_{max}$ stabiles Verhalten.

Weitere Parameterstudien und entsprechende Simulationsvergleiche wurden durchgeführt mit dem Ergebnis, daß die approximativ berechneten Werte in guter Übereinstimmung zu exakt berechenbaren oder simulativen Ergebnissen stehen. Wegen der Vielzahl der Parameter wurde hier nur der Einfluß der wichtigsten Größen auf die mittlere Durchlaufzeit betrachtet und graphisch dargestellt.

4.5.3 Numerische Ergebnisse für das dritte Teilmodell GO/C/∞

Wie im Abschnitt 4.4 beschrieben, läßt sich das dritte Teilmodell nur dann analytisch berechnen, wenn zur Untersuchung des zweiten Teilmodells die Methode der eingebetteten Markoff-Kette verwendet wurde. Nur hierbei ist der Ausgangsprozess des zweiten Teilmodells, der den Eingangsprozess des dritten Teilmodells bildet, explizit bestimmbar. Damit sind nur Ergebnisse für spezielle Ankunftsprozeßtypen des ersten und/oder Bedienungsprozeßtypen des zweiten Teilmodells analytisch berechenbar. Für H₂-VF wird im folgenden auf die Simulation zurückgegriffen.

In der Tabelle 3 ist die mittlere Wartezeit in der Ausgabewarteschlange, im Bild 4.24 der Variationskoeffizient c_K^D der Gruppengrößen-Verteilung des taktmäßigen Ausgangsprozesses für das dritte Teilmodell in Abhängigkeit des Angebots dargestellt.

Angebot ρ	M-Bed.-Prozeß Ankunftsprozeßtypen			M-Ank.-Prozeß Bedienungsprozeßtypen	
	E ₂	M	H ₂	E ₂	H ₂
0.1	8.691	8.507	8.353 ± 0.02	8.500	8.591 ± 0.01
0.3	7.759	7.563	7.284 ± 0.03	7.523	7.704 ± 0.02
0.5	6.858	6.710	6.444 ± 0.04	6.629	6.861 ± 0.02
0.7	6.043	5.956	5.769 ± 0.03	5.867	6.092 ± 0.04
0.9	5.324	5.297	5.224 ± 0.04	5.252	5.330 ± 0.02
→1.0	→5.0	→5.0	→5.0	→5.0	→5.0

Tabelle 3 Mittlere Wartezeit in der Ausgabewarteschlange für verschiedene Ankunftsprozeßtypen (1. Teilmodell) und Bedienungsprozeßtypen (E₂-M-VF EMK, H₂-VF Simulation, T=10, T_{OH}=0)

Die Parameter für beide Darstellungen sind:

- Ankunftsprozeßtypen im ersten Teilmodell
- Bedienungsprozeßtypen im zweiten Teilmodell
- Taktdauer und Verwaltungszeit.

Die Ergebnisse der Tabelle 3 zeigen, daß die verschiedenen Ankunfts- und/oder Bedienungsprozeßtypen keinen wesentlichen Einfluß auf die mittlere Wartezeit haben, insbesondere im unteren und oberen Angebotsbereich, speziell für ρ → 1. Für ρ → 0 besteht die eintreffende Gruppe im zweiten Teilmodell aus höchstens einer Anforderung und das System ist kurz vor dem Takt mit einer gegen "1" gehenden Wahrscheinlichkeit leer. Damit folgt für die maximal mögliche Wartezeit im dritten Teilmodell $w_{max} = T - T_{OH} - h$. Geht dagegen ρ → 1, ist die Bedienungseinheit praktisch während der gesamten zur Verfügung stehenden Zeitdauer $T_S = T - T_{OH}$ belegt. Der Ausgangsprozess des zweiten Teilmodells wird nun ausschließlich durch den Bedienungsprozeß beeinflusst, und der Eingangsprozess des dritten Teilmodells läßt sich durch einen rekurrenten Erneuerungsprozeß annähern. Unter dieser Voraussetzung können die für das erste Teilmodell durchgeführten Überlegungen übernommen werden und man erhält die minimale Wartezeit $w_{min} = (T - T_{OH})/2$. Dieser Wert kann nur für deterministische Bedienungs-VF bei Synchronisation zwischen Bedienungsprozeß und Taktzeitpunkt unterschritten werden. In erster Näherung läßt sich die Abhängigkeit zwischen der Wartezeit w und dem Angebot ρ durch einen linearen Zusammenhang darstellen.

Für kleinere Takt Dauern stimmen die Werte für die mittleren Wartezeiten noch besser überein, da der maximale Schwankungsbereich $w_{max} - w_{min}$ sich hierbei ebenfalls reduziert.

Im Gegensatz zur mittleren Wartezeit wird der Variationskoeffizient c_K^D erheblich von den verschiedenen Ankunfts- und/oder Bedienungsprozeßtypen der beiden anderen Teilmodelle beeinflusst. Wegen dieser Abhängigkeit ist eine approximative Berechnung kaum möglich. Lediglich im unteren Angebotsbereich bis ρ ≈ 0,5 und im Grenzfall ρ → 1 läßt sich c_K^D näherungsweise, d. h. ohne Verwendung

einer eingebetteten Markoff-Kette, bestimmen. Die Darstellung im Bild 4.24 zeigt, daß bis $\rho \approx 0.5$ der Variationskoeffizient c_K^D des Ausgangsprozesses im dritten Teilmodell näherungsweise mit dem Variationskoeffizienten c_K der Gruppengrößen-Verteilung des taktmäßigen Ausgangsprozesses des ersten Teilmodells übereinstimmt. Wegen der zuvor beschriebenen Besonderheit für $\rho \rightarrow 1$ können der Mittelwert $E[K_D]$, die Varianz $\text{Var}[K_D]$ und der Variationskoeffizient c_K^D mit den Beziehungen nach Abschnitt 4.2.2 berechnet werden, wenn dort λ durch μ und T durch $T_S = T - T_{OH}$ ersetzt wird.

4.5.4 Mittlere Gesamtdurchlaufzeit

Die mittlere Gesamtdurchlaufzeit ergibt sich aus der Addition der mittleren Wartezeiten in der Ein- und Ausgabewarteschlange sowie der mittleren Durchlaufzeit im zweiten Teilmodell. Da diese drei Mittelwerte, die aufgrund der "isolierten" Betrachtung der einzelnen Teilmodelle getrennt berechnet wurden, eine gute Übereinstimmung mit der Simulation zeigen, die das gesamte dreistufige Wartemodell nachbildet, ergibt sich für die mittlere Gesamtdurchlaufzeit ebenfalls eine gute Übereinstimmung mit den simulierten Werten und braucht hier nicht weiter betrachtet zu werden. Allgemein läßt sich jedoch für die mittlere Gesamtdurchlaufzeit die folgende Bedingung angeben:

$$t_{fG} \geq 1.5 \cdot T \tag{4.113}$$

Die Gl. (4.113) ergibt sich aus der Überlegung, daß eine beliebige Anforderung eine Wartezeit $T/2$ in der Eingabewarteschlange erfährt und sich anschließend noch mindestens eine Taktperiode im System befinden muß.

Mit der Betrachtung der mittleren Gesamtwartezeit wird die Validierung der approximativ berechneten Ergebnisse für die einzelnen Teilmodelle abgeschlossen. Für die weiteren Betrachtungen, insbesondere im Hinblick auf Dimensionierungsfragen, wird auf den letzten Abschnitt des folgenden Kapitels verwiesen.

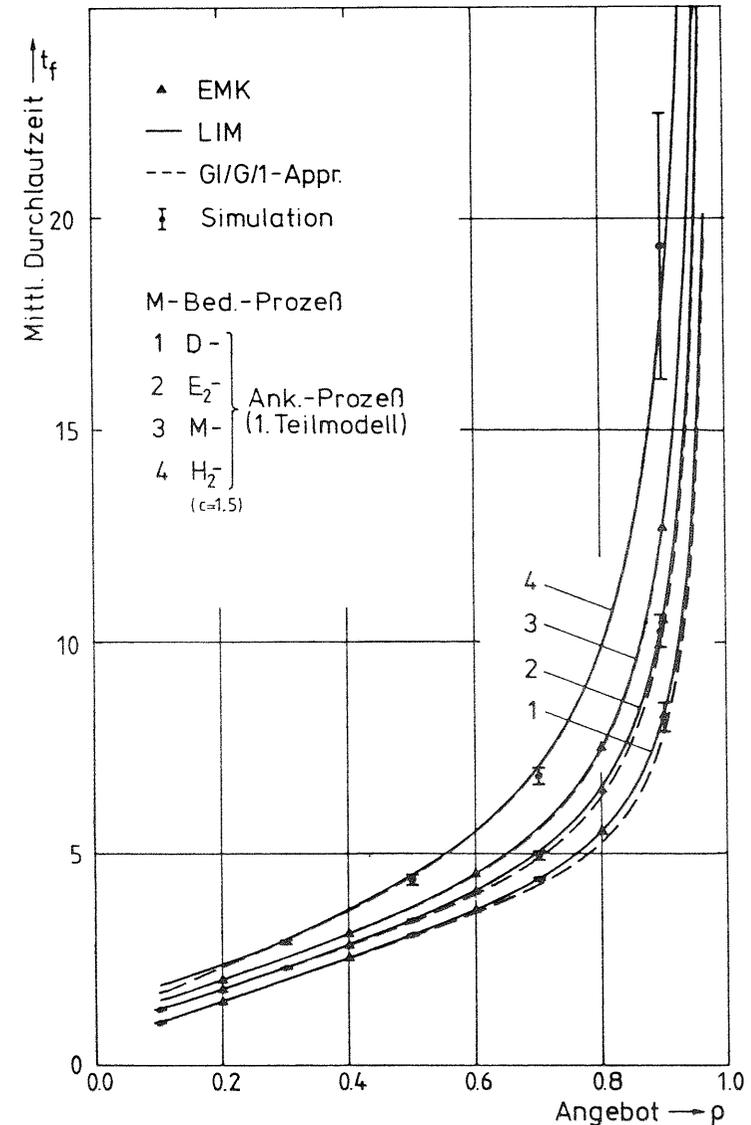


Bild 4.18. Einfluß des Ankunftsprozeßtyps im ersten Teilmodell auf die mittlere Durchlaufzeit des zweiten Teilmodells
Parameter: $T = 10, T_{OH} = 0$

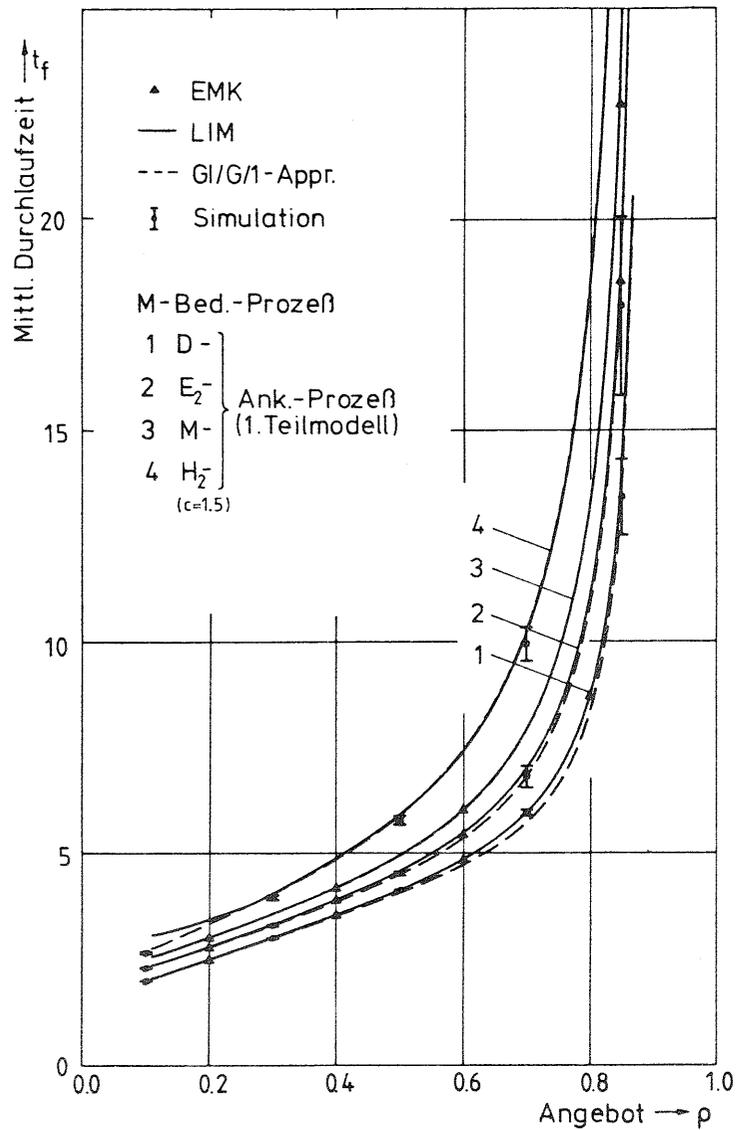


Bild 4.19. Einfluß des Ankunftsprozeßtyps im ersten Teilmodell auf die mittlere Durchlaufzeit des zweiten Teilmodells
 Parameter: $T = 10, T_{OH} = 1$

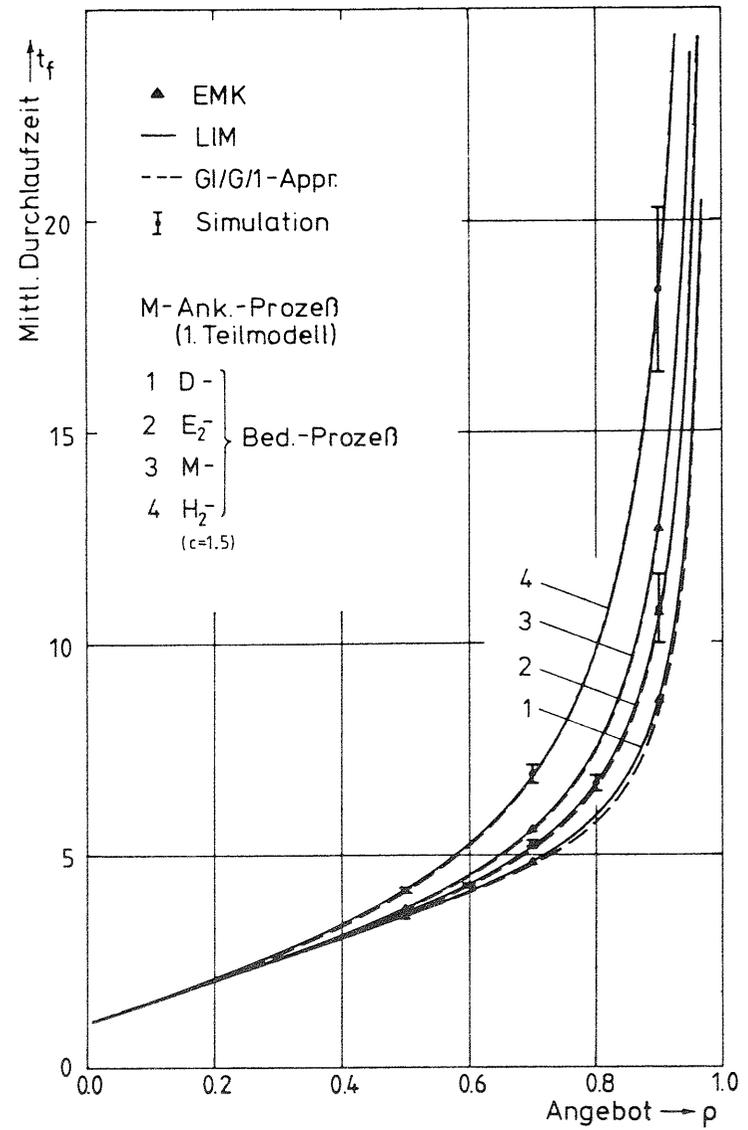


Bild 4.20. Einfluß des Bedienungsprozeßtyps auf die mittlere Durchlaufzeit des zweiten Teilmodells
 Parameter: $T = 10, T_{OH} = 0$

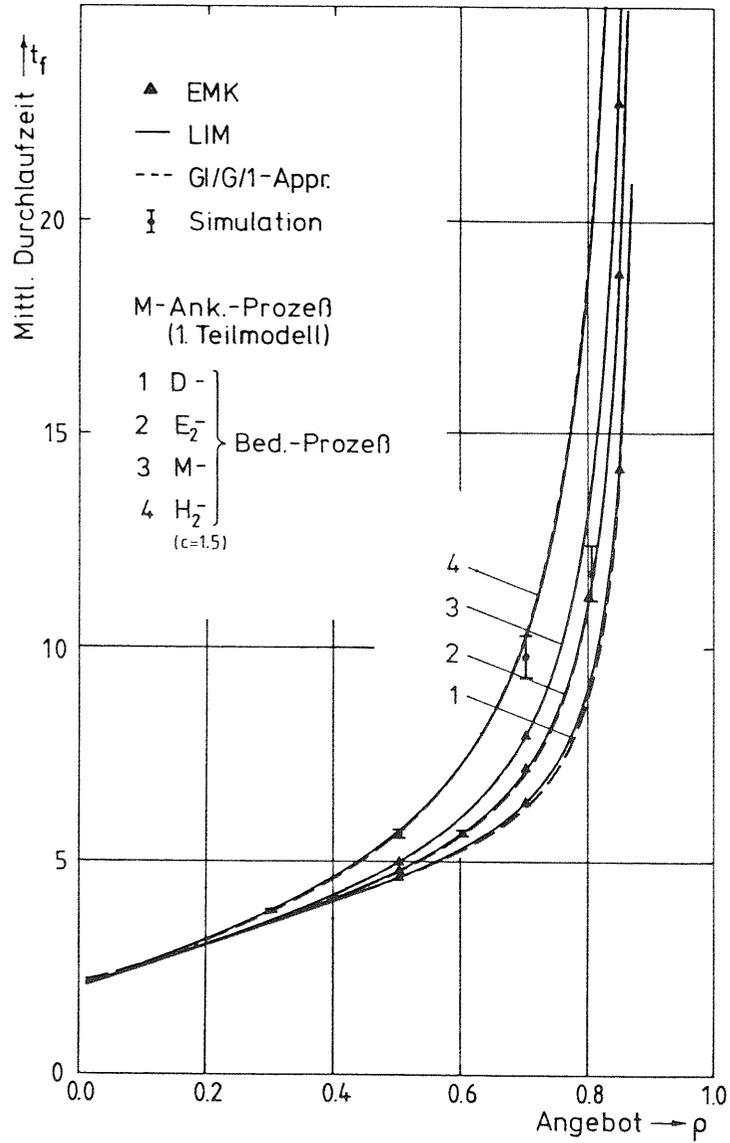


Bild 4.21. Einfluß des Bedienungsprozeßtyps auf die mittlere Durchlaufzeit des zweiten Teilmodells
Parameter: $T=10$, $T_{OH}=1$

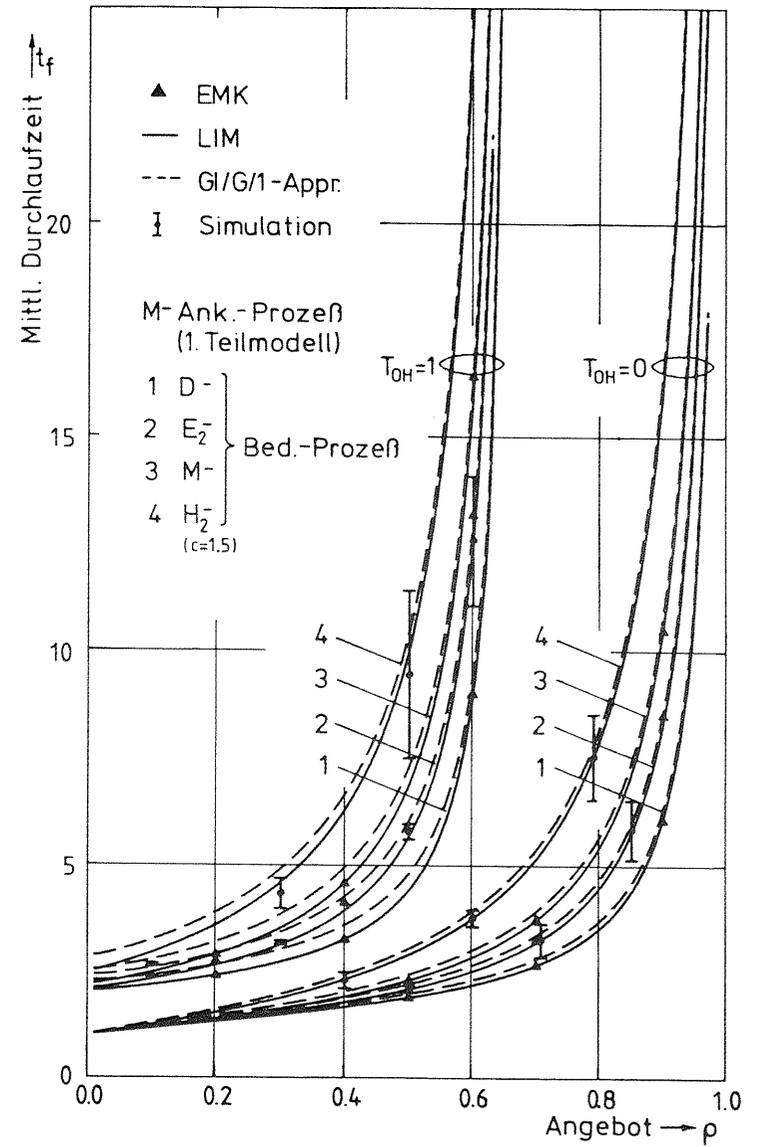


Bild 4.22. Einfluß des Bedienungsprozeßtyps auf die mittlere Durchlaufzeit des zweiten Teilmodells
Parameter: $T=3$

