Queueing Systems, 5 (1988) 345–368

HOW LARGE DELAYS BUILD UP IN A GI/G/1 QUEUE *

V. ANANTHARAM

School of Electrical Engineering, Phillips Hall, Cornell University, Ithaca, NY 14853, U.S.A.

Received 1 April 1988; revised 12 June 1989

Abstract

Let W_k denote the waiting time of customer k, $k \geqslant 0$, in an initially empty GI/G/1 queue. Fix $a > 0$. We prove weak limit theorems describing the behaviour of W_k/n, $0 \leqslant k \leqslant n$, given $W_n > na$. Let X have the distribution of the difference between the service and interarrival distributions. We consider queues for which Cramer type conditions hold for X, and queues for which X has regularly varying positive tail.

The results can also be interpreted as conditional limit theorems, conditional on large maxima in the partial sums of random walks with negative drift.

Keywords: Weak limit theorems, Cramer type conditions, random walks with negative drift, waiting time.

1. Introduction

Consider an initially empty GI/G/1 queue. Customer 0 arrives at time 0 and customer k at time $A_1 + \ldots + A_k$, where A_i, $i = 1, 2, \ldots$ are i.i.d.. Let B_i, $i = 1, 2, \ldots$ be i.i.d. service times, B_{k+1} being the service time of customer k. If we let $X_k = B_k - A_k$, then X_i, $i = 1, 2, \ldots$ are i.i.d.. Let $EX_1 = -\mu$, $\mu > 0$. The waiting time W_k of customer k before receiving service is given by the equations:

$$W_0 = 0,$$

$$W_{k+1} = \max(0, W_k + X_{k+1}),\ k > 0. \tag{1.1}$$

Let C denote $C[0, 1]$, the space of continuous functions on the unit interval with the topology of uniform convergence, and D denote $D[0, 1]$, the space of right continuous functions with left limits on the unit interval, with the topology of uniform convergence. For each $n \geqslant 1$, let $w_n(\cdot) \in C$ denote the polygonal path, and $\hat{w}_n(\cdot) \in D$ denote the piecewise constant right continuous path, constructed from the points

$$\left(\frac{k}{n}, \frac{W_k}{n}\right),\ 0 \leqslant k \leqslant n.$$

* Research supported by the NSF under Grant NCR 8710840 and under the PYI Award NCR 8857731.

Fix $a > 0$. Let P_n denote the conditional distribution $(w_n(\cdot) \mid W_n > na)$, and $\hat{P}_n$ the conditional distribution $(\hat{w}_n(\cdot) \mid W_n > na)$. P_n (resp. $\hat{P}_n$) is a probability distribution on C (resp. D).

In section 2, thm. 2.1, we identify the weak limit of $\{P_n\}$ under the following moment conditions on X_1:

(C) Let $m(t) = E \exp tX_1$. There is an interval $S = (t_-, t_+)$ containing 0, such that:

$$m(t) < \infty \text{ for all } t \in S,$$

$$\frac{m'(t)}{m(t)} \to \infty \text{ as } t \to t_+,$$

$$\frac{m'(t)}{m(t)} \to -\infty \text{ as } t \to t_-.$$

This situation covers the $M/M/1$ queue and a large class of GI/G/1 queues with rapidly decreasing tail distributions for X_1.

In section 3, thm. 3.1, we identify the weak limit of $\{\hat{P}_n\}$ under the following moment conditions on X_1:

(D) There is $q > 0$ and a slowly varying function $L(\cdot)$ such that:

$$EX_1^2 < \infty,$$

$$P(X_1 > x) = x^{-q}L(x).$$

This situation covers a large class of GI/G/1 queues where the service time distribution has a fat tail.

2. Cramer type conditions

Throughout this section we assume the moment conditions (C). Then $m(\cdot)$ is an analytic, strictly convex function on S, as is $\log m(\cdot)$. Note that $m'(0)/m(0) = -\mu$. Let $t_0 > 0$ be the unique point where $m(t_0) = 1$, and let z_0 denote $m'(t_0)/m(t_0)$. See fig. 1.

The main results of this section is the following:

THEOREM 2.1

Assume the moment conditions (C). If $a \leqslant z_0$, let $t(a) = 1 - (a/z_0)$ and $p_a(\cdot) \in C$ be defined by:

$$p_a(t) = 0, \qquad 0 \leqslant t \leqslant t(a),$$

$$p_a(t) = z_0(t - t(a)), \qquad t(a) \leqslant t \leqslant 1.$$

If $a \geqslant z_0$, let $p_a(\cdot) \in C$ be defined by:

$$p_a(t) = ta, \qquad 0 \leqslant t \leqslant 1.$$

See fig. 2.

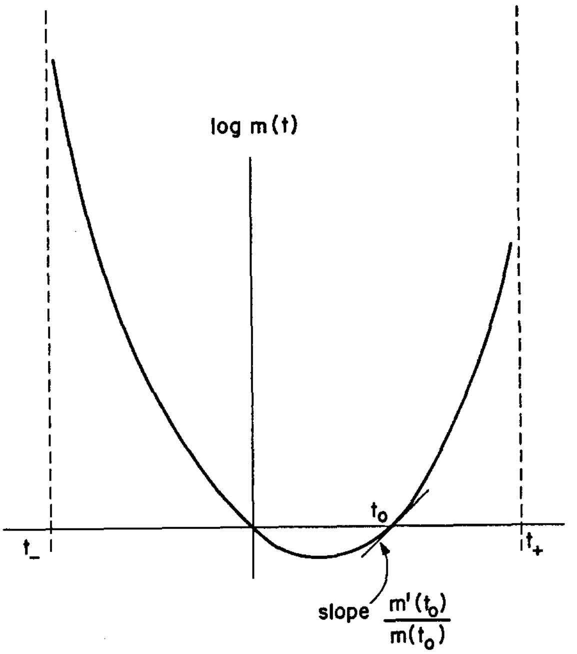

Fig. 1.

We have

$$P_n \xrightarrow{w} \delta_{p_a},$$

where δ_{p_a} denotes the probability distribution concentrated on $p_a(\cdot)$, and $\xrightarrow{w}$ denotes weak convergence of probabilities on C. □

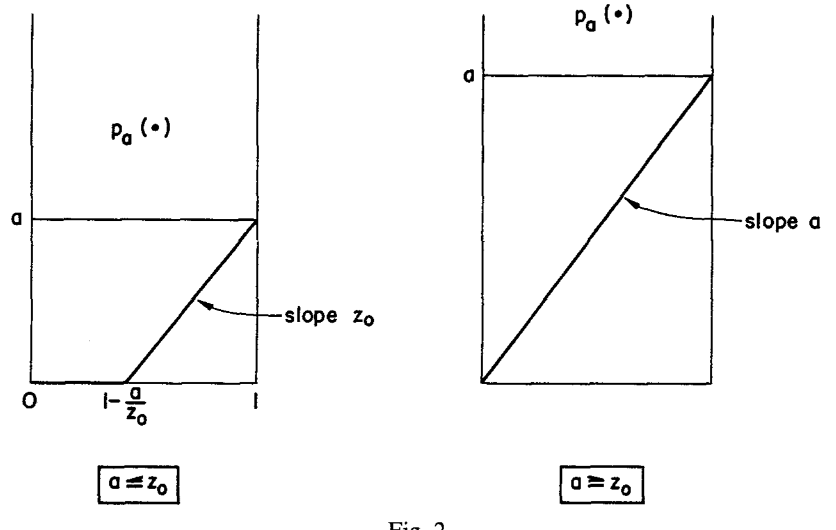

Fig. 2.

