Schwankungsbeschränkte Ströme

Claude-Joachim Hamann
TU Dresden, Fakultät Informatik
Institut für Betriebssysteme, Datenbanken und Rechnernetze

Zusammenfassung

Ereignisfolgen, die durch eine konstante Rate – die im Rahmen gewisser Grenzen schwanken kann – gekennzeichnet sind, gewinnen durch Multimedia- und durch Echtzeit-Anwendungen ständig an Bedeutung. Für derartige Folgen gibt es eine Reihe von Beschreibungsformen (z.B. Tenet-Protocol-Suite, Übertragungsprozeduren in ATM-Netzen, linear beschränkte Prozesse bei der Übertragung kontinuierlicher Multimedia-Daten). Besonderes Interesse finden dabei Bursts (kurzzeitige Datenübertragung mit höchstmöglicher Rate). Ziel der Arbeit ist es, ein allgemeines Modell vorzustellen, das die oben genannten Fälle umfaßt und damit eine Transformation der Parametersätze unterschiedlicher Prozeßbeschreibungen ermöglicht. Weiter werden eine Reihe von Eigenschaften derartiger Ereignisströme abgeleitet, so die vollständige Beschreibung des Zeitverhaltens und eine scharfe Grenze für die Mindestgröße eines Puffers, der zur verlustlosen Bearbeitung erforderlich ist.

1. Einleitung

In letzter Zeit gewinnen durch die Zunahme von multimedialen Anwendungen und den Einsatz von ATM-Technik mehr und mehr Prozesse an Interesse, die folgende Struktur haben: An einer bestimmten Schnittstelle eines verteilten Systems treten von einem Startzeitpunkt an Ereignisse nacheinander ein, und zwar eigentlich in einem konstanten Abstand, jedoch kann dieses Eintreten in festen Grenzen schwanken, nämlich eine gewisse Zeit zu spät oder zu früh erfolgen; dabei darf allerdings ein bestimmter Mindestabstand nicht unterschritten werden.

Ein anschauliches Beispiel ist das entfernte Abspielen eines Videos, wobei die auf einem Server gespeicherte Videoaufzeichnung mittels eines ATM-Netzes zu einem Client mit Videowiedergabe übertragen wird. Die Bildübertragung soll dabei mit einer konstanten Frequenz geschehen, durch eine Reihe von Einflüssen (beispielsweise Bildgröße und Netzbelastung) sind jedoch Schwankungen des Eintreffens der einzelnen Bilder gegenüber dem theoretisch konstanten Ankunftszeitabstand möglich.

Die betrachtete Schnittstelle kann dabei in den verschiedensten Schichten des verteilten Systems liegen, und entsprechend können die Ereignisse unterschiedlicher Natur sein. So kann es sich um das Aufrufen eines Dienstes zum Senden von Daten unterschiedlicher Größe (Videoframes) in der Anwendungsschicht handeln, ebenso aber auch um das Eintreffen von Zellen an einem ATM-Switch. In letzterem Fall ist klar, daß die Kapazität des Übertragungskanals, d.h. die Bitrate des zugrundeliegenden Netzes einen Mindestabstand zwischen dem Eintreffen der Zellen impliziert.

Eines der wesentlichen Beispiele für einen derartigen Prozeß ist die Folge von Ereignissen, die in einem ATM-Netz der sog. *Generic Cell Rate Algorithm* [ATM] erzeugt. Mit Hilfe dieses Algorithmus wird die oben beschriebene Eigenschaft erreicht, daß Zellen in einem gewissen Rahmen – der durch eine Spitzenübertragungsrate und eine maximale Verfrühung gegeben ist – vorzeitig gegenüber einer „mittleren“ Rate eintreffen können (Genaueres s. Abschn. 4.1.). Die Möglichkeit des verfrühten Sendens (bzw. Eintreffens) von Daten kann dabei zur sog. Burst-Bildung genutzt werden, d.h. zum Anhäufen von Daten, die dann „hintereinander“ mit der Spitzenrate gesendet werden.

Von diesem Gedanken der Burst-Bildung geht ANDERSON mit seiner Betrachtung von *Linear Bounded Arrival Processes* (LBAP) aus [Ande]. Ein derartiger Prozeß ist eine Folge von Ereignissen, die das Eintreffen von Bursts unterschiedlicher Größe bedeuten, wobei eine maximale Größe nicht überschritten werden darf und die Bursts nicht „zu dicht“ auftreten dürfen, so daß gleichfalls eine bestimmte „mittlere“ Rate eingehalten wird. Zu den wesentlichen Aufgaben gehört hier, die Größe des Puffers zu bestimmen, den die vorzeitig eintreffenden Nachrichten erfordern.

Schließlich sei an dieser Stelle noch das Tenet-Projekt (FERRARI et al., [BFMMVZ]) erwähnt, in dem Protokolle für eine Echtzeit-Kommunikation in heterogenen Netzwerken implementiert wurden und worin die auftretenden Nachrichtenströme durch folgende Parameter beschrieben werden: minimaler und mittlerer Zeitabstand zwischen zwei Nachrichten, maximale Nachrichtengröße und maximaler Jitter; auf deren Grundlage erfolgen u.a. Aussagen über maximale Ende-zu-Ende-Verzögerungen und Verlustwahrscheinlichkeiten.

Mit der vorliegenden Arbeit werden zwei Ziele verfolgt. Zum ersten wird ein Modell vorgestellt, das die drei oben beschriebenen Fälle umfaßt. Zum zweiten wird – aufbauend auf wichtigen Modellaussagen – ein Parametersatz vorgeschlagen, der eine einheitliche Beschreibung wesentlicher Komponenten eines Multimedia-Systems (wie Dateisystem oder Audio-/Videokomponenten) ermöglicht und der die Basis eines allgemeinen Schedulingverfahrens für das in Entwicklung befindliche Dresdner Echtzeit-Betriebssystem darstellt.

Daraus resultiert die Struktur der Arbeit: Im 2. Abschnitt erfolgt die Modellbeschreibung. Der 3. Abschnitt beinhaltet eine Reihe von Aussagen, die sich mit diesem Modell treffen lassen. Als wichtigste zählen dazu die Bestimmung der maximalen Burstgröße, die vollständige Chrakterisierung von Burst-Strömen und die Bestimmung der Mindestgröße eines Puffers zur verlustfreien Verarbeitung eintreffender Daten. Mit Hilfe dieser Ergebnisse erfolgt im 4. Abschnitt die Einordnung obiger Beispiele (ATM, LBAP, Tenet), und insbesondere wird die Transformation zwischen den unterschiedlichen Parametersätzen beschrieben. Schließlich wird in einem Ausblick auf die nun zu lösenden Aufgaben (vor allem Untersuchung von Systemen derartiger Prozesse und Einbeziehung stochastischer Betrachtungen) eingegangen. Eine auf die Ergebnisse konzentrierte Fassung dieser Arbeit stellt [Hama] dar.

2. Modellbeschreibung

Gegenstand dieser Arbeit sind Ereignisströme (Folgen von Ereignissen, interpretiert als das Senden oder Empfangen von Dateneinheiten gleicher Größe) der folgenden Art: Von einem Startzeitpunkt t_0 an treten die Ereignisse nacheinander theoretisch in konstantem Abstand T (d.h. mit konstanter Rate $R = 1/T$) ein, können aber in festen Grenzen schwanken, nämlich um die Zeit τ verfrüht oder um τ' verspätet sein, wobei ein Mindestabstand $D < T$ einzuhalten ist (s. Abb. 1); T ist damit der mittlere Abstand zweier aufeinanderfolgender Ereignisse. O.B.d.A. wird für das weitere $t_0 = 0$ gesetzt.

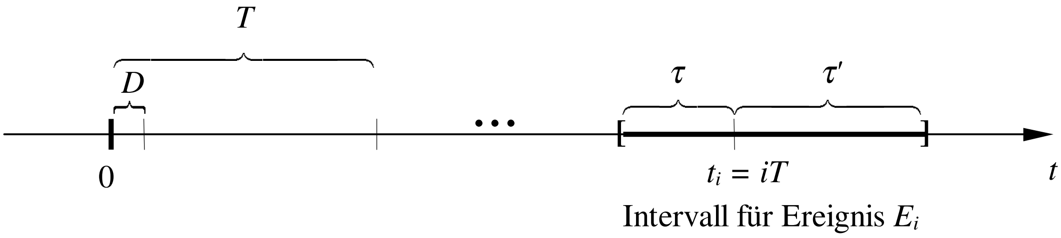

Abb. 1. Schwankungsbeschränkter Strom mit konstanter Rate

Definition 1. Gegeben seien

$$D, T, \tau, \tau' \in \mathbb{R}^{\#} \quad \text{mit} \quad T > D > 0, \quad \tau, \tau' \geq 0, \tag{1}$$

sowie $i \in \mathbb{N}$. Ein *(τ,τ')-schwankungsbeschränkter Strom mit konstanter Rate $R = T^{1}$ und Mindestabstand D* (kurz: *schwankungsbeschränkter Strom*) ist dann eine Folge $(E_i)_{i=0,1,...}$von Ereignissen, die in den Zeitpunkten

$a_i \in [t_i - \tau,\ t_i + \tau'] \subseteq \mathbb{R}$ tatsächliche (aktuelle) Ereigniszeit

mit

$t_i = iT$ theoretische Ereigniszeit

eintreten und deren Abstände

$$d_{i+1} = a_{i+1} - a_i$$

der Bedingung

$$d_{i+1} \geq D \qquad \forall i \in \mathbb{N}$$

genügen.

