SRC TR 87-206r1

Bounding Functions Of Markov Processes And The Shortest Queue Problem

by

J.A. Gubner, B. Gopinath and S.R.S. Varadhan

BOUNDING FUNCTIONS OF MARKOV PROCESSES AND THE SHORTEST QUEUE PROBLEM

JOHN A. GUBNER,* *University of Maryland*
B. GOPINATH,** *Bell Communications Research*
S. R. S. VARADHAN,*** *Courant Institute of Mathematical Sciences*

Abstract

We prove a theorem which can be used to show that the expectation of a nonnegative function of the state of a time-homogeneous Markov process is uniformly bounded in time. This is reminiscent of the classical theory of nonnegative supermartingales, except that our analog of the supermartingale inequality need not hold almost surely. Consequently, the theorem is suitable for establishing the stability of systems that evolve in a stabilizing mode in most states, though from certain states they may jump to a less stable state. We use this theorem to show that "joining the shortest queue" can bound the expected sum of the squares of the differences between all pairs among N queues, *even under arbitrarily heavy traffic.*

STABILITY; LYAPUNOV FUNCTION; SUPERMARTINGALE; LOAD BALANCING; PARALLEL QUEUES

---

*Present address: Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706, U.S.A.

**Postal address: Bell Communications Research, 435 South Street, Morristown, NJ 07960, U.S.A.

***Postal address: Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, U.S.A.

This work was supported by Bell Communications Research. The research of the first author was also supported in part by the University of Maryland Systems Research Center under NSF grant OIR-85-00108.

1. Introduction

Controlled Markov processes often arise when modeling the behavior of asynchronous algorithms, queuing systems, and computer and communication networks. In these models it is often the case that the control policy considers only the immediate future of the process in an attempt to stabilize it. While in general, a myopic policy will not guarantee the long-term stability of the process, in this paper we shall give sufficient conditions on the one-step behavior of a process to guarantee its long-term stability.

Suppose we are given a time-homogenous Markov process, $\{X_n\}$. Consider an analysis in which we choose a nonnegative function u such that if the expected value of $u(X_n)$ is suitably bounded, then the process is stable in some sense. For example, in Section 3 our process will be the state of a system of N parallel queues. Our control policy will direct the traffic arriving at time $n+1$ to the queue whose length was shortest at time n. We will show that this policy can stabilize the system in the sense that the expected sum of the squares of the differences between all pairs among the N queues is uniformly bounded for all time.

To motivate our results, consider the following example of a single-server queue. Let $\{Y_n, n = 0, 1, 2, \ldots\}$ be a discrete-time Markov chain whose transition probabilities are as follows. For $i = 0$,

$$\mathsf{P}(Y_{n+1} = j \mid Y_n = i) \;=\; \begin{cases} \lambda, & \text{if } j = 1, \\ 1-\lambda, & \text{if } j = 0. \end{cases}$$

For $i \geq 1$,

$$\mathsf{P}(Y_{n+1} = j \mid Y_n = i) \;=\; \begin{cases} \lambda, & \text{if } j = i+1, \\ 1-(\lambda+\mu), & \text{if } j = i, \\ \mu, & \text{if } j = i-1. \end{cases}$$

We take $0 \leq \lambda < \mu \leq 1$ with $\lambda + \mu \leq 1$. Our concern here is the long-term behavior of $\mathsf{E}[\,Y_n \mid Y_0 = i\,]$. In particular, we would like to know if

$$\varlimsup_{n\to\infty} \mathsf{E}[\,Y_n \mid Y_0 = i\,] \;<\; \infty.$$

In this simple example, since $\lambda < \mu$, it is very easy to compute the stationary distribution (see Billingsley (1979), Theorem 8.6, p. 106),

$$\lim_{n\to\infty} \mathsf{P}(Y_n = j \mid Y_0 = i) \;=\; \left(\frac{\lambda}{\mu}\right)^j \left(1 - \frac{\lambda}{\mu}\right).$$

Recalling that for $|z| < 1$, $\sum_{j=1}^{\infty} j z^{j-1} = (1-z)^{-2}$, it is not hard to compute that indeed,

$$\lim_{n\to\infty} \mathsf{E}[\, Y_n \mid Y_0 = i\,] \;=\; \frac{\lambda}{\mu - \lambda} \;<\; \infty,$$

assuming that we can take the limit inside the expectation.

How could we study the behavior of $\mathsf{E}[\, Y_n \mid Y_0 = i\,]$ if the stationary distribution were not available to us? Recall that in the field of differential equations, a "Lyapunov function" can sometimes be used to study the behavior of solutions to a differential equation without actually solving it. In this spirit, our paper generalizes the preceding example, and gives conditions on the one-step behavior of a process that will yield useful information about its long-run (and short-run) behavior.

We now make a precise statement of our model. Let $\{X_n, n = 0, 1, 2, \ldots\}$ be a time-homogeneous Markov process defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, and taking values in a measurable space $(\mathsf{X}, \mathcal{X})$. Let $u\colon \mathsf{X} \to [0, \infty)$ be a given measurable function. (Here $[0,\infty)$ is equipped with the usual Borel σ-field.) For $x \in \mathsf{X}$, let

$$\hat{u}_n(x) \;\triangleq\; \mathsf{E}[\, u(X_n) \mid X_0 = x\,], \qquad n = 0, 1, 2, \ldots. \tag{1.1}$$

In this paper we give sufficient conditions to answer the following question: "Does there exist a positive, finite constant D such that

$$\overline{\lim_{n\to\infty}}\, \hat{u}_n(x) \;\leq\; D$$

for all $x \in \mathsf{X}$?" To answer this question we will construct a sequence of functions $\{u_n\}$ such that for all $n \geq 0$,

$$\hat{u}_n(x) \;\leq\; u_n(x), \quad \forall x \in \mathsf{X}, \tag{1.2}$$

and such that $\{u_n\}$ is easy to analyze. To suggest how one might establish (1.2), consider the following result.

Proposition 1. *Set $u_0(x) \triangleq u(x)$, and let $\{u_n\}_{n\geq 1}$ be given. If one can show that for all $n \geq 0$, and all $x \in \mathsf{X}$,*

$$\mathsf{E}[\, u_n(X_1) \mid X_0 = x\,] \;\leq\; u_{n+1}(x), \tag{1.3}$$

then $\hat{u}_n(x) \leq u_n(x)$ for all $n \geq 0$, and all $x \in \mathsf{X}$.

