The Foreground-Background Queue
Nuijens, M.F.M.

Citation for published version (APA):
Nuijens, M. F. M. (2004). The Foreground-Background Queue

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
In this chapter we examine the effect of more variability in the service-time distribution on the expected queue length in the stationary FB queue. We show that in general a more variable service-time does not necessarily lead to a smaller expected queue length, but that for certain classes of service times it does.

6.1 Introduction

In this chapter we study the effect of a more variable service-time distribution on the mean queue length in the stationary M/G/1 FB queue. The queue length is understood as the total number of customers in the system. To measure the variability of a distribution we use the convex order. Recall from Definition 4.1 that a random variable $X$ is convexly smaller than a random variable $Y$ (notation $\leq_{cx}$), if $Eh(X) \leq Eh(Y)$ for all convex functions $h$ for which the expectations exist. If
X \leq_{sv} Y$, then we say that $X$ is more variable than $Y$.

In the stationary $M/G/1$ FIFO queue the expected queue length increases in the variability of the service times. By the Pollaczek-Khintchin mean value formula, see e.g. Kleinrock [31], the stationary queue length has mean $\rho + \lambda^2EB^2/(2(1 - \rho))$. For $EB$ and $\lambda$ fixed, by the arguments given after Definition 4.1, a more variable service-time distribution has a larger variance, and hence the expected queue length in the FIFO queue is larger.

The distribution of the queue length in the stationary $M/G/1$ PS queue depends on the service-time distribution only through its first moment. The PS discipline is in some sense in between FIFO and FB: it does not give priority to any type of customer, whereas FIFO gives priority to old customers and FB to young ones. Hence one might wonder whether the stationary queue length in the FB queue decreases if the variability of the service-time distribution increases. Yashkov [69] states that this is indeed the case, though without proof or explanation. A conjecture of the same flavour is found on page 189 of Coffman and Denning [12].

In this chapter we show that the expected queue length does not necessarily increase when the variability of the service times increases. However, if the service times are picked from certain classes of distributions, then such a relation does hold, in the following sense. Consider the classes DMRL and IMRL of random variables with a decreasing and an increasing mean residual life respectively, see Definition 2.1. Examples of IMRL distributions are Pareto distributions, certain Gamma distributions, and certain Weibull distributions. Certain other Gamma distributions, uniform distributions and deterministic distributions are in the class DMRL. For these classes the following theorem holds.

**Theorem 6.1** Consider four stationary $M/G/1$ FB queues with arrival rate $\lambda$ and the same expected service time. The service times have an IMRL distribution, an exponential distribution, a DMRL distribution, and a deterministic distribution. Let $Q_{IMRL}$, $Q_{exp}$, $Q_{DMRL}$ and $Q_{det}$ denote the stationary queue lengths in the four queues. Then

$$EQ_{IMRL} \leq EQ_{exp} \leq EQ_{DMRL} \leq EQ_{det} = \frac{(2 - \rho)\rho}{2(1 - \rho)^2}.$$  

Below we show that IMRL distributions are more variable than DMRL distributions. Hence in this situation the more variable service-time distribution yields the smaller expected stationary queue length.

The chapter is organised as follows. Section 6.2 contains the proof of Theorem 6.1. In Section 6.3 we examine the effect of making an arbitrary service-time distribution
6.2 Comparing queue lengths: two classes

In this section we compare the value of $EQ$ for two classes of service-time distributions, namely the classes DMRL and IMRL described above. Assume $\lambda$ and $EB$ are fixed. For proving the main theorem of this section, we first cite two well-known theorems. The first theorem may be found on page 11 in Kelly [29]. Note that there the service discipline is not mentioned.

**Theorem 6.2** Consider an $M/M/1$ queue, empty at time $0^-$, with service discipline $\pi$. Let $Q(t)$ be the queue length at time $t$. If $\pi \in D$, then the distribution of $Q(t)$ does not depend on $\pi$.

**Proof** Since $\pi \in D$, the remaining service time of a customer is unknown. Since the service times are exponential, the residual lifes of all customers present in the queue are exponentially distributed as well. Hence the queue length $Q(t)$ is a birth and death process with transition rates that do not depend on the service discipline. $\square$

The following theorem states that the distribution of the number of customers in the stationary $M/G/1$ PS queue only depends on the distribution of the generic service time $B$ through its first moment.