Im strengen Sinne handelt es sich um einen stochastischen eindimensionalen Punktprozeß; da stochastische Betrachtungen in dieser Arbeit unberücksichtigt bleiben, wird auf eine entsprechende Formalisierung verzichtet.

Von Interesse sind nun Ströme, bei denen die Ereignisse im kürzestmöglichen Abstand aufeinander folgen können. Dabei werden im weiteren die Ereignisse $\boldsymbol{E}_i$ mit dem Zeitpunkt a_i ihres Eintretens identifiziert. Eine „Ereignisfolge der Länge k+1“ ist damit eine Folge $a_ia_{i+1}...a_{i+k}$ $(i,\ k \in \mathbb{N})$, wobei die Elemente der Folge den in Def. 1 genannten Bedingungen genügen; unter $k = 0$ werde die einelementige Folge a_i verstanden.

Definition 2. Eine Ereignisfolge $a_ia_{i+1}...a_{i+k}$ heiße *dicht* $(i,k \in \mathbb{N})$, wenn

$$d_{i+j} = D \qquad \forall j = 1,...,k.$$

[#] $\mathbb{R}$ bezeichne die Menge der reellen Zahlen, $\mathbb{N}$ die der natürlichen Zahlen einschließlich 0, $\mathbb{N}_1$ entsprechend ohne 0.

Dies entspricht fast dem Begriff „Burst“, der ein Datenpaket und damit eine Folge von Dateneinheiten darstellt, das kurzzeitig mit der größtmöglichen Datenrate übertragen wird.

Definition 3. Ein *Burst der Länge l* (kurz: *l-Burst* oder *Burst*), $l \in \mathbb{N}_1$, ist eine dichte Ereignisfolge $B = a_i a_{i+1} \ldots a_{i+l-1}$, wobei keine der Folgen $a_i a_{i+1} \ldots a_{i+l-1} a_{i+l}$ gemäß Def. 1 dicht ist.

(Ein l-Burst soll wirklich nur die Länge l haben, danach soll eine „Pause“ kommen.)

Offenbar ist ein Burst B eindeutig durch seinen Startzeitpunkt a_i (bezeichnet mit $a(B)$) und seine Länge (bezeichnet mit $l(B)$) bestimmt. $\mathbf{B}_l$ bezeichne die Menge aller l-Bursts, und $\mathbf{B} = \bigcup_l \mathrm{B}_l$ die Menge aller Bursts. Diese Mengen können überabzählbar sein. Als Beispiel diene $T = 4$, $D = 1$, $\tau = \tau' = 1$; dann enthält $\mathbf{B}_1$ u.a. alle einelementigen Folgen (a_0) mit $a_0 \in [-1,1]$, während $\mathbf{B}_l = \varnothing$ für alle $l > 1$ gilt. Man beachte, daß $\mathbf{B}$ die Menge aller „potentiellen“, aller überhaupt möglichen Bursts bei gegebenen Parameterwerten nach (1) darstellt; ein konkreter Strom kann im betrachteten Beispiel innerhalb des Intevalls $[-1,1]$ selbstverständlich nur einen einzigen 1-Burst haben (vgl. Bsp. 2).

Trivialerweise läßt sich für beliebige Werte der Parameter $D,\ldots,\tau$ von Def. 1 stets eine Folge (a_i) konstruieren, die eine Folge von „entarteten“ 1-Bursts enthält (z.B. durch $a_i = t_i$), d.h., $\mathbf{B}_1 \neq \varnothing$ und damit $\mathbf{B} \neq \varnothing$. Andererseits sind Bursts in ihrer Länge begrenzt (dies ist offensichtlich und resultiert aus $T > D$ und der archimedischen Eigenschaft der reellen Zahlen).

Definition 4. Ein Burst heißt *maximal*, wenn er maximale Länge L besitzt mit $L = \max_{B \in \mathbf{B}} l(B)$.

Schließlich sei

$$b^l := \min_{B \in \mathbf{B}_l} (a(B))$$

frühester Startzeitpunkt eines l-Burst; Zeitpunkt, zu dem erstmals – ggf. auch vor $t = 0$ – ein Burst der Länge l auftreten kann (trivialerweise ist $\min_{B \in \mathrm{B}} (a(B)) = -\tau$);

$b_f := b^L$ frühester Startzeitpunkt eines maximalen Bursts;

$$b_s := \max_{t \geq b^L \,\wedge\, \forall t' \in \left[b^L, t\right] \exists B \in B_L :\, t' = a(B)} (t \in R)$$

spätester Startzeitpunkt eines maximalen Bursts (ein spätester Startzeitpunkt von l-Bursts mit $l<L$ existiert nur bei diskreter Zeitachse).

Von Interesse sind nun zwei Arten von Strömen:

- **Burst-Ströme**, d.h. Ströme, die nur aus Bursts maximaler Länge bestehen (Daten sollen in möglichst großen Paketen übertragen werden, die Pausen zwischen den Bursts können im Rahmen von Def. 1 schwanken);
- **dichte Ströme**, d.h. Ströme, in denen die schnellstmögliche Übertragung in unterschiedlich großen Mengen, aber insgesamt so dicht als möglich geschieht.

Für diese beiden Arten von Strömen sollen – soweit jeweils relevant – im folgenden die wichtigsten Bewertungsgrößen berechnet werden:

– maximale Burstgröße L

– b^l, b_f, b_s

– Abstände zwischen Bursts, insbesondere Mindest- und Höchstabstand I_u, I_o zwischen maximalen Bursts

– minimale Puffergröße P (maximale Anzahl zu puffernder Dateneinheiten)

– Anzahl $N(t)$ der bis zur Zeit t eingetretenen Ereignisse

– Bestimmung der Schwankungsparameter τ, τ' aus L.

Dabei spielt der ganzzahlige Teil einer reellen Zahl eine große Rolle. Für $x \in \mathbb{R}$ bezeichne $\lfloor x \rfloor$ die größte ganze Zahl, die x nicht überschreitet, also

$$\lfloor x \rfloor \le x < \lfloor x \rfloor + 1, \quad \lfloor x \rfloor \text{ ganzzahlig.} \tag{2}$$

Entsprechend sei $\lceil x \rceil$ die kleinste ganze Zahl, die x nicht unterschreitet.

Es gilt u.a. für $n \in \mathbb{N}$, $x \in \mathbb{R}$:

$$\lfloor n + x \rfloor = n + \lfloor x \rfloor, \ \lfloor -x \rfloor = -\lceil x \rceil, \quad \text{somit} \ \lfloor n - x \rfloor = n - \lceil x \rceil, \ -\lfloor x \rfloor = \lceil -x \rceil. \tag{3}$$

Zwei Beispiele sollen nun die eingeführten Begriffe und gleichzeitig die Beweismethoden erläutern. Bei allen Beispielen ist eine einheitliche Zeiteinheit unterstellt, die deshalb weggelassen wird.

Beispiel 1. Burst-Strom mit $D = 1$, $T = 4$, $\tau = \tau' = 7$.

Dann ist (s. Abb. 2) $L = 5$, $b_f = 5$, $b_s = 7$, $I_u = 14$, $I_o = 18$

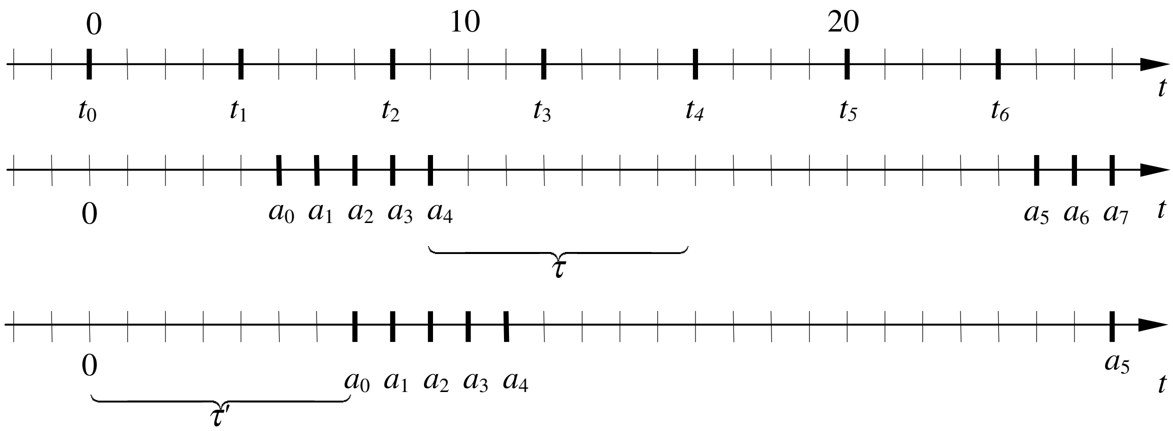

Abb. 2. Beispiel für zwei Ströme mit maximalen Bursts

Beispiel 2. Dichter Strom mit mindestens einem 3-Burst, $D = 1$, $T = 4$, $\tau = \tau' = 7$.

