A Duality Relation for Busy Cycles in GI/G/1 Queues

Shun-Chen Niu
School of Management
The University of Texas at Dallas
P. O. Box 830688
Richardson, Texas 75083-0688

Robert B. Cooper
Department of Computer Science and Engineering
Florida Atlantic University
P. O. Box 3091
Boca Raton, Florida 33431-0991

October 1988
First Revision: November 1989
Second Revision: May 1990[1]

---

[1] *Queueing Systems*, 8 (1991), pp. 203–210.

Abstract

Using a generalization of the classical ballot theorem, Niu and Cooper [7] established a duality relation between the joint distribution of several variables associated with the busy cycle in $M/G/1$ (with a modified first service) and the corresponding joint distribution of several related variables in its dual $GI/M/1$. In this note, we generalize this duality relation to $GI/G/1$ queues with modified first services; this clarifies the original result, and shows that the generalized ballot theorem is superfluous for the duality relation.

AMS 1980 subject classification. Primary: 90B22; Secondary: 60K25.
IAOR 1973 subject classification. Main: Queues.
OR/MS Index 1978 subject classification. Primary: 681 Queues.

Key words. $GI/G/1$ queue, duality, busy cycle, modified first service.

1 The Duality Relation

Consider a $GI/G/1$ queue with interarrival-time distribution F and service-time distribution G, where we assume that the arrival rate $\lambda \equiv 1/\int_0^\infty [1 - F(x)]\,dx$ and the service rate $\mu \equiv 1/\int_0^\infty [1 - G(x)]\,dx$ are positive and finite, possibly with $\lambda > \mu$. Corresponding to this queue, which we call the "primal", one can define a "dual" queue where the roles of F and G are interchanged (Feller [5], Chapter XII). In this note, we derive a duality relation between the joint distribution of several variables associated with the busy cycle in the primal queue and that of several related variables in the dual busy cycle, generalizing an earlier duality relation in Niu and Cooper [7].

For the primal queue, let $K(F, G)$ and $B(F, G)$ be, respectively, the number of customers served in and the duration of a busy period with a modified first service of duration $\tilde{S}_1$, distributed as

$$P\{\tilde{S}_1 \le t\} = \mu \int_0^t [1 - G(x)]\,dx\,, \qquad t \ge 0\,. \tag{1}$$

Also, let $I(F, G)$ be the idle period that follows this modified busy period, if it terminates; if the modified busy period does not terminate, we leave $I(F, G)$ undefined, since our theorem below does not involve such a situation. We will interpret the variable $\tilde{S}_1$ as the stationary forward recurrence time of a renewal process with interevent-time distribution G, and denote its corresponding backward recurrence time by $A(G)$ (i.e., if $A(G) + \tilde{S}_1$ has distribution $\hat{G}$, then $\hat{G}(t) = \mu \int_0^t x\,dG(x)$).

Similarly for the dual queue, let $K(G, F)$, $B(G, F)$, and $I(G, F)$ be the corresponding variables associated with a busy cycle with a modified first service $\tilde{S}_1^*$, whose distribution is defined by replacing μ by λ and G by F in the right-hand side of (1). The corresponding backward recurrence time of $\tilde{S}_1^*$ will be denoted by $A(F)$.

We are now prepared to state our duality relation:

Theorem For $z \ge 0$, $n \ge 1$, $t \ge 0$, and $r \ge 0$,

$$\begin{aligned} &\lambda\, P\{A(G) \le z,\, K(F, G) = n,\, B(F, G) \le t,\, I(F, G) \le r\} \\ &\qquad = \mu\, P\{A(F) \le r,\, K(G, F) = n,\, B(G, F) \le t,\, I(G, F) \le z\}\,. \end{aligned} \tag{2}$$

When F is exponential, our theorem specializes, after letting $r \to \infty$, to the duality relation obtained in [7], p. 284, equation (12), because then the forward recurrence time of a service time in the dual queue is also distributed as F. (Applications of this result to $M/G/1$ and $GI/M/1$ queues, including a new constructive derivation of Beneš's formula for the waiting-time distribution in $M/G/1$-FIFO as well as derivations of formulas for the distribution of idle periods in $GI/M/1$, are discussed in that paper.) The present generalization is symmetric, in that (a) both the primal and the dual busy cycles have a modified first service, and (b) the variables $A(\cdot)$ and $I(\cdot)$ appear on both sides of (2); thus, it helps to clarify the nature of the earlier duality relation. In particular, we will show that the existence of this duality property does not depend on the generalized ballot theorem in [7], pp. 281–282, Lemma.