For the meaning of weak convergence, see Billingsley, [5]. Heuristically, thm. 2.1 states that the most likely way in which the waiting time of customer n builds up to na as follows: If $a \leqslant z_0$, customers k, $k \leqslant [nt(a)]$, encounter essentially typical behaviour in the queue, after which the waiting time of customers builds up essentially linearly. If $a \geqslant z_0$, waiting time starts building up right away and builds up essentially linearly.

Let $Y_1, Y_2, \ldots$ be i.i.d. with the distribution of X_1 and let:

$$S_0 = 0,$$

$$S_k = Y_1 + Y_2 + \ldots + Y_k.$$

From the points

$$\left(\frac{k}{n}, \frac{S_k}{n}\right), 0 \leqslant k \leqslant n,$$

we construct the piecewise linear, continuous path $s_n(\cdot) \in C$. Let

$$M_n = \sup_{0 \leqslant k \leqslant n} S_k,$$

and let Q_n denote the conditional distribution of $(s_n(\cdot) \mid M_n > na)$. Let $\Psi : C \to C$ be given by

$$\Psi(\phi)(t) = \left[\sup_{1-t \leqslant s \leqslant 1} \phi(s)\right] - \phi(1-t).$$

Note that Ψ is continuous.

LEMMA 2.2

For each $n = 1, 2, \ldots,$ we have

$$P_n = \Psi_*(Q_n),$$

where $\Psi_*(Q_n)$ denotes the distribution induced on C by Ψ, i.e., for any Borel set $B \subseteq C$, $\Psi_*(Q_n)(B) = Q_n(\Psi^{-1}(B))$.

Proof

From eqns. (1.1), we see that

$$W_k = \max(0, X_k, X_k + X_{k-1}, \ldots, X_k + X_{k-1} + \ldots + X_1).$$

The claim follows easily. □

Lemma 2.2 reduces the limit behaviour of $\{P_n\}$ to that of $\{Q_n\}$. In fact, by Billingsley, Thm. 5.1, [5], if we can show that $Q_n \xrightarrow{w} Q$ for some probability distribution Q on C, then $P_n \xrightarrow{w} \Psi_*(Q)$. Thus thm. 2.1 is a direct consequence of the following:

THEOREM 2.3

Assume the moment conditions (C). If $a \leqslant z_0$, let $\tau(a) = a/z_0$ and let $q_a(\cdot) \in C$ be defined by:

$$q_a(t) = tz_0, \qquad 0 \leqslant t \leqslant \tau(a),$$
$$q_a(t) = a - \mu(t - \tau(a)), \quad \tau(a) \leqslant t \leqslant 1.$$

If $a \geqslant z_0$, let $q_a(\cdot) \in C$ be defined by:

$$q_a(t) = ta, \quad 0 \leqslant t \leqslant 1.$$

See fig. 3.

We have

$$Q_n \xrightarrow{w} \delta_{q_a}. \qquad \square$$

We proceed to prove thm. 2.3. We first introduce the convex dual of log $m(\cdot)$,

$$I(z) = \sup_{t \in S} zt - \log m(t), \; z \in R$$

$I(\cdot)$ is nonnegative, finite, and strictly convex. If $\bar{t}$ denotes the unique point at which $m'(\bar{t})/m(\bar{t}) = z$, then

$$I(z) = z\bar{t} - \log m(\bar{t}).$$

It is easily checked that $I(-\mu) = 0$. It is also easy to see that log $m(\cdot)$ is the convex dual of $I(\cdot)$, i.e.,

$$\log m(t) = \sup_{z \in R} tz - I(z), \quad t \in S.$$

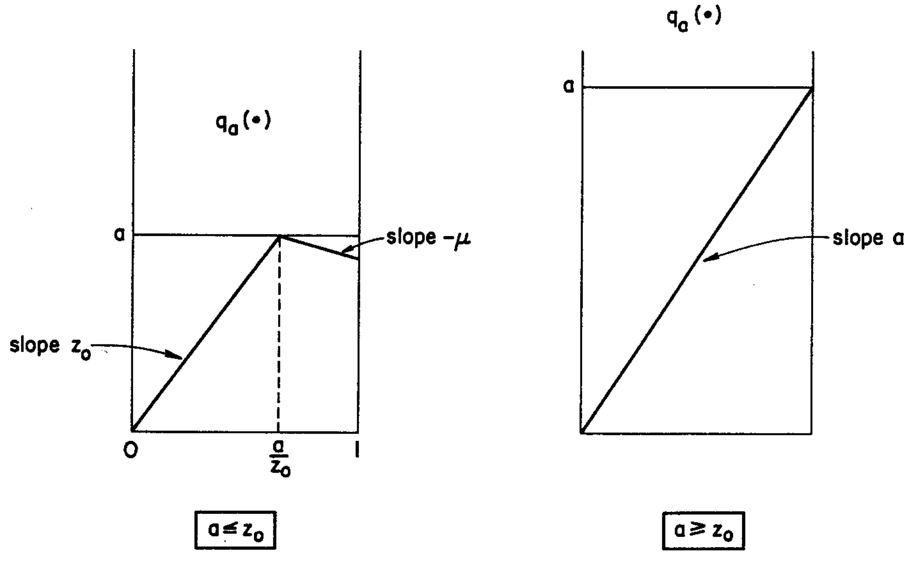

Fig. 3.

The Skorohod metric on C is defined by

$$d_S(f, g) = \inf_{\lambda,\eta}\left\{ \sup_{t\in[0,1]} |f(\lambda(t)) - g(\eta(t))| + \sup_{t\in[0,1]} |\lambda(t) - \eta(t)|\right\}$$

where $f(\cdot)$, $g(\cdot) \in C$ and the infimum is taken over all continuous, strictly increasing functions $\lambda(t)$, $\eta(t)$, $0 \leqslant t \leqslant 1$, such that $\lambda(0) = \eta(0) = 0$ and $\lambda(1) = \eta(1) = 1$. In future we will refer to such functions as "time changes". It is well known that the Skorohod metric induces the topology of uniform convergence on C, see, e.g., Bergstrom, thm. 3, pg. 169, [4]. Let E denote the completion of C in the Skorohod metric. We also recall the standard notation

$$\| f \|_\infty = \max_{t\in[0,1]} | f(t) |, \quad f \in C.$$

Let $\Gamma \subseteq C$ denote the set of all continuous, piecewise linear functions with finitely many break points, and starting at 0. If $f(\cdot) \in \Gamma$, with slope f_i' on the interval (t_{i-1}, t_i), $1 \leqslant i \leqslant N$, where $0 = t_0 < t_1 < t_2 < \dots < t_N = 1$, let

$$h(f) = \sum_{i=1}^{N} (t_i - t_{i-1}) I(f_i').$$

Extend the functional h to E by setting

$$h(f) = \inf \liminf_{n\to\infty} h(f_n), \quad f \in E, \tag{2.1}$$

where the infimum is taken over all sequences $\{f_n\} \subseteq \Gamma$ such that $d_S(f_n, f) \to 0$ as $n \to \infty$. Appealing to the convexity of $I(\cdot)$, it can be easily shown that this extension does not change the definition of h on Γ. For details, see [9]. It can also be shown that for $f(\cdot) \in C$, $h(f) = \infty$ unless $f(0) = 0$ and $f(\cdot)$ has finite total variation.

For any subset $G \subseteq E$ let

$$H(G) = \inf_{f\in G} h(f). \tag{2.2}$$

Our main tool in the proof of thm. 2.3 is the following special case of theorem 2 of Mogulskii, [9]:

THEOREM 2.4

Assume the moment conditions (C). For an arbitrary Borel subset $G \subseteq C$, let G^0 denote its interior and $\overline{G}^E$ its closure in E. If $H(G^0) = H(\overline{G}^E)$ then

$$\log P(s_n(\cdot) \in G) = -nH(G) + o(n). \qquad \square$$

The following sequence of lemmas will be used in the proof of thm. 2.3. Let

$$X = \left\{ f \in C \mid \sup_{t\in[0,1]} f(t) > a \right\}.$$

Note that X is open, and $Q_n(X) = 1$ for all n.