Dann ist (s. Abb. 3) 3-Burst jede Folge $(a, a+1, a+2)$ mit $a \in [-1,2)$; $a = 2$ entfällt, denn dann entsteht ein 4-Burst. Ab $t = 4$ ist die restliche Folge eindeutig bestimmt, weitere Bursts sind unmöglich! Ferner ist $b^3 = -1$ und

$$N(t) = \begin{cases} 0 & \text{für} \quad t < a \\ 1 + \lfloor t - a \rfloor & \text{"} \quad t \in [a, a+2] \\ 3 & \text{"} \quad t \in [a+2,\ 5) \\ 1 + \left\lfloor \dfrac{t+7}{4} \right\rfloor & \text{"} \quad t \geq 5 \end{cases}$$

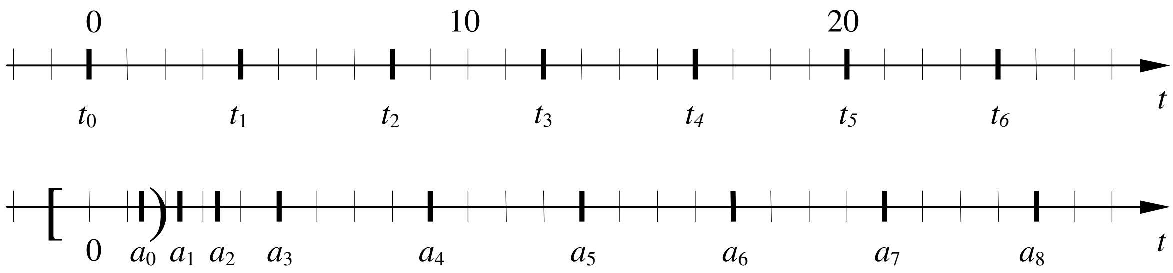

Abb. 3. Dichter Strom mit einem 3-Burst

Das Beispiel zeigt auch, warum es i.a. keinen spätesten Startzeitpunkt für l-Bursts mit $l < L$ gibt.

3. Ergebnisse

3.1. Burst-Ströme

Definition 5. Ein schwankungsbeschränkter Strom heißt *(maximaler) Burst-Strom*, wenn er eine Folge $(B_i)_{i=0,1,...}$ von Bursts B_i darstellt mit

$$l(B_i) = L \qquad \forall i \in \mathbb{N}, \quad L \text{ gemäß Def. 4.}$$

Ein Burst maximaler Länge L wird offenbar gewiß dann erzeugt, wenn der Startzeitpunkt $a(B)$ des Bursts möglichst spät liegt, d.h. bei $a(B) = b_s$ mit

$$b_s = \tau'. \tag{4}$$

Dann gilt (s. Abb. 4) mit

$$K = L - 1 \tag{5}$$

$$K \quad = \max_{b_s + kD \geq kT - \tau} (k \in \mathbb{N}) = $$
$$= \max_{k(T-D) \leq b_s - \tau} (k \in \mathbb{N}) =$$
$$= \max_{k \leq \frac{\tau+\tau'}{T-D}} (k \in \mathbb{N})$$

also

$$K = \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor, \tag{6}$$

und wegen (5) folgt

$$L = 1 + \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor. \tag{7}$$

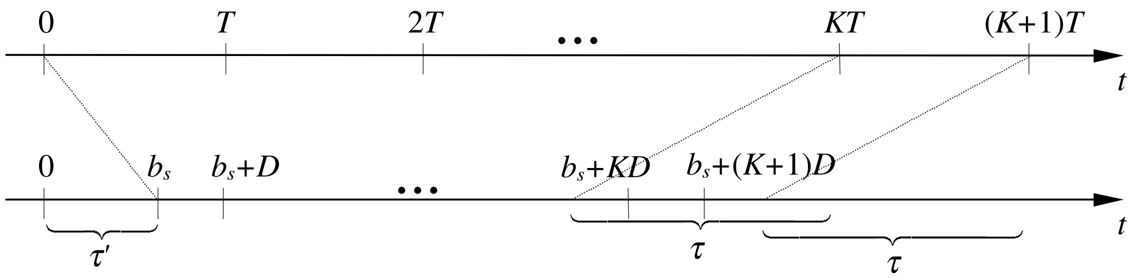

Abb. 4. Berechnung der maximalen Burstlänge

Der früheste Startzeitpunkt b_f eines maximalen Burst B der Länge L muß so liegen, daß das letzte Ereignis von B genau zum Zeitpunkt $KT - \tau$ eintritt. Damit führt eine ähnliche Betrachtung wie in Abb. 4 zu

$$KT - \tau = b_f + KD,$$

also

$$b_f = K(T - D) - \tau \; = \; (T - D)\left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor - \tau. \tag{8}$$

Bemerkenswert ist, daß b_f – im Gegensatz zu b_s – negativ werden kann; so führt $D = 1$, $T = 4$, $\tau = 7$, $\tau' = 1$ zu $b_f = -1$. Allerdings gilt erwartungsgemäß

$$0 \leq \; b_s - b_f \; < \; T - D. \tag{9}$$

Denn es ist

$$b_s - b_f \;\; = \;\; \tau' - K(T - D) + \tau \; =$$
$$= \;\; \tau + \tau' - (T - D) \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor \; =$$
$$= \;\; a - b\left\lfloor \frac{a}{b} \right\rfloor \qquad \text{(nach geeigneter Substition, wobei } a \geq 0,\ b > 0\text{)}$$

und dies entspricht der Formel für den Rest bei Division ganzer Zahlen a, b (man beachte, daß die Parameter in (1) zwar nichtnegativ, jedoch nicht notwendig ganzzahlig sind). Anschaulich bedeutet der rechte Teil von (9), daß eine weitere Verfrühung des Starts von B „am Ende“ nicht mehr zu einem Burst maximaler Länge führt. Für das weitere beachte man auch, daß zwar nicht notwendig $b_s + KD \leq KT$, wohl aber nach Definition bzw. (6)

$$b_s + KD < (K + 1)T$$

gilt.

Ferner zeigt Abb. 4 die eingangs erwähnte Tatsache, daß die Burstlänge beschränkt ist. Andernfalls gälte

$$b_s + kD \geq kT - \tau \qquad \forall\, k \in \mathbb{N} \qquad (*)$$

mithin

$$k(T - D) \leq \tau + \tau' \qquad \forall\, k \in \mathbb{N},$$

und das steht im Widerspruch dazu, daß es für zwei beliebige reelle Zahlen $a > 0$, $b \geq 0$ stets eine natürliche Zahl k_0 gibt mit $k_0 a > b$. Für die Beschränktheit der Burstlänge ist dabei zu berücksichtigen, daß (*) mindestens für $k = 0$ gilt.

Mit (4) und (8) läßt sich in einfacher Weise der Mindestabstand I_u zwischen zwei aufeinanderfolgenden Bursts bestimmen (s. Abb. 5):

$$\begin{aligned} I_u &= (K + 1)T + b_f - (b_s + KD) = \\ &= K(T - D) + T - (b_s - b_f) = \\ &= K(T - D) + T - \tau' + K(T - D) - \tau\,, \end{aligned}$$

also

$$I_u = 2K(T - D) + T - (\tau + \tau')\,. \tag{10}$$

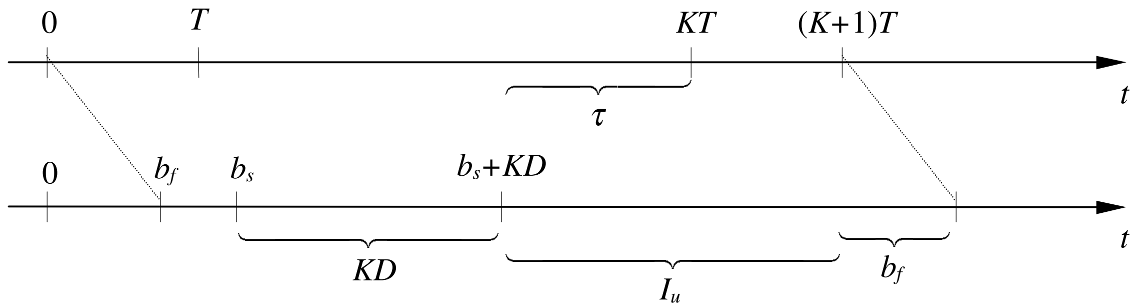

Abb. 5. Minimale Burst-Zwischenzeit I_u

In noch einfacherer Weise folgt für den größtmöglichen Abstand I_o zweier aufeinanderfolgender Bursts (Differenz zwischen dem Ende des frühestmöglichen Bursts und dem spätesten Beginn des nachfolgenden Bursts) sowie für den Abstand $I_f = I_s$ zwischen zwei frühestmöglichen bzw. zwei spätestmöglichen Bursts:

$$I_o = T + \tau + \tau'\,. \tag{11}$$

$$I_f = I_s = K(T - D) + T = L(T - D) + D \tag{12}$$

Abb. 5 zeigt auch, daß ein Burst-Strom in einem erweiterten Sinne periodisch mit der Periode LT ist, denn der Punkt $(K+1)T = LT$ spielt genau die gleiche Rolle wie der Punkt 0 auf der t-Achse. Genauer: Ist $(B_i)_{i=0,1,...}$ ein Strom von maximalen Bursts, so ist

$$a(B_i) \in [b_f + iLT,\ b_s + iLT] \ . \tag{13}$$

Damit gelten alle bisher getroffenen Aussagen (z.B. über I_u, I_f und I_o) für den gesamten Strom.

Untersucht werde schließlich der Fall, daß es sich bei dem betrachteten Strom um den Ankunftsprozeß an einer Bedieneinrichtung handelt, die in der Lage ist, ankommende Daten schwankungsfrei mit derselben Rate wie der Ankunftsprozeß (d.h. mit der konstanten Zeit T) zu bedienen. Die Daten sollen dabei in Datenelementen gleicher Größe ankommen, so daß das Eintreffen eines Elements genau ein Ereignis im obigen Sinne darstellt. Entscheidend ist nun die Forderung nach Verlustfreiheit. Dazu sind die vorzeitig (gegenüber der theoretischen Ankunftszeit bzw. der deterministisch ablaufenden Bedienung) ankommenden Datenelemente zu puffern. Daher soll nun die Frage beantwortet werden, wie groß die maximale Füllung eines unbegrenzt angenommenen Puffers werden kann oder besser: Es ist P zu bestimmen mit

P: minimale Puffergröße, so daß keine Verluste auftreten.