**Theorem 6.3** In a stationary $M/G/1$ queue with the PS service discipline, the stationary number of customers $Q$ has the distribution

$$P(Q = k) = (1 - \rho)\rho^k, \quad k = 0, 1, \ldots$$

where $\rho = \lambda EB$.

For the proof of this theorem see e.g. Cohen [15], Kelly [29] or Yashkov [66].

Theorem 6.3 yields that in two $M/G/1$ PS queues with the same arrival rate and expected service time, the distributions of the stationary queue length are equal. By combining this idea with Theorem 2.2, we prove Theorem 6.1.

**Proof of Theorem 6.1** Let $Q_{IMRL}^{PS}$, $Q_{exp}^{PS}$ and $Q_{DMRL}^{PS}$ be the queue lengths in queues with the same IMRL, exponential and DMRL service-time distribution as
the queues in the statement of the theorem, but with the PS discipline instead of the FB discipline. Then Theorems 2.2, 6.3 and 6.2 yield

$$EQ_{IMRL} \leq EQ_{PS}^{IMRL} = EQ_{exp}^{PS} = \frac{\rho}{1 - \rho} = EQ_{exp}.$$ 

Further, Theorem 6.3. Theorem 2.2 and Corollary 9.8 below yield

$$EQ_{exp}^{PS} = EQ_{DMRL}^{PS} \leq EQ_{DMRL} \leq \frac{\rho(2 - \rho)}{2(1 - \rho)^2}.$$

which finishes the proof.

To interpret Theorem 6.1, we show that a IMRL distribution is more variable than an DMRL distribution in the sense of the convex order, given by Definition (4.1). If a random variable $X \in DMRL$, then certainly $EX \geq E[X - t | X > t]$ for all $t \geq 0$. Proposition 6.1.2 in Stoyan [63] yields that there exists an exponentially distributed random variable $Z$ with mean $EX$, such that $X \leq_{cr} (\geq_{cr}) Z$. Since the relation $\leq_{cr}$ is transitive, which readily follows from Definition 4.1, the following lemma holds.

**Lemma 6.4** If $X \in DMRL, Y \in IMRL$ and $EX = EY$, then $X \leq_{cr} Y$.

Hence Theorem 6.1 states that if the service-time distributions in two queues are from the classes IMRL and DMRL, then the more variable service times yield a smaller expected queue length. Furthermore, many heavy-tailed distributions belong to the class IMRL, and many light-tailed distributions belong to the class DMRL. Hence Theorem 6.1 is another situation where the FB queue behaves better when the tail of the service-time distribution is heavier.

Next we show that Theorem 6.1 is a refinement of a conjecture by Coffman and Denning [12] that was proven to be false. Let $X$ be a random variable with $EX \neq 0$. Then $C_X = \sqrt{\text{Var}(X)}/EX$ is called the coefficient of variation of $X$.

**Corollary 6.5** If $X \in DMRL, Y \in IMRL$ and $EX = EY$, then $C_X \leq C_Y$.

Coffman and Denning [12] conjectured that a service-time distribution whose coefficient of variation is larger than 1, yields a smaller expected sojourn time than distributions with the same mean and coefficient of variation smaller than 1. Wierman et al. [65] showed that this conjecture does not hold. From Corollary 6.5 and Theorem 6.1 it follows that the queue with the larger coefficient of variation has the smaller expected queue length. By using Little’s law $EQ = \lambda EV$, this may be considered as a refinement of the conjecture.

---

1 The proof of Theorem 2.2 seems to contain an error, see Aalto et al. [1].
6.3 Convexly disturbed service times

Let $Q$ be the stationary queue length in the $M/G/1$ FB queue with service-time distribution $F$. In this section we examine the effect on $EQ$ of making the service-time distribution $F$ more variable. Assume $F$ has a density $f$. To make $F$ more variable we add to the density $f$ a disturbance function $\eta$ around a given point $y > 0$ at which the density is positive. The disturbance function has the form

$$\eta(x) = \eta_{y, \delta, \varepsilon}(x) = \varepsilon \mathbb{I}_{(y-2\delta, y-\delta)} - \varepsilon \mathbb{I}_{(y-\delta, y+\delta)} + \varepsilon \mathbb{I}_{(y+\delta, y+2\delta)},$$  

(6.1)

for some $0 < \delta < y/2$ such that $f$ is strictly positive on $[y - \delta, y + \delta]$. Here $\varepsilon > 0$ is so small that $f + \eta \geq 0$. Quantities with a tilde ("~") refer to the queue with service-time density $f + \eta$. It follows immediately that $\tilde{f} - f = \eta$ changes sign twice.

