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.


Chapter 7

Sojourn time asymptotics

In this chapter the tail behaviour of the sojourn time $V$ is studied, both for heavy-tailed and light-tailed service times. For light-tailed distributions we show that the sojourn time and the busy period length have the same decay rate. The decay rate is defined in (7.1). For a class of heavy-tailed service-time distributions, Núñez Queija [41] showed that the tail of the sojourn-time distribution and the service-time distribution are of the same type, in the sense of Theorem 7.14. We prove this theorem for a broader class of distributions. The first part of this chapter is based on the joint paper [36] with Mandjes.

7.1 Introduction

The sojourn time of a customer, i.e. the time between his arrival and departure, is an often used performance measure for queues. In this chapter we study the tail of the sojourn-time distribution of the stationary M/G/1 queue with the Foreground-Background (FB) discipline.
Let $V$ denote the sojourn time of a customer in the stationary $M/G/1$ FB queue. Núñez Queija [41] showed that for service-time distributions with regularly varying tails of index $\eta \in (1, 2)$, the distribution of $V$ satisfies
\[
P(V > x) \sim P(B > (1 - \rho)x), \quad x \to \infty.
\] (7.1)
where $\rho < 1$ is the load of the system and $B$ is the generic service time. Using Núñez Queija’s method, in the second part of this chapter we prove (7.1) under weaker assumptions. In case of regularly varying service times the tail behaviour of $V$ under other disciplines, like FIFO, LIFO, PS and SRPT, has been found to be the same or worse, see Borst, Boxma, Núñez Queija and Zwart [10].

Additional support for the optimality of FB under heavy tails is given by Righter and Shanthikumar [50, 51, 53]. They show that for certain classes of service times the FB discipline minimises the queue length, measured in number of customers. For light-tailed service times, however, the FB discipline does not perform so well. For exponential service times (and also for a subclass of Gamma-distributed service times) FB still minimises the queue length, but for service times with a log-concave density the queue shows opposite behaviour and the queue length is maximal, see Righter and Shanthikumar [50], [51], [53].

This undesirable behaviour of the FB discipline is very pronounced for deterministic service times. In this extreme case in the FB queue all customers stay till the end of the busy period, and the sojourn time under the FB discipline is maximal in the class of all work-conserving disciplines. In the first part of this chapter, we consider the (asymptotic) decay rate of the sojourn time, defined as follows.

**Definition 7.1 (Decay rate)** The (asymptotic) decay rate of a random variable $X$ is
\[
- \lim_{x \to \infty} x^{-1} \log P(X > x).
\]

In the M/D/1 FB queue the sojourn time of a customer is stochastically larger than the length of a busy period. Hence the decay rate of the sojourn time is smaller than that of the busy-period length. In fact, as we shall see below, the decay rate of the sojourn time is equal to that of the busy period length.

It turns out that for the M/G/1 FB queue this decay-rate property holds true for all service-time distributions with an exponentially fast decreasing tail, in the following sense. Assume that the service times have a finite exponential moment, or equivalently, the Laplace transform is analytic in a neighbourhood of zero. The main theorem in the first part of this chapter is then the following.

**Theorem 7.2** Let $V$ be the sojourn time of a customer in the stationary $M/G/1$ FB queue, and let $L$ be the length of a busy period. If the service-time distribution
has a finite exponential moment, then the following limits exist and

\[ \lim_{x \to \infty} \frac{1}{x} \log P(V > x) = \lim_{x \to \infty} \frac{1}{x} \log P(L > x). \]  

(7.2)

It is shown below that the decay rate of the sojourn time in an M/G/1 queue with any work-conserving discipline is bounded by the decay rate of the residual life of a busy period. For service times with an exponential moment the latter decay rate is equal to that of a normal busy period. Hence (7.2) is the lowest possible decay rate for a work-conserving discipline. Using the decay rate of \( V \) as a criterion to measure the performance of a service discipline then leads to the following conclusion. For service times with an exponential moment, the FB discipline is the worst discipline in the class of work-conserving disciplines w.r.t. the decay rate of the sojourn time.

The chapter is organised as follows. In Section 7.2 we present the notation, some preliminaries, and show that the decay rate of the sojourn time under FB is pessimal, as described above. In Section 7.3 Theorem 7.2 for the decay rate is proved. Section 7.4 discusses the result and the decay rate of the sojourn time in queues operating under other service disciplines. In Section 7.5 a stronger version of Nuñez Queija’s result (7.1) is stated and discussed. Section 7.6 provides the necessary tools for proving this stronger result, which is done in Section 7.7. Section 7.8 contains some technicalities that are used in Section 7.6.

### 7.2 Preliminaries

Throughout this first part of the chapter we assume that the generic service time \( B \) with distribution function \( F \) in the M/G/1 queue satisfies the following assumption.

**Assumption 7.3** The generic service time \( B \) has an exponential moment, i.e., \( E \exp(\gamma B) < \infty \) for some \( \gamma > 0 \).

Let in addition the stability condition \( \rho = \lambda B < 1 \) hold, where \( \lambda \) is the rate of the Poisson arrival process. The proofs in this chapter rely on some properties of the busy-period length \( L \) and related random variables, which we derive in this section.

Under assumption 7.3, Cox and Smith [17] have shown that the tail of \( L \) satisfies \( P(L > x) \sim b x^{-3/2} e^{-c x} \) for certain constants \( b, c > 0 \). In particular, \( L \) has decay rate \( c \). In fact, by expression (46) on page 154 of Cox and Smith [17], \( c = \lambda - \zeta - \lambda g(\zeta) \), where \( g \) is the Laplace transform of the service-time distribution, and \( \zeta < 0 \) is such that \( g'(\zeta) = -\lambda^{-1} \). Hence \( \zeta \) is the root of the derivative of the function \( m(x) = \lambda - x - \lambda g(x) \). Since \( m(x) \) attains its maximum in the point \( \zeta \), we may write
c in terms of the Legendre transform of $B$ as follows.

$$c = \lim_{x \to \infty} \frac{1}{x} \log P(L > x) = \sup_{\theta} \{\theta - \lambda(Ee^{\theta B} - 1)\}. \quad (7.3)$$

**Remark** Consider a Poisson stream, with intensity $\lambda$, of i.i.d. jobs, where every job is distributed according to the random variable $B$. Let $A(x)$ denote the amount of work generated in an arbitrary time window of length $x$. It is an easy corollary of Cramér’s theorem on large deviations that

$$\lim_{x \to \infty} \frac{1}{x} \log P(A(x) > x) = - \sup_{\theta} \{\theta - \log Ee^{\theta A(1)}\}. \quad (7.4)$$

Noting that

$$Ee^{\theta A(1)} = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} (Ee^{\theta B})^k = \exp(\lambda(Ee^{\theta B} - 1)),$$

we observe from (7.3) and (7.4) and $P(A(x) > x)$ have the same decay rate. This is somewhat surprising, as the event $\{L > x\}$ relates to the arrival pattern within an interval of length $x$ (and, in fact, also the jobs already present at time 0), rather than just $A(x)$.