Weiter werde angenommen, daß die Bedienung verzögerungsfrei in folgendem Sinne ablaufe: Stimmen Ankunftszeitpunkt eines Datenelements und Bedienungsbeginn (Entnahmezeitpunkt aus dem Puffer) überein, so benötige das ankommende Datenelement keinen Pufferplatz. Betrachtet werden zwei Entnahmestrategien:

(A) Das Leeren des Puffers wird so organisiert, daß die Bedieneinrichtung stets arbeiten kann (kontinuierliche Arbeit, kein „Verhungern“).

(B) Trifft ein Datenelement auf eine leere Bedienungseinrichtung, so beginnt die Bedienung sofort.

Seien B, B' der 1. und 2. maximale Burst eines Burst-Stromes. Bei Strategie (A) ist sofort ersichtlich, daß die Entnahme frühestens bei b_s beginnen kann. In diesem Fall sind die L Elemente von B genau zur Zeit $b_s + LT$ bedient, und zu dieser Zeit trifft aufgrund von (13) auch spätestens der Burst B' ein, der dann eine leere Bedienungseinheit vorfindet, so daß dessen Bedienung sofort beginnt. Trifft B' eher ein, so muß dies wegen (9) und $T - D < T$ später als $b_s + (L-1)T$ geschehen, und zu diesem Zeitpunkt ist der Puffer bereits leer. Zur Berechnung von P genügt es also, die Situation im Zeitraum $[0, LT]$ zu betrachten. Damit ist offenbar L eine obere Schranke für P. Für eine genauere Berechnung ist zu beachten, daß während des Eintretens der L Ereignisse von B bereits die Bedienung und damit die Entnahme aus dem Puffer erfolgen kann (s. Abb. 6). Für $b_s > b_f + KD$ ist $P = L$ evident. Sei

$$b_s \le b_f + KD \ . \qquad (*)$$

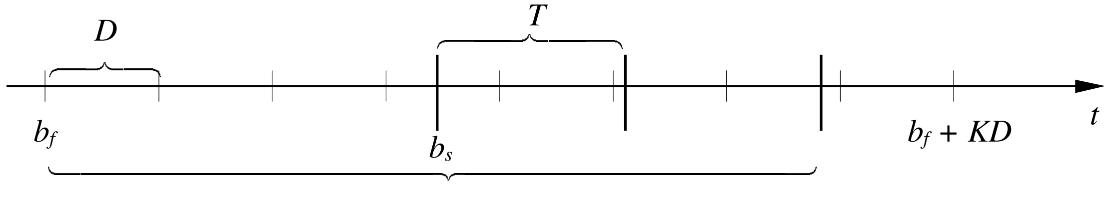

Abb. 6. Füllung des Puffers ab $t = b_f$, Entnahme ab $t = b_s$

Mit

N: maximale Anzahl der von $t = b_s$ bis einschließlich $t = b_f + KD$ möglichen Ereignisse mit Abstand T

ist dann

$$P = L - N$$

(ist $a(B) > b_f$, so kann N höchstens zunehmen, P also nur abnehmen), wobei

$$N = \left\lfloor \frac{b_f + KD - b_s}{T} \right\rfloor + 1\,,$$

so daß mit (8), (4), (5) und schließlich (3) folgt:

$$P = L - 1 - \left\lfloor \frac{KT - \tau - \tau'}{T} \right\rfloor = K - \left\lfloor K - \frac{\tau + \tau'}{T} \right\rfloor = \left\lceil \frac{\tau + \tau'}{T} \right\rceil\,.$$

Bedingung (*) ist wegen (8) gleichbedeutend mit $\tau \le KT - \tau'$, also

$$\frac{\tau + \tau'}{T} \le \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor.$$

Diese Bedingung ist zwar „häufig“ erfüllt (u.a. bei dem praktisch wichtigen Fall einer großen Burstlänge, so daß $KD \ge T$), aber nicht immer, wie z.B. $D = 1$, $T = 10$, $\tau = 5$, $\tau' = 6$ zeigt. Dieser Fall werde nun genauer betrachtet. Sei also

$$\frac{\tau + \tau'}{T} > \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor, \qquad P = L = \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor + 1.$$

Dann folgt

$$P - 1 < \frac{\tau + \tau'}{T}$$

sofort. Weiter ist wegen (1) und (2)

$$\frac{\tau + \tau'}{T} \le \frac{\tau + \tau'}{T - D} \le \left\lfloor \frac{\tau + \tau'}{T - D} \right\rfloor + 1 = P\,.$$

Nun gilt aber analog zu (2) für $x \in \mathbb{R}$, $x \ge 0$: $\lceil x \rceil$ ist diejenige eindeutig bestimmte Zahl $a \in \mathbb{N}$, für die gilt: $a - 1 < x \le a$. Damit folgt

$$P = \left\lceil \frac{\tau + \tau'}{T} \right\rceil.$$

Da bei Strategie (B) die Entnahme zu einem früheren Zeitpunkt als bei (A) beginnen kann, ist eine Vergrößerung von P gegenüber (A) ausgeschlossen. Da weiter die ungünstigste Situation bzgl. Puffergröße für $a(B) = b_s$ vorliegt, tritt zur Zeit $t = b_s + LT$ genau der Fall ein, der in Strategie (A) betrachtet wurde; das gilt analog für jeden anderen Zeitpunkt des Bedienungsbeginns im Intervall $[b_f, b_s]$ und mit der Periode LT. Damit folgt:

Satz 1. Für schwankungsbeschränkte Ströme, die ausschließlich aus Bursts maximaler Länge bestehen, ist die zur verlustlosen Bedienung erforderliche Mindest-Puffergröße unabhängig vom Zeitpunkt, zu dem die Bedienung (Entnahme aus dem Puffer) beginnt, sofern dies spätestens zur jeweils spätestmöglichen Ankunftszeit eines Bursts geschieht. Es ist ein Puffer der Mindestgröße

$$P = \left\lceil \frac{\tau + \tau'}{T} \right\rceil \tag{14}$$

erforderlich.

Für praktische Belange ist nicht die Ermittlung der maximalen Burstlänge aus den Parametern in (1) von Interesse, sondern umgekehrt die Bestimmung der Schwankungsparameter τ, τ' bei gegebener Burstgröße L. Aus (7) folgt mit (2)

$$(T - D)(L - 1) \leq \tau + \tau' < L$$

oder

$$\tau + \tau' \in \big[(L-1)(T-D),\, L(T-D)\big). \tag{15}$$

Ohne Herleitung sei noch angegeben, daß die Anzahl $N(t)$ der Ereignisse, die in einem Burst-Strom bis einschließlich zum Zeitpunkt t eintreten, der Abschätzung

$$N_u(t) \leq N(t) \leq N_o(t) \tag{16}$$

genügt mit

$$N_u(t) = \begin{cases} 0 & \text{für} \quad t < t_f \\ iL + 1 + \left\lfloor \frac{t - t_f^i}{D} \right\rfloor & \text{"} \quad t \in \left[t_f^i,\, t_f^i + KD\,\right] \\ (i+1)L & \text{"} \quad t \in \left[t_f^i + KD,\, t_f^i + LT\right) \end{cases}$$

$$N_o(t) = \begin{cases} 0 & \text{für} \quad t < t_s \\ iL + 1 + \left\lfloor \frac{t - t_s^i}{D} \right\rfloor & \text{"} \quad t \in \left[t_s^i,\, t_s^i + KD\,\right] \\ (i+1)L & \text{"} \quad t \in \left[t_s^i + KD,\, t_s^i + LT\right) \end{cases}$$

wobei

$$t_f^i = t_f + iLT, \quad t_s^i = t_s + iLT, \quad i = \left\lfloor \frac{t}{LT} \right\rfloor .$$

Dabei gilt $N(t) = N_u(t)$ für alle $t \in \mathbb{R}$, wenn alle Bursts des Stroms zum spätestmöglichen Zeitpunkt eintreffen; entsprechend gilt $N(t) = N_o(t)$ bei frühestem Eintreffen.

3.2. Dichte Ströme

Betrachtet werden nunmehr schwankungsbeschränkte Ströme, in denen Bursts beliebiger (selbstverständlich durch L begrenzter) Länge auftreten können, allerdings in dichtester Reihenfolge der Ereignisse gemäß Def. 1. Bsp. 2 zeigt bereits, daß die scheinbare Großzügigkeit dennoch sehr restriktiv ist: ein derartiger Strom ist spätestens ab $LT - \tau$ eindeutig bestimmt, und ab diesem Zeitpunkt treten nur „entartete“ 1-Bursts auf. Die folgenden Ergebnisse sind daher als Grundlage einer anschließenden Verallgemeinerung anzusehen. Die Darlegungen dieses Abschnitts sind im wesentlichen auf die Ergebnisse beschränkt; die Beweismethodik ist die gleiche wie im vorangegangenen Abschnitt.

Von den in Abschn. 2 genannten Bewertungsgrößen sind lediglich

b^l frühester Startzeitpunkt eines l-Burst

$N(t)$ Anzahl der bis zum Zeitpunkt t eingetretenen Ereignisse

P minimale Puffergröße, um Verlustfreiheit zu garantieren

von Interesse.