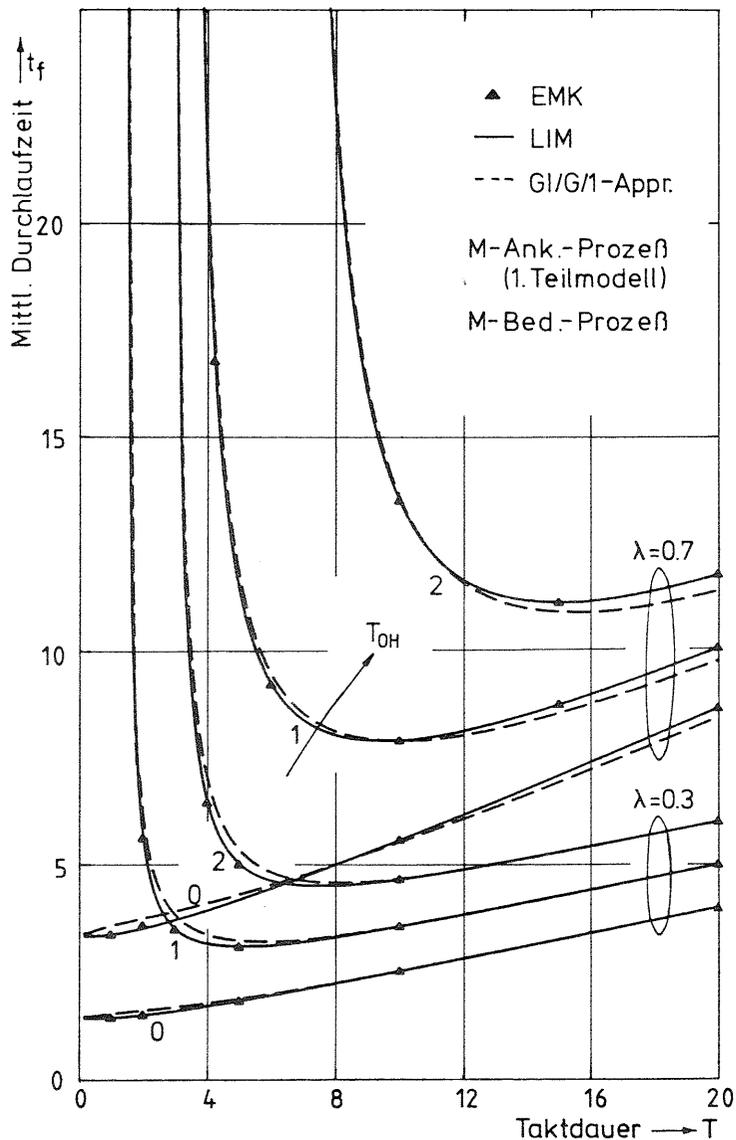
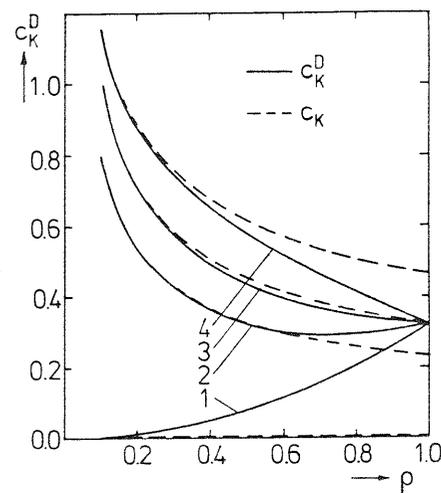


Bild 4.23. Einfluß der Verwaltungszeit auf die mittlere Durchlaufzeit des zweiten Teilmodells

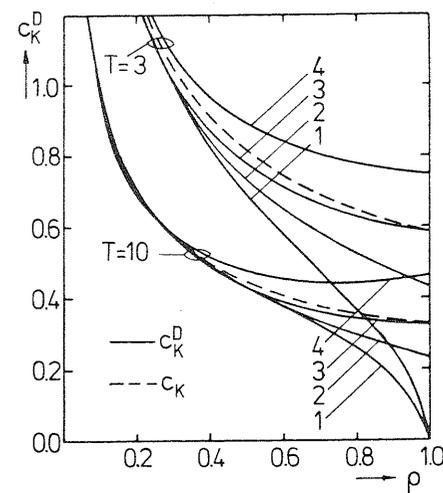


a) Einfluß des Ankunftsprozeß-typs im ersten Teilmodell

- 1 D -
 - 2 E₂ -
 - 3 M -
 - 4 H₂ -
- } Ank.-Prozeß
(c=1.5)

M - Bed.-Prozeß im zweiten Teilmodell

Parameter: T = 10, T_{OH} = 0



b) Einfluß des Bedienungsprozeß-typs im zweiten Teilmodell

- 1 D -
 - 2 E₂ -
 - 3 M -
 - 4 H₂ -
- } Bed.-Prozeß
(c=1.5)

M - Ank.-Prozeß im ersten Teilmodell

Parameter: T = 3, 10, T_{OH} = 0

Bild 4.24. Variationskoeffizient der Gruppengrößen-Verteilung des taktmäßigen Ausgangsprozesses im dritten Teilmodell (c_k ist der Variationskoeffizient der Gruppengrößen-Verteilung des taktmäßigen Ausgangsprozesses im ersten Teilmodell)

5. ANALYSE DES DREISTUFIGEN WARTEMODELLS MIT TAKTGESTEUERTER EIN-/AUSGABE, VERWALTUNGSZEIT UND P PRIORITÄTSKLASSEN

5.1 Struktur und Betriebsweise

Das im Bild 5.1 dargestellte Gesamtmodell stellt eine Erweiterung des im Kap. 4 untersuchten Modells dar. Während dort nur eine Anforderungsklasse betrachtet wurde, werden hier insgesamt P Prioritätsklassen berücksichtigt. Damit lassen sich die Wartezeiten der verschiedenen Anforderungen eines realen Systems, die entsprechend ihrer Dringlichkeit den einzelnen Prioritätsklassen zugeordnet werden, bestimmen. Die hier betrachtete Struktur entspricht damit einer realistischeren Nachbildung des realen Systemverhaltens.

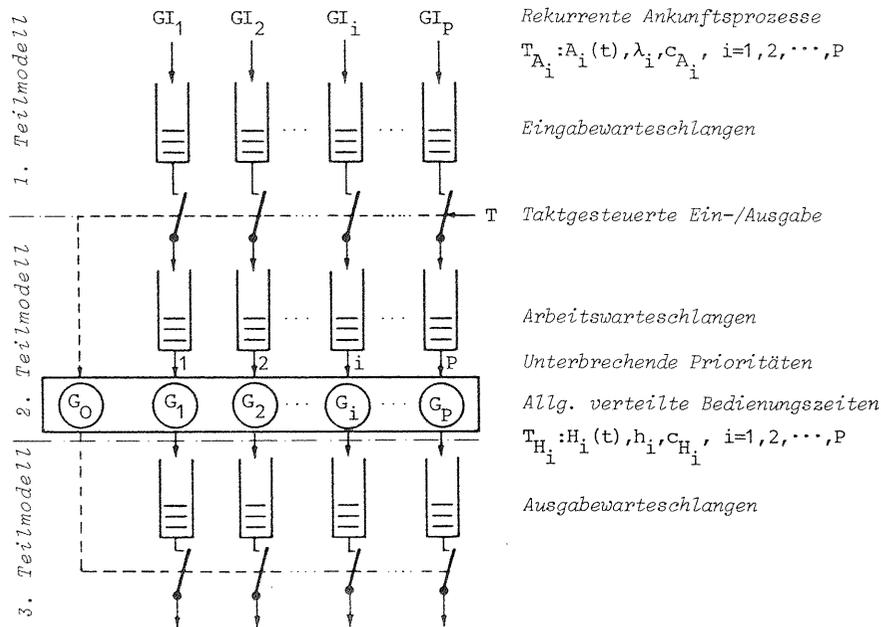


Bild 5.1. Dreistufiges Wartemodell mit taktgesteuerter Ein-/Ausgabe, Verwaltungszeit und P Prioritätsklassen

Die in der Peripherie entstehenden Anforderungen werden entsprechend ihrer Priorität in einer Eingabewarteschlange zwischengespeichert. Der Ankunftsprozeß GI_i der i-ten Prioritätsklasse hat allgemein verteilte Ankunftsabstände mit dem Mittelwert $1/\lambda_i$.

Analog zum Modell im Kap. 4 werden zu jeweils äquidistanten Taktzeitpunkten alle wartenden Anforderungen von den Eingabewarteschlangen in die zugehörigen Arbeitswarteschlangen übertragen. Die Bearbeitungszeit für diesen Transfer wird wiederum durch eine konstante Verwaltungszeit T_{OH} nachgebildet. Eine zum Taktzeitpunkt in Bedienung befindliche Anforderung wird sofort unterbrochen. Die Fortsetzung der Bedienung nach Ablauf der Overhead-Phase erfolgt nach Prioritätsklassen, d. h. eine unterbrochene Anforderung wird erst dann weiterbedient, wenn keine Anforderungen höherer Priorität mehr in den Arbeitswarteschlangen vorhanden sind. Innerhalb der Prioritätsklassen werden die Anforderungen nach der Disziplin FIFO abgearbeitet.

Jede Anforderung durchläuft eine ihrer Priorität entsprechende Bedienungsphase G_i mit der mittleren Bedienungszeit h_i . Die Bedienungs-VF und die mittlere Bedienungszeit können individuell für jede Klasse gewählt werden.

Die bei der Bearbeitung der Anforderungen erzeugten Rückmeldungen (Befehle) für die Peripherie werden prioritätsabhängig in den Ausgabewarteschlangen zwischengespeichert. Jede eintreffende Anforderung im System erzeugt genau eine Rückmeldung. Prioritätswechsel sind in der analytischen Untersuchung nicht berücksichtigt, bei der Simulation wurden sie jedoch mit einbezogen. Der Transfer zwischen den Ausgabewarteschlangen und der Peripherie erfolgt wie die Übernahme der Anforderungen in die Arbeitswarteschlangen während der Overhead-Phase und damit ebenfalls taktgesteuert.

Zur Untersuchung des Gesamtmodells wird erneut auf die im Abschnitt 3.1.2.1 angegebenen Zerlegungsmethoden zurückgegriffen. Analog zur Untersuchung des einklassigen Modells wird das Gesamtmodell in drei Teilmodelle zerlegt (vgl. Bild 5.1). Das erste Teilmodell

umfaßt die Ankunftsprozesse, die Eingabewarteschlangen und die taktmäßige Umspeicherung (Bedienung) der Anforderungen.

Das zweite Teilmodell umfaßt die Bedienungseinheit mit den Arbeitswarteschlangen. Dieses Teilmodell berücksichtigt zusätzlich die Overhead-Phase G_0 . Die Fortsetzung der Bedienung nach der Taktunterbrechung erfolgt nach Prioritäten.

Das dritte Teilmodell, das im Prinzip wiederum dem ersten Teilmodell entspricht, berücksichtigt die Ausgabewarteschlangen und die Entnahme aller wartenden Anforderungen zu den Taktzeitpunkten.

Da die einzelnen Teilmodelle wiederum isoliert untersucht werden sollen, sind die Schnittstellen und damit die Schnittstellenverkehre genauer zu betrachten.

Betrachtet werde zunächst die Schnittstelle zwischem dem ersten und zweiten Teilmodell. Unter der Voraussetzung, daß die Arbeitswarteschlangen unbegrenzt sind (reines Wartesystem), können zum Taktzeitpunkt stets alle Anforderungen übertragen werden. Da die Eingabewarteschlangen also zu jedem Taktzeitpunkt völlig entleert werden, wird in keiner Weise ein Zustand des ersten Teilmodells durch irgendeinen Zustand des zweiten Teilmodells beeinflußt (rückwirkungsfrei). Deshalb können beide Teilmodelle "isoliert" untersucht werden. Der Gruppenankunftsprozeß für die einzelnen Prioritätsklassen des zweiten Teilmodells ist identisch der Anzahl von wartenden Anforderungen in der jeweiligen Eingabewarteschlange kurz vor dem Taktzeitpunkt. Wegen der völligen Rückwirkungsfreiheit kann nun jede Eingabewarteschlange für sich betrachtet und analog zu Abschnitt 4.2 durch ein Warteschlangenmodell des Typs $GI/C/\infty$ nachgebildet werden (vgl. Bild 4.2). Mit den Ergebnissen nach Abschnitt 4.2.1 bzw. Abschnitt 4.2.2 sind somit die Wahrscheinlichkeiten g_{k_i} der Verteilung bzw. der Mittelwert $E[K_i]$ und die Varianz $Var[K_i]$ des taktmäßigen Gruppenausgangsprozesses für jede Prioritätsklasse berechenbar.

Für die zweite Schnittstelle, die zwischen der Bedienungseinheit und den Ausgabewarteschlangen liegt, können die Schnittstellenverkehre für allgemein verteilte Bedienungszeiten nur approximativ

berechnet werden. Das gleiche gilt für die Gruppenausgangsprozesse des dritten Teilmodells.

Wie oben dargestellt, kann das erste Teilmodell in P Einzelmodelle des Typs $GI/C/\infty$ zerlegt werden. Eine ausführliche Untersuchung dieses Modelltyps wurde bereits im Abschnitt 4.2 durchgeführt.

Im folgenden wird deshalb nur das zweite und dritte Teilmodell näher betrachtet. Für ein Gesamtmodell mit 2 Prioritätsklassen und negativ exponentiell verteilten Bedienungszeiten wird eine exakte Analyse mit Hilfe einer eingebetteten Markoff-Kette durchgeführt. Anschließend wird eine approximative Analysemethode vorgestellt, die ein Modell mit P Prioritätsklassen in P Modelle mit einer Klasse von Anforderungen zerlegt. Zum Abschluß dieses Kapitels werden die berechneten Ergebnisse diskutiert. Approximativ berechnete Ergebnisse werden hierbei grundsätzlich mit der Simulation verglichen.

5.2 Das zweite Teilmodell $C^{[x]}/G/1$ mit P Prioritätsklassen

Das zweite Teilmodell, welches die taktmäßigen Gruppenankunftsprozesse, die Arbeitswarteschlangen und die Bedienungseinheit mit den unterschiedlichen Bedienungsphasen für jede Prioritätsklasse berücksichtigt, wird durch ein Warteschlangenmodell des Typs $C^{[x]}/G/1$ mit Prioritäten nachgebildet. Bild 5.2 zeigt die zugehörige Modellstruktur.

Der Ankunftsprozeß entspricht dem Ausgangsprozeß des ersten Teilmodells. Jeweils zu den Taktzeitpunkten trifft für die i -te Prioritätsklasse eine Gruppe mit einer zufälligen Anzahl $k_i \geq 0$ Anforderungen mit der Wahrscheinlichkeit g_{k_i} ein. Für die synchrone Übernahme aller Anforderungen wird im Modell eine konstante Verwaltungszeit T_{OH} berücksichtigt. Eine zum Taktzeitpunkt in Bedienung befindliche Anforderung wird sofort unterbrochen und die Bedienungseinheit wird durch das Ein-/Ausgabe Programm belegt. Die Fortsetzung der Bedienung nach Ablauf der Overhead-Phase erfolgt nach Prioritätsklassen, d. h. eine unterbrochene Anforderung wird erst dann weiterbedient, wenn keine Anforderung höherer

Priorität mehr in den Arbeitswarteschlangen warten.

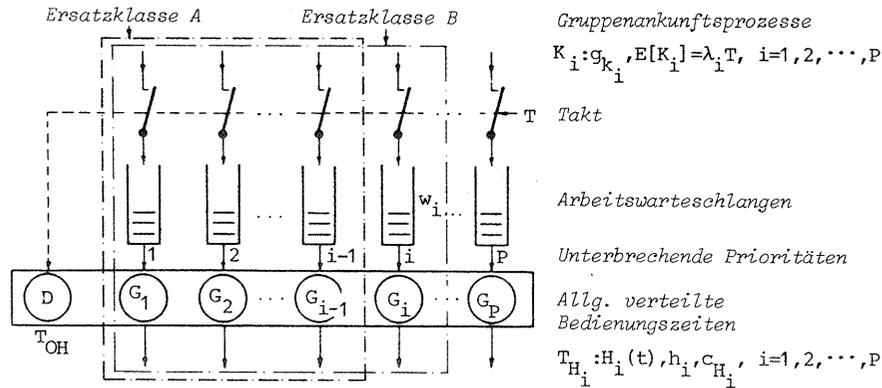


Bild 5.2. Das zweite Teilmodell $C^{[x]}/G/1$ mit P Prioritätsklassen

Zu jedem Taktzeitpunkt wird zum einen wartende Systemarbeit übernommen, zum anderen die Bedienungseinheit der höchsten Prioritätsklasse zur Verfügung gestellt. Diese Vorgehensweise wird, wie im Kap. 1 beschrieben, bei modernen rechnergesteuerten Vermittlungssystemen häufig verwendet.

Für das allgemeine Modell $C^{[x]}/G/1$ mit Prioritäten ist keine analytische Lösung bekannt. Ein Modell des Typs $C^{[x]}/D/1$ mit mehreren Klassen und gleicher Bedienungszeit für jede Klasse wurde in [9] behandelt. Hierbei werden alle ankommenden Anforderungen der einzelnen Klassen hintereinander in eine Arbeitswarteschlange abgelegt und bearbeitet. Dies entspricht quasi einer Priorisierung der einzelnen Klassen. In [14] wird ein Modell des Typs $C^{[x]}/M/1$ mit mehreren Klassen unter Einbeziehung einer Overhead-Phase untersucht und auf ein Ersatzmodell des Typs $D/G/1$ zurückgeführt. Bei beiden Untersuchungen werden Poisson-verteilte Gruppengrößen vorausgesetzt. Eine approximative Berechnung unter allgemeineren Voraussetzungen wurde in [13] durchgeführt. Hierbei wurde ein modifiziertes Modell betrachtet, bei dem die konstante Taktperiode

durch eine exponentielle-VF mit gleichem Mittelwert ersetzt wurde. Mit Hilfe der Momentenmethode und einer heuristischen "Ähnlichkeitseigenschaft" verschiedener Wartesysteme werden die charakteristischen Größen berechnet und mit der Simulation verglichen.

Im Prinzip läßt sich das Modell $C^{[x]}/G/1$ mit P Prioritätsklassen mit Hilfe einer eingebetteten Markoff-Kette analysieren. Für "unterbrechende" Prioritäten, d. h. Fortsetzung der Bedienung nach Ablauf der Verwaltungszeit mit der Anforderung höchster Priorität, benötigt man hierfür eine P -dimensionale Zustandsbeschreibung $p(x_1, x_2, \dots, x_p)$. $\{x_i\}$, $i=1, 2, \dots, P$, ist der Zustand der i -ten Klasse, d. h. die Zahl der kurz vor dem Takt angetroffenen Anforderungen der i -Klasse im System.

Sowohl die Berechnung der Übergangswahrscheinlichkeiten für ein einklassiges Modell mit Erlang- k -verteilten Bedienungsdauern (Abschnitt 4.3.1.1) als auch die iterative Berechnung der Zustandswahrscheinlichkeiten ist mit vertretbarem numerischem Rechenaufwand nicht mehr durchführbar. Selbst für ein Modell mit 2 Prioritätsklassen und dem einfachen Fall exponentiell-verteilter Bedienungszeiten ist die Berechnung recht aufwendig, wie im Abschnitt 5.2.1 gezeigt wird. Im Abschnitt 5.2.2 wird deshalb ein approximatives Analyseverfahren diskutiert, das von einer klassenweisen Zerlegung des gegebenen Modells ausgeht. Mit Hilfe eines "Erhaltungsgesetzes" für Prioritätssysteme können die Warte- und Durchlaufzeiten für die Anforderungen der einzelnen Prioritätsklassen approximativ berechnet werden. Die erzielten Ergebnisse zeigen gute Übereinstimmung mit der Simulation.