Proof. We proceed by induction. Clearly,

$$\hat{u}_0(x) \triangleq \mathsf{E}[\, u(X_0) \mid X_0 = x\,] \;=\; u(x) \;=\; u_0(x),$$

and so (1.2) holds for $n = 0$. Suppose (1.2) holds for some $n \geq 0$. Then

$$\begin{aligned}
\hat{u}_{n+1}(x) &\triangleq \mathsf{E}[\, u(X_{n+1}) \mid X_0 = x\,] \\
&= \mathsf{E}[\, \mathsf{E}[\, u(X_{n+1}) \mid X_1, X_0\,] \mid X_0 = x\,] \\
&= \mathsf{E}[\, \mathsf{E}[\, u(X_{n+1}) \mid X_1\,] \mid X_0 = x\,], \qquad \text{by the Markov property,} \\
&= \mathsf{E}[\, \hat{u}_n(X_1) \mid X_0 = x\,], \qquad \text{by time homogeneity,} \\
&\leq \mathsf{E}[\, u_n(X_1) \mid X_0 = x\,], \qquad \text{by the induction hypothesis,} \\
&\leq u_{n+1}(x), \qquad \text{by (1.3).}
\end{aligned}$$

The difficulty is to establish conditions under which (1.3) holds. If we let $p(x, B)$ denote a regular version of $\mathsf{P}(X_1 \in B \mid X_0 = x)$ for $B \in \mathcal{X}$, then we can rewrite (1.3) as

$$\int_{\mathsf{X}} u_n(y)\, p(x, dy) \;\leq\; u_{n+1}(x). \tag{1.4}$$

Remark. The careful reader will see later that since $u_n(x)$ will have the form $u_n(x) = f_n(u(x))$, in order to derive the results in this paper, we do not need a regular version of $\mathsf{P}(X_1 \in B \mid X_0 = x)$ for $B \in \mathcal{X}$, but only a regular version of $\mathsf{P}(u(X_1) \in C \mid X_0 = x)$ for C a Borel subset of $[0, \infty)$. Since u is real-valued, such a regular version always exists; see Billingsley (1979), Theorem 33.3, p. 390.

In Section 2 we give sufficient conditions under which (1.4) will hold. These conditions are summarized in Theorem 2.

In Section 3 we describe the problem of "joining the shortest queue." We then apply Theorem 2 to show that the servers are equally loaded, even in heavy traffic.

In Section 4 we prove Theorem 2.

In Section 5 we present an extention of Theorem 2.

In Section 6 we compare our results to those of Hajek (1982) for problems in which both apply.

2. Statement of Main Result

As pointed out in Section 1, the problem of finding a sequence $\{u_n\}$ satisfying (1.2) can be reduced to the problem of finding a sequence satisfying (1.4). We now present our main result, Theorem 2, which gives conditions under which this can be done. The proof is given in Section 4.

Theorem 2. Fix $b \geq 0$ and $0 < \varepsilon < 1$. Suppose that there exist positive constants C and ν satisfying

$$\frac{(b+1)(b+2)}{2} \cdot \frac{(1+\varepsilon)^b}{(1-\varepsilon)^{b+3}} \cdot C \ \leq \ 2\nu, \tag{2.1}$$

and such that for some L,

$$|u(y) - u(x)| \ \leq \ \varepsilon\, u(x), \quad p(x,\cdot)\text{-a.s.}, \quad \text{if } u(x) > L, \tag{2.2}$$

$$\mathsf{E}[\, u(X_{n+1}) - u(x) \mid X_n = x\,] \ \leq \ -\nu, \qquad \text{if } u(x) > L, \tag{2.3}$$

and

$$\mathsf{E}[\, |u(X_{n+1}) - u(x)|^2 \mid X_n = x\,] \ \leq \ C\, u(x), \qquad \text{if } u(x) > L. \tag{2.4}$$

If one can find a finite constant A, depending on b and L, such that

$$\mathsf{E}[\, u(X_{n+1})^{b+2} \mid X_n = x\,]^{1/(b+2)} \ \leq \ A, \qquad \text{if } 0 \leq u(x) \leq L, \tag{2.5}$$

then the sequence $\{u_n\}$ given by

$$u_n(x) \ \triangleq \ \frac{u(x)^{b+2}}{[u(x) + \frac{\nu}{b+1} n]^{b+1}} + A + \frac{A^{b+2}}{(\frac{\nu}{b+1})^{b+1}} \sum_{k=1}^{n-1} \frac{1}{k^{b+1}}, \tag{2.6}$$

satisfies (1.4), and by Proposition 1,

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \ \leq \ u_n(x), \qquad n \geq 0,\ x \in \mathsf{X}. \tag{2.7}$$

Of course, when $n = 0$, equation (2.6) is to be interpreted as $u_0(x) \triangleq u(x)$, and when $n = 1$, it is to be interpreted as

$$u_1(x) \ \triangleq \ \frac{u(x)^{b+2}}{[u(x) + \frac{\nu}{b+1}]^{b+1}} + A.$$

Remark. When $b = 1$, (2.6) becomes

$$u_n(x) \;=\; \frac{u(x)^3}{[u(x) + \frac{\nu}{2}n]^2} + A + 4\frac{A^3}{\nu^2}\sum_{k=1}^{n-1}\frac{1}{k^2}. \tag{2.8}$$

In fact, since $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$ (See Papadimitriou (1973) for an elementary proof of this fact. See Stark (1974) for more general series), it is easy to see that

$$\lim_{n\to\infty} u_n(x) \;=\; A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2), \tag{2.9}$$

and that

$$u_n(x) \;\le\; u(x) + A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2). \tag{2.10}$$

Note that the limit in (2.9) is independent of x, whereas the bound in (2.10) is independent of n.

Remark. The inequality (2.3) is very similar to the definition of a Markovian supermartingale. To see this, recall that if $\mathcal{F}_n \triangleq \sigma(X_0, \ldots, X_n)$, then $\{u(X_n)\}$ is an $\{\mathcal{F}_n\}$-supermartingale if and only if for all $n \ge 0$,

$$\mathsf{E}[\, u(X_{n+1}) \mid \mathcal{F}_n \,] \;\le\; u(X_n). \tag{2.11}$$

Since $\{X_n\}$ is Markovian, this can be rewritten as

$$\mathsf{E}[\, u(X_{n+1}) \mid X_n \,] \;\le\; u(X_n),$$

or

$$\mathsf{E}[\, u(X_{n+1}) - u(x) \mid X_n = x \,] \;\le\; 0, \qquad \forall x \in \mathsf{X}. \tag{2.12}$$

To compare (2.3) with (2.12), observe that (2.3) is a stronger inequality, but holds only when $u(x) > L$. On the other hand, (2.12) is a weaker inequality, but holds for *every* $x \in \mathsf{X}$. Now, observe that if (2.11) were to hold, then we could immediately write

$$\mathsf{E}[\, u(X_{n+1}) \mid X_0 \,] \;\le\; \mathsf{E}[\, u(X_n) \mid X_0 \,],$$

from which it would follow by induction that for all $n \ge 0$,

$$\mathsf{E}[\, u(X_n) \mid X_0 \,] \;\le\; u(X_0),$$

which we can rewrite as

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; u(x).$$

Now, if Theorem 2 were to hold with $b = 1$, substituting (2.10) into (2.7) would yield

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; u(x) + A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2).$$

Clearly, (2.12) (that is, the Markov version of (2.11)) is a stronger condition than (2.1) – (2.5) combined.