Für den frühesten Startzeitpunkt ergibt sich mit $k = l - 1$ (s. Abb. 7):

$$b^l + kD = kT - \tau,$$

also

$$b^l = k(T - D) - \tau, \qquad k \in \mathbb{N}.$$

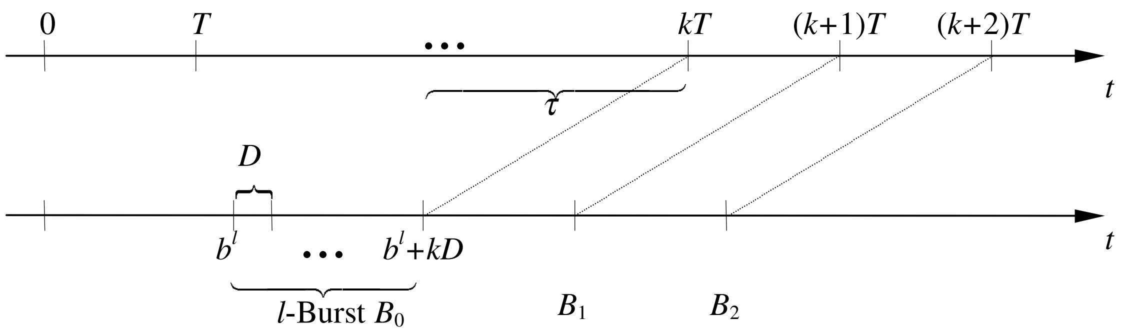

Abb. 7. Frühester Beginn eines l-Bursts

Die außerdem geltende Bedingung $b^l \leq \tau'$ führt dann gerade zu $k \leq K$ bzw. zu $l \leq L$ (K bzw. L gemäß (6) bzw. (7)), so daß sich zusammengefaßt ergibt:

$$b^l = (l - 1)(T - D) - \tau, \qquad l = 1,...,L. \tag{17}$$

Eine unmittelbare Folgerung daraus ist

$$b^L - b^1 = (L - 1)(\mathrm{T} - \mathrm{D}), \qquad b^l - b^{l-1} = T - D \quad \text{für } l = 1,...,L. \tag{18}$$

Abb. 7 zeigt außerdem sofort, daß ein dichter Strom folgende Struktur hat: Er besteht aus einem einzigen l-Burst B_0, gefolgt von beliebig vielen 1-Bursts B_i, wobei mit $l \in \{1,...,L\}$ gilt:

$$a(B_0) \in [(l - 1)(T - D) - \tau,\ l(T - D) - \tau) \tag{19}$$

$$a(B_i) = (l - 1 + i)T - \tau, \qquad i \in \mathbb{N}_1. \tag{19'}$$

Sei

$N_b^l(t)$: Anzahl der Ereignisse bis einschließlich Zeitpunkt t, wenn $B_0 \in \mathbf{B}_l$ und $b = a(B_0)$ gemäß (19).

Dann ist

$$N_b^l(t) = \begin{cases} 0 & \text{für} \quad t < b \\ 1 + \left\lfloor \dfrac{t-b}{D} \right\rfloor & \text{"} \quad t \in [b,\ b+(l-1)D) \\ l & \text{"} \quad t \in [b+(l-1)D,\ lT-\tau) \\ 1 + \left\lfloor \dfrac{t+\tau}{T} \right\rfloor & \text{"} \quad t \geq lT - \tau \end{cases} \tag{20}$$

wobei $[x, x) := \varnothing$ für $x \in \mathbb{R}$ sei. Bemerkt werde, daß $1 + \left\lfloor \frac{t-b}{D} \right\rfloor = l$ für $t = b+(l-1)D$ gilt.

Bei der Bestimmung der minimalen Puffergröße P eines dichten Stroms – die unter den gleichen Annahmen wie in Abschn. 3.1 erfolgen soll – ist zu berücksichtigen, daß bereits mit dem frühestmöglichen Eintreffen eines Bursts eine ununterbrochene Bedienung möglich ist, denn es gilt $b^l + lT > lT - \tau$. Damit sind die Strategien (A) und (B) identisch, und man errechnet leicht:

$$P = l - \left(1 + \left\lfloor \frac{(l-1)D}{T} \right\rfloor\right) = (l-1) - \left\lfloor \frac{(l-1)D}{T} \right\rfloor ,$$

also mit (3)

$$P = \left\lceil \frac{(l-1)(T-D)}{T} \right\rceil . \tag{21}$$

Als Verallgemeinerung sollen abschließend Ströme betrachtet werden, die aus Bursts unterschiedlicher Länge bestehen und deren Startzeitpunkt $a(B_i)$ jeweils so früh als möglich (im Rahmen von Def. 1) liegt.

Definition 6. Ein schwankungsbeschränkter Strom heiße *dichter Burst-Strom*, wenn er eine Folge $(B_i)_{i=0,1,\ldots}$ von Bursts darstellt mit

$$l_i := l(B_i) \leq L \qquad L \text{ gemäß (7)},$$

$$a_i := a(B_i) = t_i\, T + b^{l_i} \quad \text{mit} \quad t_i = \sum_{j=1}^{i} l_j\,, \quad b^{l_i} \text{ gemäß (17)} \tag{22}$$

für alle $i \in \mathbb{N}$ (l_i sind die freien Parameter der Folge).

Beispiel 3. $D = 1$, $T = 4$, $\tau = \tau' = 7$. Dann ist $L = 5$ (s. Bsp. 2).
Folge von Bursts der Länge 2,4,1,1,5,2:

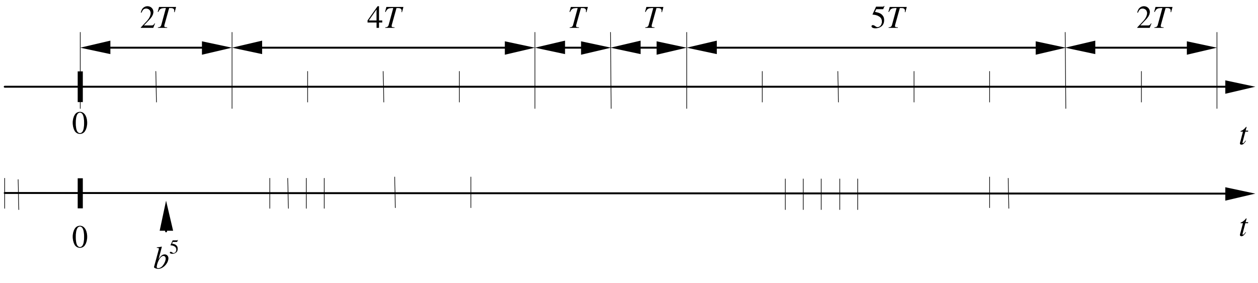

Abb. 8. Dichter Strom von Bursts unterschiedlicher Länge

Die Anzahl $M(t)$ der Ereignisse, die in einem dichten Burst-Strom bis einschließlich zum Zeitpunkt t eintreten können, läßt sich einfach (und damit recht grob) aufgrund der Tatsache abschätzen, daß ein dichter Strom von 1-Bursts die frühestmöglichen Ereignisse repräsentiert und ein dichter Strom von L-Bursts die spätestmöglichen (unter der in diesem Abschnitt grundsätzlich getroffenen Voraussetzung, daß alle Bursts möglichst früh beginnen): es gilt für alle $t \in \mathbb{R}$

$$M_u(t) \leq M(t) \leq M_o(t) \tag{23}$$

mit

$$M_u(t) = \begin{cases} 0 & t < b^L = (L-1)(T-D) \\ \left\lfloor \dfrac{T+\tau+t}{LT} \right\rfloor \cdot L & \text{sonst} \end{cases}$$

$$M_o(t) = \begin{cases} 0 & t < -\tau \\ 1 + \left\lfloor \dfrac{t+\tau}{T} \right\rfloor & \text{sonst} \end{cases}$$

Ähnliche Überlegungen wie zu Beginn dieses Abschnitts führen zu einer Aussage über den Mindestabstand, in dem Bursts in einem schwankungsbeschränkten Strom auftreten können. Für einen dichten Burst-Strom $(B_i)_{i=0,1,\ldots}$ sei $I(B_i, B_{i+1})$ ("Interburstness", Mindestabstand zwischen zwei aufeinanderfolgenden Bursts)

$$I(B_i, B_{i+1}) = a(B_{i+1}) - (a(B_i) + (l_i - 1)D).$$

Dann gilt (o. B.; zur Veranschaulichung s. Abb. 8):

$$I(B_i, B_{i+1}) = T + (l_{i+1}-1)(T-D) = l_{i+1}T - (l_{i+1}-1)D. \tag{24}$$

Interessanterweise ist damit die Zeit zwischen dem Ende eines Bursts bis zum nächstmöglichen Beginn von der Länge des nachfolgenden, nicht aber von der Länge des bisher letzten Bursts abhängig. Damit läßt sich nun auch das zeitliche Verhalten eines dichten Burst-Stroms vollständig charakterisieren.

Satz 2. Sei $(B_i)_{i=0,1,\ldots}$ ein dichter Burst-Strom, $a_i = a(B_i)$ Anfangszeit, $l_i = l(B_i)$ Länge von Burst B_i. Dann gilt:

$$a_0 = -\tau + (l_i - 1)(T-D)$$

$$a_{i+1} = a_i + l_{i+1}(T-D) + l_iD = a_i + l_{i+1}T - (l_{i+1} - l_i)D \qquad \forall i \in \mathbb{N}. \tag{25}$$

Beweis. a_0 resultiert aus (17), und für a_{i+1} gilt offenbar

$a_{i+1} = a_i + D_i + I(B_i,B_{i+1})$ mit $D_i = (l_i-1)D$ Dauer von B_i. Der Rest ergibt sich aus (24).