5.2.1 Analyse mit Hilfe einer eingebetteten Markoff-Kette kurz vor den Taktzeitpunkten (2 Prioritätsklassen, exponentiell verteilte Bedienungszeiten)

Die Analyse ist prinzipiell vergleichbar zu derjenigen des Abschnitts 4.3.1 für den Fall einer Anforderungsklasse. Da hier jedoch 2 Prioritätsklassen berücksichtigt werden, benötigt man eine zweidimensionale Zustandsbeschreibung. Entsprechend Bild 5.3 wird das System zu zwei aufeinanderfolgenden Taktzeitpunkten beobachtet. Die Wahrscheinlichkeit, daß sich das System kurz vor dem Taktzeitpunkt im Zustand $\{x,y\}$ befindet, wird als Zustandswahrscheinlichkeit $p(x,y,T)$ bezeichnet. Der Zustand $\{x,y\}$ gibt an, daß sich kurz vor dem Taktzeitpunkt genau x Anforderungen erster Priorität und y Anforderungen zweiter Priorität im System befinden. Wegen der Voraussetzung unterbrechender Prioritäten ist für $x \geq 1$ die Bedienungseinheit durch eine Anforderung der ersten Prioritätsklasse und für $x=0$ und $y \geq 1$ durch eine Anforderung der zweiten Prioritätsklasse belegt.

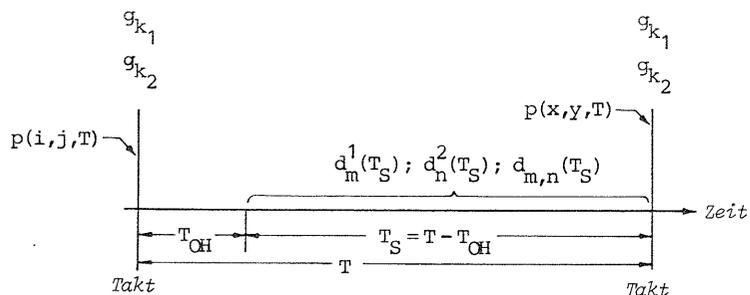


Bild 5.3. Beobachtungszeitpunkte kurz vor dem Takt

Die unabhängigen Gruppenankunftsprozesse für beide Klassen werden hier als bekannt vorausgesetzt. Die Berechnung der Wahrscheinlichkeiten g_{k_i} , $i=1,2$, wurde im Abschnitt 4.2.1 durchgeführt.

Zur Bedienung der wartenden Anforderungen steht die Zeitspanne $T_S = T - T_{OH}$ zur Verfügung. Da das System nur zu den Taktzeitpunk-

ten beobachtet wird, kann in dieser Zeitspanne (abhängig vom Systemzustand nach dem Takt) mit der Wahrscheinlichkeit

- $d_m^1(T_S)$ eine Gruppe von m Anforderungen erster Priorität,
- $d_n^2(T_S)$ eine Gruppe von n Anforderungen zweiter Priorität,
- $d_{m,n}(T_S)$ eine Gruppe von $m+n$ ($m > 0, n \geq 0$) Anforderungen, die sich aus m Anforderungen erster Priorität und n Anforderungen zweiter Priorität zusammensetzt,

bedient werden.

Wegen der Annahme exponentiell verteilter Bedienungszeiten ergibt sich für $d_m^1(T_S)$ und $d_n^2(T_S)$ eine Poisson-Verteilung mit μ_1 bzw. μ_2 als Bedienrate für die erste bzw. zweite Prioritätsklasse. Die Berechnung der Wahrscheinlichkeiten $d_{m,n}(T_S)$ wird im folgenden Abschnitt durchgeführt.

5.2.1.1 Berechnung der Wahrscheinlichkeiten $d_{m,n}(t)$

Zur Berechnung der Wahrscheinlichkeiten $d_{m,n}(t)$, daß sich die Anzahl der abgeschlossenen Bedienungen innerhalb eines Intervalls $(0,t)$ aus m Anforderungen erster Priorität und n Anforderungen zweiter Priorität zusammensetzt ($m \geq 1, n \geq 0$), wird das Bild 5.4 betrachtet.

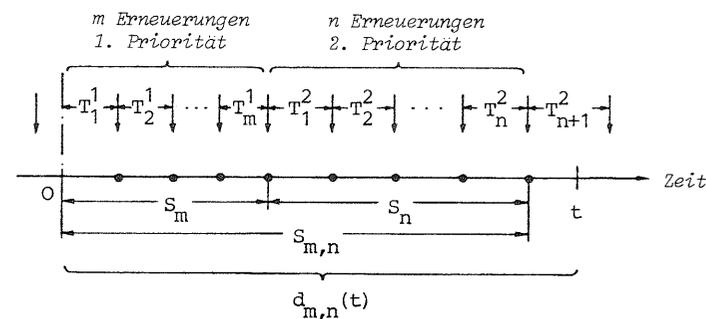


Bild 5.4. Zeitlicher Verlauf des Bedienungsprozesses

Aufgrund der unterbrechenden Prioritäten und negativ exponentiell verteilten Bedienungszeiten, haben die ZV T_i^1 , $i = 1, 2, \dots, m$, die VDF $h_1(t)$, die ZV T_j^2 , $j = 1, 2, \dots, n$, die VDF $h_2(t)$. Die VDF $f_{m,n}(t)$ der Summen-ZV

$$S_{m,n} = \sum_{i=1}^m T_i^1 + \sum_{j=1}^n T_j^2$$

ergibt sich damit aus folgender Faltungsoperation:

$$\begin{aligned} f_{m,n}(t) &= h_1^{(m)}(t) * h_2^{(n)}(t) \\ &= \int_{\tau=0}^t h_1^{(m)}(\tau) \cdot h_2^{(n)}(t-\tau) d\tau. \end{aligned} \quad (5.1)$$

$h_1^{(m)}(t)$ ist die m-fache Faltung der VDF $h_1(t)$, $h_2^{(n)}(t)$ die n-fache Faltung der VDF $h_2(t)$.

Die Wahrscheinlichkeit $d_{m,n}(t)$ ist mit Hilfe der Gl. (4.5) wie folgt berechenbar:

$$d_{m,n}(t) = \int_{u=0}^t [f_{m,n}(u) - f_{m,n+1}(u)] du. \quad (5.2)$$

Setzt man die Gl. (5.1) in die Gl. (5.2) ein und berücksichtigt, daß sich aufgrund der exponentiell verteilten Bedienungszeiten für die Summen-ZV

$$S_m = \sum_{i=1}^m T_i^1 \quad \text{bzw.} \quad S_n = \sum_{j=1}^n T_j^2$$

eine VDF des Erlang-k-Typs ergibt, so erhält man

$$\begin{aligned} d_{m,n}(t) &= \int_{u=0}^t \left[\int_{\tau=0}^u \mu_1^m \frac{\tau^{m-1}}{(m-1)!} e^{-\mu_1 \tau} \cdot \mu_2^n \frac{(u-\tau)^{n-1}}{(n-1)!} e^{-\mu_2(u-\tau)} d\tau \right] du \\ &\quad - \int_{u=0}^t \left[\int_{\tau=0}^u \mu_1^m \frac{\tau^{m-1}}{(m-1)!} e^{-\mu_1 \tau} \cdot \mu_2^{n+1} \frac{(u-\tau)^n}{n!} e^{-\mu_2(u-\tau)} d\tau \right] du. \end{aligned}$$

Durch Vertauschen der Integration und Auswertung des Integrals über u folgt nach kurzer Zwischenrechnung

$$d_{m,n}(t) = \frac{\mu_1^m}{(m-1)!} e^{-\mu_2 t} \frac{\mu_2^n}{n!} \int_{\tau=0}^t \tau^{m-1} (t-\tau)^n e^{(\mu_2 - \mu_1)\tau} d\tau.$$

Für das Integral gilt nach [46] die Beziehung

$$B(n+1, m) \cdot t^{m+n} \cdot \Phi[m, m+n+1; (\mu_2 - \mu_1) t].$$

$B(x, y)$ ist die Betafunktion, $\Phi(a, \gamma; z)$ bezeichnet man als degenerierte hypergeometrische Funktion. Für die Betafunktion erhält man wegen der ganzzahligen Werte von m und n [47]

$$B(n+1, m) = \frac{n!(m-1)!}{(m+n)!}.$$

Die Funktion $\Phi(a, \gamma; z)$ hat keine geschlossene Lösung, sie läßt sich nur als unendliche Reihe darstellen. Damit ergibt sich schließlich

$$d_{m,n}(t) = \frac{(\mu_1 t)^m \cdot (\mu_2 t)^n}{(m+n)!} e^{-\mu_2 t} \left[1 + \sum_{\ell=1}^{\infty} \frac{[(\mu_2 - \mu_1) t]^\ell}{\ell!} \prod_{v=0}^{\ell-1} \frac{m+v}{m+n+1+v} \right], \quad (5.3a)$$

mit $\mu_2 > \mu_1$.

$$\begin{aligned} m &= 1, 2, \dots \\ n &= 0, 1, \dots \end{aligned}$$

Ist die Differenz $\mu_2 - \mu_1 < 0$, dann erhält man unter Verwendung der Beziehung

$$\Phi(a, \gamma; z) = e^z \Phi(\gamma - a, a; -z)$$

die folgende Darstellung für $d_{m,n}(t)$:

$$d_{m,n}(t) = \frac{(\mu_1 t)^m \cdot (\mu_2 t)^n}{(m+n)!} e^{-\mu_1 t} \left[1 + \sum_{\ell=1}^{\infty} \frac{[(\mu_1 - \mu_2) t]^\ell}{\ell!} \prod_{v=0}^{\ell-1} \frac{n+1+v}{m+n+1+v} \right], \quad (5.3b)$$

mit $\mu_1 > \mu_2$.

$$\begin{aligned} m &= 1, 2, \dots \\ n &= 0, 1, \dots \end{aligned}$$

Für den Spezialfall $\mu_1 = \mu_2 = \mu$ ergibt sich erwartungsgemäß für beide Gleichungen eine Poisson-Verteilung, d. h.:

$$d_{m,n}(t) = \frac{(\mu t)^{m+n}}{(m+n)!} e^{-\mu t} .$$

5.2.1.2 Bestimmung der Zustandswahrscheinlichkeiten kurz vor den Taktzeitpunkten

a) Gleichungssystem der Zustandswahrscheinlichkeiten

Nach Bild 5.3 und unter Verwendung der Gl. (3.15) gilt für die Zustandswahrscheinlichkeiten $p(x,y,T)$ kurz vor den Taktzeitpunkten im stationären Gleichgewichtszustand folgendes Gleichungssystem

$$p(x,y,T) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p(i,j,T) q_{ix,jy}(T) , \quad x,y = 0,1,\dots \quad (5.4)$$

$q_{ix,jy}(T) = q_{i \rightarrow x, j \rightarrow y}(T)$ ist die Übergangswahrscheinlichkeit vom Zustand $\{i,j\}$ in den Zustand $\{x,y\}$ innerhalb einer Taktperiode T . Aufgrund der Komplexität der Übergangswahrscheinlichkeiten ist keine rekursive Berechnung der Zustandswahrscheinlichkeiten möglich. Für andere Lösungsverfahren (z. B. iterative Berechnung) muß der unbegrenzte Zustandsraum geeignet "abgeschnitten" werden. Wie im Abschnitt 4.3.1.1 beschrieben, werden deshalb die Arbeitswarteschlangen für beide Klassen auf s_1 bzw. s_2 Speicherplätze begrenzt. Da sich nun maximal $N_1 = s_1 + 1$ Anforderungen erster Priorität und $N_2 = s_2 + 1$ Anforderungen zweiter Priorität im System befinden können, ergibt sich aus der Gl. (5.4)

$$p(x,y,T) = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} p(i,j,T) q_{ix,jy}(T) , \quad \begin{matrix} x = 0,1,\dots,N_1 \\ y = 0,1,\dots,N_2 \end{matrix} \quad (5.5)$$

Zur Bestimmung der Übergangswahrscheinlichkeiten wird der zweidimensionale Zustandsraum nach Bild 5.5 genauer betrachtet.

Wegen der Voraussetzung unterbrechender Prioritäten werden nach dem Takt zunächst alle Anforderungen erster Priorität bedient,

danach beginnt die Bedienung der Anforderungen zweiter Priorität.

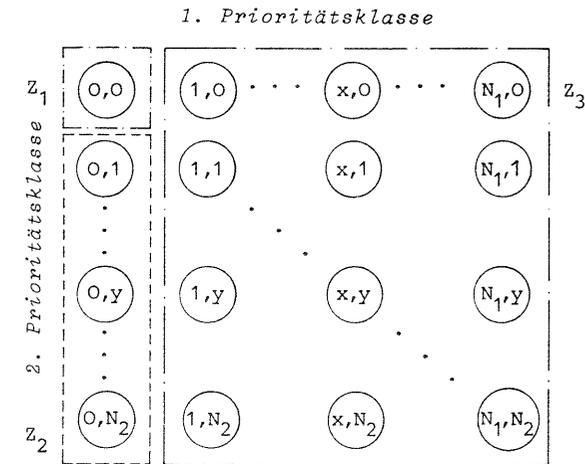


Bild 5.5. Zweidimensionaler Zustandsraum mit den Teilräumen Z_1, Z_2, Z_3

Dies hat zur Folge, daß ein Zustand $\{x,y\}$, $x = 1,2,\dots,N_1$, $y = 0,1,\dots,N_2$, nur von den Zuständen $\{i,y\}$, $i = 0,1,\dots,N_1$, der gesamten betrachteten Zeile und allen Zuständen $\{i,j\}$, $i = 0,1,\dots,N_1$, $j = 0,1,\dots,y-1$, der darüberliegenden Zeilen erreicht werden kann. Dagegen kann ein Zustand $\{0,y\}$, $y = 0,1,\dots,N_2$, der ersten Spalte von jedem Zustand des gesamten Zustandsraumes erreicht werden.

Aufgrund dieser Überlegungen läßt sich der Zustandsraum in drei Teilräume zerlegen. Diese sind nach Bild 5.5:

1. Z_1 mit dem Zustand $\{0,0\}$
2. Z_2 mit den Zuständen $\{0,y\}$, $y = 1,2,\dots,N_2$
3. Z_3 mit den Zuständen $\{x,y\}$, $x = 1,2,\dots,N_1$, $y = 0,1,\dots,N_2$.

Entsprechend der Aufteilung des Zustandsraumes erfolgt die Bestimmung der Übergangswahrscheinlichkeiten. Bei der Bestimmung von $q_{i0,j0}(T)$ ist analog zum einklassigen Modell zu berücksichtigen, daß der Zustand $\{0,0\}$ dadurch erreicht wird, daß mindestens

alle wartenden Anforderungen (beider Prioritätsklassen) nach dem Taktzeitpunkt innerhalb eines Taktintervalls bedient werden müssen. Aufgrund der Begrenzung auf s_1 bzw. s_2 Speicherplätze und des dadurch entstandenen kombinierten Warte-Verlust-Systems stellen die Zustände $\{N_1, j\}$, $j = 0, 1, \dots, N_2$ und $\{i, N_2\}$, $i = 0, 1, \dots, N_1 - 1$, ebenfalls Sonderfälle dar. Sie werden dadurch erreicht, daß die jeweils ankommenden Gruppen mindestens soviel Anforderungen enthalten, daß alle freien Warteplätze in der jeweiligen Prioritätsklasse (im Zustand $\{N_1, N_2\}$ für beide Klassen) aufgefüllt werden. Mit Hilfe dieser Überlegungen können die Übergangswahrscheinlichkeiten aus kombinatorischen Überlegungen bestimmt werden und man erhält entsprechend der Aufteilung des Zustandsraumes:

1. Z_1 mit dem Zustand $\{0,0\}$

Da der Zustand $\{0,0\}$ von jedem Zustand des gesamten Zustandsraumes erreicht werden kann, existieren alle Übergangswahrscheinlichkeiten $q_{i0,j0}(T)$, $i = 0, 1, \dots, N_1$, $j = 0, 1, \dots, N_2$. Bei dem im folgenden beschriebenen Lösungsverfahren werden alle Zustände der beiden anderen Teilräume Z_2 und Z_3 auf den Zustand $\{0,0\}$ zurückgeführt. Dadurch läßt sich $p\{0,0,T\}$ unmittelbar aus der Normierungsbedingung (Summe über alle Zustandswahrscheinlichkeiten gleich "1") berechnen. Da man die Größen $q_{i0,j0}(T)$ für das Lösungsverfahren nicht benötigt, werden sie hier nicht weiter betrachtet.

2. Z_2 mit den Zuständen $\{0,y\}$, $y = 1, 2, \dots, N_2$

$$q_{i0,jy}(T) = \sum_{k_1=0}^{N_1-i} g'_{k_1} \sum_{k_2=\max(y-j,0)}^{N_2-j} g'_{k_2} d_{i+k_1, j+k_2-y}^1(T_S) \quad (5.6a)$$

$$i = 0, 1, \dots, N_1$$

$$j = 0, 1, \dots, N_2$$

mit

$$g'_{k_1} \text{ bzw. } g'_{k_2} \text{ nach Gl. (5.6c)}$$

und

$$d_{0,n}(T_S) = d_n^2(T_S)$$

$$d_{m,n}(T_S) \text{ nach Gl. (5.3a) bzw. Gl. (5.3b) .}$$

3. Z_3 mit den Zuständen $\{x,y\}$, $x = 1, 2, \dots, N_1$, $y = 0, 1, \dots, N_2$

$$q_{ix,jy}(T) = g'_{(y-j)_2} \sum_{k_1=\max(x-i,0)}^{N_1-i} g'_{k_1} d_{i+k_1-x}^1(T_S) \quad (5.6b)$$

$$i = 0, 1, \dots, N_1$$

$$j = 0, 1, \dots, Y$$

Einheitlich ist bei der Bestimmung der Übergangswahrscheinlichkeiten wegen der Begrenzung auf s_1 und s_2 Speicherplätze für g'_{k_1} bzw. g'_{k_2} einzusetzen:

$$g'_{k_1} = \begin{cases} g_{k_1} & \text{für } k_1 < N_1 - i \\ g_{\geq(N_1-i)_1} & \text{für } k_1 = N_1 - i \end{cases} \quad (5.6c)$$

$$g'_{k_2} = \begin{cases} g_{k_2} & \text{für } k_2 < N_2 - j \\ g_{\geq(N_2-j)_2} & \text{für } k_2 = N_2 - j \end{cases}$$

Im Prinzip könnten die Übergangswahrscheinlichkeiten berechnet, in die Gl. (5.4) eingesetzt und die Zustandswahrscheinlichkeiten unter Einbeziehung der Normierungsbedingung

$$\sum_{x=0}^{N_1} \sum_{y=0}^{N_2} p(x,y,T) = 1 \quad (5.7)$$

iterativ bestimmt werden. Da hierzu jedoch alle Übergangswahrscheinlichkeiten auf einmal zur Verfügung gestellt werden müssen, ergeben sich bereits für kleine Werte von N_1 und N_2 (z. B. $N_1 = 20$, $N_2 = 50$) Speicherplatzprobleme. Außerdem wäre eine derartige Bestimmung der Zustandswahrscheinlichkeiten sehr rechenintensiv.

b) Rekursives Lösungsverfahren zur Bestimmung der Zustandswahrscheinlichkeiten

Im folgenden werden die wesentlichen Lösungsschritte zur rekursiven Bestimmung der Zustandswahrscheinlichkeiten abgeleitet und am Ende des Abschnitts in Form eines Algorithmus kurz zusammenge-

faßt. Zur weiteren Diskussion siehe auch [48].