By Theorem 2.A.17 in Shaked and Shanthikumar [60], the distribution $\tilde{F}$ is more variable (in the convex order) than the distribution $F$. The disturbance function $\eta$ is chosen to be as in (6.1) because of its simple shape.

In the remainder of this chapter we prove the following proposition, which describes the effect of the disturbance on the stationary queue length.

**Proposition 6.6** For every $\chi \in \{-1, 1\}$ there is a service-time distribution $F$ such that

$$\lim_{\delta, \varepsilon \to 0} \text{sign}(EQ - E\tilde{Q}) = \chi.$$

Hence the queue with the more variable service-time distribution does not necessarily have a larger expected queue length.

6.3.1 Calculus

For proving Proposition 6.6 the following technicalities are needed. We examine the integrals in the expressions for $EV(x)$ and $E\tilde{V}(x)$ given by (2.11). Note that $F$ and $\tilde{F}$ coincide on $\mathbb{R} - (y - 2\delta, y + 2\delta)$ and that $F(y) = \tilde{F}(y)$. The $O$-symbols in this section hold for $\varepsilon, \delta \downarrow 0$. Let $\rho_k(x) = \lambda E(B \land x)^k$ and $\rho_1(x) = \rho(x)$.

**Lemma 6.7** For $x \notin (y - 2\delta, y + 2\delta)$ we have $\rho(x) = \tilde{\rho}(x)$. Further $\rho(y) - \tilde{\rho}(y) = \lambda \varepsilon \delta^2$ and

$$\rho_2(x) - \tilde{\rho}_2(x) = \begin{cases} 
0 & x < y - 2\delta \\
2\lambda \varepsilon \delta^2(y - \delta) & x = y \\
-4\lambda \varepsilon \delta^3 & x > y + 2\delta.
\end{cases}$$
**Proof** Since $\tilde{F}(y) = F(y)$, $k \geq 1$ we write out

$$
\lambda^{-1}(\rho_k(y) - \tilde{\rho}_k(y)) = 
= \int_0^y t^k f(t)dt + y^k (1 - F(y)) - \int_0^y t^k (f(t) + \eta(t))dt - y^k (1 - \tilde{F}(y))
= -\int_0^y t^k \eta(t)dt = -\varepsilon \int_{y-2\delta}^{y-\delta} t^k dt + \varepsilon \int_{y-\delta}^{y} t^k dt.
$$

The equalities $\rho(y) - \tilde{\rho}(y) = \lambda \varepsilon \delta^2$ and $\rho_2(y) - \tilde{\rho}_2(y) = 2\lambda \varepsilon \delta^2 (y - \delta)$ then follow from easy calculations.

Now let $x > y + 2\delta$. Then $\tilde{F}(x) = F(x)$ and

$$
\varepsilon^{-1}\lambda^{-1}(\rho_k(x) - \tilde{\rho}_k(x)) = -\varepsilon^{-1} \int_0^x t^k \eta(t)dt = -\int_{y+2\delta}^{y+\delta} t^k dt + \int_{y-\delta}^{y} t^k dt - \int_{y-2\delta}^{y-\delta} t^k dt.
$$

The equality $\rho_2(x) - \tilde{\rho}_2(x) = -4\lambda \varepsilon \delta^3$ follows from computing the integrals. The proof is concluded by noting that $\rho(x) - \tilde{\rho}(x) = 0$ for $x < y - 2\delta$ since $F$ and $\tilde{F}$ coincide on $(0, y - 2\delta)$.