Prompted (apparently) by a pre-publication copy of [7], Fakinos [4] (p. 81, Theorem 2) obtained, independently, a related generalization of our duality relation in [7]. Our theorem differs from Fakinos's Theorem 2 in [4] in the following respects: Ours is transient and constructive, whereas his is derived by assuming that the primal queue is already in steady state at time 0, even when it is unstable. He also assumes that the queue discipline in the primal is preemptive-resume LIFO, to make (an interesting) connection to results in Fakinos [3], p. 763, Corollary 2, and Niu [6], p. 166, part (ii) of Proposition 1; our argument below is valid for all work-conserving queue disciplines.

2 Proof of Theorem

We consider first the primal queue. We assume that the first customer, with service time $\tilde{S}_1$, arrives at time 0, finding an empty system; and we denote the subsequent service times by $S_2,\ S_3,\ \cdots$ (iid with distribution G, independent of $\tilde{S}_1$). Successive interarrival times after time 0 will be denoted by $T_1,\ T_2,\ \cdots$ (iid with distribution F, independent of the service times).

The starting point of our proof is an expression for the joint probability $P\{A(G) \le z,\ K(F,\ G) = n,\ B(F,\ G) \le t,\ I(F,\ G) \le r\}$, obtained by successive conditioning, as follows: Consider the time interval $[0,\ t]$, $t > 0$ (the case $t = 0$ is trivial), and observe that in order for the event $\{K(F,\ G) = n,\ B(F,\ G) \le t\}$ to occur, we must have $\sum_{i=1}^{n-1} T_i < t$ and $\sum_{i=2}^{n} S_i \le t$. Therefore, we shall first condition on $\sum_{i=1}^{n-1} T_i = x$ and $\sum_{i=2}^{n} S_i = y$, with $0 \le x < t$ and $0 \le y \le t$. Next, given that $\sum_{i=1}^{n-1} T_i = x$ and $\sum_{i=2}^{n} S_i = y$, we must have $x < \tilde{S}_1 + y \le t$, in order for the n^{th} potential departure epoch (at time $\tilde{S}_1 + y$) to be after the n^{th} arrival epoch (at time x) and to be no later than time t; hence, we shall further condition on $\tilde{S}_1 = u$, with $(x - y)^+ < u \le t - y$, where $(x - y)^+$ is defined to be $\max[0,\ x - y]$. Lastly, given that $\sum_{i=1}^{n-1} T_i = x$, $\sum_{i=2}^{n} S_i = y$, and $\tilde{S}_1 = u$, we must have $u + y - x < T_n \le u + y - x + r$, in order for the busy period to contain no more than n service times and for $I(F,\ G)$ to not exceed r; and also note that the variable $A(G)$ is conditionally independent of $K(F,\ G)$ and $B(F,\ G)$, given $\tilde{S}_1 = u$. In light of these considerations, we now have

$$P\{A(G) \le z,\ K(F,\ G) = n,\ B(F,\ G) \le t,\ I(F,\ G) \le r\}$$

$$= \int_{x=0}^{t} \int_{y=0}^{t} \int_{u=(x-y)^+}^{t-y} \beta_n(x,\, y,\, u,\, u+y-x;\, F,\, G)\, P\{u+y-x < T_n \leq u+y-x+r\} \tag{3}$$

$$\times P\{A(G) \leq z \mid \tilde{S}_1 = u\}\, dP\{\tilde{S}_1 \leq u\}\, dG^{[n-1]}(y)\, dF^{[n-1]}(x)$$

