# On Deterministic Traffic Regulation and Service Guarantees: A Systematic Approach by Filtering

Cheng-Shang Chang*
Dept. of Electrical Engineering
National Tsing Hua University
Hsinchu 30043
Taiwan, R.O.C.
Email: cschang@ee.nthu.edu.tw

(Jan. 1997; revision: Sept. 1997)

**Abstract**

In this paper, we develop a filtering theory for deterministic traffic regulation and service guarantees under the $(\min, +)$-algebra. We show that traffic regulators that generate $f$-upper constrained outputs can be implemented *optimally* by a linear time invariant filter with the impulse response $f^*$ under the $(\min, +)$-algebra, where $f^*$ is the subadditive closure defined in the paper. Analogous to the classical filtering theory, there is an associate calculus, including feedback, concatenation, "filter bank summation" and performance bounds. The calculus is also applicable to the recently developed concept of service curves that can be used for deriving deterministic service guarantees. Our filtering approach not only yields easier proofs for more general results than those in the literature, but also allows us to design traffic regulators via systematic methods such as concatenation, filter bank summation, linear system realization, and FIR-IIR realization. We illustrate the use of the theory by considering a window flow control problem and a service curve allocation problem.

Keywords: traffic regulation, leaky buckets, filtering, min-plus algebra, service curves

---

*This research is supported in part by the National Science Council, Taiwan, R.O.C., under Contracts NSC-86-2221-E007-036 and NSC 86-2115-M007-025. The conference version that contains part of the results in this paper was presented in INFOCOM'97.

# 1 Introduction

Future high speed digital networks aim to serve integrated traffic, such as voice, video, fax, and so forth. To control interaction among traffic generated by different sources, traffic regulation seems inevitable. In the paper [10], Cruz proposed the following deterministic traffic characterization. For an increasing sequence $A \equiv \{A(t), t = 0, 1, 2, \dots\}$ (with $A(0) = 0$), it is $f$-upper constrained for some function $f$ if

$$A(t_2) - A(t_1) \leq f(t_2 - t_1), \quad \forall t_1 \leq t_2. \tag{1}$$

Based on this characterization, he went on to develop a network calculus, including multiplexing and demultiplexing, so that deterministic performance, such as bounded delay and bounded queue length, can be guaranteed. To use the calculus, one needs to address the fundamental question of traffic regulation:

*How does one regulate traffic* optimally *such that the constraint in (1) can be satisfied?*

In [10, 11], Cruz addressed the traffic regulation problem with the $(\sigma, \rho)$ regulator (leaky bucket) and the FIFO queue. Both only generate traffic that satisfies a linear constraint ($f$ is a linear function) in (1). For $f(t) = \sigma + \rho t$, Ananthram and Konstantopoulos [1, 17] recently showed that the leaky bucket is the best causal traffic regulator to generate such a constraint in terms of maximizing the total number of departures at any moment in time (an equivalent statement in terms of delay under FIFO is previously shown in Cruz [11]). Motivated by these important works, we propose in this paper a general filtering approach for the traffic regulation problem. Our approach is also applicable to the concept of service curve recently developed by Parekh and Gallager [19], Cruz [12] and Sariowan [21].

Our approach is built upon filtering under the $(\min, +)$-algebra. In such an algebra, one replaces the usual addition operator by the min operator and the usual multiplication operator by the addition operator. It is known (see e.g., [3]) that such an algebra is a commutative semifield that possesses properties such as associativity, commutativity, and distributivity. These two operations min and + are also applicable to matrix operations. In view of this, we consider two binary operations for sequences under the $(\min, +)$ algebra: the pointwise minimum (denoted by $\oplus$) and the convolution (denoted by $\star$). For these two binary operations on sequences, they also have associativity, commutativity, and distributivity as the min and + operators in the $(\min, +)$-algebra. More importantly, the convolution operation has monotonicity that allows us to consider the limit of self convolution. The limit is shown to be a unitary operation, called the subadditive closure, as the limit is always a subadditive sequence. Many properties of subadditive closures are developed in this paper.

Based on the properties developed for subadditive closures, we show that $A$ is $f$-upper constrained if and only if $A$ is $f^*$-upper constrained, where $f^*$ is the subadditive closure of $f$. In view of this, we find that *traffic regulation in (1) can be achieved* optimally *by a linear time invariant filter with the impulse response* $f^*$ *under the* $(\min, +)$*-algebra.* We call such a filter the *maximal* $f$*-regulator* as it is also the best causal traffic regulator that one can implement in terms of maximizing the total number of departures from the regulator at any moment in time. Analogous to the classical filtering theory, we develop the calculus for maximal regulators that includes feedback, concatenation, "filter bank summation", and performance bounds. In particular, we show that a concatenation of two maximal regulators yields another maximal regulator, independent of the order of the two maximal regulators. We also give a condition for "filter bank summation" of maximal regulators to yield another maximal regulator. Under such a condition, "filter bank summation" is equivalent to concatenation.

The filtering approach not only yields easier proofs for more general results than those in the literature, but also allows us to realize traffic regulators under the $(\min, +)$-algebra via systematic methods such as concatenation, filter bank summation, linear system realization, and FIR-IIR realization. When specialized to leaky buckets, we find a new linear system realization and a new FIR-IIR realization for a concatenation of leaky buckets. These realizations are obviously suitable for VLSI implementation. For a traffic regulator with a periodic constraint function, we also derive a FIR-IIR realization for such a regulator.

Recently, Parekh and Gallager [19], Cruz [12] and Sariowan [21] developed an important concept on service guarantees. The concept, called service curve, allows one to compute tighter end-to-end service guarantees than those in [10, 11]. Phrased in terms of the $(\min, +)$-algebra, a server that guarantees service curve $f$ for an input $A$ if its output $B$ is not less than the output from a linear time invariant filter with the input $A$ and the impulse response $f$ under the $(\min, +)$-algebra. Such a server is called an $f$-server in the paper. In view of this analogy to the maximal regulators, the calculus developed for the maximal regulators, including feedback, concatenation, "filter bank summation", and performance bounds, is also applicable to $f$-servers. We illustrate the use of the theory by considering a window flow control problem in [2]. We also show that maximal regulators, in conjunction with the maximum delay guarantee, guarantee shifted-subadditive service curves in [21]. Based on this, we provide a couple of rules for service curve allocation among a concatenation of servers.

We note that the importance of the role of the $(\min, +)$-algebra in deterministic traffic regulation and service guarantees is also recognized by Agrawal and Rajan [2], Cruz and Okino [15] and Le Boudec [4]. In particular, a "division" operator in the $(\min, +)$-algebra is defined in [2, 15]. Such an operator provides a simple representation of performance bounds and output burstiness. Extensions of such an operator can be found in Chang [9] and Cruz [13]. For the window flow control problem, Cruz and Okino [15] considered a more detailed model than the one in Agrawal and Rajan [2]. They obtained a tighter bound than the one in this paper. See also Le Boudec and Thiran [5].

The paper is organized as follows: in Section 2, we review the $(\min, +)$-algebra and develop the concept and the properties of subadditive closure. We then use these properties to show equivalent statements of Cruz's traffic characterization in Section 3. Based on these equivalent statements, we develop the properties of maximal regulators via viewing them as filters under the $(\min, +)$-algebra. The design issues of maximal regulators under the $(\min, +)$-algebra are addressed in Section 4. The analogy of maximal regulators and $f$-servers is discussed in Section 5. We conclude the paper in Section 6 by discussing possible extensions of the filtering approach.

# 2 Min-plus algebra and subadditive closure

## 2.1 Min-plus algebra

In this section, we briefly introduce the (min,+)-algebra (for more details of such an algebra, we refer to [3]). In such an algebra, one replaces the usual addition operator by the *min* operator and the usual multiplication operator by the addition operator. Denote by $\oplus$ the min operator and $\otimes$ the $+$ operator. It is known that $(I\!R \cup \{\infty\}, \oplus, \otimes)$ is a commutative semifield that possesses properties such as

**1.** (Associativity) $\forall a, b, c \in I\!R \cup \{\infty\}$,

$$(a \oplus b) \oplus c = a \oplus (b \oplus c)$$
$$(a \otimes b) \otimes c = a \otimes (b \otimes c)$$

**2.** (Commutativity) $\forall a, b \in I\!R \cup \{\infty\}$,

$$a \oplus b = b \oplus a$$
$$a \otimes b = b \otimes a$$

**3.** (Distributivity) $\forall a, b, c \in I\!R \cup \{\infty\}$,

$$(a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$$

Taking distributivity for example, $\min(a,b)+c = \min(a+c, b+c)$. It turns out that $(\min,+)$-algebra is also applicable to matrix operations (without the commutative property for matrix "multiplication"). To familiarize the readers with such operations, in the following we provide an example of matrix multiplication and an example for matrix addition under the $(\min,+)$-algebra. Denote by $\otimes$ and $\oplus$ the operators for matrix multiplication and matrix addition under the $(\min,+)$-algebra.

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{bmatrix} G_{1,1} & G_{1,2} \\ G_{,2,1} & G_{2,2} \end{bmatrix} \otimes \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{2}$$

is equivalent to

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \min[G_{1,1}+y_1, G_{1,2}+y_2] \\ \min[G_{2,1}+y_1, G_{2,2}+y_2] \end{pmatrix}, \tag{3}$$

and

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \oplus \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \tag{4}$$

is equivalent to

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \min[y_1, u_1] \\ \min[y_2, u_2] \end{pmatrix}. \tag{5}$$

### 2.2 Subadditive closure

Now we turn our attention to operations for sequences (or functions) indexed by $t = 0, 1, 2, \ldots$, under the $(\min,+)$-algebra. Let $\mathcal{F}$ (resp. $\mathcal{F}_0$) be the set of increasing sequences with $f(0) \geq 0$ (resp. $f(0) = 0$). That is, a sequence $f \equiv \{f(t), t = 0, 1, 2, \ldots\} \in \mathcal{F}$ (resp. $\mathcal{F}_0$) satisfies $f(0) \geq 0$ ( resp. $f(0) = 0$) and $f(s) \leq f(t)$ for all $s \leq t$. We say two sequences are equal, denoted by $f = g$, if $f(t) = g(t)$ for all $t$. Similarly, $f \leq g$ if $f(t) \leq g(t)$ for all $t$. Define the following operations for $f$ and $g$ in $\mathcal{F}$.