Aufgrund der speziellen Struktur der Übergangswahrscheinlichkeiten, bedingt durch die Aufteilung des Zustandsraumes, können die Zustandswahrscheinlichkeiten $p(x,y,T)$ des Teilraumes Z_3 durch die Zustandswahrscheinlichkeiten $p(O,O,T)$ und $p(O,y,T)$ der beiden anderen Teilräume Z_1 und Z_2 ausgedrückt werden. Da ein Zustand $\{x,y\}$ innerhalb einer Zeile des Teilraumes Z_3 nur von den Zuständen $\{i,y\}$, $i=0,1,\dots,N_1$, dieser Zeile und allen Zuständen $\{i,j\}$, $i=0,1,\dots,N_1$, $j=0,1,\dots,y-1$, der darüber liegenden Zeilen erreicht werden kann, stellen die Zustände $\{x,0\}$, $x=1,2,\dots,N_1$, der ersten Zeile den einfachsten Sonderfall dar. Somit kann ein Zustand $\{x,0\}$ der ersten Zeile nur von den Zuständen $\{i,0\}$, $i=0,1,\dots,N_1$, dieser Zeile, jedoch von keinem anderen Zustand erreicht werden. Damit erhält man

$$p(x,0,T) = \sum_{i=0}^{N_1} p(i,0,T) q_{ix,00}(T) \quad , \quad x = 1,2,\dots,N_1 \quad .$$

Durch den Ansatz

$$p(x,0,T) = a_{x0}(0) p(O,0,T)$$

lassen sich diese Zustandswahrscheinlichkeiten durch die noch unbekannte Wahrscheinlichkeit $p(O,0,T)$ ausdrücken, und man erhält ein lineares, inhomogenes Gleichungssystem für die Unbekannten $a_{x0}(0)$ der Form

$$a_{x0}(0) [1 - q_{xx,00}(T)] - \sum_{\substack{i=1 \\ i \neq x}}^{N_1} a_{i0}(0) q_{ix,00}(T) = q_{Ox,00}(T) \quad , \\ x = 1,2,\dots,N_1 \quad .$$

Ein Zustand $\{x,1\}$, $x=1,2,\dots,N_1$, der zweiten Zeile kann nur von den Zuständen $\{i,1\}$, $i=0,1,\dots,N_1$, dieser Zeile und den Zuständen $\{i,0\}$, $i=0,1,\dots,N_1$, der ersten Zeile erreicht werden. Man erhält unmittelbar

$$p(x,1,T) = \sum_{i=0}^{N_1} [p(i,0,T) q_{ix,01}(T) + p(i,1,T) q_{ix,11}(T)] \quad , \\ x = 1,2,\dots,N_1 \quad .$$

Durch den Ansatz

$$p(x,1,T) = a_{x1}(0) p(O,0,T) + a_{x1}(1) p(O,1,T)$$

können die Zustandswahrscheinlichkeiten $p(x,1,T)$, $x=1,2,\dots,N_1$, durch eine Linearkombination der unbekannt Wahrscheinlichkeiten $p(O,0,T)$ und $p(O,1,T)$ ausgedrückt werden, und man erhält nach entsprechender Aufspaltung zwei lineare, inhomogene Gleichungssysteme für die Unbekannten $a_{x1}(0)$ bzw. $a_{x1}(1)$ der Form

$$a_{x1}(0) [1 - q_{xx,11}(T)] - \sum_{\substack{i=1 \\ i \neq x}}^{N_1} a_{i1}(0) q_{ix,11}(T) = \sum_{i=1}^{N_1} a_{i1}(0) q_{ix,01}(T) + q_{Ox,01}(T) \quad , \\ a_{x1}(1) [1 - q_{xx,11}(T)] - \sum_{\substack{i=1 \\ i \neq x}}^{N_1} a_{i1}(1) q_{ix,11}(T) = q_{Ox,11}(T) \quad , \\ x = 1,2,\dots,N_1 \quad .$$

Durch Verallgemeinerung des Ergebnisses für die zweite Zeile ergibt sich für eine beliebige Zeile des Teilraumes Z_3

$$p(x,y,T) = \sum_{i=0}^{N_1} \sum_{j=0}^y p(i,j,T) q_{ix,jy}(T) \quad , \quad x = 1,2,\dots,N_1 \quad (5.8a) \\ y = 0,1,\dots,N_2 \quad .$$

Durch den allgemeinen Ansatz

$$p(x,y,T) = \sum_{v=0}^y a_{xy}(v) p(O,v,T) \quad (5.8b)$$

können die Zustandswahrscheinlichkeiten $p(x,y,T)$, $x=1,2,\dots,N_1$, $y=0,1,\dots,N_2$, durch eine Linearkombination der unbekannt Wahrscheinlichkeiten $p(O,v,T)$, $v=0,1,\dots,y$, ausgedrückt werden, und man erhält schließlich y+1 lineare, inhomogene Gleichungssysteme für die Unbekannten $a_{xy}(v)$ der Form

$$a_{xy}(v) [1 - q_{xx,yy}(T)] - \sum_{\substack{i=1 \\ i \neq x}}^{N_1} a_{iy}(v) q_{ix,yy}(T) = \sum_{i=1}^{N_1} \sum_{j=v}^{y-1} a_{ij}(v) q_{ix,jy}(T) + q_{Ox,vy}(T) \quad , \\ (5.8c)$$

mit

$$\sum_{i=1}^j (\cdot) = 0 \quad \text{für } j < i ,$$

$$\begin{aligned} x &= 1, 2, \dots, N_1 \\ y &= 0, 1, \dots, N_2 \\ v &= 0, 1, \dots, y . \end{aligned}$$

Die Gl. (5.8c) ist eine Rekursionsformel, aus der (beginnend mit der ersten Zeile $y=0$) die Unbekannten $a_{xy}(v)$ bestimmt werden können. Mit jedem Rekursionsschritt erhöht sich zwar die Zahl der zu berechnenden linearen, inhomogenen Gleichungssysteme um eins, sie unterscheiden sich jedoch nur durch die rechten Seiten. Das zugrundeliegende Gleichungssystem hat den Rang N_1 und läßt sich einheitlich in Matrix-Schreibweise darstellen:

$$\underline{Q} \underline{A}_v = \underline{B}_v , \quad v = 0, 1, \dots, y, \quad (5.9)$$

mit

$$\underline{A}_v = (a_{xy}(v)) , \quad \underline{B}_v = (b_{xy}(v)) .$$

Da \underline{Q} eine quadratische Matrix ist und damit die inverse Matrix \underline{Q}^{-1} gebildet werden kann, hat die Gl. (5.9) die Lösung

$$\underline{A}_v = \underline{Q}^{-1} \underline{B}_v , \quad v = 0, 1, \dots, y . \quad (5.10)$$

Zur Berechnung aller \underline{A}_v , $v = 0, 1, \dots, y$, eines Rekursionsschrittes benötigt man somit die Übergangswahrscheinlichkeitsmatrix \underline{Q} mit den Elementen $q_{ix,yy}(T)$, $i, x = 1, 2, \dots, N_1$ und die Spaltenvektoren \underline{B}_v , $v = 0, 1, \dots, y$, der rechten Seiten nach Gl. (5.8c).

Alle Zustandswahrscheinlichkeiten $p(x, y, T)$, $x = 1, 2, \dots, N_1$, $y = 0, 1, \dots, N_2$, des Teilraumes Z_3 sind damit als Funktion der noch unbekannt Wahrscheinlichkeiten $p(0, y, T)$, $y = 0, 1, \dots, N_2$, der ersten Spalte bekannt (vgl. Gl. (5.8b)).

Wie zuvor beschrieben, können die Zustände $\{0, y\}$, $y = 1, 2, \dots, N_2$, des Teilraumes Z_2 von jedem Zustand des gesamten Zustandsraumes erreicht werden, und man erhält folgenden Zusammenhang:

$$p(0, y, T) = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} p(i, j, T) q_{i0, jy}(T) , \quad y = 1, 2, \dots, N_2 . \quad (5.11a)$$

Mit dem Ansatz

$$p(0, y, T) = c_y p(0, 0, T) \quad (5.11b)$$

und unter Verwendung der Gl. (5.8b) können die Zustandswahrscheinlichkeiten des Teilraumes Z_2 durch die letzte noch unbekannt Zustandswahrscheinlichkeit $p(0, 0, T)$ ausgedrückt werden, und man erhält das lineare, inhomogene Gleichungssystem:

$$\begin{aligned} c_y \left[1 - q_{00, yy}(T) - \sum_{i=1}^{N_1} \sum_{j=y}^{N_2} a_{ij}(y) q_{i0, jy}(T) \right] \\ - \sum_{j=1}^{N_2} c_j \left[q_{00, jy}(T) + \sum_{i=1}^{N_1} \sum_{v=j}^{N_2} a_{iv}(j) q_{i0, vy}(T) \right] \\ = \sum_{i=1}^{N_1} \sum_{j=0}^{N_2} a_{ij}(0) q_{i0, jy}(T) + q_{00, 0y}(T) , \quad y = 1, 2, \dots, N_2 , \end{aligned} \quad (5.11c)$$

oder in Matrix-Schreibweise

$$\underline{Q} \cdot \underline{C} = \underline{B} \quad (5.12)$$

mit dem Rang N_2 .

Das Gleichungssystem hat die gleiche Form wie das nach Gl. (5.9) und kann durch Aufruf desselben Unterprogramms im Programm direkt berechnet werden.

Damit sind alle Zustandswahrscheinlichkeiten in Abhängigkeit von $p(0, 0, T)$ bekannt. Durch Einsetzen der Ansätze nach Gl. (5.8b) und Gl. (5.11b) in die bisher noch nicht verwendete Normierungsbedingung nach Gl. (5.7), ergibt sich für die Wahrscheinlichkeit $p(0, 0, T)$ nach kurzer Umformung

$$p(0, 0, T) = \frac{1}{1 + \sum_{y=1}^{N_2} c_y \left[1 + \sum_{x=1}^{N_1} \sum_{v=y}^{N_2} a_{xv}(y) \right] + \sum_{x=1}^{N_1} \sum_{y=0}^{N_2} a_{xy}(0)} . \quad (5.13)$$

Setzt man $p(0,0,T)$ rückwärts in die Gl. (5.11b) und Gl. (5.8b) ein, dann erhält man die Zustandswahrscheinlichkeiten für die Teilräume Z_2 und Z_3 . Damit sind alle Zustandswahrscheinlichkeiten des gesamten Zustandsraumes kurz vor den Taktzeitpunkten bekannt. Aus Gründen der Übersichtlichkeit werden die Lösungsschritte in einem Algorithmus zusammengefaßt.

Algorithmus

1. Rekursive Berechnung der $\underline{A}_v = (a_{xy}(v))$, $x = 1, 2, \dots, N_1$, $y = 0, 1, \dots, N_2$, $v = 0, 1, \dots, Y$, für jede Zeile des Teilraumes Z_3 nach Gl. (5.10) bzw. Gl. (5.8c).
2. Berechnung von c_y , $y = 1, 2, \dots, N_2$, des Teilraumes Z_2 nach Gl. (5.12).
3. Bestimmung der Zustandswahrscheinlichkeit $p(0,0,T)$ nach Gl. (5.13).
4. Endnormierung der Zustandswahrscheinlichkeiten $p(0,y,T)$, $y = 1, 2, \dots, N_2$, des Teilraumes Z_2 mit Gl. (5.11b) und $p(x,y,T)$, $x = 1, 2, \dots, N_1$, $y = 0, 1, \dots, N_2$, des Teilraumes Z_3 mit Gl. (5.8b).

Das hier dargestellte Lösungsverfahren hat, verglichen mit einer direkten iterativen Bestimmung aller Zustandswahrscheinlichkeiten des gesamten Zustandsraumes, folgende Vorteile:

1. Geringerer Speicherplatzbedarf; aufgrund der Zerlegung des Zustandsraumes und der dadurch möglichen rekursiven Berechnung (für jeden Rekursionsschritt benötigt man nur einen Teil der Übergangswahrscheinlichkeiten).
2. Kürzere Rechenzeiten, bedingt durch die direkte Lösungsmöglichkeit der linearen, inhomogenen Gleichungssysteme.

Im folgenden Abschnitt wird die Berechnung der charakteristischen Größen für beide Prioritätsklassen aus den hier berechneten Zustandswahrscheinlichkeiten durchgeführt.

5.2.1.3 Berechnung der charakteristischen Größen

I. Erste Prioritätsklasse

Wegen der Voraussetzung unterbrechender Prioritäten verhält sich das Bedienungssystem so, als sei die zweite Prioritätsklasse nicht vorhanden. Befindet sich zum Taktzeitpunkt z. B. eine Anforderung zweiter Priorität in Bedienung, so wird die Bedienung sofort unterbrochen und nach Ablauf der Verwaltungszeit steht die Bedienungseinheit der ersten Prioritätsklasse zur Verfügung. Da N_2 voraussetzungsgemäß genügend groß gewählt werden muß, sind die Makrozustandswahrscheinlichkeiten $p(x,.,T)$ kurz vor den Taktzeitpunkten

$$p(x,.,T) = \sum_{y=0}^{N_2} p(x,y,T) , \quad x = 0, 1, \dots, N_1 \quad (5.14)$$

identisch den Zustandswahrscheinlichkeiten $p(x,T)$, $x = 1, 2, \dots, N_1$, nach Gl. (4.35) für ein Modell mit einer Anforderungsklasse. Aufgrund dieser Identität können alle Ergebnisse des einklassigen Modells übernommen und die charakteristischen Größen für die erste Prioritätsklasse mit Hilfe der Gleichungen des Abschnitts 4.3.1.2 direkt berechnet werden. Hierbei müßten die Größen prinzipiell noch mit dem Index 1 gekennzeichnet werden, z. B. t_{f1} , W_{01} usw.

II. Zweite Prioritätsklasse

Für die Anforderungen der zweiten Prioritätsklasse wird die mittlere Gesamtwarezeit w_{2G} , die alle Warteanteile berücksichtigt, die mittlere Durchlaufzeit t_{f2} und die Wartewahrscheinlichkeit W_{02} berechnet.

a) Mittlere Gesamtwarezeit w_{2G}

Die mittlere Gesamtwarezeit berücksichtigt die Wartezeit in der Arbeitswarteschlange der zweiten Prioritätsklasse und die durch die Taktunterbrechung verursachten Overhead-Phasen bzw. wegen der Voraussetzung unterbrechender Prioritäten die zwischengeschobenen Bedienungen von Anforderungen der ersten Prioritätsklasse. Ausgangspunkt zur Berechnung von w_{2G} ist deshalb die Gesamtwarebelastung Ω_{2G} . Ist Ω_{2G} bekannt, dann gilt nach dem Little'schen

Theorem

$$w_{2G} = \frac{\Omega_{2G}}{\lambda_2} . \quad (5.15)$$

Analog zur Gl. (4.40) kann Ω_{2G} durch Mittelwertbildung über das gesamte Taktintervall T berechnet werden.

$$\Omega_{2G} = \frac{1}{T} \int_0^T \Omega_{2G}(t) dt . \quad (5.16)$$

Die Wartebelastung $\Omega_{2G}(t)$ ist aus den zweidimensionalen Zustandswahrscheinlichkeiten $p(x,y,t)$ zum Zeitpunkt t entsprechend der Aufteilung des Zustandsraumes (Bild 5.5) wie folgt berechenbar:

$$\Omega_{2G}(t) = \sum_{y=2}^{N_2} (y-1) p(0,y,t) + \sum_{x=1}^{N_1} \sum_{y=1}^{N_2} y p(x,y,t) \quad (5.17)$$

Der erste Anteil berücksichtigt die Zustandswahrscheinlichkeiten des Teilraumes Z_2 . Da hier keine Anforderung der ersten Klasse im System ist, befindet sich eine Anforderung der zweiten Klasse in Bedienung. Der zweite Anteil ergibt sich aus den Zustandswahrscheinlichkeiten des Teilraumes Z_3 . Wegen $x \geq 1$ ist die Bedienungseinheit durch eine Anforderung erster Priorität belegt. Da Ω_{2G} alle Warteanteile berücksichtigt, müssen sich alle Anforderungen zweiter Priorität einschließlich einer zum Taktzeitpunkt unterbrochenen Anforderung in der Arbeitswarteschlange befinden.

Da während der Overhead-Phase keine Anforderungen bedient werden, sind die Zustandswahrscheinlichkeiten $p(x,y,t)$ gleich den Zustandswahrscheinlichkeiten $p(x,y,0)$ kurz nach dem Takt.

Im Bereich $0 < t < T_{OH}$ gilt deshalb

$$p(x,y,t) = p(x,y,0) = \sum_{i=0}^x \sum_{j=0}^y p(i,j,T) g'_{(x-i)_1} g'_{(y-j)_2} , \quad (5.18)$$

mit

$$g'_{k_1} \text{ bzw. } g'_{k_2} \text{ nach Gl. (5.6c),} \quad \begin{matrix} x = 0, 1, \dots, N_1 \\ y = 0, 1, \dots, N_2. \end{matrix}$$

Für $T_{OH} < t < T$ erhält man aus kombinatorischen Überlegungen

$$p(x,y,t) = \sum_{i=x}^{N_1} p(i,y,0) d_{i-x}^1(t-T_{OH}) , \quad \begin{matrix} x = 1, 2, \dots, N_1 \\ y = 1, 2, \dots, N_2 , \end{matrix} \quad (5.19a)$$

$$p(0,y,t) = \sum_{i=0}^{N_1} \sum_{j=y}^{N_2} p(i,j,0) d_{i,j-y}(t-T_{OH}) , \quad y = 1, 2, \dots, N_2 \quad (5.19b)$$

mit

$$d_{0,j-y}(t-T_{OH}) = d_{j-y}^2(t-T_{OH}) .$$

Durch Aufteilung der Integration entsprechend der beiden Zeitschnitte und nach Einsetzen der Gl. (5.18) bzw. der Gl. (5.19) ergibt sich nach Vertauschung der Summation

$$\begin{aligned} \Omega_{2G} = E[X_2,0] \frac{T_{OH}}{T} + \frac{1}{T} \sum_{i=0}^{N_1} \sum_{j=2}^{N_2} p(i,j,0) \sum_{y=2}^j (y-1) d_{i,j-y}^I(T_S) \\ + \frac{h_1}{T} \sum_{i=1}^{N_1} \sum_{y=1}^{N_2} y p(i,y,0) \sum_{x=0}^{i-1} d_{\geq(x+1)}^1(T_S) \end{aligned} \quad (5.20)$$

mit

$$E[X_2,0] = \sum_{j=1}^{N_2} j \sum_{i=0}^{N_1} p(i,j,0) = \sum_{j=1}^{N_2} j p(\cdot, j, 0)$$

und

$$d_{0,j-y}^I(T_S) = h_2 d_{\geq(j-y+1)}^2(T_S) .$$

$E[X_2,0]$ ist die mittlere Anzahl von Anforderungen zweiter Priorität kurz nach dem Takt im System, $p(\cdot, j, 0)$ sind die Makrozustandswahrscheinlichkeiten der zweiten Klasse. Die Integration über die Poisson-Therme $d_m^1(t)$ ist im Anhang 3, die Berechnung von $d_{m,n}^I(T_S)$ (Integration über $d_{m,n}(t)$) im Anhang 4 durchgeführt. Ist das Angebot der ersten Klasse identisch Null, dann sind alle Zustandswahrscheinlichkeiten $p(i,j,0) \equiv 0$ für $i \geq 1$ (Teilraum Z_3) und die Gl. (5.20) entspricht der Gl. (4.42a) für das einklassige Modell. Setzt man Ω_{2G} in die Gl. (5.15) ein, so erhält man die gesuchte

mittlere Gesamtwartezeit w_{2G} .

b) Mittlere Durchlaufzeit t_{f2}

Da die mittlere Gesamtwartezeit w_{2G} bereits alle Warteanteile einer Anforderung zweiter Priorität berücksichtigt, gilt für die mittlere Durchlaufzeit analog zur Gl. (4.48)

$$t_{f2} = w_{2G} + h_2 \quad (5.21)$$

c) Wartewahrscheinlichkeit W_{O2}

W_{O2} gibt an, mit welcher Wahrscheinlichkeit eine Anforderung zweiter Priorität nach Ablauf der Verwaltungszeit warten muß. Die erste Anforderung einer Gruppe von Anforderungen der zweiten Klasse muß nicht warten, wenn das System kurz vor dem Takt leer ist und keine Anforderung der ersten Klasse eintrifft. Für diese komplementäre Wartewahrscheinlichkeit $1 - W_{O2}$ gilt

$$1 - W_{O2} = p(0,0,T) g_{O1} q_2 \cdot$$

q_2 ist die Wahrscheinlichkeit, daß eine beliebige Anforderung einer Gruppe mit $k_2 > 0$ Anforderungen die erste Anforderung ist. Es gilt (vgl. Abschnitt 4.3.1.2e)

$$q_2 = \frac{1}{E[K_2 | K_2 > 0]} = \frac{1 - g_{O2}}{E[K_2]}$$

und damit für die gesuchte Wartewahrscheinlichkeit

$$W_{O2} = 1 - p(0,0,T) g_{O1} \frac{1 - g_{O2}}{E[K_2]} \quad (5.22)$$

5.2.2 Approximative Berechnung des Modells $C^{[x]}/G/1$ mit P Prioritätsklassen

Das approximative Analyseverfahren beruht auf der Rückführung des Modells $C^{[x]}/G/1$ mit P Prioritätsklassen auf P Ersatzmodelle mit einer Anforderungsklasse. Entsprechend Bild 5.2 wird die i-te Klasse betrachtet. Wegen der Voraussetzung "unterbrechender" Prioritäten wird sie von allen Klassen höherer Priorität (1 bis $i - 1$), jedoch nicht von den Klassen niedrigerer Priorität ($i + 1$ bis P) beeinflusst. Faßt man alle Klassen höherer Priorität zu einer gemeinsamen Ersatzklasse zusammen, dann kann zur Berechnung jeder Prioritätsklasse ein zweiklassiges Modell zugrundegelegt werden. Wie im Abschnitt 5.2.1 dargestellt, ist die Berechnung eines zweiklassigen Modells bereits für den einfachen Fall exponentiell verteilter Bedienungszeiten sehr aufwendig. Deshalb wird hier eine approximative Berechnungsmethode für das zweiklassige Modell vorgestellt, die bei der Variation aller Systemparameter gute Ergebnisse verglichen mit der Simulation liefert. Grundidee des Approximationsverfahrens ist die Verwendung eines "Erhaltungssatzes", der für eine vergleichbare Modellstruktur mit nichtgetakteten Einzelankünften Gültigkeit hat.