Für die Bestimmung der minimalen Puffergröße P wird zunächst wieder Strategie (A) betrachtet. Dann führt die Wahl von $t = b^L$ als Beginn der Bedienung (1. Entnahme aus dem Puffer, s. Abb. 8) zu einer kontinuierlichen Arbeit der Bedieneinrichtung. Denn in diesem Fall ist die Bedienung von B_0 genau zum Zeitpunkt $b^L{}_f + l_0T = t_0T + b^L{}_f$ beendet, und spätestens zu dieser Zeit trifft B_1 ein. Auch bei früherem Eintreffen (wenn $l_1 < L$) wird B_1 erst jetzt bedient; dies ist nach der Zeit l_1T geschehen, also zum Zeitpunkt $b^L{}_f + l_0T + l_1T = t_1T + b^L{}_f$ usw.

Zur Berechnung von P werde als erstes der Fall $b^l + kD \geq b^L$ betrachtet (der mindestens für maximale Bursts vorliegt). In ähnlicher Weise wie in Abschn. 3.1 folgt dann:

$$P = \left\lceil \frac{(L-1)(T-D)}{T} \right\rceil. \tag{26}$$

Andernfalls sind mindestens alle Ereignisse zu puffern, die *vor* dem Zeitpunkt b^L eintreffen. Dies sind nach (23) N Ereignisse mit

$$N \leq \frac{(L-1)(T-D)}{T} \quad \text{für} \quad \frac{(L-1)(T-D)}{T} \in \mathbb{N}$$
$$N \leq 1 + \left\lfloor \frac{(L-1)(T-D)}{T} \right\rfloor \quad \text{sonst.}$$

In beiden Fällen gilt aber $N \leq P$ mit P aus (26). Ab b^L kann sich N nicht mehr vergrößern. Denn nach Voraussetzung ist der l-Burst vorher vollständig eingetroffen, und von da an ist T der geringstmögliche Abstand des Eintretens von Ereignissen, und in diesem Abstand werden die Ereignisse nun dem Puffer auch entnommen.

Die Überlegungen zeigen außerdem, daß genau die maximalen Bursts auf eine leere Bedienungsanlage treffen. Dies hat für Strategie (B) sofort die Konsequenz, daß auch hier der in (26) angegebene Wert die minimale Puffergröße darstellt. Zusammengefaßt ergibt sich analog zu Satz 1:

Satz 3. Für dichte Burst-Ströme, d.h. (τ,τ')-schwankungsbeschränkte Ströme mit konstanter Rate T^{-1} und Mindestabstand D, die aus Bursts (unterschiedlicher Größe) bestehen, die jeweils zum frühestmöglichen Zeitpunkt eintreffen, ist die zur verlustlosen Bedienung eforderliche Mindest-Puffergröße unabhängig vom Zeitpunkt, zu dem die Bedienung beginnt, sofern dies spätestens zum Zeitpunkt

$$b^L = (L-1)(T-D) - \tau, \qquad L = 1 + \left\lfloor \frac{\tau + \tau'}{T-D} \right\rfloor$$

geschieht. Es ist ein Puffer der Mindestgröße

$$P = \left\lceil \frac{(L-1)(T-D)}{T} \right\rceil$$

erforderlich.

Mit Blick auf den angestrebten Vergleich zwischen verschiedenen Modellen aus der Literatur ist insbesondere die in (23) angegebene obere Schranke für $M(t)$ zu grob. Es sei im Hinblick auf Abschn. 4 $\tau' = 0$. Dann folgt zunächst für den Anfangszeitpunkt $b = a(B_0)$ des 1. Bursts eines dichten Burst-Stromes $b \leq 0$. Weiter werde nun eine Zeittransformation dergestalt vorgenommen, daß die Zeitzählung mit dem Auftreten des 1. Ereignisses beginnt. Damit läßt sich die für Abschn. 4 grundlegende Aussage beweisen:

Lemma 1. Für die Anzahl $M(t)$ der Ereignisse, die in einem dichten Burst-Strom bis einschließlich zum Zeitpunkt t eintreten (Zeit gezählt ab dem 1. Eintreten eines Ereignisses des Stroms), gilt für alle $t \in \mathbb{R}$: $M(t) \leq \overline{M}(t)$ mit

$$\overline{M}(t) = \begin{cases} 0 & t < 0 \\ 1 + \left\lfloor \dfrac{t}{D} \right\rfloor & t < LT \\ 1 + \left\lfloor \dfrac{t+\tau}{T} \right\rfloor & t \geq LT \end{cases} \tag{27}$$

Beweis. Sei $(B_0 B_1 \ldots)$ ein dichter Burst-Strom (s. Abb. 9). Dann ist offenbar

$M(t) \leq N_b^l(t+b)$ mit N_b^l gemäß (20), $l = l(B_0)$, $b = a(B_0) = (l-1)(T-D) - \tau$ nach (19)

(zu Beginn ein Burst der Länge l, danach kann frühestens zu den Zeiten ein Ereignis eintreten, zu denen ein 1-Burst auftritt). In mehreren Fallunterscheidungen – die aus den nicht vollständig übereinstimmenden Grenzen der Definitionsbereiche resultieren – wird gezeigt:

$N_b^l(t+b) \leq \overline{M}(t)$ für alle $t \in \mathbb{R}$. (*)

Aus (20) folgt mit $k = l - 1$:

$$N_b^l(t+b) = \begin{cases} 0 & \text{für} \quad t < 0 & \text{(a)} \\ 1 + \left\lfloor \dfrac{t}{D} \right\rfloor & \text{"} \quad t \in [0, kD) & \text{(b)} \\ l & \text{"} \quad t \in [kD, kD+T) & \text{(c)} \\ 1 + \left\lfloor \dfrac{t+\tau+b}{T} \right\rfloor & \text{"} \quad t \geq kD+T & \text{(d)} \end{cases}$$

Dabei resultiert in (c) die obere Intervallgrenze aus $lT - \tau - b = lT - (l-1)(T-D)$.

Für (*) sind die Fälle (a) und (b) evident (direkte Übereinstimmung).

Sei nun $t \in [kD, kD+T)$. Dann liegt wegen $k \leq L-1$, $D < T$ und damit $kD < LT-T$ bzgl. $\overline{M}(t)$ der Fall $t < LT$ vor. Weiter ist hier $t \geq kD$, wegen $k \in \mathbb{N}$ folgt

$$k \leq \left\lfloor \frac{t}{D} \right\rfloor \text{ und damit } l = k + 1 \leq 1 + \left\lfloor \frac{t}{D} \right\rfloor .$$

Schließlich sei einerseits $t \geq kD+T$. Nach kurzer Rechnung ergibt sich dann

$$\left\lfloor \frac{t + k(T-D)}{T} \right\rfloor \leq \left\lfloor \frac{t}{D} \right\rfloor ,$$

und dies gilt insbesondere für $t < LT$. Andererseits sei $t \geq LT$. Dann folgt

$$\left\lfloor \frac{t + k(T-D)}{T} \right\rfloor \leq \left\lfloor \frac{t+\tau}{D} \right\rfloor ,$$

und dies auch für $t \geq kD+T$. Damit ist (*) vollständig bewiesen.

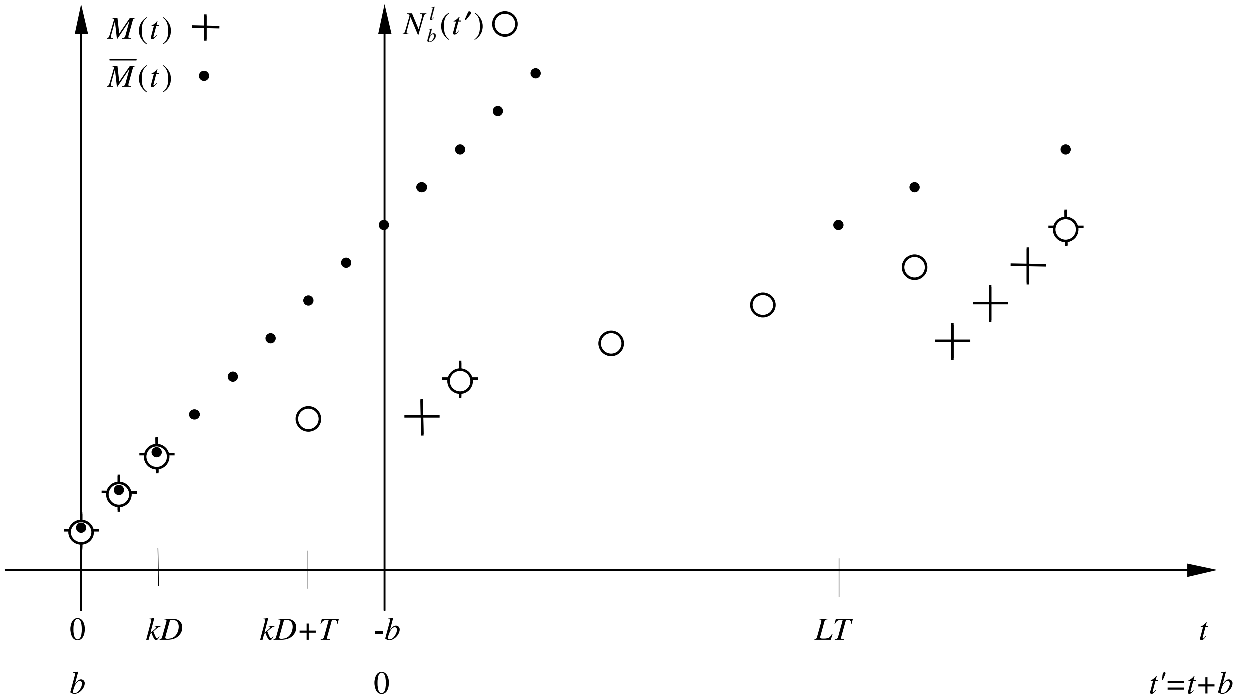

Abb. 9. Abschätzung der Zahl von Ereignissen für einen dichten Burst-Strom (Beispiel +: mit den Längen $l_0 = 3$, $l_1 = 2$, $l_2 = 4$)

Eine Betrachtung der Annahmen für Lemma 1 legt den Gedanken nahe, daß die Einführung *zweier* Parameter τ, τ' in Def. 1 nicht wesentlich ist. Genauer:

Definition 7. Zwei Modelle (D, T, τ_1, τ'_1), (D, T, τ_2, τ'_2) gem. Def. 1 heißen *äquivalent*, wenn sie beide zu derselben maximalen Burstlänge L führen.