Example. Let us apply Theorem 2 to the single-server queue described in Section 1. Since Y_n takes values in $[0, \infty)$, we take $u(Y_n) = Y_n$. Now, set $b = 1$ and fix $0 < \varepsilon < 1$. Make the following observations. First, it is easy to calculate

$$\mathsf{E}[\, Y_{n+1} \mid Y_n = i\,] \;=\; \begin{cases} \lambda, & \text{if } i = 0, \\ i - (\mu - \lambda), & \text{if } i > 0. \end{cases}$$

Rewrite this as

$$\mathsf{E}[\, Y_{n+1} - i \mid Y_n = i\,] \;=\; \begin{cases} \lambda, & \text{if } i = 0, \\ -(\mu - \lambda), & \text{if } i > 0. \end{cases} \tag{2.13}$$

Clearly, we should set $\nu = \mu - \lambda > 0$. The next step is to observe that if $Y_n = i$,

$$|Y_{n+1} - i| \;\le\; 1, \tag{2.14}$$

and so

$$\mathsf{E}[\, |Y_{n+1} - i|^2 \mid Y_n = i\,] \;\le\; 1. \tag{2.15}$$

Now, having set $\nu = \mu - \lambda$, let

$$C \;=\; \frac{2}{3}\nu \cdot \frac{(1-\varepsilon)^4}{(1+\varepsilon)}.$$

To put (2.14) and (2.15) into the form of (2.2) and (2.4), observe that in order to have

$$1 \;\le\; \varepsilon\, i \qquad \text{and} \qquad 1 \;\le\; C\, i,$$

we must have

$$i \;\ge\; \max\Big\{\frac{1}{\varepsilon},\, \frac{3(1+\varepsilon)}{2\nu(1-\varepsilon)^4}\Big\}. \tag{2.16}$$

Let L denote the right-hand side of (2.16). Finally, note that since (2.14) implies

$$Y_{n+1} \le i+1,$$

we have

$$\begin{aligned} \mathsf{E}[\,Y_{n+1}^3 \mid Y_n = i\,]^{1/3} &\le i+1 \\ &\le L+1, \qquad \text{if } i \le L. \end{aligned}$$

It should now be clear that Theorem 2 applies to the single-server queue if $\mu > \lambda$. In this case,

$$\overline{\lim_{n\to\infty}}\, \mathsf{E}[\,Y_n \mid Y_0 = i\,] \le (L+1) + \frac{2}{3}\left(\frac{\pi}{\mu-\lambda}\right)^2 (L+1)^3 < \infty.$$

We remark here that the preceding bound depends on L, which by (2.16), depends on ε. To reduce L we should try to choose $0 < \varepsilon < 1$ so that

$$\frac{1}{\varepsilon} = \frac{3(1+\varepsilon)}{2\nu(1-\varepsilon)^4},$$

or equivalently,

$$2\nu(1-\varepsilon)^4 - 3\varepsilon(1+\varepsilon) = 0.$$

In other words, to reduce L, we must find the zeros of a fourth-degree polynomial in ε. Observe that there is at least one zero in $(0,1)$.

3. Joining the Shortest Queue

In this section we will investigate the balancing of N queues to N identical servers under a range of traffic conditions. Consider a system of N identical servers, each with its own infinite waiting room, indexed by $i = 1, \ldots, N$. Suppose also that we are given $N+1$ arrival processes indexed by $i = 1, \ldots, N, N+1$. Customers from arrival processes $i = 1, \ldots, N$ are assigned to the corresponding waiting room for server i. Customers from the $(N+1)$'st arrival process may in general be assigned to any waiting room; however, we shall only consider what happens when we employ the policy of assigning these customers to the waiting room with the fewest waiting customers. In case of a tie, the waiting room of smallest index is selected.

Let $Q_n^{(i)}$ denote the size of the queue for server i at time $n = 0, 1, 2, \ldots$. This includes customers in the waiting room as well as the customer being served. Let

$$Q_n \triangleq (Q_n^{(1)}, \ldots, Q_n^{(N)}).$$

We want to show that the policy of joining the shortest queue will ensure, even under arbitrarily heavy traffic conditions, that for all $i \neq j$,

$$\mathsf{E}[\,|Q_n^{(j)} - Q_n^{(i)}|\mid Q_0 = q\,] \tag{3.1}$$

is uniformly bounded for all time. To apply the theory of Section 2, we need to find a suitable function u. Below we show that the function defined in (3.2) will suffice.

Before proceeding, we point out that the continuous-time analog of this model has been studied by a number of authors, though, to our knowledge, none has investigated the behavior of the continuous-time analog of (3.1). For example, Haight (1958), Kingman (1961), Flatto and McKean (1977), and Conolly (1984) characterized the stationary distribution of the state of the system. Foschini and Salz (1978) studied a diffusion approximation. Blanc (1987) outlined a numerical method for calculating the state probabilities and moments of the queue-length distribution. Brumelle (1971) bounded the expected waiting time. Wolff (1977) bounded moments of the delay distribution (see also Wolff (1987) for corrections and comments). Halfin (1985) bounded the probability distribution of the number of customers in the system, and its expected value in equilibrium. Other researchers, Winston (1977) and Weber (1978), proved that the policy of "joining the shortest queue" is optimal in the sense of maximizing the customer throughput, while Ephremides, Varaiya, and Walrand (1980) proved that this policy is optimal in the sense of minimizing the expected total time for completing service on all customers which arrive before a fixed time.

We now make a precise statement of our model. Using the N-fold cartesian product of the nonnegative integers as the state space for $\{Q_n\}$, we take $\{Q_n\}$ to be a Markov process with transition probabilities defined as follows. Suppose that

$$Q_n = q = (q^{(1)}, \ldots, q^{(N)}),$$

where the $q^{(\cdot)}$ are nonnegative integers. Let $\mu, \lambda, \lambda_1, \ldots, \lambda_N$ be a sequence of elements from the interval $[0, 1]$ such that $\lambda + \sum_{j=1}^{N} \lambda_j + N\mu \leq 1$. Suppose that the N-vector Δq_j is defined by $(0,\dots,0,1,0,\dots,0)$, where the 1 occupies the j'th coordinate. Let i denote the smallest index such that $q^{(i)} \le q^{(j)}$ holds for $j = 1,\dots,N$. Let $\delta_i(j) = 1$ if $j = i$, and 0 otherwise. Let $\mu(t) = \mu$ if $t \ge 1$ and $\mu(t) = 0$ if $t = 0$. We assume that

$$\mathsf{P}(Q_{n+1} = q' \mid Q_n = q) \;=\; \begin{cases} \lambda_j + \lambda\delta_i(j), & \text{if } q' = q + \Delta q_j, \\ \mu(q^{(j)}), & \text{if } q' = q - \Delta q_j, \\ 1 - \lambda - \sum_{j=1}^{N}[\lambda_j + \mu(q^{(j)})], & \text{if } q' = q. \end{cases}$$

Let

$$u(q) \triangleq \sum_{\alpha=1}^{N-1}\sum_{\beta=\alpha+1}^{N} |q^{(\beta)} - q^{(\alpha)}|^2. \tag{3.2}$$