5.2.2.1 Ein Erhaltungssatz für Prioritätssysteme

Der Erhaltungssatz basiert auf der Tatsache, daß der in einem System zu einem beliebigen Zeitpunkt anstehende Arbeitsumfang nicht von der Bedienungsreihenfolge abhängt, solange das System "arbeits-erhaltend" ("work conservative") ist. Unter arbeitserhaltend versteht man, daß die vorgegebene Reihenfolge der Bedienungen (z. B. nach Prioritäten) weder zusätzliche Arbeit im System verursacht noch die anstehende Systemarbeit reduziert. Zum Beispiel ein Modell mit unterbrechenden Prioritäten, bei dem die unterbrochene Bedienung einer Anforderung von Neuem beginnt (die bereits abgelaufene Teilbedienungszeit muß wiederholt werden, preemptive repeat), verhält sich somit nicht arbeitserhaltend. Dagegen erfüllt immer das System mit nichtunterbrechenden Prioritäten (non-preemptive) und/oder unterbrechenden Prioritäten mit Fortsetzung

der Bedienung ab der Unterbrechungsstelle (preemptive resume) das Arbeitserhaltungsprinzip [18].

Zur Herleitung des Erhaltungssatzes wird das Modell nach Bild 5.6a betrachtet. Die Anforderungen werden entsprechend ihrer Priorität in P parallele Warteschlangen abgelegt. Als Abfertigungsstrategie wird "preemptive resume" vorausgesetzt.

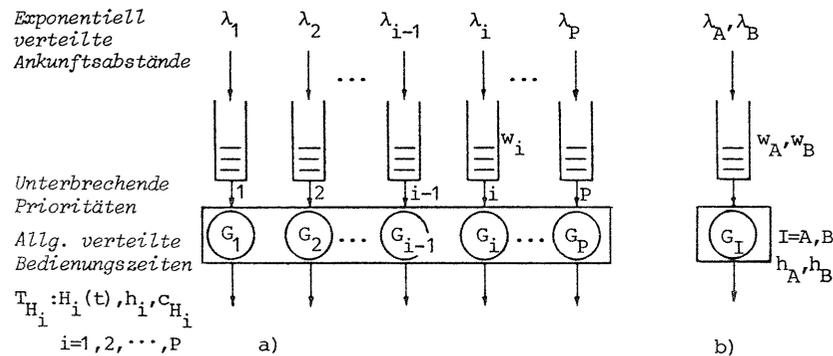


Bild 5.6. a) Warteschlangenmodell mit P Prioritätsklassen
b) Einklassiges Ersatzmodell durch Zusammenfassung der Klassen 1 bis i-1 (A) bzw. 1 bis i (B)

Die Berechnung der mittleren Wartezeit für die Anforderungen der einzelnen Prioritätsklassen kann exakt durchgeführt werden. Das Modell wird z. B. in [49,18,23] ausführlich untersucht. Die dort angegebenen Beziehungen sind jedoch für die approximative Berechnung des zu untersuchenden Modells $C^{[x]}/G/1$ mit Prioritäten nicht geeignet. Im folgenden wird deshalb eine Berechnungsmethode vorgestellt, die einen Zusammenhang zwischen der gesuchten Wartezeit w_i der i-ten Klasse und den Wartezeiten w_A bzw. w_B der im Bild 5.6b dargestellten einklassigen Ersatzmodelle herstellt.

Wegen der gewählten Abfertigungsstrategie wird die i-te Klasse wiederum nur von den Klassen höherer Priorität (1 bis i-1) beeinflusst. Aus der Sicht der i-ten Prioritätsklasse können die Klassen höherer Priorität zu einer gemeinsamen Ersatzklasse (In-

dex A) zusammengefaßt werden und man erhält ein zweiklassiges Modell mit der Ersatzklasse A und der Klasse i. Wegen der Annahme unterbrechender Prioritäten wird die Wartezeit w_A für die Anforderungen der zusammengesetzten Ersatzklasse A nicht durch die Klasse i beeinflusst und entspricht deshalb der Wartezeit des einklassigen Modells nach Bild 5.6b. Den gleichen Modelltyp erhält man durch die Zusammenfassung aller Klassen einschließlich der betrachteten i-ten Klasse; gültig sind hier die mit B gekennzeichneten Größen.

Die Ersatzankunftsraten λ_A und λ_B ergeben sich direkt aus der Überlagerung der Ankunftsrate der beteiligten Prioritätsklassen.

$$\lambda_A = \sum_{j=1}^{i-1} \lambda_j, \quad \lambda_B = \sum_{j=1}^i \lambda_j. \quad (5.23)$$

Aus der gewichteten Summation der einzelnen Bedienungs-VF ergibt sich für die Ersatzbedienungs-VF $H_A(t)$ bzw. $H_B(t)$ nach [49]

$$H_A(t) = \sum_{j=1}^{i-1} \frac{\lambda_j}{\lambda_A} H_j(t), \quad H_B(t) = \sum_{j=1}^i \frac{\lambda_j}{\lambda_B} H_j(t). \quad (5.24)$$

Aus der Sicht der Bedienungseinheit ist die anstehende, auf Bedienung wartende Arbeit in der Ersatzklasse A und der Klasse i des zweiklassigen Modells identisch der wartenden Arbeit des einklassigen Ersatzmodells (B), das beide Klassen zusammenfaßt. Betrachtet man nun zu einem beliebigen Zeitpunkt das zwei- und einklassige Modell, so gilt folgender Erhaltungssatz (conservation law)

$$\Omega_A h_A + \Omega_i h_i + x_i = \Omega_B h_B. \quad (5.25)$$

Die Größe x_i berücksichtigt den wartenden Arbeitsanteil der i-ten Klasse, der sich aufgrund von Unterbrechungen durch die Ersatzklasse A noch im System befindet. Es gilt die Beziehung

$$x_i = \rho_A \frac{\rho_i}{1 - \rho_A} h_{iR}. \quad (5.26)$$

h_{iR} ist die Vorwärtsrekurrenzzeit eines stationären Erneuerungsprozesses und kann nach Gl. (4.4b) aus der VF $H_i(t)$ berechnet werden. Der Faktor $\rho_A \rho_i / (1 - \rho_A)$, der der Wahrscheinlichkeit für das Antreffen einer unterbrochenen Anforderung der i -ten Klasse zu einem beliebigen Zeitpunkt entspricht, läßt sich wie folgt bestimmen. Während der mittleren Gesamtbedienungszeit t_C (completion time) der i -ten Klasse treffen im Mittel $\lambda_A t_C$ Anforderungen der Klasse A ein und werden sofort bedient. Damit gilt der Zusammenhang $t_C = h_i + \lambda_A t_C h_A$. Die mittlere Gesamtunterbrechungszeit t_U ergibt sich damit aus der Beziehung

$$t_U = t_C - h_i = \frac{\rho_A}{1 - \rho_A} h_i .$$

Betrachtet man das System über einen langen Zeitraum $\tau (\tau \rightarrow \infty)$, dann ist $\tau_U = \tau \lambda_i t_U$ die mittlere Zeitdauer, in der sich eine unterbrochene Anforderung der i -ten Klasse im System befindet. Bildet man den Quotienten τ_U / τ , so erhält man die gesuchte Wahrscheinlichkeit.

Unter Verwendung des Little'schen Theorems läßt sich die Gl. (5.25) wie folgt darstellen

$$\rho_A w_A + \rho_i w_i + \rho_A \frac{\rho_i}{1 - \rho_A} h_{iR} = \rho_B w_B . \quad (5.27)$$

Setzt man die Wartezeiten w_A bzw. w_B der einklassigen Ersatzmodelle als bekannt voraus, so ergibt sich schließlich für die gesuchte Wartezeit der Anforderungen der i -ten Klasse

$$w_i = \frac{\rho_B w_B - \rho_A w_A}{\rho_i} - \frac{\rho_A}{1 - \rho_A} h_{iR} , \quad i = 1, 2, \dots, P . \quad (5.28)$$

Diese Beziehung ist eine Rekursionsformel, aus welcher (beginnend mit der höchsten Priorität $i = 1$) die mittleren Wartezeiten für jede Klasse bestimmt werden können, wobei stets nur ein einklassiges Ersatzmodell zu berechnen ist. Die Untersuchung des Modells mit P Prioritätsklassen ist damit auf die Berechnung von P einklassigen Modellen zurückgeführt worden.

Entsprechend der Herleitung ist w_i die Wartezeit bis zum Bedienungsbeginn. Addiert man zu w_i die mittlere Gesamtunterbrechungszeit t_U , so erhält man die effektive Gesamtwarezeit für die i -te Klasse

$$w_{iG} = w_i + \frac{\rho_A}{1 - \rho_A} h_i . \quad (5.29a)$$

Für den Sonderfall exponentiell verteilter Bedienungszeiten erhält man wegen $h_{iR} = h_i$ für die Gesamtwarezeit der i -ten Klasse

$$w_{iG} = \frac{\rho_B w_B - \rho_A w_A}{\rho_i} . \quad (5.29b)$$

Die hier durchgeführten Überlegungen für ein Prioritätsmodell mit unterbrechenden Prioritäten werden im folgenden zur approximativen Berechnung der Wartezeiten für die Anforderungen der einzelnen Prioritätsklassen des zweiten Teilmodells verwendet. Das bei der Zusammenfassung mehrerer Klassen entstehende einklassige Ersatzmodell des Typs $C^{[x]}/G/1$ wurde im Abschnitt 4.3 unter allgemeinen Voraussetzungen untersucht. Zur Berechnung des Ersatzmodells muß die durch die Zusammenfassung der Klassen entstehende Gruppengrößen-Verteilung des Ersatzankunftsprozesses und die Bedienungs-VF des Ersatzbedienungsprozesses bestimmt werden.

5.2.2.2 Gruppengrößen-Verteilung des Ersatzankunftsprozesses

Wegen der synchronen Taktung kann die Gruppengrößen-Verteilung bei der Zusammenfassung mehrerer Prioritätsklassen aus der Faltung der einzelnen Gruppengrößen-Verteilungen bestimmt werden. Ist $g_k^{(r)}$ die Wahrscheinlichkeit dafür, daß die resultierende Gesamtgruppe, die aus der Überlagerung von insgesamt r Klassen entsteht, genau $K = k$ Anforderungen enthält, dann gilt

$$g_k^{(r)} = P\{K = k\} = \sum_{k_1 + k_2 + \dots + k_r = k} g_{k_1} g_{k_2} \dots g_{k_r} . \quad (5.30)$$

Sind die erzeugenden Funktionen $G_{k_v}(z)$, $v = 1, 2, \dots, r$, für die Ver-

teilungen der einzelnen Klassen bekannt, so erhält man für die erzeugende Funktion $G_k^{(r)}(z)$ der Gesamtverteilung

$$G_k^{(r)}(z) = \prod_{v=1}^r G_{k_v}(z) . \quad (5.31)$$

Hieraus können wie folgt die ersten beiden Momente abgeleitet werden.

$$E[k] = G'^{(r)}(z) \Big|_{z=1} = \sum_{v=1}^r E[K_v] , \quad (5.32a)$$

$$E[k^2] = G''^{(r)}(z) \Big|_{z=1} + E[K] . \quad (5.32b)$$

Aus diesen Beziehungen ergibt sich für die Varianz der Gesamtverteilung

$$\text{Var}[K] = \sum_{v=1}^r \text{Var}[K_v] . \quad (5.33)$$

Entsprechen die einzelnen Gruppengrößen-Verteilungen einer Poisson-Verteilung, dann ergibt sich für die Gesamtverteilung ebenfalls eine Poisson-Verteilung.

Mit der Gesamtverteilung und der Taktdauer T liegt somit der Ersatzankunftsprozeß für das einklassige Modell fest. Für allgemein verteilte Gruppengrößen wird die Gesamtverteilung durch die ersten beiden Momente approximiert (vgl. Abschnitt 4.3.2.5b).

5.2.2.3 Bedienung-VF des Ersatzbedienungsprozesses

Zur Berechnung dieser Bedienung-VF soll wiederum von der Überlagerung bzw. Zusammenfassung von r Prioritätsklassen ausgegangen werden. Analog zur Betrachtung im Abschnitt 5.2.2.1 ergibt sich dann für die Bedienung-VF $H_r(t)$ des Ersatzbedienungsprozesses durch gewichtete Summation der Bedienung-VF der einzelnen Klassen

die Beziehung

$$H_r(t) = \sum_{v=1}^r \frac{\lambda_v}{\lambda_r} H_v(t) \quad (5.34)$$

mit

$$\lambda_r = \sum_{v=1}^r \lambda_v .$$

Für die ersten beiden Momente erhält man hieraus

$$h_r = \sum_{v=1}^r \frac{\lambda_v}{\lambda_r} h_v , \quad (5.35a)$$

$$h_r^{(2)} = \sum_{v=1}^r \frac{\lambda_v}{\lambda_r} h_v^{(2)} . \quad (5.35b)$$

Im allgemeinen Fall wird die Bedienung-VF des Ersatzbedienungsprozesses durch die ersten beiden Momente approximiert.

Damit sind alle Größen bekannt, die zur Berechnung des einklassigen Ersatzmodells $C^{[x]}/G/1$ benötigt werden. Mit Hilfe der Gl. (5.29) läßt sich die gesuchte mittlere Gesamtwartezeit w_{iG} für die i -te Klasse approximativ berechnen. Für die Wartezeiten w_A bzw. w_B in dieser Gleichung sind die Gesamtwartezeiten der zugehörigen einklassigen Ersatzmodelle einzusetzen, die den Einfluß der Verwaltungszeiten mitberücksichtigen.

Mit der Beziehung

$$t_{fi} = w_{iG} + h_i , \quad i = 1, 2, \dots, P \quad (5.36)$$

kann schließlich die mittlere Durchlaufzeit der i -ten Klasse bestimmt werden. Sie wird bei der Diskussion der Ergebnisse hauptsächlich betrachtet.

5.2.2.4 Algorithmus

Das in diesem Kapitel entwickelte approximative Analyseverfahren für das Modell $C^{[x]}/G/1$ mit P Prioritätsklassen wird aus Gründen der Übersichtlichkeit in einem Algorithmus zusammengefaßt. Die wesentlichen Schritte des Algorithmus werden im folgenden dargestellt. Ein entsprechendes Programm wurde in der Programmiersprache FORTRAN IV implementiert.

1. Parameter für alle Prioritätsklassen eingeben.
2. $w_A = 0$, $\rho_A = 0$ und Laufindex $i = 0$ setzen.
3. Laufindex i um 1 erhöhen ($i = i + 1$).
4. Alle Klassen 1 bis i zur Ersatzklasse B zusammenfassen. Für das entstehende einklassige Ersatzmodell des Typs $C^{[x]}/G/1$ folgende Größen berechnen:
 - Gruppengrößen-Verteilung des taktmäßigen Ersatzankunftsprozesses nach Abschnitt 5.2.2.2
 - Ersatzbedienungsprozeß nach Abschnitt 5.2.2.3
 - Angebot $\rho_B = \lambda_B h_B$
 - Mittlere Gesamtwarezeit w_B mit Hilfe eines der drei im Abschnitt 4.3 entwickelten Verfahren.
5. Mittlere Wartezeit w_i bzw. w_{iG} für die i -te Prioritätsklasse mit Hilfe des Erhaltungssatzes nach Abschnitt 5.2.2.1 bestimmen.
6. $w_A = w_B$ und $\rho_A = \rho_B$ setzen.
7. Mittlere Durchlaufzeit $t_{if} = w_{iG} + h_i$ für die i -te Klasse berechnen.
8. Im Falle $i < P$ nach Schritt 3 springen und nächste Prioritätsklasse betrachten; sonst berechnete Ergebnisse für alle Klassen ausgeben.

5.3 Das dritte Teilmodell mit zwei Prioritätsklassen

Das dritte Teilmodell berücksichtigt die Ausgabewarteschlangen und die taktmäßige Bedienung. Der Eingangsprozeß wird durch den Ausgangsprozeß (GO_1 bzw. GO_2) des zweiten Teilmodells gebildet. Nach Ablauf der Verwaltungszeit werden zunächst die Anforderungen der ersten Klasse bedient und in der zugehörigen Ausgabewarteschlange zwischengespeichert. In der verbleibenden Zeit bis zum nächsten Takt steht die Bedienungseinheit der zweiten Prioritätsklasse zur Verfügung. Dies bedeutet, daß die beiden Ankunftsprozesse GO_1 und GO_2 nicht gleichzeitig aktiv sind. Bild 5.8 zeigt die Modellstruktur für ein Modell mit zwei Prioritätsklassen.

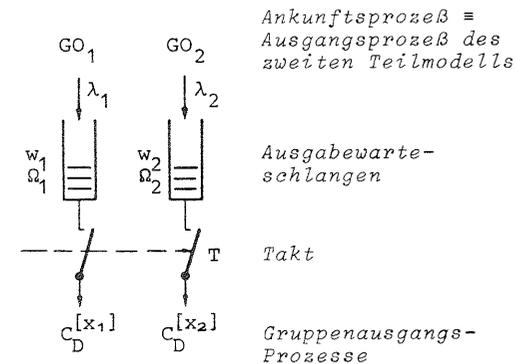


Bild 5.8. Das dritte Teilmodell mit zwei Prioritätsklassen

Die zu berechnenden charakteristischen Verkehrsgrößen sind die mittlere Wartezeit, die mittlere Warteschlangenlänge und der taktmäßige Gruppenausgangsprozeß der beiden Prioritätsklassen.

Wurde das zweite Teilmodell mit Hilfe einer eingebetteten Markoff-Kette (Abschnitt 5.2.1) untersucht, dann können, ausgehend von den Zustandswahrscheinlichkeiten $p(x,y,0)$ des zweiten Teilmodells kurz nach dem Takt, die gesuchten charakteristischen Größen und die Gruppengrößen-Verteilung der beiden Ausgangsprozesse berechnet werden. Dagegen kann bei Verwendung des approximativen Analyseverfahren nach Abschnitt 5.2.2 der Ausgangsprozeß des zweiten Teilmodells nicht explizit bestimmt werden, und das dritte Teil-

modell läßt sich ebenfalls nur approximativ berechnen.

5.3.1 Berechnung der charakteristischen Größen bei Verwendung einer eingebetteten Markoff-Kette

Entsprechend Bild 5.9 wird das System über eine Taktperiode T beobachtet.

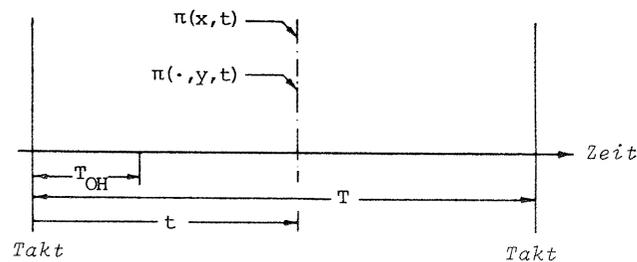


Bild 5.9. Zufälliger Beobachtungszeitpunkt

$\pi(x,t)$ und $\pi(\cdot, y, t)$ sind die Wahrscheinlichkeiten, daß zum Zeitpunkt t nach dem Takt genau x Anforderungen erster Priorität bzw. y Anforderungen zweiter Priorität in der zugehörigen Ausgabewarteschlange zwischengespeichert sind. Da jeweils zu den Taktzeitpunkten alle Anforderungen entnommen werden, entspricht $\pi(x,t)$ der Wahrscheinlichkeit für x Bedienungen erster Priorität seit dem letzten Taktzeitpunkt, $\pi(\cdot, y, t)$ der Wahrscheinlichkeit für y Bedienungen zweiter Priorität unter der Bedingung, daß bereits alle Anforderungen erster Priorität bedient sind. Während der Overhead-Phase sind beide Warteschlangen leer, da in der Zeitspanne $0 < t < T_{OH}$ keine Anforderungen bedient werden.

Im Bereich $0 < t < T_{OH}$ gilt deshalb

$$\begin{aligned} \Omega_1(t) &= 0 \\ \Omega_2(t) &= 0 . \end{aligned} \tag{5.36}$$

Im Bereich $T_{OH} < t < T$ wird die Berechnung für die beiden Prioritätsklassen getrennt durchgeführt.

I. Erste Prioritätsklasse

Wie im Abschnitt 5.2.1.3 erläutert, beeinflusst die zweite Klasse die erste Klasse in keiner Weise, und die Makrozustandswahrscheinlichkeiten $p(x, \cdot, T)$ nach Gl. (5.14) sind identisch den Zustandswahrscheinlichkeiten $p(x, T)$ nach Gl. (4.35) für ein Modell mit einer Anforderungsklasse. Damit können die charakteristischen Größen für die erste Prioritätsklasse mit Hilfe der Gleichungen des Abschnitts 4.4.2 direkt berechnet werden. Prinzipiell müßten die Größen mit dem Index 1 (1. Prioritätsklasse) gekennzeichnet werden, d. h. w_1 und Ω_1 .