LEMMAS 2.5
$H(X) = h(q_a)$.

Proof

First we show that $H(X) \geqslant h(q_a)$. Let $f \in X \cap \Gamma$. We may assume that $f(0) = 0$, else $h(f) = \infty$. Let $\bar{t} = \inf\{t \in [0, 1] \mid f(t) > a\}$. Then $f(\bar{t}) = a$. Denote by r_t the piecewise linear path connecting (0, 0) to (t, a) and then dropping off with slope $-\mu$. By convexity of $I(\cdot)$ and the fact that $I(-\mu) = 0$, we have

$$h(f) \geqslant h(r_{\bar{t}}) = \bar{t} I\left(\frac{a}{\bar{t}}\right).$$

Note that z_0 satisfies

$$I'(z_0) = \frac{I(z_0)}{z_0}.$$

Further, $I(z)/z$ decreases with z for $z \in (0, z_0)$ and increases with z for $z \in (z_0, \infty)$, see fig. 4. From this, it follows easily that

$$\inf_{t \in [0,1]} tI\left(\frac{a}{t}\right) = I(a) \text{ if } a > z_0,$$

and

$$\inf_{t \in [0,1]} tI\left(\frac{a}{t}\right) = \frac{a}{z_0} I(z_0) \text{ if } a \leqslant z_0,$$

i.e.,

$$\inf_{t \in [0,1]} h(r_t) = h(q_a).$$

Hence, $h(f) \geqslant h(q_a)$ for all $f \in X \cap \Gamma$.

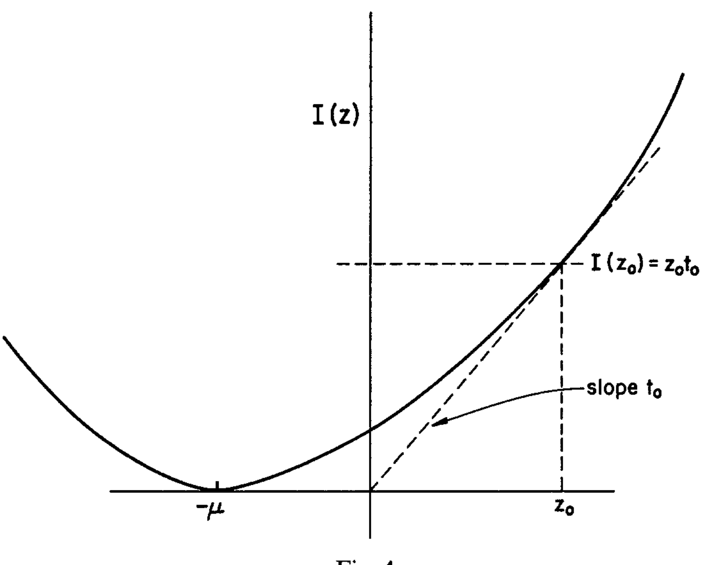

Fig. 4.

Now, pick arbitrary $g \in X$ and $\{f_n\} \subseteq \Gamma$ such that $d_S(f_n, g) \to 0$ as $n \to \infty$. Since the Skorohod metric induces the topology of uniform convergence on C we have $\| f_n - g \|_\infty \to 0$ as $n \to \infty$. Since X is open, we see that $\{f_n\}$ eventually lies in $X \cap \Gamma$. So, liminf $h(f_n) \geqslant h(q_a)$. Varying $\{f_n\}$, we have $h(g) \geqslant h(q_a)$, by Defn. (2.1). Since $g \in X$ was arbitrary we have $H(X) \geqslant h(q_a)$, by defn. (2.2).

Finally, the sequence

$$q_{a+1/n}(\cdot) \in X \cap \Gamma$$

converges uniformly to $q_a(\cdot)$. Hence $H(X) \leqslant h(q_a)$, completing the proof.

LEMMA 2.6

$H(X) = H(\overline{X}^E)$.

Proof

Let $g \in \overline{X}^E$ and $\{f_n\} \subseteq \Gamma$ such that $d_S(f_n, g) \to 0$ as $n \to \infty$. We claim that $\liminf_{n \to \infty} \| f_n \|_\infty \geqslant a$. Since $g \in \overline{X}^E$, there is $\{g_n\} \subseteq X \cap \Gamma$ such that $d_S(g_n, g) \to 0$ as $n \to \infty$. Suppose $\liminf_{n \to \infty} \| f_n \|_\infty < a - \epsilon$, $\epsilon > 0$. We can find $N(\epsilon)$ so large such that $d_S(g_m, f_n) < \epsilon/2$ for all m, $n \geqslant N(\epsilon)$. Let $N > N(\epsilon)$ be such that $\| f_N \|_\infty < a - \epsilon$. Choose $m > N(\epsilon)$ and let $\bar{t} \in [0, 1]$ be such that $g_m(\bar{t}) > a$. For any time changes $\lambda(\cdot)$ and $\eta(\cdot)$, we have,

$$\sup_{t \in [0,1]} | g_m(\lambda(t)) - f_N(\eta(t)) | \geqslant | g_m(\bar{t}) - f_n(\eta(\lambda^{-1}(\bar{t}))) | > \epsilon/2.$$

This contradiction establishes the claim.

Since $\liminf_{n \to \infty} \| f_n \|_\infty \geqslant a$, given $\epsilon > 0$, we can find $N(\epsilon)$ so large that $\| f_n \|_\infty > a - \epsilon$ for all $n \geqslant N(\epsilon)$. By lemma 2.5 it follows that $\liminf_{n \to \infty} f_n \geqslant h(q_{a-\epsilon})$. Since this applies to every $\{f_n\} \subseteq \Gamma$ converging to g, we have $h(g) \geqslant h(q_{a-\epsilon})$. Since $g \in \overline{X}^E$ was arbitrary, we have $H(\overline{X}^E) \geqslant h(q_{a-\epsilon})$. Letting $\epsilon \to 0$ establishes the lemma. □

Let

$$\overline{B}(q_a, \epsilon) = \left\{ f \in C \mid \sup_{t \in [0,1]} | f(t) - q_a(t) | \leqslant \epsilon \right\},$$

and

$$B(q_a, \epsilon) = \left\{ f \in C \mid \sup_{t \in [0,1]} | f(t) - q_a(t) | < \epsilon \right\}.$$

Let $Y_\epsilon = X - \overline{B}(q_a, \epsilon)$. Then Y_ϵ is open in C. We also have:

LEMMA 2.7

Given $\epsilon > 0$, there is $\eta > 0$ such that $H(Y_\epsilon) \geqslant h(q_a) + \eta$.

Proof

It suffices to find $\eta > 0$ such that $h(f) > h(q_a) + \eta$ for all $f \in Y_\epsilon \cap \Gamma$. For, if $g \in Y_\epsilon$ and $\{f_n\} \subseteq \Gamma$ is such that $d_S(f_n, g) \to 0$ as $n \to \infty$, then, since Y_ϵ is open, $\{f_n\}$ eventually consists of elements of $Y_\epsilon \cap \Gamma$, from which it follows that $h(g) \geqslant h(q_a) + \eta$.