Theorem 3. *Without loss of generality, assume that* $\lambda_1 \le \cdots \le \lambda_N$. *If*

$$\lambda \;>\; \sum_{k=1}^{N-1}(\lambda_N - \lambda_k), \tag{3.3}$$

then there exist finite positive constants ν *and* A *such that*

$$\mathsf{E}[\,u(Q_n) \mid Q_0 = q\,] \;\le\; \frac{u(q)^3}{[u(q) + \frac{\nu}{2}n]^2} + A(1 + 4(\tfrac{A}{\nu})^2 \sum_{k=1}^{n-1} \frac{1}{k^2}),$$

from which it follows that

$$\mathsf{E}[\,u(Q_n) \mid Q_0 = q\,] \;\le\; u(q) + A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2), \tag{3.4}$$

and

$$\varlimsup_{n\to\infty} \mathsf{E}[\,u(Q_n) \mid Q_0 = q\,] \;\le\; A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2) \;<\; \infty.$$

Remarks. (*i*) If $\lambda > (N-1)(\lambda_N - \lambda_1)$ then (3.3) holds. (*ii*) Since the inequality in (3.3) does not depend on μ, the interpretation of (3.4) is that the policy of "joining the shortest queue" will ensure that even under arbitrarily heavy traffic, $\mathsf{E}[\,u(Q_n) \mid Q_0 = q\,]$ does not "blow up." By Jensen's inequality, (3.4) also implies that for every $i \ne j$,

$$\begin{aligned} \mathsf{E}[\,|Q_n^{(j)} - Q_n^{(i)}| \mid Q_0 = q\,] &\le \mathsf{E}[\,|Q_n^{(j)} - Q_n^{(i)}|^2 \mid Q_0 = q\,]^{1/2} \\ &\le \mathsf{E}[\,u(Q_n) \mid Q_0 = q\,]^{1/2} \end{aligned}$$

is bounded for all time. In this sense, each queue has "approximately" the same number of customers; i.e., the queues are "equally loaded" or "balanced." The reader may find it interesting to contemplate various special cases such as $\mu = 0$, $\lambda_N = \lambda_1 > 0$, or $\lambda_N = \lambda_1 = 0$.

Proof. We will show below that (3.3) implies

$$|u(q \pm \Delta q_j) - u(q)| \;\leq\; (N-1)(1 + 2\sqrt{u(q)}\,), \tag{3.5}$$

$$\mathsf{E}[\,u(Q_{n+1}) - u(q) \mid Q_n = q\,] \;\leq\; (N-1) - 2\Big(\lambda - \sum_{k=1}^{N-1}(\lambda_N - \lambda_k)\Big)\sqrt{\frac{u(q)}{2N-1}}, \tag{3.6}$$

and

$$\mathsf{E}[\,|u(Q_{n+1}) - u(q)|^2 \mid Q_n = q\,] \;\leq\; (N-1)^2(1 + 4\sqrt{u(q)} + 4u(q)). \tag{3.7}$$

First, however, we will use (3.3) and (3.5) – (3.7) to show that the hypotheses of Theorem 2 are satisfied. So, set $b = 1$, and fix any $0 < \varepsilon < 1$. Now make the following observations. Since

$$\lim_{v\to\infty} \frac{(N-1)(1+2\sqrt{v}\,)}{v} \;=\; 0,$$

it is clear that for all sufficiently large v,

$$(N-1)(1+2\sqrt{v}\,) \;\leq\; \varepsilon\, v.$$

Next, since

$$\lim_{v\to\infty} \frac{(N-1)^2(1+4\sqrt{v}+4v)}{v} \;=\; 4(N-1)^2,$$

if we fix any $C > 4(N-1)^2$, then for all sufficiently large v,

$$(N-1)^2(1+4\sqrt{v}+4v) \;\leq\; C\, v.$$

So, fix any $C > 4(N-1)^2$, and then choose (recall (2.1))

$$\nu \;\geq\; \frac{3}{2}C \cdot \frac{(1+\varepsilon)}{(1-\varepsilon)^4}. \tag{3.8}$$

Now, since (3.3) holds, we have

$$\lim_{v\to\infty} \left\{(N-1) - 2\Big(\lambda - \sum_{k=1}^{N-1}(\lambda_N - \lambda_k)\Big)\sqrt{\frac{v}{2N-1}}\right\} \;=\; -\infty.$$

Thus, for all sufficiently large v,

$$(N-1) - 2\Big(\lambda - \sum_{k=1}^{N-1}(\lambda_N - \lambda_k)\Big)\sqrt{\frac{v}{2N-1}} \;\leq\; -\nu.$$

It should now be clear that if $b = 1$, $0 < \varepsilon < 1$, $C > 4(N-1)^2$, and ν satisfies (3.8), then there is some L such that (2.2) – (2.4) hold. To find a finite constant A satisfying (2.5), observe that if $u(q) \leq L$, then (3.5) implies

$$|u(q \pm \Delta q_j) - u(q)| \;\leq\; (N-1)(1 + 2\sqrt{L}),$$

and so

$$u(q \pm \Delta q_j) \;\leq\; (N-1)(1 + 2\sqrt{L}) + L.$$

Hence, we may take

$$A \;=\; (N-1)(1 + 2\sqrt{L}) + L.$$

We conclude that if (3.3) holds, (3.5) – (3.7) are sufficient to apply Theorem 2.

We now establish (3.5) – (3.7). Clearly, (3.5) implies (3.7). To establish (3.5) and (3.6), we proceed as follows. We denote by i the smallest integer in $\{1, \ldots, N\}$ such that $q^{(i)} \leq q^{(j)}$ holds for $j = 1, \ldots, N$. Setting $\ell_j \triangleq q^{(j)} - q^{(i)}$, we can write

$$q \;=\; (q^{(i)} + \ell_1, \ldots, q^{(i)} + \ell_N).$$

Note that each $\ell_j \geq 0$, and $\ell_i = 0$. Fixing any $j \in \{1, \ldots, N\}$, we can write

$$\begin{aligned} u(q) \;&=\; \sum_{\beta \neq j} |q^{(\beta)} - q^{(j)}|^2 + \sum_{\substack{\alpha=1 \\ \alpha \neq j}}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ \beta \neq j}}^{N} |q^{(\beta)} - q^{(\alpha)}|^2 \\ &=\; \sum_{\beta \neq j} |\ell_\beta - \ell_j|^2 + \sum_{\substack{\alpha=1 \\ \alpha \neq j}}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ \beta \neq j}}^{N} |\ell_\beta - \ell_\alpha|^2, \end{aligned} \tag{3.9}$$

and

$$\begin{aligned} u(q \pm \Delta q_j) \;&=\; \sum_{\beta \neq j} |q^{(\beta)} - (q^{(j)} \pm 1)|^2 + \sum_{\substack{\alpha=1 \\ \alpha \neq j}}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ \beta \neq j}}^{N} |q^{(\beta)} - q^{(\alpha)}|^2 \\ &=\; \sum_{\beta \neq j} |(\ell_\beta - \ell_j) \mp 1)|^2 + \sum_{\substack{\alpha=1 \\ \alpha \neq j}}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ \beta \neq j}}^{N} |\ell_\beta - \ell_\alpha|^2, \end{aligned}$$

so that

$$u(q \pm \Delta q_j) - u(q) \;=\; \sum_{\beta \neq j} (1 \mp 2(\ell_\beta - \ell_j)) \,. \tag{3.10}$$

It follows from (3.9) that for any $\beta \neq j$,

$$|\ell_\beta - \ell_j| \;\le\; \sqrt{u(q)}.$$

This with (3.10) implies (3.5).