In renewal theory the notion of residual life, also known as excess or forward-recurrence time, is standard. Let $\tilde{L}$ be the residual life of a busy period, cf. Section 7.9. Then $P(\tilde{L} > x) = (EL)^{-1} \int_x^\infty P(L > y)dy$, see for instance Cox [16], or Section 7.9. Using standard calculus we find

$$\lim_{x \to \infty} \frac{1}{x} \log P(\tilde{L} > x) = \lim_{x \to \infty} \frac{1}{x} \log \int_x^\infty y^{-3/2} e^{-cy} dy = -c. \quad (7.5)$$

Hence $\tilde{L}$ has the same decay rate as $L$.

Another ingredient used in the proofs below is the M/G/1 queue with truncated generic service time $B \wedge \tau$, $\tau > 0$. Call this the $\tau$-queue and let $L(\tau)$ denote the length of a busy period (a $\tau$-busy period) in this queue. Let $\tilde{L}(\tau)$ be its residual life and define $L^*(\tau)$ to be the length of a $\tau$-busy period that starts with a customer with service time at least $\tau$, i.e. $L^*(\tau) = [L(\tau) | B_1 \geq \tau]$, where $B_1$ is the first service time in the busy period. These random variables satisfy the following relation.

**Lemma 7.4** For $\tau > 0$, the random variables $L(\tau), \tilde{L}(\tau)$ and $L^*(\tau)$ have the same decay rate $c(\tau)$, where

$$c(\tau) = \lim_{x \to \infty} \frac{1}{x} \log P(L(\tau) > x) > 0.$$
Let $B_1$ denote the first service time in the busy period, hence $B_1 \overset{d}{=} B$. Then for $\tau$ such that $P(B \geq \tau) > 0$, 

\[
P(B_1 \geq \tau)P(L^*(\tau) > x) = P(L(\tau) > x \mid B_1 \geq \tau)P(B_1 \geq \tau) \\
= P(L(\tau) > x, B_1 \geq \tau) \leq P(L(\tau) > x) \\
\leq P(L(\tau) > x \mid B_1 \geq \tau) = P(L^*(\tau) > x),
\]

where the second inequality follows from the observation that the value of $B_1$ is maximal in the $\tau$-queue, since all service times are bounded by $\tau$. Since (7.5) holds as well for the queue with truncated (generic) service time $B \wedge \tau$, the decay rates of $L(\tau)$ and $\hat{L}(\tau)$ are equal. Hence for every $\tau$ such that $P(B \geq \tau) > 0$, we have by (7.5) and (7.6) that

\[
0 < c(\tau) = -\lim_{x \to \infty} \frac{1}{x} \log P(L(\tau) > x) = -\lim_{x \to \infty} \frac{1}{x} \log P(L^*(\tau) > x) \\
= -\lim_{x \to \infty} \frac{1}{x} \log P(\hat{L}(\tau) > x),
\]

which was to be shown. \qed

In this chapter we need the following lemma about the decay rate of the sum of two independent random variables.

**Lemma 7.5** If two non-negative independent random variables $X$ and $Y$ satisfy

\[
\lim_{x \to \infty} x^{-1} \log P(X > x) = -a, \quad \lim_{x \to \infty} x^{-1} \log P(Y > x) = -b
\]

for some $a, b > 0$, then

\[
\lim_{x \to \infty} x^{-1} \log P(X + Y > x) = -\min\{a, b\}.
\]

**Proof** The lower bound is obvious. For the upper bound let $n \in \mathbb{N}$ be fixed. Clearly,

\[
P(X + Y > x) \leq \sum_{i=0}^{n-1} P\left(X \geq \frac{ix}{n}\right)P\left(Y \geq \frac{(n-i-1)x}{n}\right).
\]

Fix $\varepsilon > 0$. For $x$ sufficiently large, for all $i \in \{0, \ldots, n-1\}$,

\[
P\left(X \geq \frac{ix}{n}\right)P\left(Y \geq \frac{(n-i-1)x}{n}\right) \leq \exp\left(- (a - \varepsilon)\frac{ix}{n} - (b - \varepsilon)x\frac{n-i-1}{n}\right) \\
\leq \exp\left(- (\min\{a, b\} - \varepsilon)\frac{(n-1)x}{n}\right).
\]
Hence,

\[
\limsup_{x \to \infty} \frac{1}{x} \log P(X + Y > x) \leq -\left( \min\{a, b\} - \varepsilon \right) \frac{n - 1}{n}.
\] (7.8)

Since (7.8) holds for every \( n \in \mathbb{N} \) and \( \varepsilon > 0 \), we may take the limits \( n \to \infty \) and \( \varepsilon \downarrow 0 \) and the result follows.

**Proposition 7.6** When a customer enters a stationary work-conserving queue, the time \( D \) till the system is empty again satisfies \( D \overset{d}{=} A\hat{L} + L \), where \( \hat{L} \) is distributed like the residual life of a busy period, \( L \) is a 'normal' busy period length, \( P(A = 1) = \rho = 1 - P(A = 0) \) and \( A, \hat{L} \) and \( L \) are independent.

**Proof** The time till the system is empty again is the same for all work-conserving service disciplines, and can be decomposed as follows. The customer finds the system empty with probability \( 1 - \rho \). In that case \( D \) is just the length \( L \) of the busy period started by the customer. If our customer enters a busy system, then the server may first finish all the work in the system apart from the work of our tagged customer, see Figure 3 below. The moment the remainder of the busy period, which has length \( \bar{L} \), is finished, the server starts serving our customer, and all the customers that arrive during his service. This is a sub-busy period whose length is independent of \( \bar{L} \) and distributed like a normal busy period \( L \). After this sub-busy period the system is empty again.

![Figure 3](image)

**Figure 3** A realisation of the workload process

Hence \( D \overset{d}{=} A\hat{L} + L \), where \( \hat{L} \) is the residual life of a busy period, \( L \) is a 'normal' busy period, \( P(A = 1) = \rho = 1 - P(A = 0) \) and \( A, \hat{L} \) and \( L \) are independent.
For the $\tau$-queue we have the following corollary.

**Corollary 7.7** In the stationary $\tau$-queue, the time $D$ from the entrance of a customer till the system is empty again satisfies $D \overset{d}{=} A\hat{L}(\tau) + L(\tau)$. If the customer has service time $\tau$, then $D \overset{d}{=} A(\tau)\hat{L}(\tau) + L^*(\tau)$, where $P(A(\tau) = 1) = \lambda E(B \wedge \tau)$.