Pick $0 < \delta < \min(\epsilon/3z_0, \epsilon/3\mu)$. Using the notation in the proof of lemma 2.5, we have

$$\| r_t - q_a \|_\infty < \epsilon/3 \text{ for all } t \in (\tau(a) - \delta, \tau(a) + \delta) \cap [0, 1].$$

We also observe that there is $\eta_1 > 0$ such that

$$h(r_1) > h(q_a) + \eta_1 \text{ for all } t \notin (\tau(a) - \delta, \tau(a) - \delta, \tau(a) + \delta), 0 \leqslant t \leqslant 1.$$

Given $f \in Y_\epsilon \cap \Gamma$, let $\bar{t} = \inf\{t \in [0, 1] \mid f(t) > a\}$. If $\bar{t} \notin (\tau(a) - \delta, \tau(a) + \delta)$, we directly observe that $h(f) \geqslant h(r_{\bar{t}})$. If $\bar{t} \in (\tau(a) - \delta, \tau(a) + \delta)$, we note that $\| f - r_{\bar{t}} \|_\infty > 2\epsilon/3$. Pick $t^* \in [0, 1]$ such that $| f(t^*) - r_{\bar{t}}(t^*) | > 2\epsilon/3$.

We distinguish 4 cases:

Case 1: $f(t^*) > r_{\bar{t}}(t^*) + 2\epsilon/3$ and $t^* < \bar{t}$.

Let $p(\cdot)$ be the piecewise linear path connecting (0, 0), $(t^*, f(t^*))$ and $(\bar{t}, a)$ and then dropping off with slope $-\mu$. Clearly, $h(f) \geqslant h(p)$ and $h(p) \geqslant h(r_{\bar{t}})$. We have

$$h(p) = t^* I\left(\frac{f(t^*)}{t^*}\right) + (\bar{t} - t^*) I\left(\frac{a - f(t^*)}{\bar{t} - t^*}\right).$$

Because of the constraint on $f(t^*)$, it is easy to see that there is $\eta_2 > 0$ such that $h(p) > h(r_{\bar{t}}) + \eta_2$. So, $h(f) > h(q_a) + \eta_2$.

Case 2: $f(t^*) < r_{\bar{t}}(t^*) - 2\epsilon/3$ and $t^* < \bar{t}$.

Exactly as in case 1, we can show that there is $\eta_3 > 0$ such that $h(f) > h(q_a) + \eta_3$.

Case 3: $f(t^*) > r_{\bar{t}}(t^*) + 2\epsilon/3$ and $t^* > \bar{t}$.

Let $p(\cdot)$ denote the piecewise linear path connecting the points (0, 0), $(\bar{t}, a)$ and $(t^*, f(t^*))$ and then dropping off with slope $-\mu$. Clearly $h(f) \geqslant h(p)$ and $h(p) \geqslant h(r_{\bar{t}})$. We have

$$h(p) = \bar{t} I\left(\frac{a}{\bar{t}}\right) + (t^* - \bar{t}) I\left(\frac{f(t^*) - a}{t^* - \bar{t}}\right).$$

Because of the constraint on $f(t^*)$, we see that there is $\eta_4 > 0$ such that $h(p) > h(r_{\bar{t}}) + \eta_4$. So, $h(f) > h(q_a) + \eta_4$.

Case 4: $f(t^*) < r_{\bar{t}}(t^*) + 2\epsilon/3$ and $t^* > \bar{t}$.

Exactly as in Case 3, we can show that there is $\eta_5 > 0$ such that $h(f) > h(q_a) + \eta_5$. Letting $\eta = \min(\eta_1, \eta_2, \ldots, \eta_5)$ completes the proof. □

Finally, we have the following.

LEMMA 2.8

There is a countable set $U \subseteq R_+$ such that, for $\epsilon \notin U$, $H(\overline{Y_\epsilon^E}) = H(Y_\epsilon)$.

Proof

Considered as a function of ϵ for $\epsilon > 0$, $H(Y_\epsilon)$ is nondecreasing. Hence there is a countable set $U \subseteq R_+$ such that $H(Y_\epsilon)$ is continuous in ϵ for $\epsilon \notin U$.

Fix $0 < \delta < \epsilon$. Let $g \in \overline{Y}_\epsilon^E$ and $\{f_n\} \subseteq \Gamma$ be such that $d_S(f_n, g) \to 0$ as $n \to \infty$. Let $\{g_n\} \subseteq Y_\epsilon \cap \Gamma$ be such that $d_S(g_n, g) \to 0$ as $n \to \infty$. Let $K = \max(z_0, \mu)$. Then there is $N(\epsilon, \delta)$ so large that for all $m, n \geqslant N(\epsilon, \delta)$ we have

$$d_S(g_m, f_n) < \frac{\epsilon - \delta}{K}.$$

We claim that

$$\limsup_{n\to\infty} 1(f_n \in B(q_a, \delta)) = 0,$$

i.e., f_n eventually leaves $B(q_a, \delta)$. Suppose not. Then we can find $N > N(\epsilon, \delta)$ such that $f_N \in B(q_a, \delta)$. Pick $m > N(\epsilon, \delta)$ and let t^* be such that $|g_m(t^*) - q_a(t^*)| > \epsilon$. For any time changes $\lambda(\cdot)$ and $\eta(\cdot)$, and $t \in [0, 1]$, we have,

$$|g_m(\lambda(t)) - f_N(\eta(t))| \geqslant |g_m(\lambda(t)) - q_a(\lambda(t)| - |f_N(\eta(t)) - q_a(\eta(t))| - |q_a(\lambda(t) - q_a(\eta(t))|.$$

Now, $|q_a(\lambda(t)) - q_a(\eta(t))| \leqslant K|\lambda(t) - \eta(t)|$. Hence

$$\begin{aligned} Kd_S(g_m, f_N) &\geqslant \sup_{t\in[0,1]} |g_m(\lambda(t)) - f_N(\eta(t))| + K \sup_{t\in[0,1]} |\lambda(t) - \eta(t)| \\ &\geqslant |g_m(t^*) - q_a(t^*)| - \sup_{t\in[0,1]} |f_N(t) - q_a(t)| \\ &> \epsilon - \delta. \end{aligned}$$

This contradiction establishes the claim.

From the claim we have $h(g) \geqslant h(Y_\delta)$ for every $g \in \overline{Y}_\epsilon^E$ and any $\delta < \epsilon$. If $\epsilon \notin U$, then, letting $\delta \to \epsilon$ gives $h(g) \geqslant H(Y_\epsilon)$ for every $g \in \overline{Y}_\epsilon^E$, so $h(\overline{Y}_\epsilon^E) \geqslant H(Y_\epsilon)$. It is trivial to see that $H(Y_\epsilon) \geqslant H(\overline{Y}_\epsilon^E)$, which completes the proof. □

Proof of thm. 2.3.

Let $F: C \to R$ be a bounded continuous function, and $K < \infty$ be such that $F(g) \leqslant K$ for all $g \in C$. We have to show that $\int F \, dQ_n \to F(q_a)$ as $n \to \infty$. Given $\Delta > 0$, we may choose $\epsilon > 0$, $\epsilon \notin U$, such that $|F(g) - F(q_a)| < \Delta$ for all $g \in \overline{B}(q_a, \epsilon)$. From lemmas 2.5 and 2.6, and thm. 2.4, we have

$$\log P(s_n(\cdot) \in X) = -nh(q_a) + o(n), \tag{2.3}$$

while from lemmas 2.7 and 2.8, and thm. 2.4, we have

$$\log P(s_n(\cdot) \in Y_\epsilon) \leqslant -n(h(q_a) + \eta), \tag{2.4}$$

for some $\eta > 0$. Now,

$$\begin{aligned} \left|\int F \, dQ_n - F(q_a)\right| &\leqslant \left|\int_{\overline{B}(q_a,\epsilon)} F \, dQ_n - F(q_a)\right| + \left|\int_{Y_\epsilon} F \, dQ_n\right| \\ &\leqslant \Delta Q_n(\overline{B}(q_a, \epsilon)) + KQ_n(Y_\epsilon) \\ &= \Delta \frac{P(s_n(\cdot) \in \overline{B}(q_a, \epsilon))}{P(s_n(\cdot) \in X)} + K\frac{P(s_n(\cdot) \in Y_\epsilon)}{P(s_n(\cdot) \in X)}. \end{aligned}$$

From eqns. (2.3) and (2.4), we have

$$\limsup_{n\to\infty} \left| \int F \, \mathrm{d}Q_n - F(q_a) \right| \leqslant \Delta .$$

Letting $\Delta \to 0$ completes the proof. □

3. Regularly varying positive tail

Throughout this section we assume the moment condition (D). Recall that a function $L(\cdot)$ is called slowly varying if

$$\frac{L(tx)}{L(x)} \to 1 \text{ as } x \to \infty \text{ for any } t > 0. \tag{3.1}$$

When $P(X_1 > x) \sim x^{-q} L(x)$ as $x \to \infty$, with $q > 0$ and $L(\cdot)$ slowly varying, X_1 is said to have regularly varying positive tail. For more about regularly varying distributions see Feller, secs. VIII.8 and VIII.9, [8].