II. Zweite Prioritätsklasse

Da mit der Bedienung von Anforderungen der zweiten Klasse erst dann begonnen wird, wenn alle Anforderungen der ersten Klasse bedient sind, erhält man für die Makrozustandswahrscheinlichkeiten $\pi(\cdot, y, t)$ der zweiten Klasse folgende Beziehung:

$$\begin{aligned} \pi(\cdot, y, t) &= p(0, y, 0) d_{\geq y}^2(t - T_{OH}) + \sum_{j=y+1}^{N_2} p(0, j, 0) d_Y^2(t - T_{OH}) \\ &+ \sum_{i=1}^{N_1} p(i, y, 0) \left[d_{i, \geq y}(t - T_{OH}) + \sum_{j=y+1}^{N_2} p(i, j, 0) d_{i, y}(t - T_{OH}) \right] \end{aligned} \tag{5.37}$$

mit

$$\sum_{i=1}^j (\cdot) = 0 \quad \text{für } j < i .$$

Die Bestimmung der mittleren Warteschlangenlänge Ω_2 erfolgt analog zu der des einklassigen Modells (Abschnitt 4.4.2), d. h.

$$\Omega_2 = \frac{1}{T} \int_{T_{OH}}^T \sum_{y=1}^{N_2} y \pi(\cdot, y, t) dt . \tag{5.38}$$

Die Integration über den Bereich $0 < t < T_{OH}$ liefert wegen Gl. (5.36)

keinen Beitrag.

Setzt man die Gl. (5.37) in die Gl. (5.38) ein und verwendet bei der Integration die Beziehungen (A.2) des Anhangs 3 und (A.10) des Anhangs 4, dann ergibt sich nach einigen Umformungen

$$\begin{aligned} \Omega_2 = & E[X_{2,0}] \frac{T - T_{OH}}{T} - \frac{h_2}{T} \sum_{y=1}^{N_2} y p(0,y,0) \sum_{\ell=0}^{y-1} d_{\geq(\ell+1)}(T_S) \\ & + \frac{1}{T} \sum_{y=1}^{N_2} y \sum_{i=0}^{N_1} p(i,y,0) \left[-h_1 \sum_{\ell=0}^{i-1} d_{\geq(\ell+1)}^1(T_S) + \sum_{k=0}^{y-1} (k-y) d_{i,k}^I(T_S) \right] \end{aligned} \quad (5.39)$$

mit

$$d_{0,k}^I(T_S) = h_2 d_{\geq(k+1)}(T_S)$$

und

$$\sum_{i=1}^j (\cdot) = 0 \quad \text{für } j < i .$$

$E[X_{2,0}]$ ist die mittlere Anzahl von Anforderungen der zweiten Klasse kurz nach dem Takt im System. Ist das Angebot der ersten Klasse gleich Null, dann sind die Zustandswahrscheinlichkeiten $p(i,y,0) \equiv 0$ für $i \geq 1$ und die Gl. (5.39) geht in die Gl. (4.106b) des einklassigen Modells über.

Für die mittlere Wartezeit gilt unmittelbar

$$w_2 = \frac{\Omega_2}{\lambda} . \quad (5.40)$$

5.3.2 Berechnung der Ausgangsprozesse

Die Berechnung wird wiederum für jede Prioritätsklasse getrennt durchgeführt.

I. Erste Prioritätsklasse

Der Ausgangsprozeß ist identisch dem Ausgangsprozeß eines einklassigen Modells, das nur die erste Prioritätsklasse berück-

sichtigt. Durch Einsetzen der Makrozustandswahrscheinlichkeiten $p(x, \cdot, T) = p(x, T)$ nach Gl. (5.14) in die Gl. (4.105a) erhält man die Zustandswahrscheinlichkeiten $\pi(x, t)$ für die erste Prioritätsklasse. Mit den Gleichungen des Abschnittes 4.4.3 erhält man dann die Gruppengrößen-Verteilung g_{kD}^1 und deren Momente $E[K_D^1]$ und $\text{Var}[K_D^1]$.

II. Zweite Prioritätsklasse

Setzt man in die Makrozustandswahrscheinlichkeiten der Gl. (5.37) $t = T$, so erhält man für die Wahrscheinlichkeiten g_{kD}^2 der Gruppengrößen-Verteilung analog zur ersten Klasse

$$g_{kD}^2 = \pi(\cdot, k, T) , \quad k = 0, 1, \dots, N_2 . \quad (5.41)$$

Entsprechend Gl. (4.109) und Gl. (4.110) läßt sich der Mittelwert $E[K_D^2]$ und die Varianz $\text{Var}[K_D^2]$ bestimmen. Ist N_2 voraussetzungsgemäß groß, dann ist die mittlere Gruppengröße $E[K_D^2]$ gleich der mittleren Gruppengröße $E[K_2] = \lambda_2 T$ des Ankunftsprozesses der zweiten Klasse im zweiten Teilmodell.

Mit der Betrachtung des Ausgangsprozesses wird die Analyse des dreistufigen Wartemodells $C^{[x]}/G/1$ mit taktgesteuerter Ein-/Ausgabe, Verwaltungszeit und P Prioritätsklassen abgeschlossen. Im nächsten Abschnitt werden exakte bzw. approximativ berechnete Ergebnisse diskutiert und mit Simulationsergebnissen verglichen.

5.4 Ergebnisdiskussion und Simulationsvergleich

Für das in diesem Kapitel behandelte Wartesystem (Bild 5.1) werden im folgenden anhand einiger Beispiele numerische Ergebnisse berechnet und graphisch dargestellt. Wegen der Vielzahl der beeinflussenden Parameter wie z. B.

- Anzahl der Prioritätsklassen
- Ankunftsprozeßtypen im ersten Teilmodell
- Bedienungsprozeßtypen im zweiten Teilmodell
- Taktdauer, Verwaltungszeit, Ankunfts- und Bedienraten

können die Ergebnisse nicht umfassend dargestellt werden. Hierzu müßten umfangreiche Tabellen und Diagramme für die unterschiedlichen Parameterkombinationen erstellt werden. Da die Parameter jedoch stark systemabhängig sind, z. B. vom Grad der Zentralisierung, ist es sinnvoller, für ein interessierendes System die charakteristischen Größen für die drei Teilmodelle explizit mit Hilfe der einzelnen Formeln zu berechnen. Die hier betrachteten Parameterstudien sollen deshalb hauptsächlich Systemengpässe aufzeigen und Fragen der Optimierung, wie z. B. Ermittlung der optimalen Taktdauer für minimale Durchlaufzeiten, Stabilitätsbedingungen für die einzelnen Prioritätsklassen sowie einfühend Überlastaspekte diskutieren.

Entsprechend der Zerlegung des Gesamtmodells in drei Teilmodelle erfolgt die Validierung der numerisch berechneten Ergebnisse.

5.4.1 Numerische Ergebnisse für das erste Teilmodell

Aufgrund der Rückwirkungsfreiheit zwischen dem ersten und zweiten Teilmodell und unter Erneuerungsannahmen zu den Taktzeitpunkten kann jede Eingabewarteschlange für sich betrachtet und durch ein Warteschlangenmodell des Typs GI/C/∞ nachgebildet werden (vgl. Abschnitt 5.1). Da die Ergebnisse für diesen Modelltyp bereits im Abschnitt 4.5.1 diskutiert wurden, kann hier auf eine weitere Betrachtung verzichtet werden. Es sei lediglich daran erinnert, daß die mittlere Wartezeit unter der Voraussetzung eines rekurrenten

Ankunftsprozesses stets gleich der halben Taktdauer ($w = T/2$) ist. Die Berechnung der Gruppengrößen-Verteilung des taktmäßigen Ausgangsprozesses kann, für jede Prioritätsklasse getrennt, mit den Beziehungen der Abschnitte 4.2.1 und 4.2.2 durchgeführt werden.

5.4.2 Numerische Ergebnisse für das zweite Teilmodell

Zur Untersuchung des zweiten Teilmodells wurden im Abschnitt 5.2 zwei verschiedene Analyseverfahren entwickelt:

1. Analyse mit Hilfe einer eingebetteten Markoff-Kette
2. Approximative Berechnung durch Rückführung auf einklassige Ersatzmodelle.

Da die Methode der eingebetteten Markoff-Kette bereits für zwei Prioritätsklassen und exponentiell verteilten Bedienungszeiten einen erheblichen numerischen Aufwand benötigt, wurde das Approximationsverfahren entwickelt. Es beruht auf der Rückführung des gegebenen Modells mit P Prioritätsklassen auf P einklassige Modelle. Für diesen Modelltyp wurden im Abschnitt 4.3 drei verschiedene Analyseverfahren abgeleitet. Zusätzlich wurde gezeigt, daß aufgrund der Voraussetzung unterbrechender Prioritäten zur Berechnung der Kenngrößen einer beliebigen Klasse $i, i \geq 2$, stets auf ein Modell mit 2 Prioritätsklassen zurückgegriffen werden kann. Den folgenden Betrachtungen kann deshalb ohne Einschränkung der Allgemeinheit ein zweiklassiges Modell zugrundegelegt werden. Für die Ergebnisdarstellung gelten einheitlich folgende Vereinbarungen:

- ▲ Eingebettete Markoff-Kette (EMK)
- Approximation durch klassenweise Zerlegung und Berechnung der einklassigen Ersatzmodelle mit Hilfe der Lindley'schen Integralmethode (AKZ-LIM)
- ‡ Simulation mit 95%-Vertrauensintervall und 10 Teiltests mit je 20 000 Anknunftereignissen für beide Klassen.

An einem ersten Beispiel eines zweiklassigen Modells mit gleichen Ankunfts- und Bedienungsprozeßtypen (E_2, M, H_2) soll der Einfluß der wesentlichen Parameter auf die mittleren Durchlaufzeiten t_{f1} und t_{f2} beider Klassen gezeigt und eine Validierung der durch klassenweise Zerlegung ermittelten Ergebnisse durchgeführt werden. Alle Zeitgrößen sind normiert auf die für beide Klassen gleiche mittlere Bedienungsdauer $h = 1$. Da die erste Klasse hierbei wegen der Voraussetzung unterbrechender Prioritäten exakt dem einklassigen Modell nach Abschnitt 4.3 entspricht, ist hier nur eine Überprüfung der Ergebnisse für die zweite Klasse erforderlich.

a) Einfluß des Ankunftsprozeßtyps im ersten Teilmodell

Bild 5.7 zeigt den Einfluß des Ankunftsprozeßtyps (E_2, M, H_2) im ersten Teilmodell auf die mittleren Durchlaufzeiten des zweiten Teilmodells als Funktion des Angebots. Die Bedienungszeiten sind exponentiell verteilt. Die Darstellung zeigt das typische Verhalten von Prioritätssystemen. Während die Stabilitätsgrenze für die zweite Klasse, die durch das Verhalten des Gesamtsystems (im Bild gestrichelt dargestellt) mit $\rho = \rho_1 + \rho_2$ bestimmt wird, bereits bei $\rho = 1$ ($\rho_1 = \rho_2 = 0,5$) erreicht wird, gilt für die erste Klasse als Stabilitätsgrenze $\rho_1 = 1$. Die mit Hilfe der eingebetteten Markoff-Kette berechneten Ergebnisse für die M- und E_2 -VF und die Simulationsergebnisse für die H_2 -VF zeigen, daß die klassenweise Betrachtung des Modells in Verbindung mit der Lindley'schen Integralmethode zu sehr guten Ergebnissen führt. Da sich bei Verwendung der GI/G/1-Approximation zur Berechnung der einklassigen Ersatzmodelle größere Abweichungen ergeben, wird sie nicht weiter betrachtet.

Es sei besonders vermerkt, daß der Schnittstellenverkehr zwischen dem ersten und zweiten Teilmodell bei der Berechnung durch klassenweise Zerlegung für E_k - und H_2 -verteilte Ankunftsprozeßtypen im ersten Teilmodell (allgemein verteilte Gruppengrößen des Ankunftsprozesses im zweiten Teilmodell) nur durch die ersten beiden Momente approximiert wird. Die Berechnung dieser Momente, d. h. die Bestimmung von Mittelwert und Varianz der Gruppengrößen-Verteilungen, erfolgt nach Abschnitt 4.2.2 unter Erneuerungsan-

nahmen zu den Taktzeitpunkten und wird für jede Klasse getrennt durchgeführt. Bei der Zusammenfassung beider Prioritätsklassen zu einem einklassigen Ersatzmodell ergibt sich der Mittelwert und die Varianz der Gruppengrößen-Verteilung des Ersatzankunftsprozesses wegen der Rückwirkungsfreiheit zwischen den beiden Teilmodellen direkt aus der Addition der einzelnen Mittelwerte bzw. Varianzen (vgl. Abschnitt 5.2.2.2). Die Ergebnisvergleiche im Bild 5.7 zeigen, daß die Schnittstellenbeschreibung durch die ersten beiden Momente eine gute Näherung darstellt.

b) Einfluß des Bedienungsprozeßtyps

Im Bild 5.8 ist der Einfluß des Bedienungsprozeßtyps auf die Durchlaufzeiten t_{f1}^* und t_{f2}^* im ersten und zweiten Teilmodell als Funktion der Taktdauer dargestellt. Da die Wartezeit in beiden Eingabewarteschlangen $w = T/2$ beträgt, gilt der Zusammenhang

$$t_{fi}^* = t_{fi} + T/2, \quad i = 1, 2. \quad (5.42)$$

Wegen der Voraussetzung Poisson-verteilter Gruppengrößen hat auch die bei der Zusammenfassung der beiden Klassen entstehende Gruppengrößen-Verteilung eine Poisson-Verteilung. Die Approximation bei der Berechnung des entstehenden einklassigen Ersatzmodells mit Hilfe der Lindley'schen Integralmethode besteht deshalb nur in der Zeitdiskretisierung der Ersatzbedienungs-VF. Hieraus ergibt sich die gute Übereinstimmung zwischen den approximativ und exakt (M-Bedienungs-VF) bzw. simulativ (E_2 -, H_2 -Bedienungs-VF) ermittelten Ergebnissen. Die Darstellung zeigt prinzipiell für beide Klassen die gleichen Charakteristiken, jedoch ist der Einfluß des Bedienungsprozeßtyps auf die zweite Klasse stärker, insbesondere im Bereich der minimalen Durchlaufzeiten.

c) Einfluß der Verwaltungszeit

Den Einfluß der Verwaltungszeit auf die mittleren Durchlaufzeiten t_{f1}^* und t_{f2}^* zeigt das Bild 5.9. Für $T_{OH} = 0$ und $T \rightarrow 0$ erhält man die gleichen Durchlaufzeiten wie für das Warteschlangenmodell des Typs M/M/1 mit unterbrechenden Prioritäten. Der Warteanteil ($w = T/2$) in den Eingabewarteschlangen geht hierbei gegen Null.

Abhängig von der Verwaltungszeit ergibt sich für jede Klasse eine unterschiedliche minimale Taktdauer, die unter stationären Voraussetzungen nicht unterschritten werden darf. Auf die Bestimmung dieser minimalen Taktdauern wird in einem zweiten Beispiel näher eingegangen.

Die Bilder 5.7 bis 5.9 zeigen, daß die klassenweise Zerlegung in Verbindung mit der Lindley'schen Integralmethode bei Variation der verschiedenen Parameter zu sehr guten Ergebnissen führt. Dies konnte auch durch weitere umfangreiche Untersuchungen und Ergebnisvergleiche bestätigt werden, bei denen insbesondere unterschiedliche Ankunfts- und/oder Bedienungsprozeßtypen mit verschiedenen Ankunfts- und/oder Bedienraten für beide Klassen berücksichtigt wurden. Für nicht Poisson-verteilte Gruppengrößen wurde hierbei stets auf die Approximation durch die ersten beiden Momente zurückgegriffen.

Mit Hilfe eines zweiten Beispiels soll die Bestimmung einer Stabilitätsbedingung für jede Klasse, Fragen der Optimierung und einfürend Überlastaspekte diskutiert werden. Für die Gruppengrößen werden Poisson-Verteilungen angenommen, die Bedienungszeiten sind exponentiell verteilt mit den mittleren Bedienungszeiten $h_1 = 0.5$ und $h_2 = 1$. Wegen diesen Voraussetzungen können die Ergebnisse direkt mit den exakt berechenbaren der eingebetteten Markoff-Kette verglichen werden. Dargestellt werden wiederum die mittleren Durchlaufzeiten t_{f1}^* und t_{f2}^* . Alle Zeitgrößen sind normiert auf die mittlere Bedienungszeit der zweiten Klasse ($h_2 = 1$, für ein reales System z. B. 1 ms).

Bild 5.10 zeigt den Einfluß der Ankunftsrate λ_2 auf die mittlere Durchlaufzeit t_{f2}^* der zweiten Klasse in Abhängigkeit von der Taktdauer T . Wie bereits den Darstellungen in Bild 5.8 und 5.9 entnommen werden kann, ergibt sich für jede Klasse eine minimale Taktdauer T_{min}^i , $i = 1, 2$, die unter stationären Voraussetzungen ($\rho < 1$) nicht unterschritten werden darf. Für die i -te Klasse eines mehrklassigen Modells ergibt sich analog zur Herleitung der Gl. (4.125) die Bedingung:

$$\lambda_i T \leq \mu_i (T - T_{OH} - \sum_{j=1}^{i-1} \lambda_j T) , \quad i = 1, 2, \dots .$$

Der auf der rechten Seite stehende Klammerausdruck gibt den im Mittel zur Bedienung der i -ten Klasse effektiv zur Verfügung stehenden Anteil der Taktdauer T an. In der Grenze erhält man als Stabilitätsbedingung für die i -te Klasse:

$$T_{min}^i = \frac{T_{OH}}{1 - \sum_{j=1}^i \rho_j} , \quad i = 1, 2, \dots . \quad (5.43)$$

Für die erste Prioritätsklasse ($i = 1$) ergibt sich hieraus unmittelbar die Beziehung nach Gl. (4.112). Die Darstellung im Bild 5.10 zeigt weiterhin, daß die Durchlaufzeiten ein ausgeprägtes Minimum aufweisen. Die zugehörigen optimalen Taktdauern sind jedoch für beide Klassen verschieden und zusätzlich abhängig von den Ankunftsraten bzw. Angeboten. So verschiebt sich z. B. das Optimum für die zweite Klasse von $T_{opt}^2 \approx 4$ bei $\lambda_2 = 0.2$ zu $T_{opt}^2 \approx 10$ bei $\lambda_2 = 0.6$. Läßt die Systemrealisierung nur die Wahl einer festen Taktdauer zu, so muß ein geeigneter Kompromiß gefunden werden. Hierbei ist jedoch zu beachten, daß die gewählte Taktdauer bei Änderungen des Angebots (Variation der Ankunftsraten) die in der Gl. (5.43) angegebene Stabilitätsbedingung für beide Klassen erfüllen muß. Wählt man etwa im betrachteten Beispiel die Taktdauer $T = 5$, so ist das Wartesystem für Ankunftsraten $\lambda_2 \geq 0.7$ für die zweite Klasse nicht mehr stabil. Es bietet sich deshalb an, insbesondere im Hinblick auf Fragen der Überlaststeuerung, die Taktdauer lastabhängig zu verändern. Ein geeignetes Kriterium für die Umschaltung des Taktes ist z. B. die Überwachung und Zählung der fertig bearbeiteten Anforderungen einer bestimmten Prioritätsklasse innerhalb einer vorgegebenen Zeitspanne. Erhöht man bei ansteigender Belastung die Taktdauer, so wird aufgrund des geringeren Ein-/Ausgabe-Verwaltungsanteils (Verhältnis T_{OH}/T) Realzeit-Kapazität zurückgewonnen. Hierbei erhöhen sich zwar auch die Durchlaufzeiten für die erste Klasse wegen des größeren Warte-

anteils in der zugehörigen Eingabewarteschlange, jedoch ist die Stabilitätsbedingung für die zweite Klasse wieder erfüllt. Bei stärkerer Überlast ist dieses Verfahren jedoch nicht mehr anwendbar, da zur Einhaltung vorgeschriebener Reaktionszeiten eine maximale Taktdauer nicht überschritten werden darf. Hier müssen andere Maßnahmen ergriffen werden, wie z. B. die Begrenzung der Anzahl übertragener Anforderungen je Takt und Prioritätsklasse. Da die Maßnahmen zur Optimierung und Überlaststeuerung stark systemabhängig sind, sollen sie hier nicht weiter diskutiert werden.