($G^{[n]}$ denotes the n-fold self-convolution of a distribution function G), where we have defined (for $0 \leq x \leq t$, $0 \leq y \leq t$, and $(x-y)^+ \leq u \leq t-y$, with $x = t$ and $u = (x-y)^+$ included, since the distribution function (1) is absolutely continuous),

$$\beta_n(x,\, y,\, u,\, u+y-x;\, F,\, G)$$

$$= P\left\{ T_1 < u \text{ and } \sum_{i=1}^{k} T_i < u + \sum_{i=2}^{k} S_i \text{ for all } k = 2,\, \cdots,\, n-1 \mid \right. \tag{4}$$

$$\left. \sum_{i=1}^{n-1} T_i = x,\, \sum_{i=2}^{n} S_i = y,\, \tilde{S}_i = u, \text{ and } u+y-x < T_n \leq u+y-x+r \right\},$$

which, in words, is (a version of) the conditional probability of the event that arrivals in the interval $[0,\, u{+}y]$ occur “fast enough” (relative to the potential departure epochs) so that the interval constitutes a busy period. Finally, substitution of $P\{u{+}y{-}x < T_n \leq u{+}y{-}x{+}r\} = F(u+y-x+r) - F(u+y-x)$, $P\{A(G) \leq z \mid \tilde{S}_1 = u\} = [G(u+z) - G(u)] \,/\, [1 - G(u)]$, and $dP\{\tilde{S}_1 \leq u\} = \mu[1 - G(u)]\, du$ into (3) yields

$$P\{A(G) \leq z,\, K(F,\, G) = n,\, B(F,\, G) \leq t,\, I(F,\, G) \leq r\}$$

$$= \int_{x=0}^{t} \int_{y=0}^{t} \int_{u=(x-y)^+}^{t-y} \beta_n(x,\, y,\, u,\, u+y-x;\, F,\, G)\, [F(u+y-x+r) - F(u+y-x)] \tag{5}$$

$$\times \frac{G(u+z) - G(u)}{1 - G(u)}\, \mu[1 - G(u)]\, du\, dG^{[n-1]}(y)\, dF^{[n-1]}(x)\,.$$

We now turn our attention to a busy cycle in the dual queue, and let the notation for its interarrival and service times carry superscript $*$. If we condition on $\sum_{i=1}^{n-1} T_i^* = y$, $\sum_{i=2}^{n} S_i^* = x$, and $\tilde{S}_1^* = v$ (with $0 \leq y < t$, $0 \leq x \leq t$, and $(y-x)^+ < v \leq t-x$), then an argument similar to that leading to (5) yields

$$P\{A(F) \leq r,\, K(G,\, F) = n,\, B(G,\, F) \leq t,\, I(G,\, F) \leq z\}$$

$$= \int_{y=0}^{t} \int_{x=0}^{t} \int_{v=(y-x)^+}^{t-x} \beta_n(y,\, x,\, v,\, v+x-y;\, G,\, F)\, [G(v+x-y+z) - G(v+x-y)] \tag{6}$$

$$\times \frac{F(v+r) - F(v)}{1 - F(v)}\, \lambda[1 - F(v)]\, dv\, dG^{[n-1]}(y)\, dF^{[n-1]}(x)\,,$$

where the conditional probability $\beta_n(y,\, x,\, v,\, v+x-y;\, G, F)$ is defined similar to (4). After substituting $v = u + y - x$ (to change variable) and interchanging the order of integration for y and x, formula (6) can be rewritten as

$$\begin{aligned} &P\{A(F) \le r,\, K(G,\, F) = n,\, B(G,\, F) \le t,\, I(G,\, F) \le z\} \\ &\quad = \int_{x=0}^{t} \int_{y=0}^{t} \int_{u=(x-y)^+}^{t-y} \beta_n(y,\, x,\, u+y-x,\, u;\, G,\, F)\, [G(u+z) - G(u)] \qquad (7) \\ &\qquad \times [F(u+y-x+r) - F(u+y-x)]\, \lambda\; du\, dG^{[n-1]}(y)\, dF^{[n-1]}(x)\,. \end{aligned}$$