The main theorem of this section is the following:

THEOREM 3.1

Under the moment conditions (D)

$$\hat{P}_n \xrightarrow{w} [J_T - \mu(\cdot - T)] 1(T < \cdot),$$

where T is distributed on [0, 1] with

$$P(T > t) = \frac{\displaystyle\int_t^1 \left(\frac{a + s\mu}{a} \right)^{-q} \mathrm{d}s}{\displaystyle\int_0^1 \left(\frac{a + s\mu}{a} \right)^{-q} \mathrm{d}s}.$$

and J_T is distributed on $[a + \mu(1-T), \infty)$ with

$$P(J_T > x) = \left(\frac{x}{a + \mu(1-T)} \right)^{-q}. \tag{3.2}$$

See fig. 5. Here $\xrightarrow{w}$ denotes weak convergence on D. □

Heuristically, thm. 3.1 says that the reason customer n incurred a waiting time of at least na is that a single (random) customer $[kT]$ incurred an enormous waiting time. The other customers preceding customer $[kT]$ encountered typical behaviour in the queue, while the customers k, $[kT] < k \leqslant n$, encountered typical behaviour in the queue but had large waiting times because of the waiting time of customer $[kT]$.

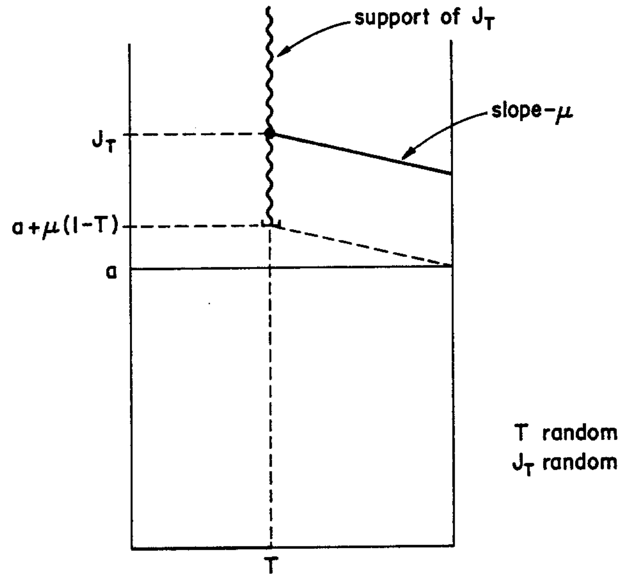

Fig. 5.

As in section 2, we can convert our problem into a problem of conditioning on large maxima in a random walk. Let $Y_1, Y_2, \ldots$ be i.i.d. with the distribution of X_1, and let:

$$S_0 = 0,$$

$$S_k = Y_1 + Y_2 + \ldots + Y_k, \quad k \geqslant 1,$$

$$M_n = \max_{0 \leqslant k \leqslant n} S_k.$$

From the points

$$\left(\frac{k}{n}, \frac{S_k}{n}\right), \quad 0 \leqslant k \leqslant n,$$

we construct the piecewise constant right continuous path $\hat{s}_n(\cdot) \in D$. Let $\hat{Q}_n$ denote the conditional distribution $(\hat{s}_n(\cdot) \mid M_n > na)$. Consider the function $\hat{\Psi}: D \to D$ given by

$$\hat{\Psi}(\phi(\cdot))(t) = \left[\left[\sup_{1-t \leqslant s \leqslant 1} \phi(s)\right] - \phi(1-t)\right]^{\text{r.c.}},$$

where $[f(\cdot)]^{\text{r.c.}}$ denotes the right continuous version of $f(\cdot)$. Then $\hat{\Psi}$ is continuous on D, and

$$\hat{P}_n = \hat{\Psi}_*(\hat{Q}_n).$$

The proof of this fact is identical to that of lemma 3.2.

By Billingsley, thm. 5.1, [5], thm. 3.1 reduces to the following:

THEOREM 3.2

Under the moment conditions (D)

$$\hat{Q}_n \xrightarrow{w} J_{1-\tau} 1(\tau < \cdot) - \mu \cdot ,$$

where τ is distributed on [0, 1] with

$$P(\tau < t) = \frac{\int_0^t \left(\frac{a + s\mu}{a} \right)^{-q} \mathrm{d}s}{\int_0^1 \left(\frac{a + s\mu}{a} \right)^{-q} \mathrm{d}s} . \qquad \square \tag{3.3}$$

The following theorem of Durrett, [7], will be used in the proof of thm. 3.2:

THEOREM 3.3

Under the moment conditions (D), for any $a > -\mu$,

$$(\hat{s}_n(\cdot) \mid S_n > na) \xrightarrow{w} J_0 1(U < \cdot) - \mu \cdot ,$$

where U is uniformly distributed on [0, 1].

Proof

This is thm. 3.1 of Durrett, [7]. □

To describe the essential idea in the proof of thm. 3.2, introduce the random variables

$$\hat{Y}_n = \max_{1 \leqslant k \leqslant n} \left(Y_k - (k-1)\mu \right),$$

and

$$N^a = \#\{ k \text{ s.t. } Y_k > na + (k-1)\mu \}.$$

Most of the work goes into showing that

$$P(M_n > na) \sim P(N^a = 1) \sim P(M_n > na,\ N^a = 1) \text{ as } n \to \infty . \tag{3.4}$$

Thus, conditioning on $\{ M_n > na \}$ is asymptotically equivalent to conditioning on $\{ N^a = 1 \}$. This reduction yields the theorem easily because the events $\{ Y_k > na + (k-1)\mu \}$, $1 \leqslant k \leqslant n$, are independent.

Before embarking on the proof, we need some preliminary lemmas. Lemmas 3.4 is Feller's lemma 2, pg. 277, [8].

LEMMA 3.4

The passage to the limit in eqn. (3.1) is uniform over finite intervals $t \in [a, b]$. □

An easy corollary of the above is the following:

COROLLARY 3.5

For any fixed C,

$$\frac{L(tx+C)}{L(x)} \to 1 \text{ as } x \to \infty \text{ for any } t > 0. \qquad \square$$

The final preliminary lemma is central to Durrett's proof of thm. 3.3, and will be used in our proof of thm. 3.2.

LEMMA 3.6

For any $a > -\mu$,

$$P(S_n > na) \sim nP(X_1 > n(a+\mu)) \text{ as } n \to \infty.$$

Proof

Let $Y_k' = Y_k - a$, $k = 1, 2, \ldots$, so $EY_k' = -\mu - a < 0$. Also,

$$P(Y_1' > x) = P(Y_1 > x + a) \sim (x+a)^{-q} L(x+a) \sim x^{-q} L(x) \text{ as } x \to \infty,$$

by cor. 3.5.