The calculus used to prove Proposition 6.6 is based partly on a Taylor series expansion of $E\tilde{V}(x)$. For evaluating the constant term we use the following relation between $E\tilde{V}(y)$ and $EV(y)$. The precision of our calculations in this section is $O(\varepsilon \delta^3)$. With use of expression (2.11) and Lemma 6.7 we calculate

$$
E\tilde{V}(y) = \frac{\tilde{\rho}_2(y)}{2(1 - \tilde{\rho}(y))^2} + \frac{y}{1 - \tilde{\rho}(y)}
= \frac{\rho_2(y) - 2\lambda \varepsilon \delta^2 y}{2(1 - \rho(y)) + \lambda \varepsilon \delta^2} + \frac{y}{1 - \rho(y)} + \lambda \varepsilon \delta^2 + O(\varepsilon \delta^3)
= \frac{\rho_2(y) - 2\lambda \varepsilon \delta^2 y}{2(1 - \rho(y))^2} \left(1 - \frac{2\lambda \varepsilon \delta^2}{1 - \rho(y)}\right) + \frac{y}{1 - \rho(y)} \left(1 - \frac{2\lambda \varepsilon \delta^2}{1 - \rho(y)}\right) + O(\varepsilon \delta^3)
= \frac{\rho_2(y)}{1 - \rho(y)} \left(1 - \frac{2\lambda \varepsilon \delta^2}{1 - \rho(y)}\right) + O(\varepsilon \delta^3).
$$

By similar but tedious calculus, it follows that the derivatives of $EV(x) - E\tilde{V}(x)$ are of order $\varepsilon \delta^2$ as well. Below we use

$$
\int_{y-2\delta}^{y+2\delta} (x - y)^k dx = \delta^{k+1} \frac{2^{k+1} - (-2)^{k+1}}{k + 1}, \quad k = 0, 1, \ldots
$$

Note that the expression (6.3) is zero if $k$ is odd. Hence, for integrating the Taylor expansion around $y$ of order two of the expression $EV(x) - E\tilde{V}(x)$, only the constant
term and the rest term are important. Using (6.2) and the remark below it, we find
\[ \int_{y-2\delta}^{y+2\delta} (EV(x) - \hat{EV}(x)) dF(x) = 4\delta(EV(y) - \hat{EV}(y)) f(y) + O(\varepsilon\delta^4) \]
\[ = \frac{8\lambda \varepsilon \delta^3}{1 - \rho(y)} f(y) EV(y) + O(\varepsilon\delta^4). \] (6.4)

For odd \( k \)
\[ \int_{y-2\delta}^{y+2\delta} (x - y)^k \eta(x) dx = \int_{-2\delta}^{2\delta} x^k \eta(x + y) dx = 0. \] (6.5)
since the integrand is an odd function. Since \( \int_{y-2\delta}^{y+2\delta} \eta(x) dx = 0 \), integrating the Taylor expansion around \( y \) of order four of \( EV(x) \) breaks down to integrating the second order term and the rest term. Since \( |x - y|^4 \leq 16\delta^4 \), straightforward integration gives
\[ \int_{y-2\delta}^{y+2\delta} EV(x) \eta(x) dx = 2\varepsilon\delta^3 \frac{d^2 EV(y)}{dy^2} + O(\varepsilon^5) = 2\varepsilon\delta^3 \frac{d^2 EV(y)}{dy^2} + O(\varepsilon\delta^4). \] (6.6)

Here the second equality follows from the remark below (6.2).

### 6.3.2 More calculus

In the next proposition we collect the pieces. Then Proposition 6.6 is proved.

**Proposition 6.8** As \( \varepsilon, \delta \to 0 \),
\[ EV - \hat{EV} = 2\varepsilon\delta^3 \left( \frac{4\lambda f(y) EV(y)}{1 - \rho(y)} \right) - \lambda \int_y^\infty (1 - \rho(x))^{-2} dF(x) - \frac{d^2 EV(y)}{dy^2} + O(\varepsilon\delta^4). \]