Comparison of the right-hand sides of (5) and (7) shows that our proof will be complete if we can establish

$$\beta_n(x,\, y,\, u,\, u+y-x;\, F,\, G) = \beta_n(y,\, x,\, u+y-x,\, u;\, G,\, F)\,. \qquad (8)$$

To prove (8), we consider a typical realization of the primal busy period, shown in Figure 1. Observe that if we look at the interval $[0,\, u+y]$ depicted in Figure 1 in reverse time and relabel each arrival epoch as a departure epoch and each departure epoch as an arrival epoch, then this (and every other) primal busy period is mapped, one-to-one, into a corresponding realization of a dual busy period (with time running backwards), shown in Figure 2; and this mapping can be taken as a proof for (8). Alternatively and more formally, if we substitute $T_1 = S_n^*,\ \cdots,\ T_{n-1} = S_2^*,\ S_2 = T_{n-1}^*,\ \cdots,$ and $S_n = T_1^*$, and replace $\tilde{S}_1 = u$ by $u < T_n^* \le u+z$ and $u+y-x < T_n \le u+y-x+r$ by $\tilde{S}_1^* = u+y-x$ (the latter replacements correspond to "relabeling" the end points of the busy period) in the right-hand side of (4), then (4) becomes

$$\begin{aligned} &\beta_n(x,\, y,\, u,\, u+y-x;\, F,\, G) \\ &\quad = P\left\{ S_n^* < u \text{ and } \sum_{i=n-k+1}^{n} S_i^* < u + \sum_{i=n-k+1}^{n-1} T_i^* \text{ for all } k = 2,\ \cdots,\ n-1 \;\middle|\right. \\ &\qquad \left. \sum_{i=2}^{n} S_i^* = x,\ \sum_{i=1}^{n-1} T_i^* = y,\ u < T_n^* \le u+z, \text{ and } \tilde{S}_1^* = u+y-x \right\}\ ; \end{aligned}$$

and since it is given that $\sum_{i=2}^{n} S_i^* = x$ and $\sum_{i=1}^{n-1} T_i^* = y$, the right-hand side of the last expression can be further rewritten as

$$\begin{aligned} &P\left\{ T_1^* < u+y-x \text{ and } \sum_{i=1}^{k} T_i^* < (u+y-x) + \sum_{i=2}^{k} S_i^* \text{ for all } k = 2,\ \cdots,\ n-1 \;\middle|\right. \\ &\qquad \left. \sum_{i=1}^{n-1} T_i^* = y,\ \sum_{i=2}^{n} S_i^* = x,\ \tilde{S}_1^* = u+y-x, \text{ and } u < T_n^* \le u+z \right\}\ , \end{aligned}$$

which, by interpreting the events in terms of the dual, is now identified with (i.e., is a version of) the right-hand side of (8). This completes our proof.

*** Figures 1 and 2 here. ***

3 Comments

(3.1) Our theorem can be used to derive a variety of relations between marginal probabilities for the primal and the dual. As examples, equations (25), (26), (14), (27), (15), and (28) (but not (16), unfortunately) in [7] generalize immediately to the $GI/G/1$ setting of this paper. In particular, we note that if $\lambda \leq \mu$ (i.e., if the dual queue is unstable), then equation (15) in [7] generalizes, by summing over n and letting $r \to \infty$ and $t \to \infty$ in our theorem (see Section 2.8 in [7]), to

$$P\{I(G,\, F) \leq z \mid B(G,\, F) < \infty\} = P\{A(G) \leq z\}\,, \tag{9}$$

where $P\{A(G) \leq z\}$ is given by the right-hand side of (1) (with z replacing t). According to (9), the distribution of the idle period (given its occurrence) in a dual busy cycle with a modified first service is insensitive to the form of F as long as $\lambda \leq \mu$, a rather remarkable conclusion (brought to our attention by Ward Whitt, and also noted in [4], p. 82, equation (14)).