Let $S_n' = Y_1' + \ldots + Y_n' = S_n - na$. Then

$$\frac{P(S_n > na)}{nP(X_1 > n(a+\mu))} = \frac{P(S_n' > 0)}{nP(Y_1' > n(a+\mu) - a)}$$

$$\sim \frac{P(S_n' > 0)}{nP(Y_1' > n(a+\mu))} \text{ as } n \to \infty,$$

by cor. 3.5.

From thm. 2.1 of Durrett, [7], we see that the last quantity tends to 1 as $n \to \infty$. $\square$

We first investigate $P\,(N^a = 1)$. We have:

$$\{N^a \geqslant 1\} = \bigcup_{1 \leqslant k \leqslant n} \{Y_k > na + (k-1)\mu\}$$

$$= \{\hat{Y}_n > na\}.$$

LEMMA 3.7

$P(N^a = 1) \sim P(\hat{Y}_n > na)$ as $n \to \infty$.

Proof

The result is a consequence of lemma 3.8 below. Indeed, we have

$$P(N^a > 1 \mid \hat{Y}_n > na) = \frac{P(N^a > 1)}{P(\hat{Y}_n > na)}$$

with

$$P(N^a > 1) \leqslant \sum_{1 \leqslant i < j \leqslant n} P\big(Y_i > na + (i-1)\mu,\ Y_j > na + (j-1)\mu\big)$$

and

$$P\big(\hat{Y}_n > na\big) \geqslant \sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big) - \sum_{1 \leqslant i < j \leqslant n} P\big(Y_i > na + (i-1)\mu,\ Y_j > na + (j-1)\mu\big).$$

From lemma 3.8, we have

$$\frac{\sum_{1 \leqslant i < j \leqslant n} P\big(Y_i > na + (i-1)\mu,\ Y_j > na + (j-1)\mu\big)}{\sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big)} \to 0 \text{ as } n \to \infty$$

from which the claim follows.

LEMMA 3.8

$$P(\hat{Y}_n > na) \sim \sum_{k=1}^{n} P(Y_k > na + (k-1)\mu) \text{ as } n \to \infty.$$

Proof

Clearly,

$$P\big(\hat{Y}_n > na\big) \leqslant \sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big). \tag{3.5}$$

Also,

$$\begin{aligned} P\big(\hat{Y}_n > na\big) &\geqslant \sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big) \\ &\quad - \sum_{1 \leqslant i < j \leqslant n} P\big(Y_i > na + (i-1)\mu,\ Y_j > na + (j-1)\mu\big) \\ &\geqslant \sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big) - \frac{n(n-1)}{2}\big[P(Y_1 > na)\big]^2. \end{aligned}$$

So,

$$\begin{aligned} \frac{P\big(\hat{Y}_n > na\big)}{\sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big)} &\geqslant 1 - \frac{\frac{n(n-1)}{2}\big[P(Y_1 > na)\big]^2}{\sum_{k=1}^{n} P\big(Y_k > na + (k-1)\mu\big)} \\ &\geqslant 1 - \frac{\frac{n(n-1)}{2}\big[P(Y_1 > na)\big]^2}{nP\big(Y_1 > n(a+\mu)\big)}. \end{aligned} \tag{3.6}$$

Now,

$$\frac{P(Y_1 > na)}{P(Y_1 > n(a+\mu))} = \left(\frac{a}{a+\mu}\right)^{-q} \frac{L(na)}{L(n(a+\mu))}$$

$$\to \left(\frac{a}{a+\mu}\right)^{-q} \text{ as } n \to \infty, \tag{3.7}$$

and

$$nP(Y_1 > na) \leqslant \frac{EY_1^2}{na^2}. \tag{3.8}$$

From (3.5), (3.6), (3.7), and (3.8), the claim follows. □

We proceed to investigate $P(M_n > na)$.

LEMMA 3.9

Fix $0 < \epsilon < a \wedge \mu$. Then,

$$P(M_n > na) \sim P(M_n > na,\ S_n > n(a - \mu - \epsilon)) \text{ as } n \to \infty.$$

Proof

Write

$$P(M_n > na) = P(M_n > na,\ S_n > n(a-\mu-\epsilon)) + P(S_n \leqslant n(a-\mu-\epsilon) \mid M_n > na) P(M_n > na).$$

We first estimate $P(S_n \leqslant n(a-\mu-\epsilon) \mid M_n > na)$. Let

$$\tau = \inf\{1 \leqslant k \leqslant n : S_k > na\},$$

and

$$q_k = P(\tau = k \mid M_n > na).$$

Then

$$\begin{aligned}
&P(S_n \leqslant n(a-\mu-\epsilon) \mid M_n > na) \\
&= \sum_{k=1}^{n} P(S_n \leqslant n(a-\mu-\epsilon) \mid M_n > na,\ \tau = k) q_k \\
&\leqslant \sum_{k=1}^{n} P(Y_{k+1} + \ldots + Y_n \leqslant -n(\mu+\epsilon)) q_k \\
&= \sum_{k=1}^{n} P(Y_{k+1} + \ldots + Y_n + (n-k)\mu \leqslant (-k\mu - n\epsilon) q_k \\
&\leqslant \sum_{k=1}^{n} \frac{(n-k)E(X_1+\mu)^2}{(k\mu + n\epsilon)^2} q_k \\
&\leqslant \frac{E(X_1+\mu)^2}{\epsilon^2} \sum_{k=1}^{n} \frac{q_k(n-k)}{(n+k)^2} \\
&\leqslant \frac{1}{n\epsilon^2} E(X_1+\mu)^2.
\end{aligned}$$

Thus we have,

$$\frac{P(M_n > na,\ S_n > n(a-\mu-\epsilon))}{1-\frac{1}{n\epsilon^2}E(X_1+\mu)^2}$$

$$\geqslant P(M_n > na) \geqslant P(M_n > na,\ S_n > n(a-\mu-\epsilon)),$$

from which the claim follows. □

From thm. 3.3 we easily get

$$P(M_n > na \mid S_n > n(a-\mu-\epsilon)) \to \int_0^1 \left(\frac{a+s\mu}{a-\epsilon}\right)^{-q} \mathrm{d}s, \tag{3.9}$$

while from lemma 3.6 we have

$$P(S_n > n(a-\mu-\epsilon)) \sim nP(Y_1 > n(a-\epsilon)) \text{ as } n \to \infty. \tag{3.10}$$

Together with eqns. (3.9) and (3.10) and lemmas (3.7)–(3.9), the following lemma completes the proof of the asymptotic equivalence of $P(N^a = 1)$ and $P(M_n > na)$.

LEMMA 3.10

$$\frac{\sum_{k=1}^{n} P(Y_k > na + (k-1)\mu)}{nP(Y_1 > n(a-\epsilon))} \to \int_0^1 \left(\frac{a+s\mu}{a-\epsilon}\right)^{-q} \mathrm{d}s \text{ as } n \to \infty.$$

Proof

We have

$$\frac{\sum_{k=1}^{n} P(Y_k > na + (k-1)\mu)}{nP(Y_1 > n(a-\epsilon))} = \frac{\frac{1}{n}\sum_{k=1}^{n}\left(a+\frac{(k-1)}{n}\mu\right)^{-q} L(na+(k-1)\mu)}{(a-\epsilon)^{-q}L(n(a-\epsilon))}.$$

From lemma 3.4, we have

$$\frac{\sup_{t\in[a,a+\mu]} L(nt)}{L(n(a-\epsilon))} \to 1 \text{ as } n \to \infty, \tag{3.11}$$

and

$$\frac{\inf_{t\in[a,a+\mu]} L(nt)}{L(n(a-\epsilon))} \to 1 \text{ as } n \to \infty. \tag{3.12}$$

Also,

$$\frac{1}{n}\sum_{k=1}^{n}\left(\frac{a+\frac{(k-1)}{n}\mu}{a-\epsilon}\right)^{-q} \to \int_0^1\left(\frac{a+s\mu}{a-\epsilon}\right)^{-q} \mathrm{d}s \text{ as } n\to\infty. \tag{3.13}$$

From (3.11), (3.12), and (3.13) the claim follows. □

With the next lemma we will have established (3.4).