**(i)** (min) the pointwise minimum of two sequences:

$$(f \oplus g)(t) = \min[f(t), g(t)].$$

**(ii)** (convolution) the convolution of two sequences under the $(\min,+)$-algebra:

$$(f \star g)(t) = \min_{0 \leq s \leq t}[f(s) + g(t-s)].$$

It is clear that these two operations are closed in $\mathcal{F}$. Let $\epsilon$ be the sequence with $\epsilon(t) = \infty$ for all $t$, and $\mathbf{e}$ be the sequence with $\mathbf{e}(0) = 0$ and $\mathbf{e}(t) = \infty$ for all $t > 0$. Then one can easily verify the following properties:

**1.** (Associativity) $\forall f, g, h \in \mathcal{F}$,

$$(f \oplus g) \oplus h = f \oplus (g \oplus h)$$
$$(f \star g) \star h = f \star (g \star h)$$

**2.** (Commutativity) $\forall f, g \in \mathcal{F}$,

$$f \oplus g = g \oplus f$$
$$f \star g = g \star f$$

**3.** (Distributivity) $\forall f, g, h \in \mathcal{F}$,

$$(f \oplus g) \star h = (f \star h) \oplus (g \star h)$$

**4.** (zero element) $\forall f \in \mathcal{F}$,

$$f \oplus \epsilon = f$$

**5.** (Absorbing zero element) $\forall f \in \mathcal{F}$,

$$f \star \epsilon = \epsilon \star f = \epsilon$$

**6.** (Identity element) $\forall f \in \mathcal{F}$,

$$f \star \mathbf{e} = \mathbf{e} \star f = f$$

**7.** (Idempotency of addition) $\forall f \in \mathcal{F}$,

$$f \oplus f = f$$

These show that $(\mathcal{F}, \oplus, \star)$ is a commutative dioid (see [3]) with the zero element $\epsilon$ and the identity element $\mathbf{e}$. Though some of the results reported in this section may be derived from known results for dioids in [3], we present our results in a self-contained manner.

Beside the properties for a dioid, we have the following monotonicity that plays an important role in our development.

**8.** (Monotonicity) $\forall f, \tilde{f}, g, \tilde{g} \in \mathcal{F}$, if $f \leq \tilde{f}$ and $g \leq \tilde{g}$, then

$$f \oplus g \leq \tilde{f} \oplus \tilde{g} \leq \tilde{f}$$
$$f \star g \leq \tilde{f} \star \tilde{g}$$

If $g$ is also in $\mathcal{F}_0$, then $f \star g \leq f$. If both $f$ and $g$ are in $\mathcal{F}_0$, then $f \oplus g \geq f \star g$.

In view of the monotonicity, if we define $f^{(n)}$ to be the self convolution of $f$ for $n$ times, i.e., $f^{(n)} = f \star f^{(n-1)}$ with $f^{(1)} = f$, then $\mathbf{e} \oplus f$ is in $\mathcal{F}_0$ and $(\mathbf{e} \oplus f)^{(n)}(t)$ is decreasing in $n$ for any fixed $t$. This implies there is a limiting sequence for $(\mathbf{e} \oplus f)^{(n)}$. This sequence turns out to be the subadditive closure defined below (the proof is in Lemma 2.2 below).

**Definition 2.1** *(Subadditive closure) For any $f \in \mathcal{F}$, define $f^*$ via the following recursive equation:*

$$f^*(0) = 0,$$
$$f^*(t) = \min[f(t), \min_{0<s<t}[f^*(s) + f^*(t-s)]], t > 0.$$

*We call $f^*$ the subadditive closure of $f$.*

To ease our presentation, we define the unitary operation $\circ$ such that $f^\circ = f \oplus \mathbf{e}$, i.e., $f^\circ(t) = f(t)$ for $t > 0$ and $f^\circ(0) = 0$. Clearly, $f^* = (f^\circ)^*$ as $f(0)$ is not used in the recursive construction of $f^*$. We also note that $f^\circ = f$ for all $f$ that are already in $\mathcal{F}_0$. Thus, $(f^\circ)^\circ = f^\circ$. From the monotonicity of $\oplus$, one has the monotonicity for $\circ$, i.e., $f^\circ \leq g^\circ$ if $f \leq g$. The operation $\circ$ that converts $f \in \mathcal{F}$ into $f^\circ \in \mathcal{F}_0$ allows us to use more monotonicity results, e.g.,

$$f^\circ \geq f^\circ \oplus g^\circ \geq f^\circ \star g^\circ.$$

In the following lemma, we derive properties for subadditive closure.

**Lemma 2.2** *Suppose that $f, g \in \mathcal{F}$.*

**(i)** *The subadditive closure $f^*$ is in $\mathcal{F}_0$ ($f^*$ is increasing with $f^*(0) = 0$). Thus, $f^* = (f^\circ)^* = (f^*)^\circ$, and the subadditive closure is closed in $\mathcal{F}$.*

**(ii)** *The subadditive closure $f^*$ is subadditive, i.e., $f^*(s) + f^*(t) \geq f^*(s+t)$, for all $s, t \geq 0$.*

**(iii)** *$f^* \leq f^\circ \leq f$.*

**(iv)** *$f^* = f$ if and only if $f$ is subadditive and $f(0) = 0$. From (ii), it follows that $(f^*)^* = f^*$.*

**(v)** *$f \star f = f$ if and only if $f$ is subadditive and $f(0) = 0$. From (ii), it follows that $f^* \star f^* = f^*$.*

**(vi)** *$(f^\circ)^{(n)} \geq (f^\circ)^{(n+1)} \geq f^*$ for all $n$, and $(f^\circ)^{(t)}(t) = f^*(t)$ for all $t$. As a result, $(f^\circ)^{(n)}(t) = f^*(t)$ for all $t \leq n$ and*

$$\lim_{n\to\infty} (f^\circ)^{(n)}(t) = f^*(t) \text{ for all } t.$$

**(vii)** *(Maximum solution)* $f^*$ *is the maximum solution of the equation* $h = (h \star f)^\circ$, *i.e., for any* $h$ *satisfying* $h = (h \star f)^\circ$, $h \le f^*$.

**(viii)** *(Monotonicity) If* $f \le g$, *then* $f^* \le g^*$.

**(ix)** *If* $f^\circ$ *and* $g^\circ$ *are subadditive, then* $f^\circ \star g^\circ$ *is also subadditive and* $(f^\circ \star g^\circ)(0) = 0$. *From (iv), it follows that* $(f^* \star g^*)^* = f^* \star g^*$.

**(x)** $(f \star g)^* \ge (f^\circ \star g^\circ)^* = f^* \star g^*$.

**(xi)** $(f \oplus g)^* = ((f \oplus g)^\circ)^* = (f^\circ \oplus g^\circ)^* = f^* \star g^*$ *(cf. Lemma 4.107 in [3]).*

**(xii)** *(Maximum subadditive solution)* $f^*$ *is the maximum subadditive solution of the equation* $h = h \oplus f^\circ$, *i.e.,* $f^*$ *is the maximum subadditive sequence in* $\mathcal{F}_0$ *that is bounded above by* $f$.

The proof for Lemma 2.2 is in Appendix A. We note that there is an equivalent view of the subadditive closure from Lemma 2.2(vi). As $\mathbf{e}$ is the identity element for $\star$, one has from distributivity and idempotency that

$$(f^\circ)^{(n)} = (\mathbf{e} \oplus f)^{(n)} = \mathbf{e} \oplus f \oplus f^{(2)} \oplus \ldots \oplus f^{(n)}.$$

Thus,

$$\begin{aligned} f^* = (\mathbf{e} \oplus f)^* &= \lim_{n\to\infty} (\mathbf{e} \oplus f)^{(n)} \\ &= \lim_{n\to\infty} (\mathbf{e} \oplus f \oplus f^{(2)} \oplus \ldots \oplus f^{(n)}), \end{aligned} \tag{6}$$

and the subadditive closure is equivalent to the star operation in [3].

**Remark 2.3** We note that over all $f \in \mathcal{F}$ (with $h \in \mathcal{F}$ fixed), there is a minimum solution of the equation $h = h \star f$. The minimum solution, denoted by $\hat{h}$, is the minimum envelope of $h$ (see [6]), i.e.,

$$\hat{h}(t) = \sup_{s \ge 0} [h(t+s) - h(s)]. \tag{7}$$

To see this, note from (7) that $h(s) + \hat{h}(t-s) \ge h(t)$ for all $s \le t$ (with the equality when $s = t$). Thus, $h \star \hat{h} = h$. On the other hand, if $f$ is a solution of $h = h \star f$, then for all $s, t \ge 0$

$$h(t+s) = \min_{0 \le \tau \le t+s} [h(\tau) + f(t+s-\tau)] \le h(s) + f(t).$$

Thus, $\hat{h}(t) \le f(t)$ for all $t$.

**Theorem 2.4** *(Linear system with feedback)*

**(i)** *(cf. Theorem 4.75(i) in [3]) For the equation (see Fig. 1)*

$$B = A \oplus (B \star f), \tag{8}$$

*$B = A \star f^*$ is the maximum solution.*

**(ii)** *If $f(0) > 0$ and $A(t) < \infty$ for all $t$, then $B = A \star f^*$ is the unique solution for (8).*

**(iii)** *Under the condition in (ii), if*

$$B \geq A \oplus (B \star f), \tag{9}$$

*then $B \geq A \star f^*$.*

One may view (8) as a linear system with feedback under the $(\min, +)$-algebra, where $A$ is the input and $B$ is the output. The condition that $f(0) > 0$ and $A(t) < \infty$ for all $t$ may be viewed as the stability condition. It ensures that the looping effect (due to feedback) eventually dies out as we shall see in the proof.