Einen Vergleich zwischen der taktgesteuerten und interruptgesteuerten Ein-/Ausgabe zeigt ein drittes Beispiel im Bild 5.11. Wie bereits im Abschnitt 2.3.3 für ein einklassiges Modell gezeigt wurde, führt die taktgesteuerte Methode bei höheren Lasten zu geringeren Durchlaufzeiten. Für ein mehrklassiges System bedeutet dies, daß insbesondere die Durchlaufzeiten der niedrigeren Prioritätsklassen erheblich reduziert werden können. Im Bild 5.11 wird deshalb zum Vergleich nur die zweite Klasse betrachtet. Zur Bestimmung der Durchlaufzeiten bei der interruptgesteuerten Methode wurde auf die von Herzog in [8] angegebene Beziehung zurückgegriffen. Eine Begründung für die Annahme gleicher Verwaltungszeiten für beide Ein-/Ausgabe-Mechanismen wurde bereits im Abschnitt 2.3.2 gegeben. Die Darstellung zeigt, daß die taktgesteuerte Methode für höhere Ankunftsrate zu geringeren Durchlaufzeiten führt. Die eingezeichneten Schnittpunkte sind von den Verkehrs- und Systemparametern abhängig. Während der Ein-/Ausgabe-Verwaltungsanteil bei der taktgesteuerten Methode nur vom Verhältnis T_{OH}/T bestimmt wird, ist er bei der interruptgesteuerten Methode von T_{OH} und den Ankunftsrate beider Klassen abhängig. Hieraus ergeben sich die unterschiedlichen Auslastungsgrenzen für beide Verfahren. Aus der Gl. (5.43) erhält man für die taktgesteuerte Ein-/Ausgabe die maximale Ankunftsrate

$$\lambda_2^T = \mu_2 \left(1 - \frac{T_{OH}}{T} - \rho_1 \right) . \quad (5.44a)$$

Für die interruptgesteuerte Methode ergibt sich aus der Beziehung

$$\rho_1 + \rho_2 + (\lambda_1 + \lambda_2) T_{OH} \leq 1$$

die Auslastungsgrenze

$$\lambda_2^I = \frac{1 - \rho_1 - \lambda_1 T_{OH}}{h_2 + T_{OH}} . \quad (5.44b)$$

Mit diesem Vergleich soll die Ergebnisdiskussion für das zweite Teilmodell abgeschlossen werden. Zusammenfassend kann festgestellt werden, daß die klassenweise Betrachtung in Verbindung mit der Lindley'schen Integralmethode gute Ergebnisse liefert.

5.4.3 Numerische Ergebnisse für das dritte Teilmodell

Der Schnittstellenverkehr zwischen dem zweiten und dritten Teilmodell ist nur dann explizit bestimmbar, wenn das zweite Teilmodell mit Hilfe einer eingebetteten Markoff-Kette analysiert wird (vgl. Abschnitt 5.3). Der Eingangsprozeß des dritten Teilmodells, der dem Ausgangsprozeß des zweiten Teilmodells entspricht, ist nicht rekurrent. Analog zu den Überlegungen des einklassigen Modells (Abschnitt 4.5.3) kann das dritte Teilmodell auch hier "isoliert" betrachtet und die Wartezeiten in den Ausgabewarteschlangen approximativ bestimmt werden. Zur Bedienung von Anforderungen der i -ten Klasse steht im Mittel nur der Anteil $T_{eff}^i = T - T_{OH} - \rho_{\leq(i-1)} T$ der gesamten Taktdauer T zur Verfügung. $\rho_{\leq(i-1)}$ ist das Gesamtangebot aller Klassen 1 bis $i-1$. Für $\rho_i \rightarrow 0$ ergibt sich die maximal mögliche Wartezeit $w_{max}^i = T_{eff}^i - h_i$. Geht dagegen $\rho_i \rightarrow 1$, dann ist die Bedienungseinheit während der gesamten zur Verfügung stehenden Zeitdauer T_{eff}^i belegt und man erhält $w_{min}^i = T_{eff}^i/2$. In erster Näherung läßt sich die Abhängigkeit zwischen der Wartezeit und dem Angebot für jede Klasse durch einen linearen Zusammenhang darstellen. Für $i=1$ ergeben sich unmittelbar die im Abschnitt 4.5.3 angegebenen Beziehungen. In der dort angegebenen Tabelle 3 wurde außerdem gezeigt, daß die verschie-

denen Ankunfts- und/oder Bedienungsprozeßtypen keinen wesentlichen Einfluß auf die mittleren Wartezeiten haben. Die Genauigkeit der Approximation wird im nächsten Abschnitt bei der Betrachtung der Gesamtdurchlaufzeiten gezeigt.

Eine approximative Berechnung der taktmäßigen Gruppenausgangsprozesse $c_D^{[x_i]}$ ist kaum möglich. Während sich bereits für die erste Klasse die im Bild 4.24 gezeigten komplexen Abhängigkeiten ergeben, werden die Ausgangsprozesse der niedrigeren Prioritätsklassen zusätzlich von allen Parametern der höheren Klassen beeinflusst. Im Einzelfall muß deshalb entweder auf die Berechnung nach Abschnitt 5.3.2 oder auf die Simulation zurückgegriffen werden.

5.4.4 Mittlere Gesamtdurchlaufzeiten.

Die mittleren Gesamtdurchlaufzeiten ergeben sich aus der Addition der mittleren Wartezeiten in den Ein- und Ausgabewarteschlangen sowie den mittleren Durchlaufzeiten im zweiten Teilmodell. Da die einzelnen Mittelwerte für die "isoliert" betrachteten Teilmodelle eine gute Übereinstimmung mit den Ergebnissen der eingebetteten Markoff-Kette für exponentiell verteilte Bedienungszeiten oder unter allgemeineren Voraussetzungen mit der Simulation zeigen, gilt dies auch für die Gesamtdurchlaufzeiten. Für jede Klasse gilt hierbei analog zur Gl. (4.113) des einklassigen Modells die Beziehung:

$$t_{FG}^i \geq 1.5 \cdot T, \quad i = 1, 2, \dots \quad (5.45)$$

Die Darstellung im Bild 5.12 zeigt die Abhängigkeit der Gesamtdurchlaufzeiten für drei Klassen in Abhängigkeit vom Angebot für zwei verschiedene Takt Dauern bei gleicher Verwaltungszeit. Im wesentlichen setzt sich die Charakteristik des zweiten Teilmodells durch (vgl. Bild 5.7). Der Warteanteil in den Ausgabewarteschlangen erhöht zwar die Gesamtdurchlaufzeiten beider Klassen (insbesondere im unteren Angebotsbereich), jedoch bewirkt er auch eine

Linearisierung des Kurvenverlaufs. Die eingezeichneten exakt berechneten Ergebnisse zeigen, daß die gewählte Approximation bei der "isolierten" Betrachtung des dritten Teilmodells zu guten Ergebnissen führt. Dies zeigten auch weitere Untersuchungen mit anderen Ankunfts- und/oder Bedienungsprozeßtypen, die hier jedoch nicht weiter diskutiert werden.

Wegen den Markoff-Voraussetzungen und der Annahme gleicher, mittlerer Bedienungsdauern für alle drei Klassen läßt sich auch das dreistufige Modell mit Hilfe einer eingebetteten Markoff-Kette exakt berechnen. Zur Berechnung der dritten Klasse werden hierzu die beiden ersten Klassen zu einer Ersatzklasse mit gleicher Bedienungs-VF wie für die einzelnen Klassen zusammengefaßt, und man erhält ein zweiklassiges Modell.

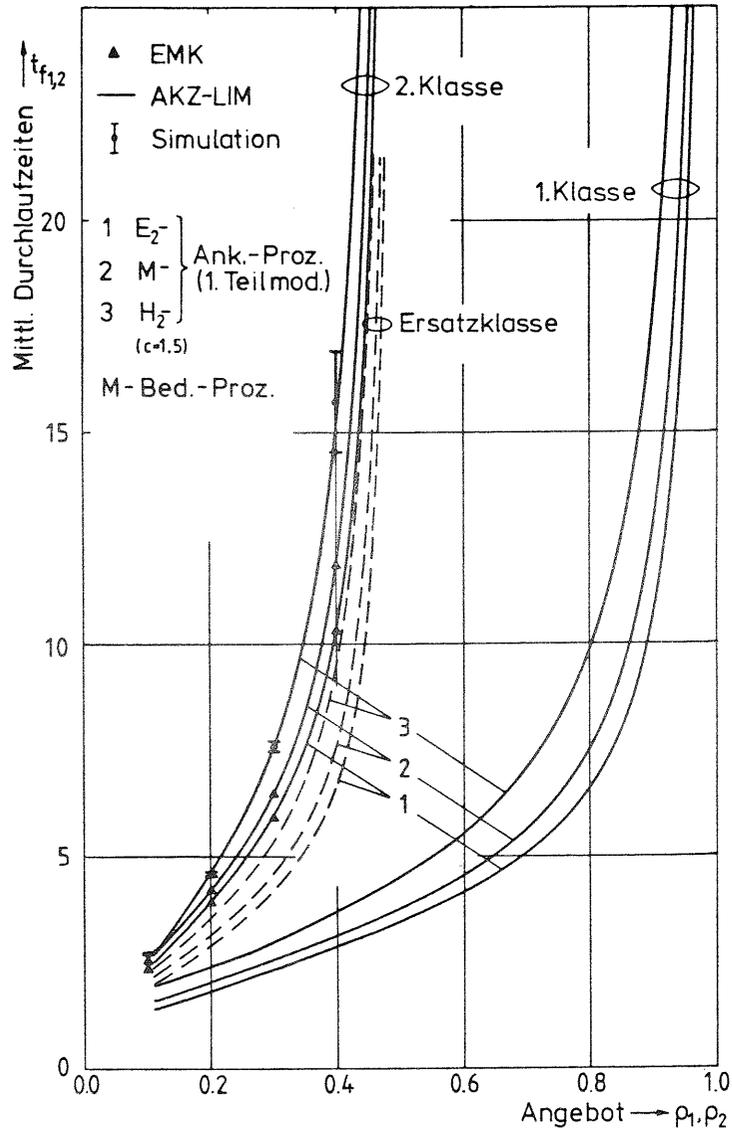


Bild 5.7. Einfluß des Ankunftsprozeßtyps im ersten Teilmodell auf die mittleren Durchlaufzeiten des zweiten Teilmodells
 Parameter: $h_1 = h_2 = 1$, $T = 10$, $T_{OH} = 0$

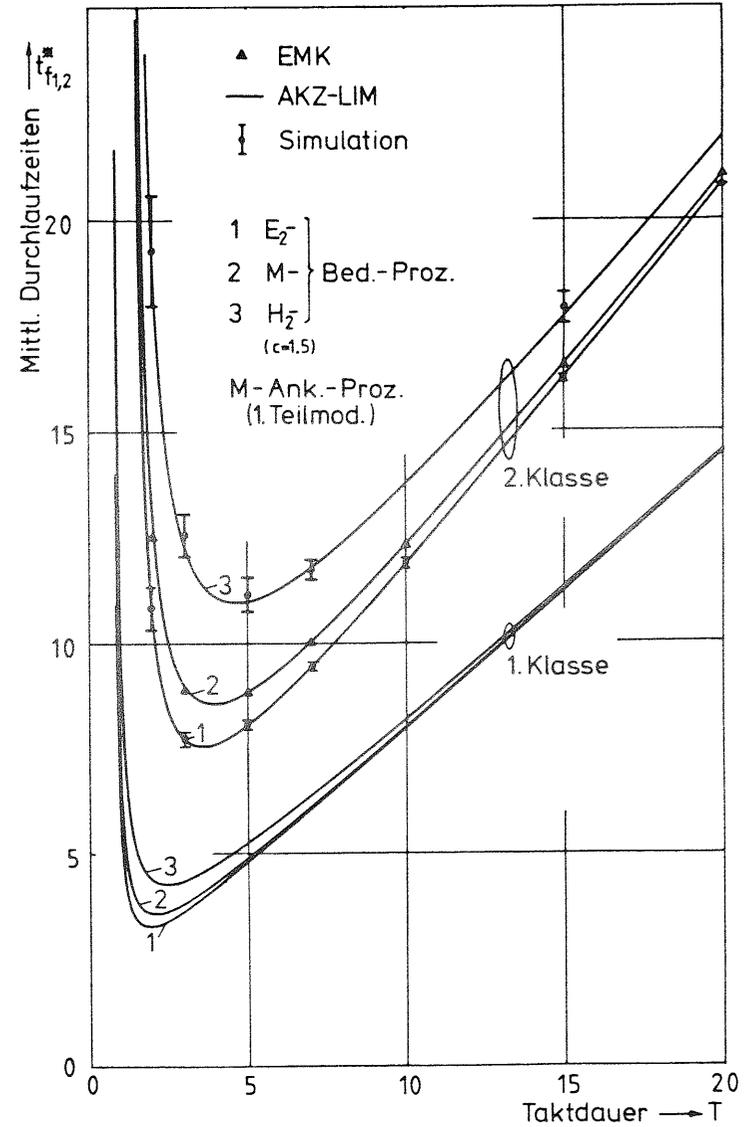


Bild 5.8. Einfluß des Bedienungsprozeßtyps auf die mittleren Durchlaufzeiten des ersten und zweiten Teilmodells
 Parameter: $\lambda_1 = \lambda_2 = 0.3$, $h_1 = h_2 = 1$, $T_{OH} = 0.5$

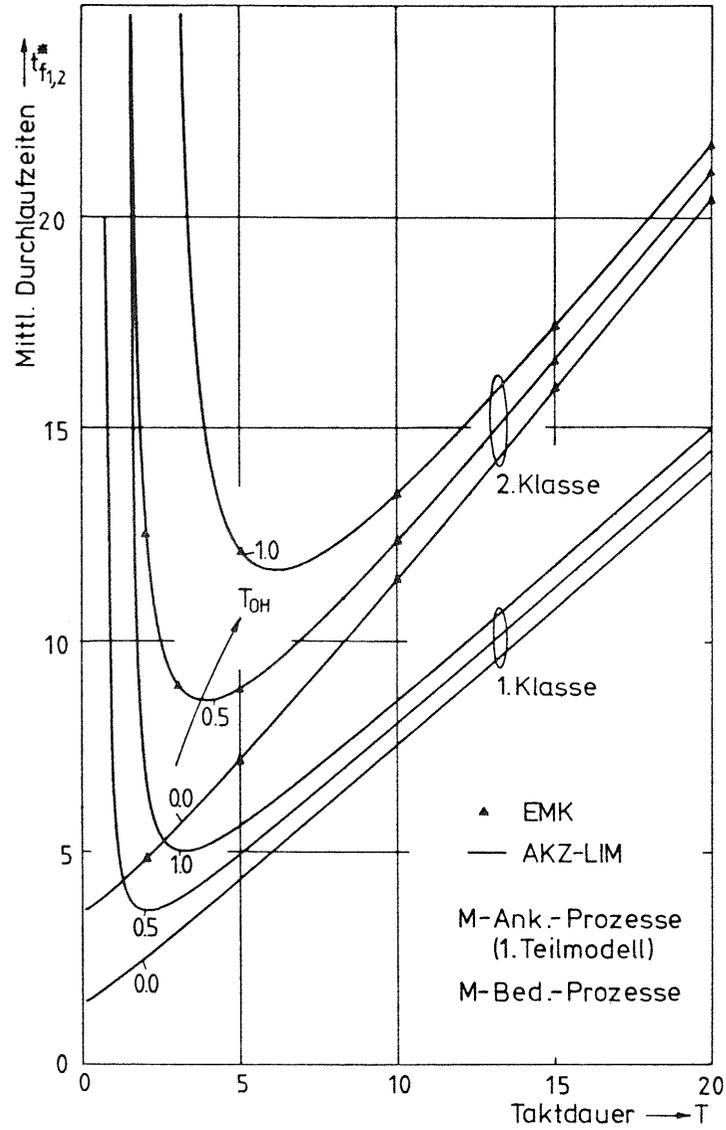


Bild 5.9. Einfluß der Verwaltungszeit auf die mittleren Durchlaufzeiten des ersten und zweiten Teilmodells
 Parameter: $\lambda_1 = \lambda_2 = 0.3$, $h_1 = h_2 = 1$

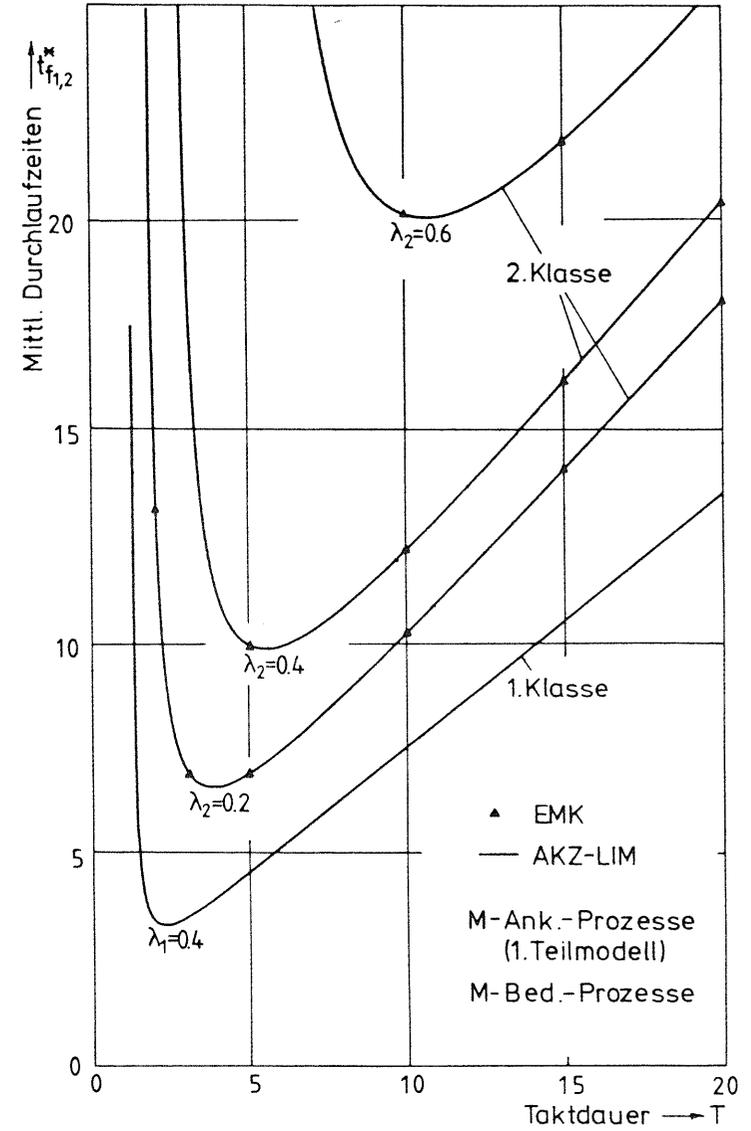


Bild 5.10. Zur Ermittlung der minimalen und optimalen Taktdauer Einfluß der Ankunftsrate λ_2 auf die mittlere Durchlaufzeit der 2. Klasse im ersten und zweiten Teilmodell
 Parameter: $h_1 = 0.5$, $h_2 = 1$, $T_{OH} = 1$

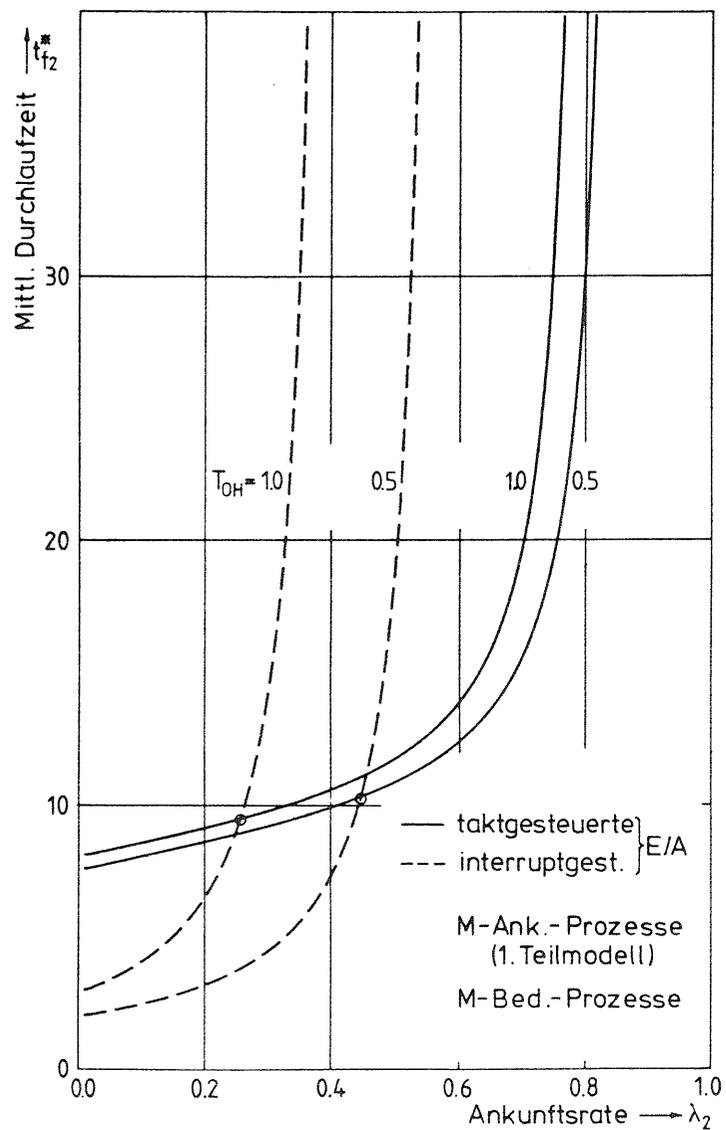


Bild 5.11. Vergleich der mittleren Durchlaufzeiten für die zweite Klasse bei takt- und interruptgesteuerter Ein-/Ausgabe
 Parameter: $\lambda_1 = 0.1, h_1 = h_2 = 1, T = 10$