Man beachte, daß dann auch fast alle Bewertungsgrößen (insbesondere die minimale Puffergröße P nach (14) bzw. (26)) übereinstimmen und daß für einen dichten Burst-Strom – gegeben durch die Burstlängen – die Burstabstände nach (24) in beiden Modellen gleich sind. Der einzige Unterschied liegt in der Zeitzählung; wird in den jeweiligen Modellen eine Zeittransformation wie oben vorgenommen (Beginn mit dem ersten Eintreten eines Ereignisses), so sind die entstehenden Ströme identisch. Insbesondere folgt aus Satz 2 unter Beachtung von (7) unmittelbar:

Corollar 1. Die Modelle (D, T, τ, τ'), $(D, T, \tau + \tau', 0)$ und $(D, T, 0, \tau + \tau')$ sind äquivalent.

4. Vergleichende Betrachtungen

4.1. ATM

Grundlage des Vergleiches zwischen Strömen gemäß Def. 1. und Strömen (Ereignisfolgen) in ATM-Netzen ist [ATM], vor allem Abschn. 3.6.2.4: Traffic Contract Parameter Specification; darauf wird besonders bezüglich aller hier nicht erläuterten Begriffe verwiesen.

Der Verkehr einer ATM-Verbindung kann durch die folgenden Verkehrsparameter beschrieben werden:

PCR	Peak Cell Rate
CDVT	Cell Delay Variation Tolerance
SCR	Sustainable Celll Rate
BT	Burst Tolerance.

Die Steuerung des Verkehrsstromes erfolgt durch den sog. *Generic Cell Rate Algorithm* GCRA(I,L), der die beiden Parameter I (Inkrement) und L (Limit) besitzt. Dieser Algorithmus erzeugt einen schwankungsbeschränkten Strom mit folgenden Besonderheiten:

(a) Es gibt keine Beschränkung der Verspätung. Tritt ein Ereignis nach seiner theoretischen Ereigniszeit TAT ein, so wird TAT auf diesen Wert weitergestellt.

(b) Die Zeitzählung beginnt mit 0 zu dem Zeitpunkt der Realzeit, zu dem das 1. Ereignis tatsächlich eintritt.

Weiter werden aus [ATM] 3.6.2.4.2 folgende Begriffe und Bezeichnungen übernommen:

- $R_p = \frac{1}{T_p}$ Peak Cell Rate PCR,

T_p: Peak Emission Interval (in [ATM] nur mit T bezeichnet!);
kleinster Abstand zwischen 2 Basisereignissen nach 3.6.2.4.2.2.;

- τ^*: Cell Delay Variation Tolerance;
  Parameter für L in GCRA, der das „Zellklumpen-Phänomen“ beschreibt;

- δ: Zeit zum Senden einer ATM-Zelle.

Dann sind die Prozesse, die durch die in 3.6.2.4.2 auftretenden GCRA-Algorithmen erzeugt werden, i.S.v. Def. 1 (τ^*, 0)-schwankungsbeschränkt mit der Rate T_p und Mindestabstand δ; dabei ist $\tau' = 0$ eine Folgerung von (a).

Damit ist die in 3.6.2.4.2.4 angegebene Formel für die maximale Anzahl konformer „back-to-back cells“ (also die maximale Länge von Bursts bei $D = \delta$)

$$N = \left\lfloor 1 + \frac{\tau}{T_p - \delta} \right\rfloor$$

offenbar identisch mit (6) bei $L = N$ (man beachte $\tau' = 0$).

Entsprechend sind in [ATM] 3.6.2.4.3 definiert:

- $R_s = \dfrac{1}{T_s}$ Sustainable Cell Rate SCR

  T_s: Parameter für I in GCRA;
  kleinster Abstand zwischen 2 Basisereignissen nach 3.6.2.4.3.2;

- τ_s: Parameter für L in GCRA (bezeichnet als Burst-Tolerance BT).

Analog sind die Prozesse, die durch die in 3.6.2.4.3. auftretenden GCRA-Algorithmen erzeugt werden, i.S.v. Def. 1 (τ_s, 0)-schwankungsbeschränkt mit der Rate T_s und dem Mindestabstand T_p.

Die in 3.6.2.4.3.3. angegebenen Formeln für die maximale Burstlänge MBS und die Bestimmung von τ_s aus MBS

$$MBS = \left\lfloor 1 + \frac{\tau_s}{T_s - T_p} \right\rfloor$$

$$\tau_s \in \left[(MBS - 1)(T_s - T_p), MBS(T_s - T_p) \right)$$

sind dann bei L = MBS identisch mit (6) und (15).

Ebenso stimmt die in 3.6.2.4.3.3. angegebene Formel für die Zahl $N(t)$ der bis zur Zeit t möglichen Ereignisse (unter Beachtung der dort nachfolgenden Bemerkung)

$$N(t) \le \begin{cases} 1 + \left\lfloor \dfrac{t}{T_p} \right\rfloor & \text{für } \; t < MBS \cdot T_s \\[2ex] 1 + \left\lfloor \dfrac{t + \tau_s}{T_s} \right\rfloor & \text{für } \; t \ge MBS \cdot T_s \end{cases}$$

bei Berücksichtigung der Zeittransformation, die aus (b) resultiert, mit (27) überein.

4.2. Linear beschränkte Ankunftsprozesse

ANDERSON betrachtet in [Ande] ein verteiltes System als ein Netz von Ressourcen, in dem Datenströme unterschiedlicher Qualität bei Einhaltung von Zeitschranken verarbeitet, gespeichert und zwischen verschiedenen Endpunkten übertragen werden. Dabei tritt an den einzelnen Schnittstellen des Systems eine Arbeitslast auf, die als Ankunftsprozeß von diskreten Nachrichten aufgefaßt und die durch drei Parameter beschrieben wird:

- M maximale Nachrichtengröße (im Rahmen dieser Arbeit nicht relevant)
- R maximale Nachrichtenrate
- W Workahead-Grenze (in [Stei] als Burstiness bezeichnet).

Dabei resultiert das *workahead* daraus, daß Prozesse oder Geräte Nachrichten-Bursts erzeugen können, die kurzzeitig mit einer R übersteigenden Rate übertragen werden bzw. daß Nachrichten an einer Schnittstelle vorfristig eintreffen. Ein derartiger Ankunftsprozeß wird *linear beschränkt* (LBAP, linear bounded arrival process) genannt, wenn er die Eigenschaft

$$N_I(t_0,t_1) \leq W + R(t_1 - t_0) \qquad \forall\ t_1 > t_0 \tag{28}$$

erfüllt, wobei $N_I(t_0,t_1)$ die Anzahl der an der Schnittstelle I im Intervall $[t_0, t_1]$ eintreffenden Nachrichten bezeichnet.

Bevor der ANDERSONsche Modellansatz mit dem hier in Abschn. 2 vorgestellten Ansatz verglichen wird, seien zwei Grenzfälle betrachtet, der bisher ausgeschlossen waren, nämlich $T = D$ und $D = 0$. Der Fall $T = D$ ist fast völlig uninteressant, denn dann gibt es weder maximale noch l-Bursts, da die einzige dichte Ereignisfolge das unbegrenzte Auftreten von Ereignissen im Abstand D ist. Anders bei $T > 0$, aber $D = 0$ (der Fall $T = D = 0$ soll naheliegenderweise ausgeschlossen bleiben): Eine genaue Analyse von Abschn. 2 und 3 zeigt, daß fast alle Begriffsbildungen und Ergebnisse gültig bleiben (allerdings kann man dann nicht mehr Ereignisse E_i mit ihren Zeiten $a(E_i)$ identifizieren). Dies trifft auch beispielsweise für (20) zu, denn wegen der Festlegung $[x,x) := \varnothing$ entfällt der Term mit D im Nenner.

Nach dieser Bemerkung überrascht die folgende Feststellung nicht: Ein LBAP-Strom mit den Parametern R und W ist ein $\left(\frac{W-1}{R},0\right)$-schwankungsbeschränkter Strom mit der Rate R und dem Mindestabstand $D = 0$. Denn für einen solchen Strom gilt mit $T = R^{-1}$ wegen (7) gerade

$$L = 1 + \left\lfloor \frac{W-1}{RT} \right\rfloor = W\,,$$

und aus (10) folgt für die Anzahl der Ereignisse bis zum Zeitpunkt t:

$$M(t) \leq 1 + \left\lfloor \frac{t}{T} + \frac{W-1}{RT} \right\rfloor = 1 + \left\lfloor \frac{t}{T} \right\rfloor + W - 1\,;$$

vereinbart man, daß im letzten Term das Operationssymbol $\lfloor\ \rfloor$ entbehrlich ist, da ohnehin $M(t)$ ganzzahlig sein muß, so folgt

$$M(t) \leq W + Rt,$$

was bis auf die Notation und die o.B.d.A. getroffene Annahme $t_0 = 0$ mit (28) übereinstimmt.