To establish (3.6), use (3.10) to write

$$\begin{aligned}
\mathsf{E}[\, u(Q_{n+1}) - u(q) \mid Q_n = q \,] \;&=\; \sum_{j=1}^{N} \Big\{ \Big(\sum_{\beta \neq j} (1 - 2(\ell_\beta - \ell_j)) \Big) (\lambda_j + \lambda \delta_i(j)) \\
&\qquad + \Big(\sum_{\beta \neq j} (1 + 2(\ell_\beta - \ell_j)) \Big) \mu(q^{(i)} + \ell_j) \Big\} \\
&=\; \sum_{j=1}^{N} \Big\{ (N-1)[\mu(q^{(i)} + \ell_j) + \lambda_j + \lambda \delta_i(j)] \\
&\qquad + 2 \sum_{\beta \neq j} (\ell_\beta - \ell_j)[\mu(q^{(i)} + \ell_j) - \lambda_j - \lambda \delta_i(j)] \Big\} \\
&\le\; \nu_0 + 2 \sum_{j=1}^{N} \sum_{\beta \neq j} (\ell_\beta - \ell_j)[\mu(q^{(i)} + \ell_j) - \lambda_j - \lambda \delta_i(j)],
\end{aligned}$$

where $\nu_0 \triangleq (N-1)[N\mu + \sum_{j=1}^{N} \lambda_j + \lambda] \le (N-1)$. For the second term, write

$$\nu_1 \;\triangleq\; \sum_{j=1}^{N} \sum_{\beta \neq j} (\ell_\beta - \ell_j)[\mu(q^{(i)} + \ell_j) - \lambda_j - \lambda \delta_i(j)].$$

Break this into two terms. In the first change the order of summation so that

$$\nu_1 \;=\; \sum_{\beta=1}^{N} \ell_\beta \sum_{j \neq \beta} [\mu(q^{(i)} + \ell_j) - \lambda_j - \lambda \delta_i(j)] - \sum_{j=1}^{N} \ell_j \sum_{\beta \neq j} [\mu(q^{(i)} + \ell_j) - \lambda_j - \lambda \delta_i(j)].$$

Using the fact that $\ell_i = 0$, that $\mu(\cdot) \le \mu$, and that the minimum value of $\sum_{j \neq \beta} \lambda_j$ occurs when $\beta = N$,

$$\begin{aligned}
\nu_1 \;&\le\; \sum_{\beta \neq i} \ell_\beta \sum_{j \neq \beta} [\mu - \lambda_j - \lambda \delta_i(j)] - \sum_{j \neq i} \ell_j [\mu(q^{(i)} + \ell_j) - \lambda_j - \lambda \delta_i(j)](N-1) \\
&=\; \sum_{\beta \neq i} \ell_\beta [(N-1)\mu - \lambda - \sum_{j \neq \beta} \lambda_j] - \sum_{j \neq i} \ell_j [\mu(q^{(i)} + \ell_j) - \lambda_j](N-1) \\
&\le\; \sum_{\beta \neq i} \ell_\beta [(N-1)\mu - \lambda - \sum_{k=1}^{N-1} \lambda_k] - \sum_{j \neq i} \ell_j [\mu(q^{(i)} + \ell_j) - \lambda_j](N-1) \\
&=\; \sum_{j \neq i} \ell_j [(N-1)\mu - \lambda - \sum_{k=1}^{N-1} \lambda_k] - \sum_{j \neq i} \ell_j [\mu - \lambda_j](N-1),
\end{aligned}$$

where the last step follows by observing that $\ell_j \mu(q^{(i)} + \ell_j) = \ell_j \mu$ for all j. Continuing, we see that

$$\begin{aligned} \nu_1 &\leq [(N-1)\lambda_N - \sum_{k=1}^{N-1} \lambda_k - \lambda] \sum_{j \neq i} \ell_j \\ &= -[\lambda - \sum_{k=1}^{N-1} (\lambda_N - \lambda_k)] \sum_{j \neq i} \ell_j. \end{aligned} \tag{3.11}$$

Now, observe that according to (3.9), if we set $j = i$, and use the fact that $\ell_i = 0$, then

$$\begin{aligned} u(q) &= \sum_{\beta \neq i} \ell_\beta^2 + \sum_{\substack{\alpha=1 \\ \alpha \neq i}}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ \beta \neq i}}^{N} |\ell_\beta - \ell_\alpha|^2 \\ &\leq \sum_{\beta \neq i} \ell_\beta^2 + \sum_{\substack{\alpha=1 \\ \alpha \neq i}}^{N-1} \sum_{\substack{\beta=\alpha+1 \\ \beta \neq i}}^{N} (\ell_\beta^2 + \ell_\alpha^2) \\ &\leq (2N-1) \sum_{\beta \neq i} \ell_\beta^2, \end{aligned}$$

and so

$$\sqrt{\frac{u(q)}{2N-1}} \leq \sum_{\beta \neq i} \ell_\beta. \tag{3.12}$$

Now, if (3.3) holds, substituting (3.12) into (3.11) yields

$$\nu_1 \leq -\Big(\lambda - \sum_{k=1}^{N-1} (\lambda_N - \lambda_k)\Big) \sqrt{\frac{u(q)}{2N-1}}. \tag{3.13}$$

Finally, we can write

$$\begin{aligned} \mathsf{E}[\, u(Q_{n+1}) - u(q) \mid Q_n = q \,] &\leq \nu_0 + 2\nu_1 \\ &\leq (N-1) - 2\Big(\lambda - \sum_{k=1}^{N-1} (\lambda_N - \lambda_k)\Big) \sqrt{\frac{u(q)}{2N-1}}, \end{aligned}$$

which is exactly (3.6). This completes the proof of Theorem 3.

We now briefly discuss the continuous-time analog of this model. Suppose that $\{Q_n\}$ is the jump chain corresponding to a continuous-time queuing process $\{\hat{Q}_t\}$ with instantaneous arrival rates $\hat{\lambda}, \hat{\lambda}_1, \ldots, \hat{\lambda}_N$, and instantaneous departure rate $\hat{\mu}$ for each of the N servers. If we let $d(q) \triangleq \hat{\lambda} + \sum_{j=1}^{N} [\, \hat{\lambda}_j + \hat{\mu}(q^{(j)}) \,]$, then the jump chain parameters are

$$\begin{aligned} \lambda &= \hat{\lambda}/d(q), \\ \lambda_j &= \hat{\lambda}_j/d(q), \\ \mu(q^{(j)}) &= \hat{\mu}(q^{(j)})/d(q). \end{aligned} \tag{3.14}$$

Clearly, the transition probabilities given that $Q_n = q$ now depend on q in a slightly different way. However, assuming that $\hat{\lambda}_1 \leq \cdots \leq \hat{\lambda}_N$, a review of the proof of Theorem 3 through (3.11) when (3.14) is in force will show that the crucial question is whether or not

$$\frac{\hat{\lambda}}{d(q)} > \frac{1}{d(q)} \sum_{k=1}^{N-1} (\hat{\lambda}_N - \hat{\lambda}_k).$$