**Proof.** (i) We first show that $B = A \star f^*$ is a solution. Observe from the associative property of $\star$, the identity element $\mathbf{e}$, the distributive property and Lemma 2.2(vii) that

$$A \oplus ((A \star f^*) \star f) = (A \star \mathbf{e}) \oplus (A \star (f^* \star f)) = A \star (\mathbf{e} \oplus (f^* \star f)) = A \star (f^* \star f)^\circ = A \star f^*.$$

Thus, $B = A \star f^*$ is indeed a solution. To see that $A \star f^*$ is the maximum solution, iterating the equation in (8) yields

$$B = A \oplus (A \star f) \oplus (A \star f^{(2)}) \oplus \ldots \oplus (A \star f^{(n)}) \oplus (B \star f^{(n+1)}). \tag{10}$$

Thus, from the monotonicity of $\oplus$ and the distributive property,

$$\begin{aligned} B &\leq A \oplus (A \star f) \oplus (A \star f^{(2)}) \oplus \ldots \oplus (A \star f^{(n)}) \\ &= A \star (\mathbf{e} \oplus f \oplus f^{(2)} \oplus \ldots \oplus f^{(n)}). \end{aligned}$$

In view of Lemma 2.2(vi), letting $n \to \infty$ yields $B \leq A \star f^*$.

(ii) Now we show that $A \star f^*$ is the unique solution under $f(0) > 0$ and $A(t) < \infty$ for all $t$. Since for any $B$ and $f$ in $\mathcal{F}$, $B \star f^{(n+1)}$ is still in $\mathcal{F}$. One has

$$(B \star f^{(n+1)})(t) \geq (B \star f^{(n+1)})(0) = B(0) + (n+1)f(0) \geq (n+1)f(0).$$

As we assume that $f(0) > 0$ and $A(t) < \infty$ for all $t$, for any fixed $t$, there exists a finite $n$ such that $A(t) \leq (n+1)f(0) \leq (B \star f^{(n+1)})(t)$. In view of (10), for each fixed $t$ there is a finite $n$ such that

$$\begin{aligned} B(t) &= \Big(A \oplus (A \star f) \oplus (A \star f^{(2)}) \oplus \ldots \oplus (A \star f^{(n)})\Big)(t) \\ &= \Big(A \star (\mathbf{e} \oplus f \oplus f^{(2)} \oplus \ldots \oplus f^{(n)})\Big)(t) \geq (A \star f^*)(t). \end{aligned}$$

As $A \star f^*$ is the maximum solution, one must have $B(t) = (A \star f^*)(t)$ for any fixed $t$. Thus, $A \star f^*$ is the unique solution.

(iii) Follow the same argument in (ii). ■

# 3 Traffic regulation

Consider a sequence $A \in \mathcal{F}_0$. Viewing $A(t)$ as the cumulative arrival of a source by time $t$, Cruz [10] introduced the following traffic characterization.

**Definition 3.1** *A sequence $A$ is $f$-upper constrained for some sequence $f \in \mathcal{F}_0$ if for all $t$*

$$A(t) - A(s) \leq f(t-s), \quad \forall 0 \leq s \leq t. \tag{11}$$

This characterization has the following equivalent statements.

**Lemma 3.2** *Suppose that $A, f \in \mathcal{F}_0$. The following statements are equivalent.*

**(i)** *The sequence $A$ is $f$-upper constrained.*

**(ii)** $A = A \star f$.

**(iii)** $A = A \star f^*$.

**(iv)** *The sequence $A$ is $f^*$-upper constrained.*

**Proof.** (i) $\Rightarrow$ (ii) One may rewrite (11) as

$$A(t) = \min_{0 \leq s \leq t} [A(s) + f(t-s)],$$

taking into account that $f(0) = 0$. Thus, $A = A \star f$.

(ii) $\Rightarrow$ (iii) By iterating $A = A \star f$, one has $A = A \star f^{(n)} = A \star (f^\circ)^{(n)}$ for all $n$ as we assume that $f \in \mathcal{F}_0$. As a result of Lemma 2.2(vi), one has $A = A \star f^*$.

(iii) $\Rightarrow$ (iv) Since $A = A \star f^*$,

$$A(t) = \min_{0 \le s \le t} [A(s) + f^*(t-s)].$$

Thus, $A(t) - A(s) \le f^*(t-s)$ for all $0 \le s \le t$.

(iv) $\Rightarrow$ (i) From Lemma 2.2(iii), $f^* \le f$. Thus, $A(t) - A(s) \le f^*(t-s) \le f(t-s)$ for all $0 \le s \le t$. ■

As discussed in the Section 1, the fundamental question is how one regulates a source *optimally* such that the output is $f$-upper constrained. The following theorem is the key to the question.

**Theorem 3.3** *Suppose that $A, f \in \mathcal{F}_0$. Construct a sequence $B = A \star f^*$, i.e.,*

$$B(t) = \min_{0 \le s \le t} [A(s) + f^*(t-s)]. \tag{12}$$

**(i)** *(Traffic regulation) $B$ is $f^*$-upper constrained and thus $f$-constrained.*

**(ii)** *(Optimality) For any $f$-upper constrained sequence $\tilde{B}$ that satisfies $\tilde{B} \le A$, one has $\tilde{B} \le B$.*

**(iii)** *(Conformity) $A$ is $f$-upper constrained if and only if $B = A$.*

**Proof.** (i) In view of Lemma 3.2(iii), it suffices to show that $B = B \star f^*$. Since $B = A \star f^*$, it then follows from the associativity of $\star$ and Lemma 2.2(v) that

$$B \star f^* = (A \star f^*) \star f^* = A \star (f^* \star f^*) = A \star f^* = B.$$

(ii) As we assume that $\tilde{B}$ is $f$-upper constrained and $\tilde{B} \le A$, it then follows from Lemma 3.2(iii) and the monotonicity of $\star$ that

$$\tilde{B} = \tilde{B} \star f^* \le A \star f^* = B.$$

(iii) This is a direct application of Lemma 3.2(iii). ■

Note that the condition $\tilde{B} \le A$ is referred to one of the causal conditions in [1] as the number of departures cannot be larger than the number of arrivals. Theorem 3.3(i) shows that for any input sequence $A$, the construction in (12) generates an $f$-upper constrained sequence $B$. Theorem 3.3(ii) shows that it is the best construction that one can implement if one would like to maximize the number of departures by time $t$. Finally, Theorem 3.3(iii) shows that if $A$ is already $f$-upper constrained, then it will pass through the construction in (12) without any change.

To ease our presentation, we make the following definition.

**Definition 3.4** *(Maximal $f$-regulator) For $f \in \mathcal{F}_0$, the construction in (12) is called the* maximal $f$-regulator *(for the input $A$).*

We note that the definition in Definition 3.4 is a generalization of that in our conference version [8] (see also [21, 2]), where the sequence $f$ is required to be subadditive. Analogous to the classical filtering theory, we may view the maximal $f$-regulator as a linear time invariant filter with the impulse response $f^*$ under the $(\min, +)$-algebra. From Theorem 2.4(i), one can also view the maximal $f$-regulator as the maximum solution of the feedback system $B = A \oplus (B \star f)$.

In the following theorem, we discuss a concatenation of maximal regulators. The result is a natural extension of its counterpart in the classical filtering theory.

**Theorem 3.5** *(Concatenation) A concatenation of the maximal $f_1$-regulator and the maximal $f_2$-regulator, independent of the order, is the maximal $f$-regulator, where $f = f_1 \star f_2$ is the convolution of $f_1$ and $f_2$ under the $(\min, +)$-algebra, i.e.,*

$$f(t) = \min_{0 \le s \le t} [f_1(s) + f_2(t - s)]. \tag{13}$$

**Proof.** Now consider an input sequence $A$ to the maximal $f_1$-regulator. Let $B_1$ be the output sequence of the maximal $f_1$-regulator, which is then fed into the maximal $f_2$-regulator. Let $B$ be the output of the maximal $f_2$-regulator. Thus, we have $B_1 = A \star f_1^*$ and $B = B_1 \star f_2^*$. It then follows from the associativity of $\star$ and Lemma 2.2(x) that

$$B = B_1 \star f_2^* = (A \star f_1^*) \star f_2^* = A \star (f_1^* \star f_2^*) = A \star (f_1 \star f_2)^*,$$

taking into account that both $f_1$ and $f_2$ are in $\mathcal{F}_0$. This shows putting the maximal $f_1$-regulator and the maximal $f_2$-regulator in tandem results in the maximal $f$-regulator. Since $f(t)$ is symmetric with respect to $f_1$ and $f_2$, the result is independent of the order of the concatenation. ∎

Now we consider feeding an $f_1$-upper constrained sequence to the maximal $f_2$-regulator.

**Theorem 3.6** *(Performance bounds) Let $A$ (resp. $B$) be the input (resp. output) of the maximal $f_2$-regulator. Let $q(t) = A(t) - B(t)$ be the queue length at the maximal $f_2$-regulator at time $t$, and $d(t) = \inf\{d \geq 0 : B(t+d) \geq A(t)\}$ be the virtual delay of the last customer that arrives at time $t$. Suppose that $A$ is $f_1$-upper constrained. Then*

**(i)** *$B$ is $f$-upper constrained, where $f = f_1 \star f_2$.*

**(ii)** *$q(t) \leq \max_{0 \leq s \leq t}[f_1^*(s) - f_2^*(s)]$ for all $t$.*

**(iii)** *$d(t) \leq \inf\{d \geq 0 : f_1^*(s) \leq f_2^*(s+d), s = 1, \ldots, t\}$.*

**Proof.** (i) Since $A$ is $f_1$-upper constrained, we have from Lemma 3.2(iii) that $A = A \star f_1^*$. As $B$ is the output from the maximal $f_2$-regulator,

$$B = A \star f_2^* = (A \star f_1^*) \star f_2^* = A \star (f_1^* \star f_2^*) = A \star (f_1 \star f_2)^* = A \star f^*,$$

where we use the associativity of $\star$ and Lemma 2.2(x) for $f_1, f_2 \in \mathcal{F}_0$. It then follows from Theorem 3.3(i) that $B$ is $f$-upper constrained.