Since the system is work-conserving, the sojourn time of a customer is not longer than $D$, where $D$ is the time till the system is empty again. Hence $V \leq_{st} D$ for every service discipline. Since $A\hat{L}$ and $L$ satisfy the conditions of Lemma 7.5, the following corollary holds.

**Corollary 7.8** For every work-conserving service discipline the sojourn time $V$ of a customer in the stationary queue satisfies

$$-\limsup_{x \to \infty} \frac{1}{x} \log P(V > x) \geq -\lim_{x \to \infty} \frac{1}{x} \log P(A\hat{L} + L > x) = c.$$ 

**Corollary 7.9** For service times with an exponential moment, the FB discipline minimises the decay rate of the sojourn time in the class of work-conserving disciplines.

In Section 7.4 it is discussed that there are service disciplines with a strictly larger decay rate, e.g. FIFO.

Interestingly, for a subclass of Gamma-distributed service times the FB discipline minimises the queue length, but the sojourn time has the smallest decay rate. This shows that optimising one characteristic in a queue may have an ill effect on other characteristics.

The existence of a finite exponential moment in the corollary is crucial: for heavy-tailed service times the tail of $V$ cannot be bounded by that of $L$. For example, in the M/G/1 FIFO queue with service times satisfying $P(B > x) = x^{-\nu}\mathcal{L}(x)$, where $\mathcal{L}(x)$ is a slowly varying function at $\infty$ and $\nu > 1$, De Meyer and Teugels [20] showed that

$$P(L > x) \sim (1 - \rho)^{-\nu-1}x^{-\nu}\mathcal{L}(x).$$

It may be seen that in this case the tail of $\hat{L}$ is one degree heavier than that of $L$. Now note that for the FIFO discipline we have $V \geq A\hat{L}$. Hence the tail of $V$ is at least one degree heavier than that of $L$, see also Borst et al. [10] for further references. In the light-tailed case this phenomenon is absent since the tails of $L$ and $\hat{L}$ have the same decay rate.
7.3 Results

The results in this section rely on the following decomposition of $V$. Let $V(\tau)$ be the sojourn time in the stationary queue of a customer with service time $\tau$. The sojourn time $V$ of an arbitrary customer in the stationary queue satisfies

$$P(V > x) = \int P(V(\tau) > x)dF(\tau). \quad (7.9)$$

Here $F$ is the service-time distribution.

Hence we may write $P(V > x) = E_B P(V(B) > x)$, where $B$ is a generic service time independent of $V(\tau)$, and $E_B$ denotes the expectation w.r.t. $B$. Theorem 7.2 is proved using this representation of $V$. In the next lemma we show that $V(\tau)$ and $L(\tau)$ have the same decay rate.

**Proposition 7.10** Let $V(\tau)$ be the sojourn time of a customer with service time $\tau$ in the stationary queue. If the service-time distribution satisfies Assumption 7.3, then for $\tau > 0$,

$$\lim_{x \to \infty} \frac{1}{x} \log P(V(\tau) > x) = -c(\tau),$$

where

$$c(\tau) = -\lim_{x \to \infty} \frac{1}{x} \log P(L(\tau) > x). \quad (7.10)$$

**Proof** By the nature of the FB discipline, the sojourn time $V(\tau)$ of a customer with service time $\tau$ who enters a stationary queue is the time till the first epoch that no customers younger than $\tau$ are present. This is the time till the end of the $\tau$-busy period that he either finds in the $\tau$-queue, or starts. By Corollary 7.7, $V(\tau)$ then satisfies

$$V(\tau) \overset{d}{=} A(\tau)\tilde{L}(\tau) + L^*(\tau),$$

where $\tilde{L}(\tau)$ is the residual life of a $\tau$-busy period, $L^*(\tau)$ is a $\tau-$busy period that starts with a customer with service time $\tau$, $P(A(\tau) = 1) = 1 - P(A(\tau) = 0) = \lambda E(B \wedge \tau)$ and $A(\tau)$, $\tilde{L}(\tau)$ and $L^*(\tau)$ are independent. By Lemma 7.4 the random variables $A(\tau)\tilde{L}(\tau)$ and $L^*(\tau)$ satisfy the condition of Lemma 7.5. Hence

$$\lim_{x \to \infty} \frac{1}{x} \log P(V(\tau) > x) = \lim_{x \to \infty} \frac{1}{x} \log P(A(\tau)\tilde{L}(\tau) + L^*(\tau) > x) = -c(\tau). \quad (7.11)$$

This finishes the proof. \qed
Having found the upper bound for the decay rate in Corollary 7.8, the following lemma provides the basis for finding the lower bound. The end-point $x_F$ of the service-time distribution $F$ is defined as $x_F = \inf\{u \geq 0 : F(u) = 1\}$.

**Lemma 7.11** Let $V$ be the sojourn time of a customer in the stationary $M/G/1$ FB queue. Suppose the service-time distribution satisfies Assumption 7.3. If $\tau_0 > 0$ and $P(B \geq \tau_0) > 0$, then

$$
\liminf_{x \to -\infty} \frac{1}{x} \log P(V > x) \geq -P(B \geq \tau_0)^{-1} \int_{[\tau_0, x_F]} c(\tau) dF(\tau). \quad (7.12)
$$

Here $F$ is the distribution function of the generic service time $B$.

**Proof** Let $B_0$ and $V$ denote the service time and the sojourn time of a customer in the stationary queue. Let $\tau_0 > 0$ be such that $P(B_0 \geq \tau_0) > 0$. Then

$$
P(V > x) \geq P(V > x, B_0 \geq \tau_0) = P(V > x \mid B_0 \geq \tau_0)P(B_0 \geq \tau_0). \quad (7.13)
$$

Using the representation (7.9), we find

$$
\log P(V > x \mid B_0 \geq \tau_0) = \log E_{B_0}[P(V > x) \mid B_0 \geq \tau_0]. \quad (7.14)
$$

Since $\log x$ is a concave function, applying Jensen's inequality to the conditional expectation in (7.14) yields

$$
\log E_{B_0}[P(V > x) \mid B_0 \geq \tau_0] \geq E_{B_0}[^\log P(V > x) \mid B_0 \geq \tau_0]. \quad (7.15)
$$

From (7.13), (7.14) and (7.15) it follows that $\Theta := \liminf_{x \to -\infty} \frac{1}{x} \log P(V > x)$ satisfies

$$
\Theta \geq \liminf_{x \to -\infty} \frac{1}{x} \int_{[\tau_0, x_F]} \log P(V(\tau) > x)dF(\tau)/(1 - F(\tau_0)). \quad (7.16)
$$

Applying Fatou's lemma to (7.16) yields

$$
\Theta \geq P(B \geq \tau_0)^{-1} \int_{[\tau_0, x_F]} \lim_{x \to -\infty} \frac{1}{x} \log P(V(\tau) > x)dF(\tau).
$$

which was to be shown. \qed

The following lemma is used to distill the lower bound for the decay rate of $V$ from Lemma 7.11.