(3.2) Formula (5), which, after letting $z \to \infty$, gives the joint distribution for the basic busy-cycle variables $K(F,\, G)$, $B(F,\, G)$, and $I(F,\, G)$, can be modified easily to yield similar formulas when the first service time in a busy period has an arbitrary distribution (i.e., not necessarily limited to the form (1)); consequently, it is perhaps of independent interest. However, its usefulness depends on whether we can evaluate explicitly the conditional probability $\beta_n(\cdot)$, a difficult task in general. For $M/G/1$ or $GI/M/1$, the joint probability $P\{K(F,\, G) = n,\, B(F, G) \leq t,\ I(F,\, G) \leq z\}$ can be evaluated using the Lemma in [7], pp. 281–282; see formulas (4) and (5) there.

(3.3) It is possible to relax some of our independence assumptions. For a fixed n, for example, our theorem remains valid as long as the variables $T_1,\, T_2,\, \cdots,\, T_{n-1}$ (possibly dependent) have the same joint distribution as that of $T_{n-1},\, T_{n-2},\, \cdots,\, T_1$, independently of all other variables (a similar statement can also be made for the service times); such assumptions are somewhat artificial, however. Our theorem does not appear to hold for $G/G/1$ queues in stationary frameworks.

(3.4) The general concepts of duality and time-reversal can be exploited to obtain other duality properties (Bhat [2], pp. 45–46; Prabhu [8], p. 31, Theorems 12 and 13; and

Takács [10], p. 97). In two recent papers (again, brought to our attention by Ward Whitt), Ramaswami [9] and Asmussen and Ramaswami [1] also used time-reversal (as well as "sign reversal") to obtain results that (loosely put) relate the distributions of first-passage times from "level 1" (i.e., having 1 customer in the system) to level 0 in a Markov renewal process embedded after departure epochs of a primal queue with phase-type arrival to the corresponding first-passage-time distributions from level 0 to level 1 in a Markov renewal process embedded prior to arrival epochs of the dual (with phase-type service). The spirit of our duality theorem differs from theirs, we believe, in the following important way: The first-passage-time distributions in the primal and the dual that we relate are both from level 1 to level 0.

Acknowledgments We are very grateful to Søren Asmussen for observing that the original version of our generalization could be improved to its present, perfectly symmetric, form; and to Ward Whitt, the Associate Editor, for reading and commenting on an earlier version of our paper.

References

[1] S. Asmussen and V. Ramaswami, "Probabilistic Interpretations of Some Duality Results for the Matrix Paradigms in Queueing Theory," preprint (1989).

[2] U. N. Bhat, *A Study of the Queueing Systems* $M/G/1$ *and* $GI/M/1$ (Springer-Verlag, 1968).

[3] D. Fakinos, "The Single-Server Queue with Service Depending on Queue Size and with the Preemptive-Resume Last-Come-First-Served Queue Discipline," *Journal of Applied Probability*, **24** (1987), 758–767.

[4] D. Fakinos, "Duality Relations for Certain Single Server Queues," *Queueing Systems*, **4** (1989), 77–83.

[5] W. Feller, *An Introduction to Probability Theory and Its Applications, Vol. II*, 2nd ed. (Wiley, New York, 1971).

[6] S.-C. Niu, "Representing Workloads in $GI/G/1$ Queues Through the Preemptive-Resume LIFO Queue Discipline," *Queueing Systems*, **3** (1988), 157–178.

[7] S.-C. Niu and R. B. Cooper, "Duality and Other Results for $M/G/1$ and $GI/M/1$ Queues, via a New Ballot Theorem," *Mathematics of Operations Research*, **14** (1989), 281–293.

[8] N. U. Prabhu, *Stochastic Storage Processes: Queues, Insurance Risk, and Dams* (Springer-Verlag, New York, 1980).

[9] V. Ramaswami, “A Duality Theorem for the Matrix Paradigms in Queueing Theory,” *Stochastic Models*, **6** (1990), 151–161.

[10] L. Takács, *Combinatorial Methods in the Theory of Stochastic Processes* (Wiley, New York, 1967).