This condition obviously holds if and only if

$$\hat{\lambda} > \sum_{k=1}^{N-1} (\hat{\lambda}_N - \hat{\lambda}_k). \tag{3.15}$$

So, if (3.15) holds, we can express (3.13) in terms of the infinitesimal rates:

$$\nu_1 \leq -\frac{\hat{\lambda} - \sum_{k=1}^{N-1} (\hat{\lambda}_N - \hat{\lambda}_k)}{d(q)} \sqrt{\frac{u(q)}{2N-1}}.$$

Since $\max_q d(q) = \hat{\lambda} + \sum_{j=1}^{N} \hat{\lambda}_j + N\hat{\mu}$,

$$\nu_1 \leq -\frac{\hat{\lambda} - \sum_{k=1}^{N-1} (\hat{\lambda}_N - \hat{\lambda}_k)}{\hat{\lambda} + \sum_{j=1}^{N} \hat{\lambda}_j + N\hat{\mu}} \sqrt{\frac{u(q)}{2N-1}}.$$

Thus, for jump chains, (3.6) should be replaced by

$$\mathsf{E}[\,u(Q_{n+1}) - u(q) \mid Q_n = q\,] \leq (N-1) - 2\frac{\hat{\lambda} - \sum_{k=1}^{N-1} (\hat{\lambda}_N - \hat{\lambda}_k)}{\hat{\lambda} + \sum_{j=1}^{N} \hat{\lambda}_j + N\hat{\mu}} \sqrt{\frac{u(q)}{2N-1}},$$

and (3.3) should be replaced by (3.15).

4. Proof of Theorem 2

Before proving Theorem 2, we introduce the following notation. Let

$$s \triangleq \frac{\nu}{b+1}, \tag{4.1}$$

and set

$$B \triangleq \frac{A^{b+2}}{s^{b+1}}$$

so that we can use (2.5) to write

$$\frac{1}{s^{b+1}} \int_{\mathsf{X}} u(y)^{b+2}\, p(x, dy) \leq B, \qquad \text{if } 0 \leq u(x) \leq L. \tag{4.2}$$

Now, for $v \geq 0$, let $\varphi_0(v) \equiv v$, and for $n \geq 1$ set

$$\varphi_n(v) \triangleq \frac{v^{b+2}}{[v+sn]^{b+1}}. \tag{4.3}$$

With this notation we can write (recall (2.6))

$$\begin{aligned} u_0(x) &= \varphi_0(u(x)), \\ u_1(x) &= \varphi_1(u(x)) + A, \\ u_n(x) &= \varphi_n(u(x)) + A + B\sum_{k=1}^{n-1}\frac{1}{k^{b+1}}, \qquad n \geq 2. \end{aligned} \tag{4.4}$$

For later reference, note that differentiating φ_n with respect to v yields

$$\varphi_n'(v) = \frac{v^{b+2} + sn(b+2)v^{b+1}}{[v+sn]^{b+2}}, \tag{4.5}$$

and

$$\varphi_n''(v) = \frac{(b+1)(b+2)s^2n^2v^b}{[v+sn]^{b+3}}. \tag{4.6}$$

Proof of Theorem 2. Fix $x \in \mathsf{X}$. We first show that (1.4) holds if $0 \leq u(x) \leq L$. Then we will show that (1.4) holds if $u(x) > L$. Suppose $0 \leq u(x) \leq L$ and $n = 0$. Use the fact that $u_0(y) = u(y)$ followed by Jensen's inequality and (2.5) to get

$$\begin{aligned} \int_{\mathsf{X}} u_0(y)\,p(x,dy) - u_1(x) &= \int_{\mathsf{X}} u(y)\,p(x,dy) - \varphi_1(x) - A \\ &\leq \left(\int_{\mathsf{X}} u(y)^{b+2}\,p(x,dy)\right)^{1/(b+2)} - \varphi_1(x) - A \\ &\leq A - \varphi_1(x) - A \leq 0. \end{aligned}$$

Now for $n \geq 1$, keep in mind (4.2) while using (4.4) to write

$$\begin{aligned} \int_{\mathsf{X}} u_n(y)\,p(x,dy) - u_{n+1}(x) &= \int_{\mathsf{X}} \varphi_n(u(y))\,p(x,dy) - \varphi_{n+1}(u(x)) - \frac{B}{n^{b+1}} \\ &= \int_{\mathsf{X}} \frac{u(y)^{b+2}}{[u(y)+sn]^{b+1}}\,p(x,dy) - \varphi_{n+1}(u(x)) - \frac{B}{n^{b+1}} \\ &\leq \frac{1}{(sn)^{b+1}}\int_{\mathsf{X}} u(y)^{b+2}\,p(x,dy) - \varphi_{n+1}(u(x)) - \frac{B}{n^{b+1}} \\ &\leq \frac{B}{n^{b+1}} - \varphi_{n+1}(u(x)) - \frac{B}{n^{b+1}} \\ &\leq 0. \end{aligned}$$

We next establish (1.4) when $u(x) > L$ and $n \geq 1$. The case $n = 0$ follows the same pattern (without Taylor's theorem, since $\varphi_0(v) \equiv v$) and is and left to the reader. Using Taylor's theorem, we can write

$$\begin{aligned} \int_{\mathsf{X}} u_n(y)\, p(x, dy) &= u_n(x) + \int_{\mathsf{X}} [\varphi_n(u(y)) - \varphi_n(u(x))]\, p(x, dy) \\ &\leq u_n(x) + \int_{\mathsf{X}} [\varphi_n'(u(x))(u(y) - u(x)) + K(u(y) - u(x))^2]\, p(x, dy), \end{aligned} \tag{4.7}$$

where K is the upper bound on $\frac{1}{2}\varphi_n''$ given by

$$K \triangleq \frac{(b+1)(b+2)}{2} \cdot \frac{(1+\varepsilon)^b}{(1-\varepsilon)^{b+3}} \cdot \frac{u(x)^b s^2 n^2}{[u(x) + sn]^{b+3}}.$$

To verify that K is an upper bound, observe that (2.2) implies that $p(x, \cdot)$-a.s.,

$$u(x)(1-\varepsilon) \leq u(y) \leq u(x)(1+\varepsilon).$$

Now, using (4.6),

$$\begin{aligned} \sup_{u(x)(1-\varepsilon) \leq v \leq u(x)(1+\varepsilon)} \tfrac{1}{2}\varphi_n''(v) &\leq \frac{(b+1)(b+2)s^2n^2u(x)^b(1+\varepsilon)^b}{2[u(x)(1-\varepsilon) + sn]^{b+3}} \\ &\leq K, \end{aligned}$$

since $[u(x)(1-\varepsilon) + sn] \geq (1-\varepsilon)[u(x) + sn]$. Returning to (4.7), we use (2.3) and (2.4) to write

$$\int_{\mathsf{X}} u_n(y)\, p(x, dy) \leq u_n(x) + \varphi_n'(u(x)) \cdot (-\nu) + KCu(x).$$