**Lemma 7.12** Let $c(\tau) = -\lim_{x \to -\infty} \frac{1}{x} \log P(L(\tau) > x)$. Then $c(\tau)$ is decreasing in $\tau$. Furthermore $c(\tau) \to c(x_F)$ as $\tau \to x_F$, where $c(x_F) = c$. 
The function \( h_\tau(\theta) = \theta - \lambda(Ee^{\theta B} - 1) \) is concave since any moment generating function is convex. Furthermore \( \lim_{\theta \to -\infty} h_\tau(\theta) = \lim_{\theta \to -\infty} h_\tau(\theta) = -\infty \). By definition of \( L(\tau) \) and (7.3), we may write \( c(\tau) = \sup_\theta \{h_\tau(\theta)\} \). Then \( c(\tau) \) is decreasing in \( \tau \), since \( h_\tau(\theta) \) is decreasing in \( \tau \). Since \( c(\tau) \geq h_\tau(0) = 0 \) for all \( \tau \) and \( c(\tau) \) is decreasing, \( c(\tau) \) converges for \( \tau \to x_F \). Now note that \( h_\tau(\theta) \) is continuous in \( \tau \) for all \( \theta \in [0, \sup\{\eta : Ee^{\eta B} < \infty\}] \), even if \( B \) has a discrete distribution. Since the supremum of \( \theta - \lambda(Ee^{\eta B} - 1) \) is attained in this interval, we have \( \lim_{\tau \to x_F} c(\tau) = c(x_F) \).

**Proposition 7.13** Let \( V \) be the sojourn time of a customer in the stationary M/G/1 FB queue. If the service-time distribution satisfies Assumption 7.3, then

\[
\liminf_{x \to \infty} \frac{1}{x} \log P(V > x) \geq -c. \tag{7.17}
\]

where

\[ c = -\lim_{x \to \infty} \frac{1}{x} \log P(L > x). \]

**Proof** If \( P(B = x_F) > 0 \), then choosing \( \tau_0 = x_F \) in (7.12) yields

\[
\liminf_{x \to \infty} \frac{1}{x} \log P(V > x) \geq -c(x_F) = -c
\]

and (7.17) holds. Let \( \varepsilon > 0 \). If \( P(B = x_F) = 0 \), then by Lemma 7.12 there exists an \( x_\varepsilon < x_F \) such that \( c(\tau) \leq c + \varepsilon \) for all \( \tau \geq x_\varepsilon \). Choosing \( \tau_0 = x_\varepsilon \) in (7.12) then yields

\[
\liminf_{x \to \infty} \frac{1}{x} \log P(V > x) \geq -P(B \geq x_\varepsilon)^{-1} \int_{x_\varepsilon}^{x_F} c(\tau) dF(\tau) \\
\geq -P(B \geq x_\varepsilon)^{-1} \int_{x_\varepsilon}^{x_F} (c + \varepsilon) dF(\tau) = -c - \varepsilon.
\]

Since \( \varepsilon > 0 \) was arbitrary, the lower bound (7.17) follows.

**Proof of Theorem 7.2** The upper bound is established in Corollary 7.8 and the lower bound in Proposition 7.13.

**7.4 Discussion**

The decay rate of the sojourn time \( V \) in the M/G/1 FB queue is the same as for the preemptive LFQ queue. Indeed, the sojourn time of a customer in the stationary M/G/1 queue under the preemptive LFQ discipline is just the length of the sub-busy
period started by that customer. From Theorem 7.2 it follows that the decay rates of the sojourn times for LIFO and FB are equal.

The sojourn time of a customer in the stationary queue under FIFO satisfies $V = B + W$, where $W$ is the stationary workload. The decay rate of $W$ is the positive root $\theta_0$ of $h(\theta) = \theta - \lambda(Ee^{\theta B} - 1)$. Since $h(0) = 0, h'(0) = 1 - \lambda EB < 1$ and $h$ is concave, we have $c_{FB} := c = \sup \theta h(\theta) < \theta_0$, see also Figure 2 below. Furthermore, $\theta_0$ is strictly smaller than the decay rate $c_B$ of the generic service time $B$, which is given by $c_B = \inf \theta : h(\theta) = -\infty \leq \infty$. An analogue of Lemma 7.5 yields that the decay rate $c_{FIFO}$ of the sojourn time in the FIFO system is strictly larger than $c$.

Mandjes en Zwart [37] consider the PS queue with light-tailed service requests. They show that the decay rate of $P(V > x)$ is $c$, under the additional requirement that, for any positive constant $k$,

$$\lim_{x \to \infty} \frac{1}{x} \log P(B > k \log x) = 0.$$  

For deterministic requests, clearly this criterion is not met. Indeed, in [37] it is shown that the decay rate in the M/D/1 queue with PS is larger than $c$.

![Figure 4 The decay rates of the sojourn time under FB and FIFO.](image)

### 7.5 Tail equivalence for heavy tails

In the second part of the chapter we consider the distribution of the sojourn time $V$ for light-tailed service times. For light-tailed service times, Theorem 2.15 by Núñez Queija [41] gives the behaviour of the tail probabilities $P(V > x)$ in terms of the tail of the service-time distribution. In the remainder of this chapter we prove that the result of his Theorem 2.15 holds under weaker conditions.
We consider the following class of distributions. The service-time distribution function is denoted by $F$. The tail $1 - F$ is said to be of intermediate regular variation at infinity if

$$\liminf_{x \to 0} \liminf_{x \to \infty} \frac{1 - F(x(1 + \varepsilon))}{1 - F(x)} = 1.$$ 

The class of intermediate regular varying tails contains the class of regularly varying tails. In the remainder of this chapter we prove the following theorem.

**Theorem 7.14** Let $V(x)$ be the sojourn time of a customer with service time $x$ in the stationary $M/G/1$ FB queue. Suppose the tail $1 - F$ is of intermediate regular variation at infinity. If there exist positive constants $\alpha$ and $\zeta$ with $\alpha > 1$ and such that $E B^\alpha < \infty$ and $E B^\zeta = \infty$, then

$$\lim_{x \to \infty} \frac{P(V > x/(1 - \rho))}{P(B > x)} = 1. \quad (7.18)$$

Theorem 7.14 may be interpreted as follows. When the service-time distribution is heavy-tailed, the 'cheapest way' for a customer to have a long sojourn time, is apparently to have a large service time. By the nature of the FB discipline, a customer with a very large demand is mostly served when the queue is empty. The fraction of time that this is the case converges to $1 - \rho$ when the service time converges to $\infty$. Theorem 7.14 indicates that for obtaining a long sojourn time, the customer does not get help from (an)other customer(s) with a large service time.