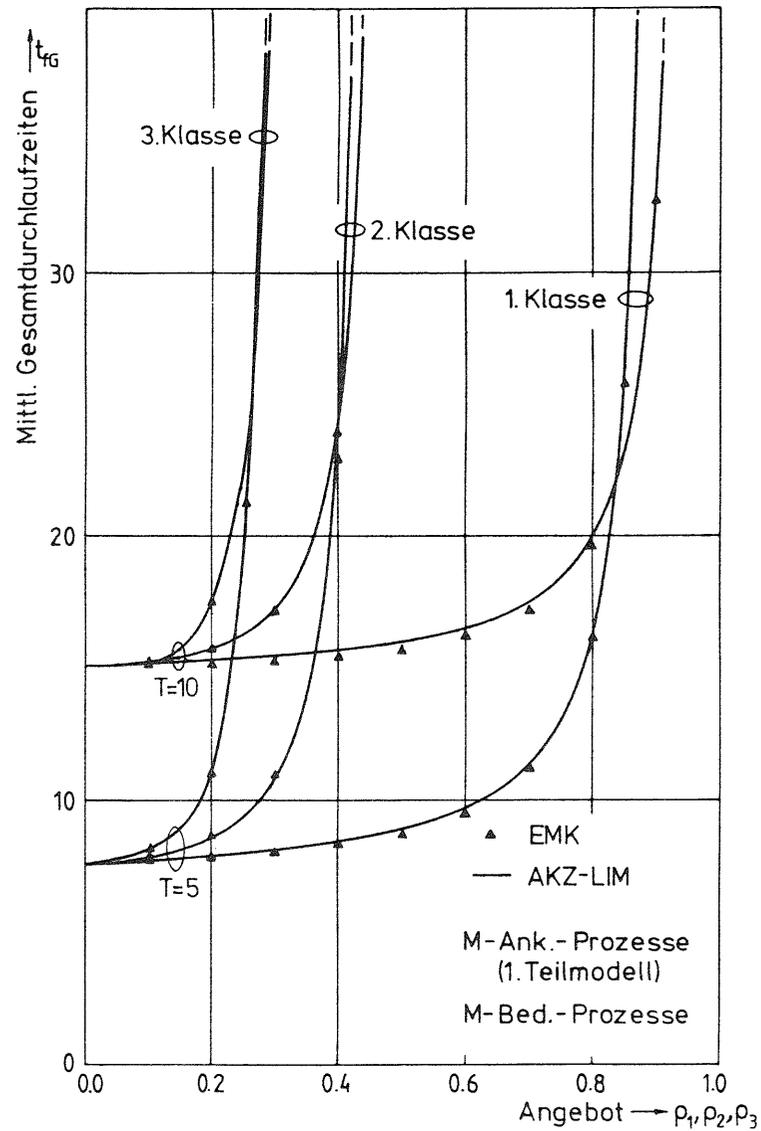


Bild 5.12. Mittlere Gesamtdurchlaufzeiten für drei Prioritätsklassen in Abhängigkeit vom Angebot
 Parameter: $h_1 = h_2 = h_3 = 1, T_{OH} = 0.5$

6. ZUSAMMENFASSUNG

Bei modernen rechnergesteuerten Vermittlungssystemen hat sich die Steuerungsstruktur in den letzten Jahren von einer zunächst rein zentralen zu einer gemischt zentral/dezentralen Struktur verändert. Der Umfang des Steuerdatenaustauschs zwischen den dezentralen Steuerungsmodulen und der zentralen Steuereinheit (Interprozessor-Kommunikation) bestimmt maßgeblich die Leistungsfähigkeit des Systems. Die Ein-/Ausgabe der Anforderungen wird bei realen Systemen häufig von der zentralen Steuereinheit durchgeführt und erfolgt in der Regel taktgesteuert, d. h. in äquidistanten zeitlichen Abständen. Die zu bearbeitenden Anforderungen werden hierbei, entsprechend ihrer Dringlichkeit, verschiedenen Prioritätsklassen zugeordnet.

Zur Untersuchung und Leistungsbewertung derartiger Systeme werden die Verkehrsflüsse durch zufallsabhängige Prozesse beschrieben und das auf ein Modell abgebildete System nach verkehrstheoretischen Gesichtspunkten analysiert. In der vorliegenden Arbeit wurde ein prinzipieller Überblick über die verkehrstheoretische Modellbildung gegeben, wobei insbesondere Modelle zur Nachbildung von Ein-/Ausgabe-Mechanismen vorgestellt wurden. Bei der Diskussion der Methoden der Modellanalyse, die in analytische und simulative Verfahren einteilbar sind, wurden schwerpunktmäßig diejenigen näher betrachtet, die im Rahmen dieser Arbeit verwendet wurden.

In den abschließenden Kapiteln wurden dreistufige Wartemodelle mit taktgesteuerter Ein-/Ausgabe, beginnend mit einer Anforderungsklasse und abschließend mit P Prioritätsklassen, betrachtet. Die Untersuchung erfolgte durch Zerlegung der Gesamtmodelle in geeignete Teilmodelle, die "isoliert" betrachtet und analysiert wurden. Bei der Betrachtung der einklassigen Modellstruktur wurde für das zweite Teilmodell des Typs $C^{[x]}/G/1$ ein neues Analyseverfahren entwickelt, das von der Lindley'schen Integralgleichung ausgeht und für allgemeine Gruppengrößen-Verteilungen und Bedienungs-VF gute Ergebnisse liefert. Die Lösung der Integralgleichung

erfolgte durch Zeitdiskretisierung der vorgegebenen Bedienungs-VF.

Die Untersuchung eines zweiklassigen Modells mit unterbrechenden Prioritäten und exponentiell verteilten Bedienungszeiten konnte mit Hilfe einer eingebetteten Markoff-Kette durch eine zweidimensionale Zustandsbeschreibung durchgeführt werden. Da dieses Verfahren jedoch numerisch sehr aufwendig und unter allgemeinen Voraussetzungen nicht mehr anwendbar ist, wurde ein approximatives Analyseverfahren entwickelt, welches das Modell $C^{[x]}/G/1$ mit P Prioritätsklassen durch sukzessive Zusammenfassung von Klassen auf P einklassige Modelle zurückführt, die dann mit Hilfe der Lindley'schen Integralmethode berechnet werden können. Bei der Bestimmung der Wartezeiten für die einzelnen Klassen wurde hierbei zusätzlich auf einen "Erhaltungssatz" zurückgegriffen, der für ein Modell des Typs M/G/1 mit unterbrechenden Prioritäten abgeleitet wurde und für dieses Modell exakte Ergebnisse liefert. Mit Hilfe eines entwickelten Simulationsprogramms, das das gesamte Wartemodell unter Einbeziehung aller gegenseitigen Abhängigkeiten nachbildet, wurde die Genauigkeit der verschiedenen Approximationsverfahren überprüft. Die numerisch gewonnenen und graphisch dargestellten Ergebnisse, die den Einfluß der wesentlichen System- und Verkehrsparameter auf die mittlere Durchlauf- bzw. Reaktionszeit des Gesamtsystems und der Teilsysteme aufzeigen, können als Grundlage zur Systemoptimierung herangezogen werden.

Die im Rahmen dieser Arbeit entwickelten und analysierten Modelle sind auch in anderen Bereichen anwendbar, z. B. zur Untersuchung des Systemverhaltens von Realzeit-Steuerungen in der Prozeßdatenverarbeitung oder zur Analyse eines verteilten Rechnersystems. Da das Ein-/Ausgabe-Verhalten bzw. die Schnittstellenverkehre bei der Analyse miteinbezogen wurden, ist auch eine Erweiterung der Analyse auf ein Warteschlangennetz mit taktgesteuerter Übergabe zwischen den einzelnen Knoten möglich, wobei eine prioritätsgesteuerte Abfertigung berücksichtigt werden kann.

ANHANG 1

Gegeben sei eine diskrete ZV X mit der Verteilung $p_i = P\{X=x\}$, $i = 0, 1, 2, \dots$. Dann heißt

$$G(z) = \sum_{i=0}^{\infty} p_i z^i$$

erzeugende Funktion der Wahrscheinlichkeiten p_i . Die wichtigsten Eigenschaften dieser Funktion sind:

1. $G(z)$ existiert in $|z| \leq 1$.
2. $G(1) = 1$.
3. $E[X] = G'(1)$, $\text{Var}[X] = G''(1) + G'(1) - [G'(1)]^2$.
4. $p_i = \frac{1}{i!} \cdot \frac{d^i G(z)}{dz^i} \Big|_{z=0}$, $i = 0, 1, 2, \dots$
5. Sind X_1 und X_2 unabhängige ZV mit den erzeugenden Funktionen $G_1(z)$ und $G_2(z)$, dann gilt für die erzeugende Funktion $G(z)$ der Summen-ZV $X = X_1 + X_2$ die Beziehung

$$G(z) = G_1(z) \cdot G_2(z)$$

$$6. F_k = \frac{d^k G(z)}{dz^k} \Big|_{z=1} \quad k\text{-tes faktorielles Moment.}$$

Beispiel: Poisson-Verteilung

$$p_i = P\{X=i\} = \frac{A^i}{i!} e^{-A}, \quad i = 0, 1, 2, \dots$$

$$G(z) = \sum_{i=0}^{\infty} \frac{A^i}{i!} e^{-A} z^i = e^{-A} \sum_{i=0}^{\infty} \frac{(Az)^i}{i!} = e^{-A(1-z)}$$

$$E[X] = G'(1) = A, \quad \text{Var}[X] = G''(1) + G'(1) - [G'(1)]^2 = A$$

ANHANG 2

Bei der Bestimmung von $q_{j0}(T)$, $j = 0, 1, \dots, N$, ist zu beachten, daß der Zustand $p(0, T)$ dadurch erreicht wird, daß mindestens alle wartenden Anforderungen nach dem Takt innerhalb eines Taktintervalls bedient werden. Die zugehörigen Wahrscheinlichkeiten werden mit $\bar{d}_{\geq i}(T_S)$ bzw. $\check{d}_{\geq i}(T_S)$ bezeichnet und sind aus den Sterbewahrscheinlichkeiten $\bar{d}_m(T_S)$ bzw. $\check{d}_m(T_S)$ wie folgt berechenbar:

$$\bar{d}_{\geq i}(T_S) = \sum_{m=i}^{\infty} \bar{d}_m(T_S), \quad \check{d}_{\geq i}(T_S) = \sum_{m=i}^{\infty} \check{d}_m(T_S)$$

Wegen der Begrenzung auf s Speicherplätze stellt der Zustand $\{N\}$ einen Sonderfall dar. Er wird immer dann erreicht, wenn die ankommende Gruppe mindestens so groß ist, daß alle freien Warteplätze aufgefüllt werden. Die zugehörigen Wahrscheinlichkeiten $g_{\geq i}$ sind aus der Gruppengrößen-Verteilung g_k berechenbar, d. h.

$$g_{\geq i} = \sum_{k=i}^{\infty} g_k$$

Nach Einführung dieser Größen erhält man für die Übergangswahrscheinlichkeiten $q_{jx}(T)$ die Beziehungen

$$q_{00}(T) = g_0 + \sum_{k=1}^{N-1} g_k \bar{d}_{\geq k}(T_S) + g_{\geq N} \bar{d}_{\geq N}(T_S)$$

$$q_{j0}(T) = \sum_{k=0}^{N-j-1} g_k \check{d}_{\geq (k+j)}(T_S) + g_{\geq (N-j)} \check{d}_{\geq N}(T_S)$$

$$j = 1, 2, \dots, N,$$

$$q_{0x}(T) = \sum_{k=x}^{N-1} g_k \bar{d}_{k-x}(T_S) + g_{\geq N} \bar{d}_{N-x}(T_S)$$

$$q_{jx}(T) = \sum_{k=\max(0, x-j)}^{N-j-1} g_k \check{d}_{k+j-x}(T_S) + g_{\geq (N-j)} \check{d}_{N-x}(T_S)$$

$$j, x = 1, 2, \dots, N,$$

$$\text{mit } \sum_{i=1}^j (\cdot) = 0 \quad \text{für } j < i \quad \text{und } T_S = T - T_{OH}$$

ANHANG 3

Zur Bestimmung der mittleren Warteschlangenlänge der Arbeits- und Ausgabewarteschlange werden im Abschnitt 4.2 und Abschnitt 4.4 für exponentiell verteilte Bedienungszeiten bzw. Erlang-k-verteilte Bedienungszeiten die Integrale über die zugehörigen Sterbewahrscheinlichkeiten $\bar{d}_m(t)$, $\bar{d}_{\geq m}(t)$ bzw. $\check{d}_m(t)$, $\check{d}_{\geq m}(t)$ über das Intervall $(0, T_S)$ benötigt.

I. Exponentiell verteilte Bedienungszeiten

Durch Integration der Poisson-Therme $\bar{d}_m(t)$ bzw. der Wahrscheinlichkeiten $\bar{d}_{\geq m}(t)$ ergibt sich

$$\bar{d}_m^I(T_S) = \int_0^{T_S} \frac{(\mu t)^m}{m!} e^{-\mu t} dt = \frac{1}{\mu} \sum_{\ell=m+1}^{\infty} \bar{d}_\ell(T_S) = \frac{1}{\mu} \bar{d}_{\geq(\ell+1)}(T_S), \quad (A.1)$$

$$\bar{d}_{\geq m}^I(T_S) = \int_0^{T_S} \bar{d}_{\geq m}(t) dt = \int_0^{T_S} \left[1 - \sum_{\ell=0}^{m-1} \bar{d}_\ell(t) \right] dt = T_S - \sum_{\ell=0}^{m-1} \bar{d}_\ell^I(T_S), \quad (A.2)$$

$m = 0, 1, \dots$

II. Erlang-k-verteilte Bedienungszeiten

Durch Integration von $\bar{d}_m(t)$ für die Erlang-k-VF und mit Hilfe der Gl. (A.1) (Integration über die Poisson-Therme) folgt:

$$\bar{d}_m^I(T_S) = \frac{1}{k\mu} \sum_{\ell=mk}^{(m+1)k-1} \sum_{i=\ell+1}^{\infty} f_i(T_S), \quad f_i(T_S) = \frac{(k\mu T_S)^i}{i!} = e^{-k\mu T_S}$$

Durch Vertauschen der Summation ergibt sich nach kurzer Umformung

$$\bar{d}_m^I(T_S) = \frac{1}{\mu} \bar{d}_{\geq(m+1)}(T_S) + (T_S - \frac{m}{\mu}) \bar{d}_m(T_S) - T_S f_{(m+1)k-1}(T_S) + \frac{m}{\mu} f_{mk}(T_S), \quad (A.3)$$

$m = 0, 1, \dots$

mit $f_j(T_S) = 0$ für $j < 0$. $1/\mu$ ist die mittlere Bedienungszeit.

Für den stationären Erneuerungsprozeß, d. h. Integration von $\check{d}_m(t)$ (vgl. Gl. (4.14)), erhält man folgende Rekursionsbeziehung:

$$\check{d}_m^I(T_S) = \sum_{\ell=0}^{m-1} [\bar{d}_\ell^I(T_S) - \check{d}_\ell^I(T_S)] + (m+1) \bar{d}_m^I(T_S) - \frac{2m-1-k}{2\mu} \bar{d}_{\geq(m+1)}(T_S) + \frac{(mk-1)m}{2k\mu} \bar{d}_m(T_S) - \frac{\mu T_S^2}{2} [\bar{d}_m(T_S) + f_{mk-1}(T_S) + f_{mk-2}(T_S) - f_{(m+1)k-1}(T_S) - f_{(m+1)k-2}(T_S)], \quad m=0, 1, \dots \quad (A.4)$$

mit $\sum_{i=1}^j (\cdot) = 0$ für $j < i$ und $f_j = 0$ für $j < 0$.

Analog zur Berechnung der Gl. (A.2) erhält man für $\bar{d}_{\geq m}^I(T_S)$ und $\check{d}_{\geq m}^I(T_S)$

$$\bar{d}_{\geq m}^I(T_S) = T_S - \sum_{\ell=0}^{m-1} \bar{d}_\ell^I(T_S), \quad m=0, 1, \dots, \quad (A.5)$$

$$\check{d}_{\geq m}^I(T_S) = T_S - \sum_{\ell=0}^{m-1} \check{d}_\ell^I(T_S), \quad m=0, 1, \dots \quad (A.6)$$

Alle abgeleiteten Beziehungen lassen eine rekursive Berechnung der einzelnen Größen zu.

ANHANG 4

Bei der Analyse des zweiklassigen Modells ist in Abschnitt 5.2.1.3 und im Abschnitt 5.3 das Integral über die Wahrscheinlichkeiten $d_{m,n}(t)$ zu bestimmen. Ausgangspunkt zur Berechnung ist die Gleichung

$$d_{m,n}(t) = \frac{\mu_1^m}{(m-1)!} \frac{\mu_2^n}{n!} e^{-\mu_2 t} \int_{\tau=0}^t \tau^{m-1} (t-\tau)^n e^{(\mu_2 - \mu_1)\tau} d\tau, \quad (A.7)$$

$m = 1, 2, \dots,$
 $n = 0, 1, \dots,$

des Abschnittes 5.2.1.1, die auf beiden Seiten über das Intervall $(0, T_S)$ integriert wird. Bezeichnet man die Integration der linken Seite mit $d_{m,n}^I(T_S)$, dann erhält man nach Vertauschen der Integration auf der rechten Seite:

$$d_{m,n}^I(T_S) = \frac{\mu_1^m}{(m-1)!} \frac{\mu_2^n}{n!} \int_{\tau=0}^{T_S} \tau^{m-1} e^{-\mu_1 \tau} \int_{u=\tau}^{T_S} (u-\tau)^n e^{-(u-\tau)\mu_2} du d\tau .$$

Unter Verwendung der Gl. (A.1) ergibt sich für die Integration über u

$$\begin{aligned} \int_{u=\tau}^{T_S} \dots du &= \frac{n!}{\mu_2^{n+1}} \sum_{k=n+1}^{\infty} \mu_2^k \frac{(T_S - \tau)^k}{k!} e^{-\mu_2(T_S - \tau)} \\ &= \frac{n!}{\mu_2^{n+1}} \left[1 - \sum_{k=0}^n \mu_2^k \frac{(T_S - \tau)^k}{k!} e^{-\mu_2(T_S - \tau)} \right] . \end{aligned}$$

Aufgrund der beiden Darstellungsmöglichkeiten ergeben sich zwei Berechnungsmöglichkeiten für $d_{m,n}^I(T_S)$.

$$d_{m,n}^I(T_S) = \begin{cases} \frac{1}{\mu_2} \sum_{k=n+1}^{\infty} \frac{\mu_1^m}{(m-1)!} \frac{\mu_2^k}{k!} e^{-\mu_2 T_S} \int_{\tau=0}^{T_S} \tau^{m-1} (T_S - \tau)^k e^{(\mu_2 - \mu_1)\tau} d\tau \\ \frac{1}{\mu_2} d_{\geq m}^1(T_S) - \frac{1}{\mu_2} \sum_{k=0}^n \frac{\mu_1^m}{(m-1)!} \frac{\mu_2^k}{k!} e^{-\mu_2 T_S} \int_{\tau=0}^{T_S} \tau^{m-1} (T_S - \tau)^k e^{(\mu_2 - \mu_1)\tau} d\tau \end{cases}$$

Der erste Term in der zweiten Darstellungsmöglichkeit folgt aus Gl. (A.1). Die beiden Ergebnisse lassen sich wie folgt durch die Gl. (A.7) ausdrücken, wenn dort $t = T_S$ gesetzt wird.

$$d_{m,n}^I(T_S) = \begin{cases} \frac{1}{\mu_2} \sum_{k=n+1}^{\infty} d_{m,n}(T_S) = \frac{1}{\mu_2} d_{m,\geq(n+1)}(T_S) \end{cases} \quad (A.8a)$$

$$d_{m,n}^I(T_S) = \begin{cases} \frac{1}{\mu_2} \left[d_{\geq m}^1(T_S) - \sum_{k=0}^n d_{m,k}(T_S) \right] , & m = 1, 2, \dots \\ n = 0, 1, \dots . \end{cases} \quad (A.8b)$$

Aus dem Vergleich der Ergebnisse der Gl. (A.8a) und Gl. (A.8b) erhält man den folgenden Zusammenhang:

$$d_{m,\geq n}(t) = d_{\geq m}(t) - \sum_{k=0}^{n-1} d_{m,k}(t) , \quad m = 1, 2, \dots \quad (A.9)$$

$n = 0, 1, \dots$

mit $\sum_I^j (\cdot) = 0$ für $j < i$.

Die Integration dieser Gleichung über das Intervall T_S ergibt unter Verwendung der Gl. (A.1) und Gl. (A.2)

$$d_{m,\geq n}^I(T_S) = T_S^{-\frac{1}{\mu_1}} \sum_{\ell=0}^{m-1} d_{\geq(\ell+1)}(T_S) - \sum_{k=0}^{n-1} d_{m,k}^I(T_S) , \quad m = 1, 2, \dots \quad (A.10)$$

$n = 0, 1, \dots$

mit $\sum_I^j (\cdot) = 0$ für $j < i$.

Alle abgeleiteten Beziehungen lassen eine rekursive Berechnung der einzelnen Größen zu.