(ii) Note from (12) that

$$q(t) = \max_{0 \leq s \leq t}[A(t) - A(s) - f_2^*(t-s)]. \tag{14}$$

Since we assume that $A$ is $f_1$-upper constrained (and thus $f_1^*$-upper constrained, i.e., $A(t) - A(s) \leq f_1^*(t-s)$), we have

$$q(t) \leq \max_{0 \leq s \leq t}[f_1^*(t-s) - f_2^*(t-s)] = \max_{0 \leq s \leq t}[f_1^*(s) - f_2^*(s)].$$

(iii) Observe that

$$A(t) - B(t+d) = \max_{0 \leq s \leq t+d}[A(t) - A(s) - f_2^*(t+d-s)].$$

As $A(t)$ is increasing and $f_2^*(t) \geq 0$, $A(t) - A(s) - f_2^*(t+d-s) \leq 0$ for $s > t$. Thus,

$$\begin{aligned} A(t) - B(t+d) \quad &\leq \max\Big[0, \max_{0 \leq s \leq t}[A(t) - A(s) - f_2^*(t+d-s)]\Big] && (15)\\ &\leq \max\Big[0, \max_{0 \leq s \leq t}[f_1^*(t-s) - f_2^*(t+d-s)]\Big] && (16)\\ &= \max\Big[0, \max_{0 \leq s \leq t}[f_1^*(s) - f_2^*(s+d)]\Big], && (17) \end{aligned}$$

where we use that $A$ is $f_1^*$-upper constrained in (16). Thus, $d(t) \leq \inf\{d \geq 0 : f_1^*(s) \leq f_2^*(s+d), s = 1, \ldots, t\}$. ■

Another important realization in filtering is the *filter bank summation* (see Fig. 2). Note that the summation in the $(\min, +)$-algebra is the min operator. As shown in Fig. 2, let $A$ be the input and $B_1$ (resp. $B_2$) be the output from the maximal $f_1$-regulator (resp. maximal $f_2$-regulator). The output from the "filter bank summation", denoted by $B$, is $B = B_1 \oplus B_2$, i.e.,

$$B(t) = \min[B_1(t), B_2(t)]. \tag{18}$$

**Theorem 3.7** *(Filter bank summation) A "filter bank summation" of the maximal $f_1$-regulator and the maximal $f_2$-regulator is the maximal $f$-regulator with*

$$f = f_1 \oplus f_2, \tag{19}$$

*if $f_1^* \oplus f_2^*$ is subadditive.*

**Proof.** Note from (19) and (12) that

$$B = B_1 \oplus B_2 = (A \star f_1^*) \oplus (A \star f_2^*) = A \star (f_1^* \oplus f_2^*),$$

where we use the distributive property. Since we assume that $f_1^* \oplus f_2^*$ is subadditive, it then follows from Lemma 2.2(iv) and (xi) that

$$B = A \star (f_1^* \oplus f_2^*)^* = A \star ((f_1^*)^* \star (f_2^*)^*) = A \star (f_1^* \star f_2^*) = A \star (f_1 \oplus f_2)^*. \tag{20}$$

■

Note from (20) in the proof of Theorem 3.7 that the realization using concatenation in Theorem 3.5 is equivalent to the realization using "filter bank summation" in Theorem 3.7 under the condition that $f_1^* \oplus f_2^*$ is subadditive.

**Lemma 3.8** *The condition that $f^* \oplus g^*$ is subadditive is satisfied if $f$ and $g$ are in the family of sequences*

$$\mathcal{F}_1 = \{f : f \in \mathcal{F}_0, \frac{f(s)}{s} \geq \frac{f(t)}{t}, \forall s \leq t\}. \tag{21}$$

**Proof.** Note that for all $f \in \mathcal{F}_1$, $-f$ is star-shaped, i.e., $-f(t)/t$ is increasing in $t$ and $f(0) = 0$ (see e.g., [18] pp. 453). Since a star-shaped function is superadditive, $f \in \mathcal{F}_1$ is subadditive. This can also be seen directly from

$$f(s) + f(t-s) \geq \frac{s}{t} f(t) + \frac{t-s}{t} f(t) = f(t), \ \ \forall s \leq t.$$

Thus, $f$ is subadditive (with $f(0) = 0$) and $f = f^*$ (Lemma 2.2(iv)) for all $f \in \mathcal{F}_1$.

It remains to verify that $f \oplus g$ is subadditive for $f, g \in \mathcal{F}_1$. By the definition in (21), it is easy to see that the operation $\oplus$ is closed in $\mathcal{F}_1$, i.e., for all $f$ and $g$ in $\mathcal{F}_1$, $f \oplus g \in \mathcal{F}_1$. As a result, $f$, $g$ and $f \oplus g$ are all subadditive when $f, g \in \mathcal{F}_1$. ■

**Corollary 3.9** *The condition that $f^* \oplus g^*$ is subadditive is satisfied if $f$ and $g$ are in the family of sequences*

$$\mathcal{F}_{icv} = \{f : f \in \mathcal{F}_0, f \textit{ is concave}\}. \tag{22}$$

**Proof.** It suffices to show that $\mathcal{F}_{icv}$ is a subclass of sequences in $\mathcal{F}_1$. To see this, note from concavity and $f(0) = 0$ that for $s \le t$

$$f(s) \ge \frac{t-s}{t} f(0) + \frac{s}{t} f(t) = \frac{s}{t} f(t).$$

■

We note that $\mathcal{F}_{icv}$ is a strict subclass of $\mathcal{F}_1$. To illustrate the difference between $\mathcal{F}_1$ and $\mathcal{F}_{icv}$, see Fig. 3, where we provide an example that is in $\mathcal{F}$, but not in $\mathcal{F}_{icv}$.

# 4 Realizations of maximal regulators under the $(\min, +)$-algebra

## 4.1 Leaky buckets

In this section, we show that leaky buckets are indeed maximal regulators and they can be realized by a systematic approach under the $(\min, +)$-algebra.

Consider a discrete-time leaky bucket with the token generation rate $\rho$ (the number of tokens generated per unit of time) and the size of the token buffer $\sigma$ (see e.g., [22]). Let $A$ be the input sequence to the leaky bucket. The buffer size for $A$ is assumed to be infinite. In the following theorem, we state that leaky buckets are maximal $f$-regulators. The proof of Theorem 4.1 is given in Appendix B.

**Theorem 4.1** *The discrete-time $(\sigma, \rho)$-leaky bucket is the maximal $f$-regulator, where $f(t) = \rho t + \sigma$, $t \ge 1$ and $f(0) = 0$. That is, the output $B$ satisfies*

$$B(t) = \min\Big[A(t), \min_{0 \le s \le t-1} [A(s) + \rho(t-s) + \sigma]\Big].$$

Note that the sequence $f$ defined in Theorem 4.1 is in $\mathcal{F}_{icv}$. Thus, it is subadditive (with $f(0) = 0$) and $f = f^*$. To gain more intuition on leaky buckets, we show that leaky buckets can be realized by the first order infinite duration impulse response (IIR) systems under the $(\min, +)$-algebra. To our best knowledge, such a realization appears to be new. Consider the following "linear system" under the $(\min, +)$-algebra.

$$x(t) = \min[\rho + x(t-1), A(t)], \quad x(0) = 0,$$
$$B(t) = \min[x(t) + \sigma, A(t)].$$

Recursive expansion yields $x(t) = \min_{0 \le s \le t}[A(s) + \rho(t-s)]$, and thus

$$B(t) = \min\Big[\min_{0 \le s \le t}[A(s) + \rho(t-s)] + \sigma, A(t)\Big],$$

which is exactly the output from the $(\sigma, \rho)$-leaky bucket. In Fig. 4, we show the "linear system" realization for the $(\sigma, \rho)$-leaky bucket.

Note that the function $f$ in Theorem 4.1 is not only subadditive, but also concave. Hence, according to Corollary 3.9, the condition in Theorem 3.7 is satisfied and one can implement "filter bank summation" for a concatenation of $(\sigma_i, \rho_i)$-leaky buckets, $i = 1, \ldots, K$. This yields the maximal $f$-regulator with

$$f = f_1 \oplus f_2 \oplus \ldots \oplus f_K = f_1 \star f_2 \star \ldots \star f_K,$$

where $f_i(t) = \rho_i t + \sigma_i$, $t \ge 1$ and $f_i(0) = 0$. Thus, $f(t) = \min_{1 \le i \le K}[\rho_i t + \sigma_i]$, $t \ge 1$ and $f(0) = 0$. This result was previously shown in [11, 1]. In fact, such a system can be realized by the following "linear system" under the $(\min, +)$-algebra (see Fig. 5):

$$x(t) = G \otimes x(t-1) \oplus O \otimes A(t), \quad x(0) = O,$$
$$B(t) = H \otimes x(t) \oplus A(t),$$

where $x(t)$ is the $K$ dimensional state vector, $G$ is the "diagonal" matrix with its $i^{th}$ diagonal element being $\rho_i$, $O$ is the $K$ dimensional column vector with all its elements being 0, and $H$ is the $K$ dimensional row vector with its $i^{th}$ element being $\sigma_i$. Here we use $\otimes$ and $\oplus$ to denote the matrix multiplication and addition under the $(\min, +)$-algebra. Once again, such a realization appears to be new. For example, if $K = 3$, then

$$G = \begin{bmatrix} \rho_1 & \infty & \infty \\ \infty & \rho_2 & \infty \\ \infty & \infty & \rho_3 \end{bmatrix}. \tag{23}$$

$$O = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{24}$$

$$H = [\sigma_1, \sigma_2, \sigma_3]. \tag{25}$$

There is another systematic way to implement the maximal $f$-regulator with

$$f(t) = \min_{1 \le i \le K} [\rho_i t + \sigma_i], t \ge 1 \text{ and } f(0) = 0.$$