For queues with the PS discipline results similar to (7.18) exist, see Guillemin et al. [23].

Theorem 7.14 is a stronger version of Theorem 2.15 by Núñez Queija [41]. In his Theorem 2.15 the condition that $E B^\zeta = \infty$ for some $1 < \zeta < 2$ is somewhat unnatural. It is caused by the fact that in literature (Yashkov [68]) only the first two moments of the sojourn time are computed.

To prove Theorem 7.14 we have to do some work. Section 7.6 is devoted to the necessary tools. En passant the following theorem is proved.

**Theorem 7.15** Consider the stationary queue with generic service time $B \wedge x$. Let $L(x)$ denote the busy period length and $W(x)$ the stationary workload. If there exists an $\alpha > 1$ such that $E B^\alpha < \infty$, then for $n \in \mathbb{N}$ as $x \to \infty$,

$$EL(x)^n = \begin{cases} o(x^{n-\alpha}) & n > \alpha \\ O(1) & n \leq \alpha. \end{cases}$$
7.66 The tools

\[ EW(x)^n = \begin{cases} 
    o(x^{n+1-\alpha}) & n > \alpha - 1 \\
    O(1) & n \leq \alpha - 1.
\end{cases} \]

Furthermore,

\[ EV(x)^n = \frac{x^n}{(1 - \rho(x))^n} + \begin{cases} 
    o(x^{n+1-\alpha}) & 3 \alpha > 1 : EB^\alpha < \infty, \\
    O(x^{n-1}) & 3 \alpha \geq 2 : EB^\alpha < \infty,
\end{cases} \]

where \( \rho(x) = \lambda E(B \land x) \). Finally, if there exists a \( \gamma > 0 \) such that \( EB^\gamma < \infty \), then \( E(B \land x)^r = o(x^{r-\gamma}) \) for all \( r > \gamma, r \in \mathbb{R} \).

This theorem is proved by combining Lemmas 7.16, 7.17 and 7.20 below, and the remark below Lemma 7.19.

7.6 The tools

In this section we develop the lemmas that are needed for proving Theorem 7.14. In proving these lemmas we use some technical lemmas that are proved in the appendix of this chapter.

Lemma 7.16 If \( \alpha > 0 \) is such that \( EB^\alpha < \infty \), then \( 1 - F(x) = o(x^{-\alpha}) \) for \( x \to \infty \), and

\[ E(B \land x)^r = o(x^{r-\alpha}), \quad r > \alpha. \]

Proof Since \( EB^\alpha = \int u^\alpha dF(u) < \infty \), we have

\[ x^\alpha (1 - F(x)) \leq \int_x^\infty u^\alpha dF(u) \to 0, \quad x \to \infty. \quad (7.19) \]

Hence \( 1 - F(x) = o(x^{-\alpha}) \). To prove the second statement, fix an \( \varepsilon > 0 \). Then there exists an \( x_\varepsilon \) such that \( 1 - F(x) \leq \varepsilon x^{-\alpha} \) for all \( x \geq x_\varepsilon \). After rewriting

\[ E(B \land x)^r = \int_0^x u^r dF(u) + x^r (1 - F(x)) \]

\[ = x^r F(x) - r \int_0^x u^{r-1} F(u) du + x^r (1 - F(x)) \]

\[ = r \int_0^x u^{r-1} (1 - F(u)) du. \]

we find that for \( r > \alpha \),

\[ \lim_{x \to \infty} x^{-r+\alpha} E(B \land x)^r = \lim_{x \to \infty} x^{-r+\alpha} \int_0^x u^{r-1} (1 - F(u)) du \]
\[ \lim_{x \to \infty} r x^{-r^{\alpha}} \int_{x^r} u^{-1} (1 - F(u)) du \leq \lim_{x \to \infty} r x^{-r^{\alpha}} \int_0^x u^{r^{1-\alpha}} du = \frac{\varepsilon r}{r - \alpha}. \]

Since \( \varepsilon \) was arbitrary, the proof is finished. \( \square \)

Kleinrock [32] provides the Laplace transforms

\[ E e^{-sW(x)} = \exp \left( -x(s + \lambda - \lambda g_x(s)) \right) E \exp \left( - (s + \lambda - \lambda g_x(s)) W(x) \right). \]  \hfill (7.20)

\[ E e^{-sW(x)} = \frac{s - s \lambda E(B \wedge x)}{s - \lambda + \lambda E e^{-s(B \wedge x)}} \]  \hfill (7.21)

and the following relation for the Laplace transform \( g - x(s) \) of \( L(x) \).

\[ g_x(s) = E e^{-sL(x)} = E \exp \left( - (s + \lambda - \lambda g_x(s)) (B \wedge x) \right). \]  \hfill (7.22)

**Lemma 7.17** If there exists an \( \alpha > 1 \) such that \( EB^\alpha < \infty \), then for \( n > \alpha - 1 \),

\[ EW(x)^n = o(x^{n+1-\alpha}), \quad x \to \infty \]

and \( EW(x)^n = O(1) \) for \( n \leq \alpha - 1 \).

**Proof** Using (7.21), and writing \( \rho(x) = \lambda E(B \wedge x) \), we have

\[ E e^{-sW(x)} = \frac{s(1 - \lambda E(B \wedge x))}{s - \lambda + \lambda E e^{-s(B \wedge x)}} = \frac{1 - \rho(x)}{1 + \lambda s^{-1} E(e^{-s(B \wedge x)} - 1)} \]

\[ = (1 - \rho(x)) \left( 1 - \lambda E \sum_{k=0}^\infty \frac{(-s)^k (B \wedge x)^{k+1}}{(k+1)!} \right)^{-1}. \]

Using Lemma 7.27 below and \( \rho(x) = O(1) \) for \( x \to \infty \), we find

\[ EW(x)^n = (-1)^n \frac{d^n}{ds^n} E e^{-sW(x)} \bigg|_{s=0} \]

\[ = (-1)^n (1 - \rho(x)) \frac{d^n}{ds^n} \left( 1 - \lambda E \sum_{k=0}^\infty \frac{(-s)^k (B \wedge x)^{k+1}}{(k+1)!} \right)^{-1} \bigg|_{s=0} \]

\[ = q_n(E(B \wedge x)^2 \ldots E(B \wedge x)^{n+1}). \]

where \( q_n(y_1, \ldots, y_n) \) is a polynomial in \( y_1, \ldots, y_n \) whose terms are of the form \( c(x) \prod_{k=1}^n y_{a_k}^{a_k} \), with \( c(x) = O(1) \) as \( x \to \infty \) and \( \sum_{k=1}^n k a_k = n \).