LEMMA 3.11

$P(M_n > na \mid N^a = 1) \to 1$ as $n \to \infty$.

Proof

Fix $b > a$. Write

$$P(M_n > na \mid N^a = 1)$$
$$> P(M_n > na \mid N^a = 1,\ \hat{Y}_n > nb) P(\hat{Y}_n > nb \mid N^a = 1). \tag{3.14}$$

Now,

$$\begin{aligned} P(\hat{Y}_n > nb \mid N^a = 1) &\geqslant \inf_{1\leqslant k\leqslant n} \frac{P(X_1 > nb + (k-1)\mu)}{P(X_1 > na + (k-1)\mu)} \\ &\geqslant \left(\frac{b}{a}\right)^{-q} \inf_{1\leqslant k\leqslant n} \frac{L(nb + (k-q)\mu)}{L(na + (k-1)\mu)} \\ &\geqslant \left(\frac{b}{a}\right)^{-q} \inf_{x\in[na, n(a+\mu)]}\ \inf_{t\in[b/(a+\mu), (b+\mu)/a]} \frac{L(tx)}{L(x)} \\ &\to \left(\frac{b}{a}\right)^{-q} \text{ as } n\to\infty, \end{aligned} \tag{3.15}$$

by lemma 3.4.

On $\{N^a = 1,\ \hat{Y}_n > nb\}$ there is a unique $\hat{k}$, $1 \leqslant \hat{k} \leqslant n$, such that $Y_{\hat{k}} > nb + (\hat{k} - 1)\mu$. Then

$$\begin{aligned} &P(M_n > na \mid N^a = 1,\ \hat{Y}_n > nb) \\ &\quad = \sum_{k=1}^{n} P(M_n > na \mid N^a = 1,\ \hat{Y}_n > nb,\ \hat{k} = k) \\ &\qquad\qquad P(\hat{k} = k \mid N^a = 1,\ \hat{Y}_n > nb). \end{aligned} \tag{3.16}$$

Define independent random variables $\tilde{Y}_k$, $1 \leqslant k \leqslant n$, with distribution

$$P(\hat{Y}_k \leqslant x) = \frac{P(Y_k \leqslant (x-\mu) \wedge (na + (k-1)\mu))}{P(Y_k \leqslant na + (k-1)\mu)},$$

and independent random variables Y_k^*, $1 \leqslant k \leqslant n$, with the distribution of $\tilde{Y}_1$. Let

$$\tilde{S}_k = \tilde{Y}_1 + \ldots + \tilde{Y}_k,$$

and

$$S_k^* = Y_1^* + \ldots + Y_k^*.$$

If we let

$$p_k = P(\tilde{S}_k \leqslant n(a-b)),$$

then we have

$$P(M_n > na \mid N^a = 1,\ \hat{Y}_n > nb,\ \hat{k} = k) \geqslant 1 - p_{k-1}, \tag{3.17}$$

because the conditional distribution of $\{Y_j + \mu,\ 1 \leqslant j \leqslant k-1\}$, given $N^a = 1$ and $\hat{k} = k$ is the distribution of $\tilde{Y}_j,\ 1 \leqslant j \leqslant k-1\}$.

Since $E(Y_1 + \mu) = 0$, we have $E\tilde{Y}_1 > (a-b)/2$ for large enough n. So,

$$\begin{aligned} p_k &\leqslant P(S_k^* \leqslant n(a-b)) = P(S_k^* - kEY_1^* \leqslant n(a-b) - kEY_1^*) \\ &\leqslant P\left(S_k^* - kEY_1^* \leqslant \frac{n(a-b)}{2}\right) \\ &\leqslant \frac{4k \text{ var } Y_1^*}{n^2(a-b)^2}. \end{aligned}$$

Since $EY_1^2 < \infty$, there is a uniform bound on var Y_1^* for all large enough n. Hence,

$$\sup_{1 \leqslant k \leqslant n-1} p_k \to 0 \text{ as } n \to \infty.$$

This, together with (3.16) and (3.17), gives

$$P(M_n > na \mid N^a = 1,\ \hat{Y}_n > nb) \to 1 \text{ as } n \to \infty. \tag{3.18}$$

Putting (3.14), (3.15) and (3.18) together, and letting $b \to a$, yields the claim. □

Proof of thm. 3.2

From (3.4), which we have established by the sequence of lemmas above, it suffices to show that

$$(\hat{s}_n(\cdot) \mid N^a = 1) \xrightarrow{w} J_{1-\tau} 1(\tau < \cdot) - \mu \cdot,$$

with τ and $J_{1-\tau}$ distributed as in (3.3) and (3.2).

On $\{N^a = 1\}$ there is a unique index $\hat{k}$, $1 \leqslant \hat{k} \leqslant n$, such that $Y_{\hat{k}} > na + (\hat{k} - 1)\mu$. Clearly,

$$\begin{aligned} P(\hat{k} \leqslant j \mid N^a = 1) &= \frac{\sum_{k=1}^{j} P(Y_k > na + (k-1)\mu)}{\sum_{k=1}^{n} P(Y_k > na + (k-1)\mu)} \\ &\leqslant \frac{\frac{1}{n}\sum_{k=1}^{j}\left(a + \frac{k-1}{n}\mu\right)^{-q}}{\frac{1}{n}\sum_{k=1}^{n}\left(a + \frac{k-1}{n}\mu\right)^{-q}} \frac{\sup_{1 \leqslant k \leqslant j} L(na + (k-1)\mu)}{\inf_{1 \leqslant k \leqslant n} L(na + (k-1)\mu)}. \end{aligned}$$

Now,

$$\frac{\sup_{1\leqslant k\leqslant j} L(na+(k-1)\mu)}{\inf_{1\leqslant k\leqslant n} L(na+(k-1)\mu)} \leqslant \sup_{x\in[na,n(a+\mu)]} \sup_{t\in[a/(a+\mu),(a+\mu)/a]} \frac{L(tx)}{L(x)},$$

which tends to 1 as $n\to\infty$, by lemma 3.4.

We get a similar lower bound for $P(\hat{k}\leqslant j \mid N^a=1)$. Observing that

$$\frac{\frac{1}{n}\sum_{k=1}^{[nt]}\left(a+\frac{k-1}{n}\mu\right)^{-q}}{\frac{1}{n}\sum_{k=1}^{n}\left(a+\frac{k-1}{n},\ u\right)^{-q}} \to \frac{\int_0^t\left(\frac{a+s\mu}{a}\right)^{-q}\mathrm{d}s}{\int_0^1\left(\frac{a+s\mu}{a}\right)^{-q}\mathrm{d}s} \quad \text{as } n\to\infty,$$

we have

$$P(\hat{k}\leqslant[nt]\mid N^a=1)\to\frac{\int_0^t\left(\frac{a+s\mu}{a}\right)^{-q}\mathrm{d}s}{\int_0^1\left(\frac{a+s\mu}{a}\right)^{-q}\mathrm{d}s}. \tag{3.19}$$

The big jump $Y_{\hat{k}}$ satisfies

$$\begin{aligned} P\left(Y_{\hat{k}}>nb\mid N^a=1,\ \hat{k}=k\right) &= \frac{P(Y_1>nb)}{P(Y_1>na+(k-1)\mu)} \\ &= \left(\frac{b}{a+\frac{k-1}{n}\mu}\right)^{-q}\frac{L(nb)}{L(na+(k-1)\mu)}. \end{aligned} \tag{3.20}$$

From (3.19) and (3.20) it follows that

$$\left(\frac{Y_{\hat{k}}}{n}1(\hat{k}\leqslant n\cdot)-\mu\cdot\mid N^a=1\right)\xrightarrow{w} J_{1-\tau}1(\tau<\cdot)-\mu\cdot,$$

with τ and $J_{1-\tau}$ distributed as in (3.3) and (3.2). To complete the proof of the theorem it suffices to establish that

$$\left(\hat{s}_n(\cdot)-\left[\frac{Y_{\hat{k}}}{n}1(\hat{k}\leqslant n\cdot)-\mu\cdot\right]\mid N^a=1\right)\xrightarrow{w}\delta_0,$$

where δ_0 denotes the probability distribution concentrated on the zero function. This is done in lemma 3.12 below.