**Proof** First, since \( F \) and \( \hat{F} \) coincide on \((0, y - 2\delta)\)
\[ \int_0^{y-2\delta} EV(x) dF(x) - \int_0^{y-2\delta} \hat{EV}(x) d\hat{F}(x) = 0. \] (6.7)
Secondly by (6.4) and (6.6)
\[ \int_{y-2\delta}^{y+2\delta} EV(x) dF(x) - \int_{y-2\delta}^{y+2\delta} \hat{EV}(x) d\hat{F}(x) = \]
\[ = \frac{8\lambda \varepsilon \delta^3}{1 - \rho(y)} f(y) EV(y) - 2\varepsilon\delta^3 \frac{d^2 EV(y)}{dy^2} + O(\varepsilon\delta^4). \] (6.8)
Finally, since \( F(x) = \hat{F}(x) \) for \( x > y + 2\delta \), we have
\[ \int_y^{\infty} EV(x) dF(x) - \int_y^{\infty} \hat{EV}(x) d\hat{F}(x) = \int_{y+2\delta}^{\infty} (EV(x) - \hat{EV}(x)) dF(x). \]
Since $\rho(x) = \rho(x)$ for $x > y + 2\delta$, by Lemma 6.7, this becomes
\[
\int_{y+2\delta}^{\infty} \frac{\rho_2(x) - \tilde{\rho}_2(x)}{2(1 - \rho(x))^2} dF(x) = -2\lambda \delta \int_{y+2\delta}^{\infty} (1 - \rho(x))^{-2} dF(x). \tag{6.9}
\]
again by Lemma 6.7. Equations (6.7), (6.8) and (6.9) yield the desired result. \hfill \square

From Proposition (6.8) we conclude that, letting $\varepsilon \downarrow 0$ and $\delta \downarrow 0$ in arbitrary order,
\[
\lim_{\varepsilon, \delta \downarrow 0} \text{sign}(EV - \tilde{EV}) = \text{sign}(\Delta). \tag{6.10}
\]
where $\Delta$ is given by
\[
\Delta = \frac{4f(y)EV(y)}{1 - \rho(y)} - \int_{y}^{\infty} \frac{1}{(1 - \rho(x))^2} dF(x) - \frac{1}{\lambda} \frac{d^2EV(y)}{dy^2}. \tag{6.11}
\]
Differentiating the derivative of $EV(x)$, which is given by (2.15), yields
\[
\frac{d^2EV(x)}{dx^2} = \frac{\lambda(1 - F(y))}{(1 - \rho(y))^2} \left(2 + 6\lambda y(1 - F(y)) + \frac{3\lambda^2(1 - F(y))\rho_2(y)}{(1 - \rho(y))^2}\right) - \frac{\lambda y f(y)}{(1 - \rho(y))^2} - \frac{\lambda^2 f(y) \rho_2(y)}{(1 - \rho(y))^3}.
\]
Since $1 - \rho \leq 1 - \rho(x) \leq 1 - \rho(y)$ for $x \in [y, \infty)$, we have
\[
\frac{1 - F(y)}{(1 - \rho(y))^2} \leq \int_{y}^{\infty} \frac{1}{(1 - \rho(x))^2} dF(x) \leq \frac{1 - F(y)}{(1 - \rho)^2}.
\]
Using the expression (2.11) for $EV(x)$ yields
\[
\Delta = \left(f(y)c_1 - (1 - F(y))(c_2 + c_3)\right) \frac{1}{(1 - \rho(y))^2}, \tag{6.12}
\]
where $c_1 = 3\lambda \rho_2(y)(1 - \rho(y))^{-1} + 5y$, $1 \leq c_3 \leq (1 - \rho(y))^2(1 - \rho)^{-2}$ and
\[
c_2 = 2 + \frac{6\lambda y(1 - F(y))}{1 - \rho(y)} + \frac{3\lambda^2(1 - F(y))\rho_2(y)}{(1 - \rho(y))^2}.
\]

**Proof of Proposition 6.6** By (6.12) it may be seen that the quantity $\Delta$ in equation (6.11) is positive when $f(y) \gg 1 - F(y)$, and negative when $f(y) \ll 1 - F(y)$. Using (6.10) and applying Little’s law completes the proof. \hfill \square

It would be interesting to find out whether Proposition 6.6 holds as well for distributions with monotone densities, or log-convex densities. We conjecture that in the latter case for two queues with convexly ordered service times, the more variable service-time distribution always produces the smaller mean queue length.