Suppose \( EB^\alpha < \infty \). Using Jensen's inequality, for \( r \leq \alpha \) we have

\[ E(B \wedge x)^r \leq (E(B \wedge x)^\alpha)^{r/\alpha} \leq (EB^\alpha)^{r/\alpha}. \quad x \in \mathbb{R}. \]
Since $\alpha - 1 > 0$, Lemma 7.16 and Lemma 7.26 below yield that for $n > \alpha - 1$

$$q_n(E(B \wedge x)^2 \ldots E(B \wedge x)^{n+1}) = o(x^{n+1-\alpha}). \quad x \to \infty$$

and for $n \leq \alpha - 1$

$$q_n(E(B \wedge x)^2 \ldots E(B \wedge x)^{n+1}) = O(1), \quad x \to \infty.$$  

This finishes the proof. \hfill \square

**Lemma 7.18** As $x \to \infty$,

$$\lambda g_x^{(1)}(0) = 1 - \frac{1}{1 - \rho(x)} = O(1).$$

**Proof** Using (7.22), we find

$$g_x^{(1)}(0) = -E(B \wedge x)(1 - \lambda g_x^{(1)}(0)).$$

Hence

$$\lambda g_x^{(1)}(0) = \frac{-\lambda E(B \wedge x)}{1 - \lambda E(B \wedge x)} = \frac{-\rho(x)}{1 - \rho(x)}.$$  

The result follows from noting $\rho(x) = \lambda E(B \wedge x) \to \lambda EB = \rho < 1$ for $x \to \infty$. \hfill \square

In the next lemmas the notation $(x)^\pm = \max\{0, x\}$ is used.

**Lemma 7.19** Write $g_n = g_x^{(n)}$. Assume that there exists an $\alpha > 1$ such that $EB^\alpha < \infty$. Then as $x \to \infty$, we have $g_n = o(x^{n-\alpha})$ for $n > \alpha$ and $g_n = O(1)$ for $n \leq \alpha$.

**Proof** By induction. Write $g(s) := g_x(s)$ and suppose that the lemma holds for $n \in \mathbb{N}$. Using (7.22) and Lemma 7.18, we have

$$g_{n+1} = \lambda g_{n+1} E(B \wedge x) + \sum_{k=1}^{n-1} E(B \wedge x)^{n+1-k} q_k(g_2, \ldots, g_{k+1}) + E(B \wedge x)^{n+1} (\lambda g_1 - 1)^{n+1}. \quad (7.23)$$

where $q_n(y_1, \ldots, y_k)$ are polynomials with terms of the form $c(x) \prod_{k=1}^n y_k^{a_k}$ with $c(x) = O(1)$ as $x \to \infty$ and $\sum_{k=1}^n k a_k = n$.

Then by the induction hypothesis and Lemma 7.26,

$$q_k(g_2, \ldots, g_{k+1}) = o(x^{k+(1-\alpha)}) \quad (7.24)$$
if \( k > \alpha - 1 \) and \( q_k(g_2, \ldots, g_{k+1}) = O(1) \) if \( k \leq \alpha - 1 \). Using Lemma 7.16 and the fact that \( \alpha > 1 \), we have from (7.23) and (7.24) (with abuse of notation)

\[
g_{n+1} = \frac{1}{1 - \rho(x)} \left( \sum_{k=1}^{n-1} O(x^{(n+1-k-\alpha)^+})O(x^{(k+1-\alpha)^+}) + O(x^{(n+1-\alpha)^+}) \right) = O(x^{(n+1-\alpha)^+}).
\]

In fact if \( n+1 > \alpha \), then by (7.24) the last \( O \)-symbol may be replaced by a \( o \)-symbol. The induction basis is established in Lemma 7.18 and the proof is complete. \( \square \)

Since \( g_x^{(n)}(0) = (-1)^n EL(x)^n \), we have \( EL(x)^n = o(x^{n-\alpha}) \) for all \( n > \alpha \) and \( EL(x)^n = O(1) \) for all \( n \leq \alpha \).

**Lemma 7.20** If \( EB^n < \infty \) for some \( \alpha > 1 \), then for every \( n \in \mathbb{N} \) as \( x \to \infty \).

\[
EV(x)^n = \frac{x^n}{(1 - \rho(x))^n} + \begin{cases} O(x^{n-1}) & \text{if } \alpha \geq 2 \\ o(x^{n+1-\alpha}) & \text{if } 1 < \alpha < 2. \end{cases} \tag{7.25}
\]

**Proof** Write again \( g_n = g_x^{(n)}(0) \). By (7.20) we have

\[
EV(x)^n = (-1)^n \frac{d^n}{ds^n} \exp \left( -x(s + \lambda - \lambda g_x(s)) \right) \exp \left( -(s + \lambda - \lambda g_x(s))W(x) \right) \bigg|_{s=0}.
\]

Differentiating this product in total \( n \) times and using that \( g_x(0) = 1 \) yields

\[
EV(x)^n = (-1)^n \sum_{i,j \geq 0}^n \binom{n}{i,j} x^j W(x)^i \left. \frac{d^n}{ds^n-i-j}(\lambda g_1(s) - 1)^{i+j} \right|_{s=0}.
\]

\[
= (1 - \lambda g_1)^n \sum_{k=0}^n \binom{n}{k} x^k W(x)^{n-k} + (-1)^n \sum_{i,j \geq 0}^{i+j < n} x^j W(x)^i q_{n-i-j}(g_2, \ldots, g_{n+1-i-j}).
\]

where, by Lemmas 7.23 and 7.18, \( q_n(y_1, \ldots, y_m) \) is a polynomial whose terms are of the form \( c(x) \prod_{k=1}^n d_k^{a_k} \) with \( c(x) = O(1) \) as \( x \to \infty \) (sic) and \( \sum_{k=1}^n ka_k = n \).

Using Lemmas 7.18, 7.17, 7.19 and 7.26 below yields (again with abuse of notation)

\[
EV(x)^n = (1 - \rho(x))^{-n} \left( x^n + \sum_{k=0}^{n-1} x^k O(x^{(n-k+1-\alpha)^+}) \right) + \sum_{i,j \geq 0}^{i+j < n} x^j O(x^{i+1-\alpha}) O(x^{(n-j-i+1-\alpha)^+}). \tag{7.26}
\]
7.7 The tail of the sojourn-time distribution

Using
\[ j + (i + 1 - \alpha) + (n - j - i + 1 - \alpha) \leq j + i + n - j - i - 1 + (2 - \alpha) = n - 1 + (2 - \alpha) \]
for the second summation in (7.26) and a similar argument for the other summation, we find
\[ EV(x)^n = \frac{x^n}{(1 - \rho(x))^n} + O(x^{n-1+(2-\alpha)^+}), \quad x \to \infty. \] (7.27)
In fact, in case \( 1 < \alpha < 2 \) the first and third \( O \)-symbol in (7.26) may be replaced by an \( o \)-symbol and hence the \( O \)-symbol in (7.27) as well. This finishes the proof. \( \square \)

## 7.7 The tail of the sojourn-time distribution

Having done all the hard work, we now first prove the last lemma, and then Theorem 7.14.