For such a function, there exists a $t_0 < \infty$ such that $f(t) = \rho(t - t_0) + f(t_0)$ for all $t \ge t_0$ and $\rho = \min_{1 \le i \le K}[\rho_i]$ (see Fig. 6). Define the following two functions

$$f_1(t) = \rho(t - t_0) + f(t_0), \quad f_1(0) = 0,$$
$$f_2(t) = \begin{cases} f(t) & \text{if } t \le t_0, \\ \infty & \text{otherwise.} \end{cases}$$

Since $f(t)$ is concave, $f_1(t)$ (as the tangent of $f(t)$ at $t_0$) provides a upper bound of $f(t)$. Thus, $f(t) = \min[f_1(t), f_2(t)]$ and we may use the "filter bank summation" method to implement the maximal $f$-regulator. To be precise, we write

$$A \star f = A \star (f_1 \oplus f_2) = (A \star f_1) \oplus (A \star f_2). \tag{26}$$

Clearly, the term $A \star f_1$ in (26) is the maximal $f_1$-regulator and can be easily implemented by a $(f(t_0) - \rho t_0, \rho)$ leaky bucket in Fig. 4. However, the term $A \star f_2$ in (26) is not the maximal $f_2$-regulator, as $f_2$ is not subadditive. Since $f_2(t) = \infty$ for all $t \ge t_0$, the term $A \star f_2$ in (26) is in fact the output from a finite duration impulse response (FIR) systems under the $(\min, +)$-algebra, i.e.,

$$(A \star f_2)(t) = \min_{0 \le s \le t} [A(s) + f_2(t - s)] = \min_{0 \le s \le t} [A(t - s) + f_2(s)]$$
$$= \min_{0 \le s \le t_0} [A(t - s) + f_2(s)],$$

with the convention that $A(t) = \infty$ for all $t < 0$. This leads to the FIR-IIR realization in Fig. 7.

## 4.2 Traffic regulation for periodic constraint functions

Suppose that the input sequence $A \in \mathcal{F}_0$ is claimed to be $f$-upper constrained for a subadditive sequence $f \in \mathcal{F}_0$ that is periodic with period $p$, i.e., $f(t + p) - f(t)$ is a constant for all $t$. How does one find an efficient design for such a regulator? Note that such functions cannot be realized by leaky buckets due to periodicity.

Define the average rate $\rho = (f(t+p) - f(t))/p$. Thus, once $f(t)$, $t = 0, \ldots, p-1$, and $\rho$ are specified, the whole function $f$ is defined via the recursive equation

$$f(t+p) = f(t) + \rho p. \tag{27}$$

It then follows that

$$\begin{aligned}
B(t) &= \min_{0\le s\le t}[A(s) + f(t-s)] \\
&= \min\Big[\min_{0\le s\le t-p}[A(s) + f(t-s)], \min_{t-p<s\le t}[A(s) + f(t-s)]\Big] \\
&= \min\Big[\min_{0\le s\le t-p}[A(s) + f(t-p-s)] + \rho p, \min_{t-p<s\le t}[A(s) + f(t-s)]\Big] \\
&= \min\Big[B(t-p) + \rho p, A(t-p+1) + f(p-1), A(t-p+2) + f(p-2), \ldots, A(t) + f(0)\Big].
\end{aligned}$$

Such a regulator can be efficiently implemented by the FIR-IIR realization in Fig. 8. One possible application for such a regulator is traffic policing for video sequence, e.g. MPEG, where a certain periodic structure exists.

# 5 Service guarantees

In the section we discuss the connections between the maximal $f$-regulator and the concept of service curves developed by Cruz [12] and Sariowan [21].

**Definition 5.1** *($f$-server) A server guarantees service curve $f \in \mathcal{F}$ for an input sequence $A \in \mathcal{F}$ if its output sequence satisfies $B \ge A \star f$, i.e.,*

$$B(t) \ge \min_{0\le s\le t}[A(s) + f(t-s)] \tag{28}$$

*for all $t$. Such a server is called an $f$-server for $A$. If the inequality in (28) is satisfied for all input sequences, then we say the $f$-server is* universal.

We note the service curve guarantee defined in [21, 12] is universal. Clearly, the maximal $f$-regulator is a *universal* $f^*$-server (with equality). Moreover, Theorems 3.5, 3.6, and 3.7. can be rephrased as follows:

**Theorem 5.2** *(Concatenation) A concatenation of an $f_1$-server for $A$ and an $f_2$-server for the output from the $f_1$-server is an $f$-server for $A$, where $f = f_1 \star f_2$ is the convolution of $f_1$ and $f_2$ under the* $(\min, +)$*-algebra, i.e.,*

$$f(t) = \min_{0\le s\le t}[f_1(s) + f_2(t-s)]. \tag{29}$$

*Moreover, if both servers are* universal, *then the $f$-server is also* universal *and the result is independent of the order.*

**Theorem 5.3** *(Performance bounds) Consider an $f_2$-server for $A$. Let $B$ be the output, $q(t) = A(t) - B(t)$ be the queue length at the server at time $t$, and $d(t) = \inf\{d \geq 0 : B(t+d) \geq A(t)\}$ be the virtual delay of the last customer that arrives at time $t$. Suppose that $A$ is $f_1$-upper constrained. Then*

**(i)** *the server is also an $f$-server for $f$ satisfying*

$$f_2 \geq f_1^* \star f.$$

**(ii)** $q(t) \leq \max_{0 \leq s \leq t}[f_1^*(s) - f_2(s)]$ *for all $t$.*

**(iii)** $d(t) \leq \inf\{d \geq 0 : f_1^*(s) \leq f_2(s+d), s = 1, \ldots, t\}$.

**Theorem 5.4** *(Filter bank summation) Consider an input sequence $A$. Let $B_1$ (resp. $B_2$) be the output from an $f_1$-server (resp. $f_2$-server) for $A$. The output from the "filter bank summation", denoted by $B$, is $B_1 \oplus B_2$, i.e.,*

$$B(t) = \min[B_1(t), B_2(t)]. \tag{30}$$

*Then the "filter bank summation" of an $f_1$-server for $A$ and an $f_2$-server for $A$ is an $f$-server for $A$, where $f = f_1 \oplus f_2$, i.e.,*

$$f(t) = \min[f_1(t), f_2(t)]. \tag{31}$$

All the proofs of these three theorems are identical to those in Theorems 3.5, 3.6, and 3.7. We note that the result in Theorem 3.6(i) is in the form of "convolution" while the result in Theorem 5.3(i) is in the form of "deconvolution." Both of them are based on the fact that an $f_1$-upper constrained sequence passes through the maximal $f_1$-regulator without any change (Lemma 3.2(iii)). We also note that equivalent statements to Theorem 5.2 and Theorem 5.3(ii) and (iii) were stated in [21]. The convolution result in Theorem 5.2 was previously derived in [12] under a stronger definition of service curve.

Using the result in Theorem 2.4(iii), one has the following service guarantee for $f$-servers with feedback.

**Theorem 5.5** *(Feedback) Consider a sequence $A$ and an $f$-server for $A_1$, where $A_1 = A \oplus B$, and $B$ is the output from the $f$-server. If $f(0) > 0$ and $A(t) < \infty$ for all $t$, then the feedback system is an $f^* \star f$-server for $A$.*

**Proof.** One has from the definition of $f$-server that $B \geq A_1 \star f$. From the monotonicity of $\oplus$, it follows that

$$A_1 = A \oplus B \geq A \oplus (A_1 \star f).$$

Applying Theorem 2.4(iii) yields $A_1 \geq A \star f^*$. Thus, it follows from the monotonicity and the associativity of $\star$ that

$$B \geq A_1 \star f \geq (A \star f^*) \star f = A \star (f^* \star f).$$

This implies the feedback system is an $f \star f^*$-server. ■

**Example 5.6 (Window flow control)** Consider the window control problem in [2]. Let $A$ be the input to a network and $B$ be the output. Suppose that the network is a *universal* $f$-server and that the network also enforces a window flow control for the input $A$ with the window size $w > 0$. Thus, the effective input, denoted by $A_1$, satisfies

$$A_1(t) = \min[A(t), B(t) + w]. \tag{32}$$

Also, as the network is an $f$-server, we have

$$B \geq A_1 \star f. \tag{33}$$

We further assume that $B \leq A_1$ (a necessary condition for causality). This implies $f(0) = 0$. Otherwise, we would have from (33) that $B(0) \geq A_1(0) + f(0) > A_1(0)$. Also, assume that $A(t) < \infty$ for all finite $t$.

To carry out the analysis for such a scheme under our algebra, observe that $B(t) + w = (B \star \mathbf{I}_w)(t)$, where $\mathbf{I}_w$ is the function that $\mathbf{I}_w(t) = \infty$ for $t > 0$ and $\mathbf{I}_w(0) = w$. One may rewrite (32) as follows:

$$A_1 = A \oplus (B \star \mathbf{I}_w). \tag{34}$$

In conjunction with (33),

$$B \geq A_1 \star f = (A \oplus (B \star \mathbf{I}_w)) \star f = (A \star f) \oplus (B \star (\mathbf{I}_w \star f)),$$

where we apply the distributive property and the associativity of $\star$. Since we assume that $A(t) < \infty$ for all $t$, we then have from Theorem 2.4(iii) that

$$B \geq A \star f \star (\mathbf{I}_w \star f)^*,$$

as $(\mathbf{I}_w \star f)(0) = w + f(0) = w > 0$. Thus, the window flow control system is an $f \star (\mathbf{I}_w \star f)^*$-server. We note that Cruz and Okino [15] considered a more detailed model than the one presented here. A tighter bound for the service guarantee of the window control problem was derived there. See also Le Boudec and Thiran [5].