Using the definition of K and (2.1) together with (4.5),

$$\begin{aligned} \int_{\mathsf{X}} u_n(y)\, p(x, dy) &\leq u_n(x) - \nu \frac{u(x)^{b+2} + sn(b+2)u(x)^{b+1}}{[u(x) + sn]^{b+2}} + \frac{2\nu u(x)^{b+1} s^2 n^2}{[u(x) + sn]^{b+3}} \\ &= u_n(x) - \frac{\nu u(x)^{b+2}}{[u(x) + sn]^{b+2}} - \frac{\nu s n u(x)^{b+1}}{[u(x) + sn]^{b+3}}[(b+2)u(x) + bsn] \\ &\leq u_n(x) - \frac{\nu u(x)^{b+2}}{[u(x) + sn]^{b+2}} \\ &= u_{n+1}(x) + u_n(x) - u_{n+1}(x) - \frac{\nu u(x)^{b+2}}{[u(x) + sn]^{b+2}}. \end{aligned} \tag{4.8}$$

Observe that

$$
\begin{aligned}
u_n(x) - u_{n+1}(x) &= \varphi_n(u(x)) - \varphi_{n+1}(u(x)) - \frac{B}{n^{b+1}} \\
&\leq \varphi_n(u(x)) - \varphi_{n+1}(u(x)) \\
&= u(x)^{b+2}\left(-\int_n^{n+1} \frac{\partial}{\partial\theta}\left(\frac{1}{[u(x)+s\theta]^{b+1}}\right) d\theta\right) \\
&= u(x)^{b+2}(b+1)s\int_n^{n+1} \frac{d\theta}{[u(x)+s\theta]^{b+2}} \\
&\leq \frac{\nu u(x)^{b+2}}{[u(x)+sn]^{b+2}}, \qquad (4.9)
\end{aligned}
$$

where the last step follows by setting $\theta = n$ in the integrand, and then using (4.1). Combining (4.8) and (4.9) yields (1.4) when $u(x) > L$.

5. Extension of Theorem 2

It is possible to weaken the hypotheses of Theorem 2 by replacing (2.2) – (2.4) with (5.1) – (5.4) below.

Suppose that for each $x \in \mathsf{X}$, there is a "bad" subset, B_x, with

$$B_x \subset \{y \in \mathsf{X} : u(y) \leq u(x)\}, \qquad (5.1)$$

and such that

$$I_{B_x^c}(y)|u(y) - u(x)| \leq \varepsilon\, u(x), \quad p(x,\cdot)\text{-a.s.}, \quad \text{if } u(x) > L, \qquad (5.2)$$

$$\mathsf{E}[\, I_{B_x^c}(X_{n+1})(u(X_{n+1}) - u(x)) \mid X_n = x\,] \leq -\nu, \qquad \text{if } u(x) > L, \qquad (5.3)$$

and

$$\mathsf{E}[\, I_{B_x^c}(X_{n+1})|u(X_{n+1}) - u(x)|^2 \mid X_n = x\,] \leq C\, u(x), \qquad \text{if } u(x) > L. \qquad (5.4)$$

Here, B_x^c denotes the complement of B_x and $I_{B_x^c}$ denotes the indicator function of the set B_x^c.

To prove that we can weaken the hypotheses as claimed, we proceed as follows. Observe that (4.5) implies $\varphi_n' \geq 0$, and so φ_n is a nondecreasing function on $[0, \infty)$. Hence, (5.1) implies

$$\int_{B_x} [\varphi_n(u(y)) - \varphi_n(u(x))]\, p(x, dy) \;\le\; 0.$$

From this and Taylor's theorem,

$$\begin{aligned}
\int_{\mathsf{X}} u_n(y)\, p(x, dy) \;&=\; u_n(x) + \int_{\mathsf{X}} [\varphi_n(u(y)) - \varphi_n(u(x))]\, p(x, dy) \\
&\le\; u_n(x) + \int_{B_x^c} [\varphi_n(u(y)) - \varphi_n(u(x))]\, p(x, dy) \qquad (5.5) \\
&\le\; u_n(x) + \int_{B_x^c} [\varphi_n'(u(x))(u(y) - u(x)) + K(u(y) - u(x))^2]\, p(x, dy).
\end{aligned}$$

If we replace (4.7) with (5.5), then the derivation given in Section 4 will prove the modification of Theorem 2 stated above.

6. Relation to Other Work

In this section we compare our results to those of Hajek (1982) for problems in which both apply. We point out that our results are restricted to nonnegative functions of time-homogenous Markov processes while Hajek's results apply to real-valued functions of non-Markov processes. The purpose of this section is to show that when considering a nonnegative function of a time-homogenous Markov process such that (2.2) holds, whenever Hajek's work applies, so does ours, and further, we can sometimes reach stronger conclusions.

To apply Hajek's results to $\{u(X_n)\}$ requires that the following hold.

There exist positive constants ν, η, ρ, and δ such that for some $L < \infty$,

$$\mathsf{E}[\, u(X_{n+1}) - u(x) \mid X_n = x\,] \;\le\; -\nu, \qquad \text{if } u(x) > L, \tag{6.1}$$

$$\mathsf{E}[\, e^{\eta(u(X_{n+1}) - u(x))} \mid X_n = x\,] \;\le\; \rho, \qquad \text{if } u(x) > L, \tag{6.2}$$

and

$$\mathsf{E}[\, e^{\eta(u(X_{n+1}) - L)} \mid X_n = x\,] \;\le\; \delta, \qquad \text{if } u(x) \le L. \tag{6.3}$$

Remarks. (i) If $\rho < 1$, applying Jensen's inequality with the natural logarithm function to (6.2) yields (6.1) with $-\nu = \frac{1}{\eta}\ln\rho$. (ii) Observe that if (6.1) and (6.2) hold for some L, they also hold for any $L' \ge L$. With regard to (6.3), rewrite (6.2) and (6.3) as

$$\mathsf{E}[\, e^{\eta u(X_{n+1})} \mid X_n = x\,] \;\le\; \rho e^{\eta u(x)}, \qquad \text{if } u(x) > L,$$

and

$$\mathsf{E}[\, e^{\eta u(X_{n+1})} \mid X_n = x\,] \;\le\; \delta e^{\eta L}, \qquad \text{if } u(x) \le L.$$

It follows that for any $L' \ge L$,

$$\mathsf{E}[\, e^{\eta u(X_{n+1})} \mid X_n = x\,] \;\le\; \max\{\delta e^{\eta L}, \rho e^{\eta L'}\}, \qquad \text{if } u(x) \le L'.$$

So, for all sufficiently large L',

$$\mathsf{E}[\, e^{\eta u(X_{n+1})} \mid X_n = x\,] \;\le\; \rho e^{\eta L'}, \qquad \text{if } u(x) \le L'.$$

In other words, if (6.1) – (6.3) hold for some L, they hold for all sufficiently large L (with $\delta = \rho$ for very large L).