**Lemma 7.21** If for some \( 1 < \alpha < 2 \) and \( n \in \mathbb{N} \)
\[ EV(x)^n = \frac{x^n}{(1 - \rho(x))^n} + o(x^{n+1-\alpha}), \quad x \to \infty \]
then for all \( m \in \mathbb{N} \)
\[ E|V(x) - EV(x)|^{2m} = o(x^{2m+1-\alpha}), \quad x \to \infty. \]

**Proof** Since \( 2m \) is even, by Lemma 7.20 we may write, as \( x \to \infty \).
\[ E|V(x) - EV(x)|^{2m} = \sum_{k=0}^{2m} \binom{2m}{k} EV(x)^k (-EV(x))^{2m-k} \]
\[ = \sum_{k=0}^{2m} \binom{2m}{k} \left( \frac{x^k}{(1 - \rho(x))^k} + o(x^{k+1-\alpha}) \right) \left( \frac{-x}{1 - \rho(x)} + o(x^{2-\alpha}) \right)^{2m-k} \]
\[ = \sum_{k=0}^{2m} \binom{2m}{k} \left( \frac{x^k}{(1 - \rho(x))^k} + o(x^{k+1-\alpha}) \right) \left( \frac{(-x)^{2m-k}}{(1 - \rho(x))^{2m-k}} + o(x^{2m-k-1+(2-\alpha)}) \right) \]
\[ = \frac{x^{2m}}{(1 - \rho(x))^{2m}} \sum_{k=0}^{2m} \binom{2m}{k} (-1)^k + \sum_{k=0}^{2m} \binom{2m}{k} o(x^{2m+1-\alpha}) = o(x^{2m+1-\alpha}), \]
since \( \sum_{k=0}^{n} \binom{n}{k}(-1)^k = 0 \) by Newton’s binomial formula. \( \square \)

**Proof of Theorem 7.14** The proof is analogous to the proof of Theorem 5.2.4
in Núñez Queija [41], with the following difference. By Lemma 7.20 and Lemma 7.21 we have that for each $m \in \mathbb{N}$
\[
\lim_{x \to \infty} \frac{E|V(x) - EV(x)|^{2m}}{x^{2m+1-\alpha}} = 0.
\]
Now take $m$ such that $2m > \zeta$, put $\kappa = 2m$ and $0 < \delta < \alpha - 1$. Then Assumption 5.1.2 in Núñez Queija [41] has been fulfilled and the result follows. \(\Box\)

Suppose that $P(B > x) \sim \eta(x)x^{-\alpha}$, where the function $\eta$ is slowly varying at $\infty$ and $\alpha > 0$. De Meyer and Teugels [20] proved that
\[
P(B > x) \sim \eta(x)x^{-\alpha} \iff P(L > x) = (1 - \rho)^{-\alpha} - 1 \eta(x)x^{-\alpha}.
\]
Since $x^{-\alpha}\eta(x)$ is intermediate regular varying at $\infty$ for any $\alpha > 0$, we have by Theorem 7.14.
\[
P(V > x) \sim P(B > (1 - \rho)x) \sim \eta((1 - \rho)x)((1 - \rho)x)^{-\alpha}
\sim \eta(x)((1 - \rho)x)^{-\alpha} \sim (1 - \rho)P(L > x).
\]
Hence the tails of $V$ and $L$ are asymptotically equal up to a multiplicative constant. This conclusion is comparable to the situation in the light-tailed case, discussed in the first part of this chapter, where the decay rates of $V$ and $L$ were equal.

Núñez Queija [41] shows that Theorem 7.14 holds for the PS and the SRPT discipline as well. It is an interesting open problem whether there are service disciplines such that in the M/G/1 queue with heavy-tailed service times the tail of the sojourn time $V$ satisfies
\[
P(V > x) \sim P(B > ax)
\]  
(7.28)
for some constant $1 - \rho < a \leq 1$. Note that the larger the value of $a$, the less likely it is that a sojourn time is (very) large. We conjecture that such disciplines do not exist. If this would indeed be true, then the queue with the FB discipline possesses the surprising quality that although large jobs are discriminated, large sojourn times are as unlikely as possible.

Finally, consider the following example about a queue with priority classes. In a queue with two priority classes, arriving customers belong to the highest priority class with probability $0 < p < 1$, independent of their service request. Customers in this class have absolute priority over the customers in the other class, i.e. customers
Appendix I: functions and polynomials

Lemma 7.22 Let \( g \) and \( h \) be positive \( L^1 \) functions. If \( g \) and \( h \) are \( o(x^{-1}) \) as \( x \to \infty \), then

\[
\int_0^x g(x-s)h(s)ds = o(x^{-1}), \quad x \to \infty.
\]

Proof Let \( \varepsilon > 0 \) and let \( x_\varepsilon \) be such that \( g(x), h(x) \leq \varepsilon/x \) for all \( x \geq x_\varepsilon \). Then for \( x \geq 2x_\varepsilon \),

\[
\int_0^x g(x-s)h(s)ds \leq \int_0^{x/2} \frac{\varepsilon}{x-s}h(s)ds + \int_{x/2}^x g(x-s)\frac{\varepsilon}{s}ds
\]

\[
\leq \frac{2\varepsilon}{x} \int_0^{x/2} h(s)ds + \frac{2\varepsilon}{x} \int_{x/2}^x g(x-s)ds
\]

\[
= \frac{2\varepsilon}{x} \int_0^{x/2} (g(s) + h(s))ds \leq \frac{2\varepsilon}{x} \int_0^{x/2} (g(s) + h(s))ds \leq \frac{\varepsilon m}{x}
\]

for some \( m \in \mathbb{R} \). Since \( \varepsilon > 0 \) was arbitrary, the statement is proved. \( \square \)

Without proof we state that the same holds for summation instead of integration.

The following lemma holds by induction.

Lemma 7.23 Assume that the function \( h(s) \) is \( n \) times differentiable. Then the expression \( e^{-h(s)}d^n e^{h(s)}/ds^n \) is a polynomial in \( h^{(1)}(s), \ldots, h^{(n)}(s) \) with terms of the form \( k \prod_{j=1}^n h^{(j)}(s)^{a_j} \) with \( \sum_{j=1}^n j a_j = n \) and \( k \in \mathbb{N} \).

Lemma 7.24 Let \( q(y_1, \ldots, y_n) \) be a polynomial in \( y_1, \ldots, y_n \) whose terms are of the form \( c(x) \prod_{k=1}^n y_k^{a_k} \) with \( \sum_{k=1}^n k a_k = n \) and \( c(x) = O(1) \) as \( x \to \infty \).