Note from the monotonicity of $\star$ that $f \star (\mathbf{I}_w \star f)^* \le f$. This is expected as adding window flow control degrades the service guarantee. Also, from the monotonicity of $\star$ and $*$, $f \star (\mathbf{I}_w \star f)^*$ is increasing in the window size $w$. One might expect that there is a window size $w$ large enough such that

$$f \star (\mathbf{I}_w \star f)^* = f \tag{35}$$

and the service guarantee of window flow control scheme is the same as that of the open loop. We claim the minimum window size is

$$w_{min} = \sup_{t \ge 0} \sup_{s \ge 0} [f(t+s) - f(s) - f(t)]. \tag{36}$$

This means that if $f$ is subadditive, then the window size $w$ could be made arbitrarily small (we still need $w > 0$ for stability). To see the claim, let $\hat{f}(t) = \sup_{s \ge 0}[f(t+s) - f(s)]$ be the minimum envelope in [6]. As discussed in Section 2.2, $\hat{f}$ is the minimum solution of $f = f \star h$. Moreover, $\hat{f}$ is subadditive and $(\hat{f})^* = \hat{f}$ from Lemma 2.2(iv). Thus, in order to have (35), one must have $(\mathbf{I}_w \star f)^* \ge \hat{f}$. It follows from the monotonicity of subadditive closure that $\mathbf{I}_w \star f \ge (\mathbf{I}_w \star f)^* \ge \hat{f}$. The condition is equivalent to

$$f(t) + w \ge \hat{f}(t) \quad \forall t$$

and this gives $w_{min}$ in (36). On the other hand, for any $w \ge w_{min}$, we have from the monotonicity of $\star$ and $*$ that

$$f \star (\mathbf{I}_w \star f)^* \ge f \star (\hat{f})^* = f \star \hat{f} = f.$$

Thus, $f \star (\mathbf{I}_w \star f)^* = f$.

Though the concept of service curves is originally developed for providing bounded end-to-end delay [12, 21], it is of the same importance to observe that a server that provides bounded delay also guarantees a service curve. In the following lemma, we show that the maximum delay guarantee is equivalent to the service curve guarantee.

**Lemma 5.7** *(Maximum delay guarantee) A server guarantees maximum delay $d$ for an input sequence $A$ if and only if it is an $\mathbf{O}_d$-server for $A$, where*

$$\mathbf{O}_d(t) = \begin{cases} 0 & \text{if } 0 \le t \le d \\ \infty & \text{otherwise} \end{cases}. \tag{37}$$

**Proof.** Let $B$ be the output from the server. Note that the server guarantees maximum delay $d$ for an input sequence $A$ if and only if

$$B(t) \geq A((t-d)^+) = \min_{0 \leq s \leq t} [A(s) + \mathbf{O}_d(t-s)]$$

with $\mathbf{O}_d$ defined in (37). ■

The following result strengthens the one-sided result for service curve specification in [21].

**Corollary 5.8** *Suppose that the input sequence $A$ is $f$-upper constrained. Then a server guarantees maximum delay $d$ for the sequence $A$ if and only if it is an $f^* \star \mathbf{O}_d$-server for $A$.*

**Proof.** To see the *only if* part, one has from Lemma 5.7 that a server that guarantees maximum delay $d$ is an $\mathbf{O}_d$-server. From the monotonicity of $\star$, it follows that $\mathbf{O}_d \geq f^* \star \mathbf{O}_d$. Thus, the server is also an $f^* \star \mathbf{O}_d$-server.

For the *if* part, we have from Theorem 5.3(i) that the server is also an $\mathbf{O}_d$-server if the server is an $f^* \star \mathbf{O}_d$-server. From Lemma 5.7, the server then guarantees maximum delay $d$. We note that another way to prove this is to apply Theorem 5.3(iii).

■

Functions of the form $f^* \star \mathbf{O}_d$ are called shifted subadditive in [21] as $f^*$ is subadditive. From Corollary 5.8 and Theorem 5.2, one may view an $f^* \star \mathbf{O}_d$-server as a concatenation of an $f^*$-server and an $\mathbf{O}_d$-server. This implies an $f^* \star \mathbf{O}_d$-server can be achieved by a concatenation of the maximal $f$-regulator and a server that guarantees maximum delay $d$. Since the output from the maximal $f$-regulator is $f$-upper constrained, bounded delay can be accomplished by proper admission control for simple scheduling policies such as the First Come First Served (FCFS) policy and the static priority policy, or more complicated deadline scheduling polices such as the Generalized Processor Sharing (GPS) and the Earliest Deadline First policy (EDF) (see [10, 6, 19, 23]). We note that deadline scheduling policies usually provide a larger admission set, but they are at the cost of design complexity, e.g. time stamping and sorting.

Corollary 5.8 shows that the best service curve allocation for a server to yield bounded delay is shifted-subadditive (if the input is $f$-upper constrained). In a network environment, a server is often a concatenation of servers. The question is then how one allocates minimum service curves among a concatenation of servers so that the convolution of these service curves guarantees a shifted-subadditive service curve.

Without loss of generality, consider the case with a concatenation of two servers. In the following, we provide a couple of rules of thumbs for service curve allocation.

**(i)** Suppose we restrict ourselves to the problem that only shifted-subadditive service curves in these two servers are allocated. For instance, assume we allocate $h_1 = g_1^* \star \mathbf{O}_{d_1}$ for the first server and $h_2 = g_2^* \star \mathbf{O}_{d_2}$ for the second server. Then a better way is to allocate the first server $\tilde{h}_1 = g_1^* \star g_2^* \star \mathbf{O}_{d_1}$ and the second server $\tilde{h}_2 = g_1^* \star g_2^* \star \mathbf{O}_{d_2}$, To see this, note that the overall service curves for the original allocation and the modified allocation are the same since

$$h_1 \star h_2 = g_1^* \star \mathbf{O}_{d_1} \star g_2^* \star \mathbf{O}_{d_2} = (g_1^* \star g_2^* \star \mathbf{O}_{d_1}) \star (g_1^* \star g_2^* \star \mathbf{O}_{d_2}) = \tilde{h}_1 \star \tilde{h}_2.$$

From the monotonicity of $\star$, $g_1^* \star g_2^* \leq g_1^*$ and $g_1^* \star g_2^* \leq g_2^*$. Thus, $\tilde{h}_i \leq h_i$, $i = 1$ and 2.

**(ii)** Suppose we restrict ourselves to the problem that only shifted-concave service curves in these two servers are allocated. As the definition for shifted-subadditive functions, we call a function $h$ shifted-concave if it is of the form $h = \tilde{h} \star \mathbf{O}_d$ for some $\tilde{h} \in \mathcal{F}_{icv}$ and some $d \geq 0$. Thus, the class of shifted-concave functions is a subset of shifted-subadditive functions. Assume our goal is to guarantee the overall service curve $f = g \star \mathbf{O}_d$ for some function $g \in \mathcal{F}_0$ and $d \geq 0$. (Without loss of generality, assume $g(1) > 0$. Otherwise, we may increase $d$.) Then the best allocation is in the following form: $\tilde{h}_1 = \tilde{g} \star \mathbf{O}_{d_1}$ and $\tilde{h}_2 = \tilde{g} \star \mathbf{O}_{d_2}$, where $d_1 + d_2 \leq d$ and

$$\tilde{g}(t) = \min_{z \in I\!R^+ \cup \{0\}} \Big[ zt - \min_{\tau \geq 0} [z\tau - (g \star \mathbf{O}_{d-d_1-d_2})(\tau)] \Big]. \tag{38}$$

Note that $\tilde{g}$ is known as the concave envelope of $g \star \mathbf{O}_{d-d_1-d_2}$ and $\tilde{g}$ is the smallest concave function that is not less than $g \star \mathbf{O}_{d-d_1-d_2}$ (see e.g. [20]). We first show such an allocation guarantees the overall service curve $g \star \mathbf{O}_d$. Note that the overall service curve is

$$\tilde{g} \star \mathbf{O}_{d_1} \star \tilde{g} \star \mathbf{O}_{d_2} = \tilde{g} \star \mathbf{O}_{d_1+d_2}.$$

Thus, the overall service curve is guaranteed as $\tilde{g} \geq g \star \mathbf{O}_{d-d_1-d_2}$. This inequality can also be seen directly from (38) by taking $\tau = t$. Now we argue such an allocation is optimal. From (i), it suffices to consider allocations of the following form: $h_1 = g_1 \star \mathbf{O}_{d_1}$ and $h_2 = g_1 \star \mathbf{O}_{d_2}$ for some $g_1 \in \mathcal{F}_{icv}$. Since the overall service curve $g \star \mathbf{O}_d$ is guaranteed, one has $g_1 \star \mathbf{O}_{d_1+d_2} \geq g \star \mathbf{O}_d$. Thus,

$$g_1((d+1-d_1-d_2)^+) = (g_1 \star \mathbf{O}_{d_1+d_2})(d+1) \geq (g \star \mathbf{O}_d)(d+1) = g(1).$$

As we assume $g(1) > 0$ and $g_1 \in \mathcal{F}_{icv}$ ($g_1(0) = 0$), one must have $d_1 + d_2 \leq d$. Recall that $\tilde{g}$ is the smallest concave function that is not less than $g \star \mathbf{O}_{d-d_1-d_2}$. This implies $g_1 \geq \tilde{g}$. Thus, $h_i \geq \tilde{h}_i$, $i = 1$ and 2.

# 6 Conclusions and future research

In this paper, we developed in Section 2.2 the concept and the properties of subadditive closure. Based on the properties developed for subadditive closures, in Section 3 we developed the filtering theory for the maximal $f$-regulator, including feedback, concatenation, "filter bank summation", and performance bounds. We discussed in Section 4 how one realizes maximal regulators, including leaky buckets, under the $(\min, +)$-algebra via systematic methods such as concatenation, filter bank summation, linear system realization, and FIR-IIR realization. In Section 5, we discussed the analogy between maximal $f$-regulators and $f$-servers. We showed that the filtering theory for $f$-servers is still applicable.