Hajek proves by a simple induction argument that (6.2) and (6.3) imply

$$\mathsf{E}[\, e^{\eta u(X_n)} \mid X_0 = x\,] \;\le\; \rho^n e^{\eta u(x)} + \frac{1-\rho^n}{1-\rho}\delta e^{\eta L}. \tag{6.4}$$

To relate this to our results, apply Jensen's inequality to (6.4) to get

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; \frac{1}{\eta}\ln[\rho^n e^{\eta u(x)} + \frac{1-\rho^n}{1-\rho}\delta e^{\eta L}], \tag{6.5}$$

which yields both

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; \frac{1}{\eta}\ln[e^{\eta u(x)} + \frac{1}{1-\rho}\delta e^{\eta L}], \qquad \text{if } \rho < 1, \tag{6.6}$$

and

$$\overline{\lim_{n\to\infty}}\, \mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; \frac{1}{\eta}\ln\Big(\frac{\delta}{1-\rho}\Big) + L, \qquad \text{if } \rho < 1. \tag{6.7}$$

When our assumptions (see Section 2) hold, our theory yields (for $b = 1$ and $0 < \varepsilon < 1$)

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; \frac{u(x)^3}{[u(x) + \frac{\nu}{2}n]^2} + A + 4\frac{A^3}{\nu^2}\sum_{k=1}^{n-1}\frac{1}{k^2}. \tag{6.8}$$

$$\mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; u(x) + A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2), \tag{6.9}$$

and

$$\overline{\lim_{n\to\infty}}\, \mathsf{E}[\, u(X_n) \mid X_0 = x\,] \;\le\; A(1 + \tfrac{2}{3}(\tfrac{\pi A}{\nu})^2) \;<\; \infty. \tag{6.10}$$

Clearly, both theories yield very similar results. We now compare the assumptions which each theory requires in order to be applied. First, equation (6.1) is precisely our equation (2.3). Second, we will show below that (6.3) implies the existence of a *finite* constant A satisfying (2.5). Third, we will show below that if (2.2) holds, then (6.2) implies our equation (2.4) with a constant C satisfying (2.1). It follows that if there exist positive constants ν, η, ρ, δ, and $0 < \varepsilon < 1$ such that for some $L < \infty$, (6.1), (6.2), and (6.3) hold, and if we have $p(x, \cdot)$–a.s.,

$$|u(y) - u(x)| \;\le\; \varepsilon\, u(x), \qquad \text{if } u(x) > L,$$

then both (6.5) and (6.8) hold. However, if $\rho \ge 1$, the uniform bounds (6.6) and (6.7) are not available, while the uniform bounds (6.9) and (6.10) still hold.

The remainder of this section is devoted to establishing the preceding claims. Suppose that (6.3) holds. Rewrite it as

$$\int_{\mathsf{X}} e^{\eta u(y)}\, p(x, dy) \;\le\; \delta e^{\eta L}, \qquad \text{if } u(x) \le L.$$

Now, for any $b = 0, 1, 2, \ldots$, since $\eta u(y) \ge 0$ implies $\dfrac{[\eta u(y)]^{b+2}}{(b+2)!} \le e^{\eta u(y)}$, we can write

$$\int_{\mathsf{X}} u(y)^{b+2}\, p(x, dy) \;\le\; \frac{(b+2)!}{\eta^{b+2}} \delta e^{\eta L}, \qquad \text{if } u(x) \le L,$$

providing a finite constant A satisfying (2.5).

Suppose that (6.2) and (2.2) hold. Rewrite (2.2) as

$$(1-\varepsilon) u(x) \;\le\; u(X_{n+1}) \;\le\; (1+\varepsilon) u(x).$$

Taylor's theorem implies

$$e^{\eta(u(X_{n+1}) - u(x))} \;=\; 1 + \eta[u(X_{n+1}) - u(x)] + \frac{1}{2}\eta^2 e^{\eta v^*} (u(X_{n+1}) - u(x))^2,$$

where $(1-\varepsilon)u(x) \;\le\; v^* \;\le\; (1+\varepsilon)u(x)$. Thus,

$$e^{\eta(u(X_{n+1}) - u(x))} \;\ge\; 1 - \eta\varepsilon u(x) + \frac{1}{2}\eta^2 e^{\eta(1-\varepsilon)u(x)} (u(X_{n+1}) - u(x))^2.$$

So, (6.2) implies

$$\mathsf{E}[\, |u(X_{n+1}) - u(x)|^2 \mid X_n = x \,] \;\le\; [\eta\varepsilon u(x) - 1 + \rho] \frac{2}{\eta^2} e^{-\eta(1-\varepsilon)u(x)}.$$

Now, setting

$$a(v) \triangleq [\eta\varepsilon v - (1-\rho)]\frac{2}{\eta^2}e^{-\eta(1-\varepsilon)v},$$

we clearly have $\lim_{v\to\infty} \frac{a(v)}{v} = 0$. If we choose any positive C satisfying (2.1) we can then take L large enough so that whenever $v > L$, $a(v) \le C\,v$.

Acknowledgement

We thank Professor A. M. Makowski for his careful reading of an earlier version of this paper and his suggestions for improving its presentation.

References

BILLINGSLEY, P. (1979) *Probability and Measure.* John Wiley & Sons, New York.

BLANC, J. P. C. (1987) A note on waiting times in systems with queues in parallel. *J. Appl. Prob.* **24**, 540-546.

BRUMELLE, S. L. (1971) Some inequalities for parallel-server queues. *Opns. Res.* **19**, 402-413.

CONOLLY, B. W. (1984) The autostrada queueing problem. *J. Appl. Prob.* **21**, 394-403.

EPHREMIDES, A., VARAIYA, P., AND WALRAND, J. (1980) A simple dynamic routing problem. *IEEE Trans. on Automat. Contr.* **25**, 690-693.

FLATTO, L. AND MCKEAN, H. P. (1977) Two queues in parallel. *Comm. on Pure and Appl. Math.* **30**, 255-263.

FOSCHINI, G. J. AND SALZ, J. (1978) A basic dynamic routing problem and diffusion. *IEEE Trans. on Comm.* **26**, 320-327.

HAIGHT, F. A. (1958) Two queues in parallel. *Biometrika.* **45**, 401-410.

HAJEK, B. (1982) Hitting-time and occupation-time bounds implied by drift analysis with applications. *Adv. Appl. Prob.* **14**, 502-525.

HALFIN, S. (1985) The shortest queue problem. *J. Appl. Prob.* **22**, 865-878.

KINGMAN, J. F. C. (1961) Two similar queues in parallel. *Ann. Math. Statist.* **32**, 1314-1323.

PAPADIMITRIOU, I. (1973) A simple proof of the formula $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$. *American Math. Monthly.* **80**, 424-425.

STARK, E. L. (1974) The series $\sum_{k=1}^{\infty} k^{-s}, s = 2, 3, 4, \cdots$, once more. *Mathematics Mag.* **47**, 197-202.

WEBER, R. R. (1978) On the optimal assignment of customers to parallel servers. *J. Appl. Prob.* **15**, 406-413.

WINSTON, W. (1977) Optimality of the shortest line discipline. *J. Appl. Prob.* **14**, 181-189.

WOLFF, R. W. (1977) An upper bound for multi-channel queues. *J. Appl. Prob.* **14**, 884-888.

WOLFF, R. W. (1987) Upper bounds on work in system for multichannel queues. *J. Appl. Prob.* **24**, 547-551.