If for some \( \beta > 0 \) we have \( y_k = o(x^{k-\beta}) \) for \( k = 1, \ldots, n \) as \( x \to \infty \), then

\[
q(y_1, \ldots, y_n) = o(x^{n-\beta}).
\]
**Proof** The terms of the polynomial \( q(y_1, \ldots, y_n) \) satisfy, with abuse of notation, in the limit as \( x \to \infty \).

\[
c(x) \prod_{k=1}^{n} y_k^{a_k} = c(x) \prod_{k=1}^{n} o(x^{k-3})^{a_k} = \prod_{k=1}^{n} o(x^{\alpha_k (k-3)})
\]

\[
= o(x^{\sum_{k=1}^{n} \alpha_k (k-3)}) = o(x^{n-3} \sum_{k=1}^{n} \alpha_k) = o(x^{n-3}).
\]

since \( \sum_{k=1}^{n} \alpha_k \geq 1 \).

\[\square\]

**Lemma 7.25** Let \( n \in \mathbb{N} \) and \( \alpha_k \in \{0, 1, 2, \ldots \} \) for \( k = 1, \ldots, n \). If \( \sum_{k=1}^{n} k \alpha_k = n \) and \( \beta \geq 0 \), then

\[
\gamma := \sum_{k=1}^{n} \alpha_k(k-\beta)^+ \leq n-\beta. \tag{7.29}
\]

**Proof** If \( \beta = 0 \) then (7.29) follows immediately. For \( \beta > 0 \) we prove (7.29) by contradiction. Suppose \( \gamma > n-\beta \). Since \( \gamma \) is of the form \( m_1 - m_2, \beta \) with \( \beta > 0 \) and \( m_1, m_2 \in \{0, 1, \ldots, n\} \), we must have that \( \gamma = n \). But then \( m_1 = n \) and \( m_2 = 0 \), which implies the desired contradiction. \[\square\]

**Lemma 7.26** Let \( q(y_1, \ldots, y_n) \) be a polynomial in \( y_1, \ldots, y_n \), whose terms are of the form \( c(x) \prod_{k=1}^{n} y_k^{a_k} \), where \( \sum_{k=1}^{n} k \alpha_k = n \) and \( c(x) = O(1) \) as \( x \to \infty \).

If for some \( \beta > 0 \) we have \( x_k = o(x^{\beta-\beta}) \) for \( k < \beta \) and \( x_k = O(1) \) for \( k \leq \beta \) as \( x \to \infty \), then

\[
q(y_1, \ldots, y_n) = \begin{cases} 
  o(x^{n-\beta}) & \beta < n \\
  O(1) & \beta \geq n. 
\end{cases} \tag{7.30}
\]

**Proof** The terms of the polynomial \( q(y_1, \ldots, y_n) \) satisfy (with this chapter's final abuse of notation).

\[
c(x) \prod_{k=1}^{n} y_k^{a_k} = O(1) \prod_{k=1}^{n} O(x^{(k-\beta)^+})^{a_k} = O(1) \prod_{1 \leq k \leq n, k>\beta} O(x^{\alpha_k (k-\beta)}). \tag{7.31}
\]

For \( \beta \geq n \) the product on the RHS of (7.31) is empty and (7.30) follows immediately. Now assume \( \beta < n \). Lemma 7.25 yields that the RHS of (7.31) is \( O(x^{n-\beta}) \). In fact, the RHS of (7.31) is \( o(x^{n-\beta}) \) if there is a \( k > \beta \) such that \( a_k > 0 \). If there is no such \( k \), then the RHS of (7.31) is \( O(1) \) and hence \( o(x^{n-\beta}) \). \[\square\]

**Lemma 7.27** Let \( b_0 = 1 \). Then for any \( n \in \mathbb{N} \)

\[
\frac{d^n}{ds^n} \left( \sum_{k=0}^{\infty} b_k s^k \right)^{-1} \bigg|_{s=0} = q_n(b_1, \ldots, b_n). \tag{7.32}
\]
where $q_n(b_1,\ldots,b_n)$ is a polynomial in $b_1,\ldots,b_n$ whose terms are of the form $c \prod_{k=1}^n b_k^{a_k}$ with $\sum_{k=1}^n k a_k = n$ and $c \in \mathbb{Z}$.

**Proof** The LHS of (7.32) is equal to $n!$ times the coefficient of $s^n$ in the formal series expansion

$$\left( \sum_{k=0}^{\infty} b_k s^k \right)^{-1} = \left( 1 - \sum_{k=1}^{\infty} (b_k) s^k \right)^{-1} = 1 + \sum_{k=1}^{\infty} (-b_k) s^k + \left( \sum_{k=1}^{\infty} (-b_k) s^k \right)^2 + \cdots.$$  

The coefficient of $s^n$ in this expression is then given by

$$\sum_{k,i_1,\ldots,i_k \geq 1} (-1)^k b_{i_1} \cdots b_{i_k} = \sum_{a_1,\ldots,a_n \geq 0} \frac{(-1)^{a_1+\cdots+a_n}}{a_1+\cdots+a_n=n} \left( a_1 + \cdots + a_n \right) b_{a_1} \cdots b_{a_n},$$

where the equality follows from putting $a_k := |\{m : i_m = k\}|$ in each term of the sum. This proves the lemma. \[\square\]

### 7.9 Appendix II: Notions in renewal theory

Consider the renewal process given by a sequence $(T_k)_{k \in \mathbb{N}}$ of i.i.d. positive random variables (called *failure times*) with distribution function $F$, density $f$ and finite expectation $\beta$. Set $S_0 = 0$ and $S_n = T_1 + \cdots + T_n$, $n \in \mathbb{N}$.

Define $N(t) = \sup\{n \geq 0 : S_n \leq t\}, t \geq 0$, and let

$$R(t) = S_{N(t)+1} - t.$$  

For the following results we refer to Cox [16]. Let $G_t$ be the distribution of $R(t)$. As $t \to \infty$, the distributions $G_t$ converge to a limiting distribution $G$. The density $g$ of $G$ satisfies

$$g(x) = (1 - F(x))/\beta. \quad (7.33)$$

A random variable $R$ with distribution $G$ is called the *residual life*\(^1\) of the generic failure time $T_1$. The distribution of $S_{N(t)+1} - S_{N(t)}$ converges to a limiting distribution $G_T$ say, as well. Let $T$ be a random variable with distribution $G_T$. Then $T$ is called a *length biased* (or *size biased*) version of $T_1$. Its density is given by

$$f_T(x) = x f(x)/\beta. \quad (7.34)$$

This may intuitively be understood as follows. Larger failure times have a larger 'probability' of containing the point $t$. The density $f_T(x)$ should be proportional to both $x$ and $f(x)$. The value of the normalising constant $\beta$ is determined by integration of $xf(x)$.

\(^1\)also called forward recurrence time or excess
Chapter 7  Sojourn time asymptotics