4.3. Tenet Protocol Suite

Die Tenet Real-Time Protocol Suite [BFMMVZ], [Ferr] wurde entwickelt, um den Forderungen gerecht zu werden, die moderne Anwendungen im Mulitmedia- und im Echtzeitbereich stellen, besonders im Hinblick auf Garantien für Durchsatz und Verzögerung. Die Tenet Protocol Suite ist ein Satz von Kommunikationsprotokollen für ein Paketvermittlungs-Netzwerk, das derartige Garantien liefert. Es wird davon ausgegangen, daß Clients Performance-Forderungen bzgl. der Übertragung oder Verarbeitung von Nachrichten an sog. Echtzeitkanäle (Übertragungskanäle eines Netzes, aber auch Prozessoren oder Rechnerknoten) stellen und dabei eine Beschreibung des zu erwartenden Verkehrs mitteilen. Diese Verkehrsbeschreibung besteht aus den Parametern

$$\left.\begin{array}{ll} X_{\min} & \text{minimaler Nachrichtenabstand} \\ X_{ave} & \text{minimaler mittlerer Nachrichtenabstand} \\ I & \text{Intervallänge für die Mittelung} \end{array}\right\} \qquad (29)$$

$S_{\max}$ maximale Nachrichtengröße.

Wie in Abschn. 4.2 ist S_{max} im Rahmen dieser Arbeit nicht relevant. Weiter ist sofort ersichtlich, daß $D = X_{min}$, $T = X_{ave}$ i.S.v. Def. 1 gilt. Bezeichnet $N(I)$ die Anzahl der in einem Intervall der Länge I auftretenden Nachrichten. so gilt unter Einschluß der Intervallgrenzen

$$N(I) \le 1 + \frac{I}{X_{ave}} .$$

Da diese Beziehung für jede Lage eines Intervalls der Länge I auf der Zeitachse gelten muß (und wegen $N(I) \in \mathbb{N}$), ist das gleichbedeutend mit

$$N(t) \le 1 + \left\lfloor \frac{t}{T} \right\rfloor$$

mit $T = X_{ave}$. Gleichzeitig ist $1 + \left\lfloor \frac{I}{X_{ave}} \right\rfloor$ die größtmögliche Burstlänge, wobei die Burstbildung im Gegensatz zu Abschn. 4.2 nicht durch vorzeitiges, sondern durch verspätetes (aufgeschobenes) Eintreffen von Nachrichten entsteht. Damit ergibt sich zusammengefaßt (unter Beachtung von (4), (7) und (10)):

Ein Strom mit den Parametern $X_{min.}$, X_{ave} und I gemäß (29) ist ein $\left(0, \left\lfloor \frac{I}{X_{ave}} \right\rfloor (X_{ave} - X_{\min})\right)$-schwankungsbeschränkter Strom mit der Rate X_{ave}^{-1} und dem Mindestabstand $X_{\min}$ im Sinne von Def. 1.

4.4. Weitere Arbeiten

In der Literatur über Echtzeit-Systeme und Kommunikationssysteme finden sich viele Ansätze und Ergebnisse zur Beschreibung der Verkehrslast in solchen Systemen; dabei nehmen Aussagen über Pufferanforderungen und das Garantieren von maximalen Ende-zu-Ende-Verzögerungen einen besonderen Platz ein. In den meisten Fällen wird mit der Entwicklung eines neuen Systems oder einer neuen Anwendung gleichzeitig ein neues Verkehrs- oder Lastmodell vorgeschlagen. So wird beispielsweise das „leaky bucket“-Modell (d.h. der Generic Cell Rate Algorithm von ATM) in [GoLV] eingeführt, Burst-Ströme werden in [LaXi] betrachtet, beides bezüglich Verzögerungsgarantien und basierend auf einem deterministischen Modell. ZHANG und FERRARI führen in [ZhFe] ein probabilistisches Modell ein, und in [KoTS], [MeMP] und [WuMa] werden – unabhängig voneinander – Puffer-Anforderungen mittels MARKOV-Ketten (stochastischen Prozessen mit diskreter Zeit und endlichem Zustandsraum) behandelt. Andererseits wird in [WiWo] ein deterministisches Modell zu dem gleichen Zweck benutzt.

Nur in wenigen Arbeiten erfolgt eine vergleichende Betrachtung verschiedener Modelle oder Verkehrsbeschreibungen, und dann meist im Hinblick auf Bedienungsdisziplinen [ZhKe] oder auf Funktionen der Verkehrsbeschränkung [KWLZ]; letztere Arbeit schließt auch das Tenet-Modell ein und stellt einen Ansatz zu einem allgemeinen deterministischen Modell dar. KHANS Vorgehen

[KhLi] ist dem hier vorgestellten ähnlich: Es wird ein sog. (L, M, T)-Mechanismus eingeführt, um das Verhalten des „leaky bucket“ zu beschreiben, wobei L den minimalen Abstand zwischen zwei ATM-Zellen, M die maximale Burstgröße und T den Burst-Abstand bezeichnen, allerdings wird der Input-Prozeß als stochastisch angenommen. Jedoch findet sich an keiner Stelle – unseres Wissens – eine quantitative Gegenüberstellung von Verkehrsbeschreibungen oder Verkehrsmodellen, die die Grundlage einer Vereinheitlichung sein könnte.

5. Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurde ein Satz von Parametern vorgeschlagen, um schwankungsbeschränkte Ströme zu beschreiben, und es wurden einige wichtige Aussagen und Kenngrößen derartiger Ströme zusammengestellt:

- maximale Burstlänge
- Mindest-Puffergröße zur Vermeidung von Datenverlust
- Mindest-Abstände zwischen Bursts
- Anzahl von Ereignissen in einem Zeitintervall der Länge t.

Als ein wesentliches Ergebnis wurde gezeigt, daß das beschriebene Modell einige wichtige ähnliche Modelle aus der Literatur einschließt. Vorrangige Aufgabe der weiteren Untersuchungen ist die Einbeziehung probabilistischer Betrachtungen (die vor allem aus der Überlagerung derartiger Ströme resultieren) und ferner die Konstruktion eines Systems, dem die geschilderte Lastbeschreibung durchgehend, d.h. in allen Systemkomponenten und auf allen Ebenen zugrunde liegt.

Literatur

[ATM] The ATM Forum: ATM User Network Interface Specification, Version 3.0. Prentice Hall PTR, 1993.

[Ande] ANDERSON, D. J.: Metascheduling for Continuous Media. ATM Transactions on Computer Systems, Vol. 11, No. 3 (August 1993), p. 226-252.

[BFMMVZ] BANERJEA, A., D. FERRARI, B. A. MAH, M. MORAN, D. C. VERMA, H. ZHANG: The Tenet Real-Time Protocol Suite: Design, Implementation, and Experiences. University of California at Berkeley and ICSI, TR-94-059, November 1994.

[Ferr] FERRARI, D.: Real-Time Communication in an Internetwork. University of California at Berkeley and ICSI, Technical Report tr-92-001.

[Hama] HAMANN, CL.-J.: On the Quantitative Specification of Jitter Constrained Periodic Streams. TU Dresden, 1996 (eingereicht zur Veröffentlichung).

[GoLV] GOYAL, P.; S. S. LAM, H. M. VIN: Determing End-to-End Delay Bounds in Heterogeneous Networks. Proceedings NOSSDAV '95, Lecture Notes in Computer Science, Vol. 1018.

[KhLi] KHAN, J.; V. O. K. LI: Traffic control in ATM networks. Computer Networks and ISDN Systems, Vol. 27(1) (1994), p. 85 - 100.

[KoTS] KONTOVASSILIS, K. P.; J.T. TSILIGARIDIS; G.I. STASSINOPOULOS: Buffer dimensioning for delay/loss-sensitive traffic. Computer Communications, Vol. 18(5) (May 1995), p. 315-328.

[KWLZ] KNIGHTLY, E. W. ; D. E. WREGE; J. LIEBEHERR; H. ZHANG: Fundamental Limits and Tradeoffs of Providing Deterministic Guarantees to VBR Video Traffic. Proceedings of ACM SIGMETRICS '95.

[LaXi] LAM, S. S.; G. G. XIE: Burst Scheduling Networks: Flow Specification and Performance Guarantees. Proceedings NOSSDAV '95, Lecture Notes in Computer Science, Vol. 1018.

[MeMP] MEYER, J. F.; S. MONTAGNA; R. PAGLINO: Dimensioning of an ATM switch with shared buffer and threshold priority. Computer Networks and ISDN Systems, Vol. 26(1) (Sep 1993) p. 95-108.

[Stei] STEINMETZ, R.: Multimedia Operating Systems: Resource Reservation, Scheduling, File Systems, and Architectures. Technical Report 43.9402, IBM ENC, Heidelberg 1994.

[WiWo] WILLIAMSON, J.; L. C. WOLF: Reducing Buffer Space Requirements for Multimedia Data Streams. IBM European Network Center, Technical Report 43.9504, 1995.

[WuMa] WU, G.-L.; J. W. MARK: Discrete time analysis of leaky-bucket congestion control. Computer Networks and ISDN Systems, Vol. 26(1) (Sep 1993), p. 79 - 94.

[ZhFe] ZHANG, H.; D. FERRARI: Rate Controlled Service Disciplines. Journal of High Speed Networking (to appear).

[ZhKe] ZHANG, H.; S. KESHAV: Comparison of Rate-Based Service Disciplines. Proceedings NOSSDAV '95, Lecture Notes in Computer Science, Vol. 1018.