LEMMA 3.12

Let $\hat{k}$, $1\leqslant\hat{k}\leqslant n$, denote the unique index on $\{N^a=1\}$ such that $Y_{\hat{k}}>na+(\hat{k}-1)\mu$. *Then,*

$$\left(\hat{s}_n(\cdot)-\left[\frac{Y_{\hat{k}}}{n}1(\hat{k}\leqslant n\cdot)-\mu\cdot\right]\mid N^a=1\right)\xrightarrow{w}\delta_0.$$

Proof

Let $\tilde{Y}_k$, $1 \leqslant k \leqslant n$, be independent random variables with distribution

$$P(\tilde{Y}_k \leqslant x) = \frac{P(Y_k \leqslant (x-\mu) \wedge (na + (k-1)\mu))}{P(Y_k \leqslant na + (k-1)\mu)}.$$

Let

$$\tilde{S}_0 = 0,$$
$$\tilde{S}_k = \tilde{Y}_1 + \cdots + \tilde{Y}_k,$$

and consider the piecewise constant, right continuous path $\tilde{s}_n(\cdot)$ constructed from the points $(k/n, \tilde{S}_k/n)$, $0 \leqslant k \leqslant n$. Let $\tilde{k}$ be independent of $(\tilde{Y}_1, \ldots, \tilde{Y}_n)$, taking values in $\{1, \ldots, n\}$, with distribution

$$P(\tilde{k} = k) = \frac{P(X_1 > na + (k-1)\mu)}{\sum_{j=1}^{n} P(X_1 > na + (j-1)\mu)}.$$

We erase the jump $\tilde{Y}_{\tilde{k}}$ and consider the resulting path

$$\tilde{s}_n(\cdot) - \frac{\tilde{Y}_{\tilde{k}}}{n} 1([n\cdot] \geqslant \tilde{k}).$$

The distribution of this path is clearly the same as the distribution

$$\left(\hat{s}_n(\cdot) - \left[\frac{Y_{\hat{k}}}{n} 1(\hat{k} \leqslant n\cdot) - \mu \frac{[n\cdot]}{n}\right] \mid N^a = 1\right).$$

Hence, to prove the lemma, it suffices to show that, for any $\epsilon > 0$,

$$P\left(\sup_{0 \leqslant t \leqslant 1} \left| \tilde{s}_n(t) - \frac{\tilde{Y}_{\tilde{k}}}{n} 1([nt] \geqslant \tilde{k}) \right| > \epsilon\right) \to 0 \text{ as } n \to \infty,$$

or, equivalently, that

$$P\left(\sup_{0 \leqslant k \leqslant n} |\tilde{S}_k - \tilde{Y}_{\tilde{k}} 1(k \geqslant \tilde{k})| > n\epsilon\right) \to 0 \text{ as } n \to \infty.$$

In our notation we suppressed the dependence of the distribution of $\tilde{Y}_k$, $1 \leqslant k \leqslant n$, on n. From $EX_1^2 < \infty$, it is easy to see that there is $K < \infty$ such that $E|\tilde{Y}_k|^2 < K$ for all $1 \leqslant k \leqslant n$ and all n. By Chebyshev's inequality, see Feller, pg. 151, [8], we have,

$$\frac{n^2\epsilon^2}{4} P\left(|\tilde{Y}_k| > \frac{n\epsilon}{2}\right) \leqslant K \text{ for all } 1 \leqslant k \leqslant n \text{ and all } n.$$

So,

$$\sup_{1 \leqslant k \leqslant n} P\left(|\tilde{Y}_k| > \frac{n\epsilon}{2}\right) \to 0 \text{ as } n \to \infty. \tag{3.21}$$

Next, note that $\tilde{S}_k^2$, $0 \leqslant k \leqslant n$, is a positive integrable supermartingale. By Doob's inequality, see Neveu, lemma IV-2-9, [10],

$$\frac{n^2\epsilon^2}{4} P\left(\sup_{1 \leqslant k \leqslant n} |\tilde{S}_k| > \frac{n\epsilon}{2} \right) \leqslant E\tilde{S}_n^2 \leqslant nK.$$

So,

$$P\left(\sup_{1 \leqslant k \leqslant n} |\tilde{S}_k| > \frac{n\epsilon}{2} \right) \to 0 \text{ as } n \to \infty. \tag{3.22}$$

The lemma follows directly from eqns. (3.21) and (3.22). □

4. Concluding remarks

Soren Asmussen has kindly pointed out that theorem 2.3 can also be directly derived by an application of Wald's identity for the exponentially transformed random walk and a central limit theorem for the time of level crossing, [1]. The techniques for such a derivation are available in [2] and [3], see especially [2], cor. 3.1. Using these techniques, it is also possible to get somewhat more information, for example when $a \leqslant z_0$ one gets a Brownian bridge as correction to the straight line.

Proofs of theorem 2.3 under the assumption that the moment generating function $m(t)$ is defined for all t are available in the papers of Borovkov, [7], and Picard and Deshayes, [11]. It appears that the technique of Asmussen, [1], will yield the result under the assumption that $m(t)$ is defined for an interval of type $[0, t_+)$, with $t_+ > 0$ and $m'(t)/m(t) \to \infty$ as $t \to t_+$, which is weaker than our condition (C).

References

[1] S. Asmussen, Private Communication, 1988.
[2] S. Asmussen, Conditioned limit theorems relating a random walk to its associate, with applications to risk reserve processes and the GI/G/1 queue, Advances in Applied Probability 14 (1982) 143–170.
[3] S. Asmussen, *Applied Probability and Queues* (John Wiley and Sons, 1987).
[4] H. Bergstrom, *Weak Convergence of Measures* (Academic Press, 1982).
[5] P. Billingsley, *Convergence of Probability Measures* (John Wiley and Sons, 1968).
[6] A.A. Borovkov, Boundary value problems for random walks and large deviations in function spaces, Theory of Probability and its Applications 12, No. 4 (1967) 575–595.
[7] R. Durett, Conditional limit theorems for random walks with negative drift, Zeit. fur Wahr. und Ver. Geb. 52 (1980) 277–287.
[8] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II (John Wiley and Sons, 1971).

[9] A.A. Mogulskii, Large deviations for trajectories of multidimensional random walks, Theory of Probability and its Applications XXI, No. 2 (1976) 300–315.
[10] J. Neveu, *Discrete Parameter Martingales* (North-Holland, 1975).
[11] D. Picard and J. Deshayes, Grandes et moyennes deviations pour les marches aleatories, in: *Grandes Deviations et applications statistiques* (Seminaire Orsay 1977-1978, Asterisque, Vol. 68, 1979) pp. 53–71.