In the following, we discuss possible extensions of the filtering theory.

**(i)** Matrix operations: one may consider square $n \times n$ matrices with entries in $\mathcal{F}$. The addition and multiplication of matrices are defined conventionally after the "addition"operator $\oplus$ and the "multiplication" operator $\star$ in $\mathcal{F}$. Use $\oplus$ and $\star$ to denote the "matrix addition" operator and the "matrix multiplication" operator. Then $(\mathcal{F}^{n\times n}, \oplus, \star)$ is still a dioid with the zero matrix $\underline{\epsilon}$ and the identity matrix $\underline{\mathbf{e}}$, where $\underline{\epsilon}$ has all its entries equal to $\epsilon$, and $\underline{\mathbf{e}}$ has its diagonal entries equal to $\mathbf{e}$ and all other entries equal to $\epsilon$. Such an extension could be used for modelling filtering with multiple inputs and outputs, e.g., nested window flow control in [2]. Results along this line is reported in [9].

**(ii)** Transfer functions under the $(\min, +)$-algebra: one may define the so called $z$-transform under the $(\min, +)$-algebra by $F(z) = \min_t[f(t) + zt]$ (see [3]). This definition corresponds to the Legendre transform (the convex transform) [20] in the continuous time setting.

**(iii)** Time varying filtering: time varying filtering is related to stochastic extension of the deterministic filtering theory developed here. It is expected such an extension would be closely related to the exponential transformation from the $(\min, +)$-algebra to the usual algebra (see e.g. [7] and references therein). Cruz [14] also noted that time varying filtering could be used for transient analysis of service guarantees.

**(iv)** Constrained filtering: In [17], Ananthram and Konstantopoulos considered traffic regulation problems with buffer and delay constraints. Though they focused on the $(\sigma, \rho)$ regulators, the technique developed there might be applied to the general setting discussed in this paper.

**Acknowledgements:**

The author would like to thank Rene Cruz for many helpful comments on this paper. The author would also like to thank the referees for their careful comments which enhanced the presentation of this work.

# A Appendix A

In this section, we prove Lemma 2.2.

(i) That $f^*$ is increasing and subadditive is shown in Chang [6], Lemma 2.1. Thus, $f^* \in \mathcal{F}_0$.

(ii) Same as (i).

(iii) By the construction of $f^*$, one has clearly $f(t) \geq f^*(t)$ for all $t$.

(iv) For the *only if* part, we have $f(0) = f^*(0) = 0$. From definition of $f^*(t)$, it follows that

$$\begin{aligned} f(t) = f^*(t) &= \min[f(t), \min_{0<s<t}[f^*(s) + f^*(t-s)]] \\ &= \min[f(t), \min_{0<s<t}[f(s) + f(t-s)]] = \min_{0\leq s\leq t}[f(s) + f(t-s)]. \end{aligned}$$

Thus, $f(t) \leq f(s) + f(t-s)$ for all $0 \leq s \leq t$. This shows that $f$ is subadditive.

We prove the *if* part by induction on $t$. Clearly, it holds for $t = 0$ and $t = 1$ as $f^*(0) = f(0) = 0$ and $f^*(1) = f(1)$. Suppose that $f(s) = f^*(s)$ for $s = 0, \ldots, t-1$. By the definition of

$f^*(t)$ and the subadditivity of $f$, we have

$$\begin{aligned}
f^*(t) &= \min[f(t), \min_{0<s<t}[f^*(s) + f^*(t-s)]] \\
&= \min[f(t), \min_{0<s<t}[f(s) + f(t-s)]] = \min_{0\le s\le t}[f(s) + f(t-s)] \\
&= f(t).
\end{aligned}$$

(v) For the *only if* part, note that $(f \star f)(0) = f(0) + f(0)$. Thus, $f \star f = f$ implies that $f(0) = 0$. Moreover,

$$f(t) = (f \star f)(t) = \min_{0\le s\le t}[f(s) + f(t-s)]. \tag{39}$$

This implies that $f(t) \le f(s) + f(t-s)$ for all $0 \le s \le t$. Thus, $f$ is subadditive.

The *if* part follows from (39) trivially.

(vi) From (iii), the monotonicity of $\star$, (ii) and (v), it follows that

$$(f^\circ)^{(2)} = f^\circ \star f^\circ \ge f^* \star f^* = f^*.$$

By induction, one has $(f^\circ)^{(n)} \ge f^*$ for all $n$.

We use induction to show that $(f^\circ)^{(t)}(t) = f^*(t)$. Clearly, it holds for $t = 1$ as $(f^\circ)^{(1)}(1) = f(1) = f^*(1)$. Now suppose that $(f^\circ)^{(s)}(s) = f^*(s)$ for $s \le t-1$. By the definition of the subadditive closure and the induction hypothesis,

$$\begin{aligned}
f^*(t) &= \min[f(t), \min_{0<s<t}[f^*(s) + f^*(t-s)]] \\
&= \min[f(t), \min_{0<s<t}[(f^\circ)^{(s)}(s) + (f^\circ)^{(t-s)}(t-s)]] \\
&\ge \min\Big[f(t), \min_{0<s<t}[\min_{0\le\tau\le t}[(f^\circ)^{(s)}(\tau) + (f^\circ)^{(t-s)}(t-\tau)]]\Big] \\
&= \min\Big[f(t), \min_{0<s<t}[((f^\circ)^{(s)} \star (f^\circ)^{(t-s)})(t)]\Big] \\
&= \min\Big[f(t), \min_{0<s<t}[(f^\circ)^{(t)}(t)]\Big] \\
&= \min[f(t), (f^\circ)^{(t)}(t)].
\end{aligned}$$

Since $f \ge f^\circ = (f^\circ)^{(1)} \ge (f^\circ)^{(t)}$, $\min[f(t), (f^\circ)^{(t)}(t)] = (f^\circ)^{(t)}(t)$. Thus, $f^*(t) \ge (f^\circ)^{(t)}(t)$. Since $f^* \le (f^\circ)^{(t)}$, we then have $f^*(t) = (f^\circ)^{(t)}(t)$.

(vii) one has from the monotonicity of $\star$ and the monotonicity of $\circ$ that

$$(f^* \star f)^\circ \ge (f^* \star f^*)^\circ = (f^*)^\circ = f^*.$$

Also, from the monotonicity of $\star$ and $\oplus$,

$$\begin{aligned}
&(f^* \star f)^\circ = (f^* \star f) \oplus \mathbf{e} \\
&\le ((f^\circ)^{(n)} \star f) \oplus \mathbf{e} = ((\mathbf{e} \oplus f \oplus f^{(2)} \oplus \ldots \oplus f^{(n)}) \star f) \oplus \mathbf{e} \\
&= (f \oplus f^{(2)} \oplus \ldots \oplus f^{(n+1)}) \oplus \mathbf{e} = (f^\circ)^{(n+1)}.
\end{aligned}$$

In view of (vi), letting $n \to \infty$ yields $(f^* \star f)^\circ \le f^*$. Thus, $f^*$ is a solution of $h = (h \star f)^\circ$. To see that $f^*$ is the maximum solution, we iterate the equation

$$
\begin{aligned}
& h = (h \star f)^\circ = \mathbf{e} \oplus (f \star h) \\
& = (\mathbf{e} \oplus f \oplus f^{(2)} \oplus \ldots \oplus f^{(n)}) \oplus (h \star f^{(n+1)}) = (f^\circ)^{(n)} \oplus (h \star f^{(n+1)}).
\end{aligned}
$$

From the monotonicity of $\oplus$, we have $h \le (f^\circ)^{(n)}$. Letting $n \to \infty$ yields $h \le f^*$.

(viii) We prove this by induction on $t$. Clearly, it holds for $t = 1$ as $f^*(1) = f(1) \le g(1) = g^*(1)$. Suppose that $f^*(s) \le g^*(s)$ for $s = 0, \ldots, t-1$. By the definition of the subadditive closure, we have from the induction hypothesis and the assumption $f \le g$ that

$$
\begin{aligned}
& f^*(t) = \min[f(t), \min_{0<s<t}[f^*(s) + f^*(t-s)]] \\
& \le \min[g(t), \min_{0<s<t}[g^*(s) + g^*(t-s)]] \\
& = g^*(t).
\end{aligned}
$$

(ix) Let $h = f^\circ \star g^\circ$, i.e.,

$$h(t) = \min_{0 \le s \le t}[f^\circ(s) + g^\circ(t-s)]. \tag{40}$$

Clearly, $h(0) = f^\circ(0) + g^\circ(0) = 0$. It remains to show that for all $s, t \ge 0$,

$$h(s) + h(t) \ge h(t+s).$$

Let $s^*$ (resp. $t^*$) be the argument that achieves the minimum in (40) for $h(s)$ (resp. $h(t)$). Thus, $h(s) = f^\circ(s^*) + g^\circ(s - s^*)$ and $h(t) = f^\circ(t^*) + g^\circ(t - t^*)$. This implies

$$
\begin{aligned}
& h(s) + h(t) = f^\circ(s^*) + f^\circ(t^*) + g^\circ(s - s^*) + g^\circ(t - t^*) \\
& \ge f^\circ(s^* + t^*) + g^\circ(s + t - s^* - t^*) \\
& \ge \min_{0 \le \tau \le t+s}[f^\circ(\tau) + g^\circ(t + s - \tau)] = h(t+s),
\end{aligned} \tag{41}
$$

where we apply subadditivity for $f^\circ$ and $g^\circ$ in the first inequality.

(x) From the monotonicity of $\star$ and $*$, clearly we have $(f \star g)^* \ge (f^\circ \star g^\circ)^*$. Now we show that $(f^\circ \star g^\circ)^* = f^* \star g^*$. From (iii), we know that $f^\circ \ge f^*$ and $g^\circ \ge g^*$. Thus, from the monotonicity of $\star$, $f^\circ \star g^\circ \ge f^* \star g^*$. As a result of and (viii) and (ix),

$$(f^\circ \star g^\circ)^* \ge (f^* \star g^*)^* = f^* \star g^*.$$

On the other hand, one has from the monotonicity of $\star$ that $f^\circ \ge f^\circ \star g^\circ$ and $g^\circ \ge f^\circ \star g^\circ$. From (viii), it follows that $f^* = (f^\circ)^* \ge (f^\circ \star g^\circ)^*$ and $g^* = (g^\circ)^* \ge (f^\circ \star g^\circ)^*$. Thus, using the monotonicity of $\star$ and (v) yields

$$f^* \star g^* \ge (f^\circ \star g^\circ)^* \star (f^\circ \star g^\circ)^* = (f^\circ \star g^\circ)^*.$$

(xi) Note that $f \oplus g \ge f^\circ \oplus g^\circ \ge f^\circ \star g^\circ$. Thus, from (viii) and (x), we have

$$(f \oplus g)^* \ge (f^\circ \star g^\circ)^* = f^* \star g^*.$$

On the other hand, one has from the monotonicity of $\oplus$ that $f \geq f \oplus g$ and $g \geq f \oplus g$. From (viii), it follows that $f^* \geq (f \oplus g)^*$ and $g^* \geq (f \oplus g)^*$. Thus, using the monotonicity of $\star$ and (v) yields

$$f^* \star g^* \geq (f \oplus g)^* \star (f \oplus g)^* = (f \oplus g)^*.$$

(xii) From the monotonicity of $\oplus$, $h \leq f^\circ$. Thus, $h(0) = 0$. Since $h$ is subadditive (with $h(0) = 0$) and $h = h \oplus f^\circ$, we have from (iv) and (xi) that

$$h = h^* = (h \oplus f^\circ)^* = (h^\circ \oplus f^\circ)^* = h^* \star f^* = h \star f^*.$$

The result then follows from the maximum solution result in (vii).

# B Appendix B

In this section, we prove Theorem 4.1.

Let $a(t) = A(t) - A(t-1)$ be the number of cells (or packets) that arrive at time $t$. Denote by $q_1(t)$ (resp. $q_2(t)$) the number of cells in the cell buffer (the number of tokens in the token buffer) at time $t$. Assume $q_1(0) = 0$ and $q_2(0) = \sigma$. Let $z(t) = q_1(t) - q_2(t)$. We first show that $z(t)$ satisfies the following recursive equation:

$$z(t+1) = \max[z(t) + a(t+1) - \rho, -\sigma]. \tag{42}$$

This equation can be verified by considering the following three cases: (i) $z(t) > 0$, (ii) $z(t) < 0$ and $z(t) = 0$.

**Case 1.** $z(t) > 0$:

In this case, one has $z(t) = q_1(t) > 0$ and $q_2(t) = 0$ as there are only cells in the buffer at time $t$. The maximum number of departures is bounded by the number of tokens that arrive at time $t$, i.e., $\rho$. If $q_1(t) + a(t+1) \geq \rho$, then $q_1(t+1) = q_1(t) + a(t+1) - \rho$ and $q_2(t+1) = 0$. Thus, (42) holds. On the other hand, if $q_1(t) + a(t+1) < \rho$, then $q_1(t+1) = 0$ and there will be tokens left in the token buffer. Since the maximum number of tokens that can be stored in the token buffer is $\sigma$, $q_2(t+1) = \min[\rho - q_1(t) - a(t+1), \sigma]$. Once again, (42) holds.

**Case 2.** $z(t) < 0$:

In this case, one has $q_1(t) = 0$ and $q_2(t) = -z(t) > 0$ as there are only tokens in the token buffer at time $t$. The maximum number of departures is bounded by the sum of the number of tokens left at time $t$ and the number of tokens that arrives at time $t$, i.e., $q_2(t) + \rho$. If $a(t+1) \geq q_2(t) + \rho$, then $q_1(t+1) = a(t+1) - q_2(t) - \rho$ and $q_2(t+1) = 0$. Thus, (42) holds. On the other hand, if $a(t+1) < q_2(t) + \rho$, then $q_1(t+1) = 0$ and there will be tokens left in the token buffer. Since the maximum number of tokens that can be stored in the token buffer is $\sigma$, $q_2(t+1) = \min[\rho + q_2(t) - a(t+1), \sigma]$. Once again, (42) holds.

**Case 3.** $z(t) = 0$:

In this case, one has $q_1(t) = 0$ and $q_2(t) = 0$ as both buffers are empty at time $t$. The maximum number of departures is bounded by $\rho$. If $a(t+1) \geq \rho$, then $q_1(t+1) = a(t+1) - \rho$ and $q_2(t+1) = 0$. Thus, (42) holds. On the other hand, if $a(t+1) < \rho$, then $q_1(t+1) = 0$ and $q_2(t+1) = \min[\rho - a(t+1), \sigma]$. Once again, (42) holds.

Note that $q_1(t) = \max[z(t), 0]$ from the three cases. Let $B$ be the output from the $(\sigma, \rho)$-leaky bucket. Then

$$B(t) = A(t) - q_1(t) = A(t) - \max[z(t), 0] = \min[A(t), A(t) - z(t)]. \tag{43}$$

Let

$$y(t) = z(t) + \sigma. \tag{44}$$

It then follows from (42) that $y(t)$ satisfies Lindley's recursion

$$y(t+1) = \max[y(t) + a(t+1) - \rho, 0], \quad y(0) = 0.$$

Thus,

$$y(t) = \max_{0 \le s \le t}[A(t) - A(s) - \rho(t-s)].$$

In conjunction with (43) and (44), one has

$$B(t) = \min\left[A(t), \min_{0 \le s \le t}[A(s) + \rho(t-s)] + \sigma\right] \tag{45}$$

$$= \min\left[A(t), \min_{0 \le s \le t-1}[A(s) + \rho(t-s) + \sigma]\right], \tag{46}$$

as $\sigma \ge 0$. Clearly, the function $f(t) = \rho t + \sigma$, $t \ge 1$ and $f(0) = 0$ is increasing and subadditive. This completes the argument.

## References


[1] V. Anantharam and T. Konstantopoulos. "Optimality and interchangeability of leaky buckets in tandem," Univ. of Texas at Austin, Tech. Report SCC-95-01, 1995. Initial version also in: Proc. 32nd Allerton Conference, pp. 235-244, Monticello, IL, October 1994.

[2] R. Agrawal and R. Rajan, "Performance bounds for guaranteed and adaptive services," *IBM RC 20649*, 1996. Also in *Proc. 34th Allerton Conf. on Comm., Cont. & Comp.*, Monticello, IL, Oct. 1996.

[3] F. Baccelli, G. Cohen, G. Oslder, and J. Quadrat, *Synchronization and Linearity: An Algebra for Discrete Event Systems,* Wiley, 1992.

[4] J.Y. Le Boudec, "Network calculus made easy," *preprint.*

[5] J.Y. Le Boudec and P. Thiran, "A note on time and space methods in network calculus," *preprint*, 1997.

[6] C.S. Chang, "Stability, queue length and delay of deterministic and stochastic queueing networks," *IEEE Transactions on Automatic Control*, Vol. 39, pp. 913-931, 1994.

[7] C.S. Chang, "On the exponentiality of stochastic linear systems under the max-plus algebra," *IEEE Transactions Automatic Control*, Vol. 41, pp. 1182-1188, 1996.

[8] C.S. Chang, "A filtering theory for deterministic traffic regulation," *IEEE INFOCOM'97*, Kobe, Japan.

[9] C.S. Chang, "Matrix extensions of the filtering theory for deterministic traffic regulation and service guarantees," preprint, 1997.

[10] R.L. Cruz, "A calculus for network delay, Part I: Network elements in isolation," *IEEE Trans. Inform. Theory*, Vol. 37, pp. 114-131, 1991.

[11] R.L. Cruz, "A calculus for network delay, Part II: Network analysis," *IEEE Trans. Inform. Theory*, Vol. 37, pp. 132-141, 1991.

[12] R.L. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE J. Selected Areas in Communications*, Vol. 13, pp. 1048-1056, 1995.

[13] R.L. Cruz, "SCED+: efficient management of quality of service guarantees," *preprint*, 1997.

[14] R.L. Cruz, *personal communication*, 1997.

[15] R.L. Cruz and C.M. Okino, "Service guarantees for window flow control," *Proc. 34th Allerton Conf. on Comm., Cont. & Comp.*, Monticello, IL, Oct. 1996.

[16] J.F.C. Kingman, *Subadditive processes* in Ecole d'ete de probabilities de Saint-Flour, edited by A. Dold and B. Eckmann, Lecture Notes in Mathematics, 539, Springer-Verlag, Berlin, pp. 165-223, 1976.

[17] T. Konstantopoulos and V. Anantharam. "Optimal flow control schemes that regulate the burstiness of traffic," IEEE/ACM Trans. Networking, Vol. 3, pp. 423-432, 1995.

[18] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications.* New York: Academic Press, 1979.

[19] A.K. Parekh and R.G. Gallager, "A generalized processor sharing approach to flow control in integrated service networks: the single-node case," *IEEE/ACM Trans. Networking*, Vol. 1, pp. 344-357, 1993.

[20] R.T. Rockafellar. *Convex Analysis.* Princeton: Princeton University Press, 1970.

[21] H. Sariowan, "A Service-Curve Approach to Performance Guarantees in Integrated-Service Networks," Ph.D. Dissertation, University of California, San Diego, 1996.

[22] J.S. Turner, "New Directions in Communications (or Which Way to the Information Age?)," *IEEE Communication Magazine*, Vol. 24, No. 10, Oct. 1986.

[23] H. Zhang, "Service displines for guaranteed performance service in packet-switching networks," *Proceedings of the IEEE*, Vol. 83, pp. 1374-1396, 1995.