Queueing Systems 44, 109–123, 2003


# Properties of the Reflected Ornstein–Uhlenbeck Process

AMY R. WARD amy@isye.gatech.edu
*School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205, USA*

PETER W. GLYNN glynn@stanford.edu
*Department of Management Science & Engineering, Stanford University, Stanford, CA 94305, USA*

Received 20 July 2002; Revised 5 November 2002

**Abstract.** Consider an Ornstein–Uhlenbeck process with reflection at the origin. Such a process arises as an approximating process both for queueing systems with reneging or state-dependent balking and for multi-server loss models. Consequently, it becomes important to understand its basic properties. In this paper, we show that both the steady-state and transient behavior of the reflected Ornstein–Uhlenbeck process is reasonably tractable. Specifically, we (1) provide an approximation for its transient moments, (2) compute a perturbation expansion for its transition density, (3) give an approximation for the distribution of level crossing times, and (4) establish the growth rate of the maximum process.

**Keywords:** diffusion approximation, Ornstein–Uhlenbeck process, reflecting diffusion, steady-state, transient moment, level crossing time, maximum process

## 1. Introduction

The reflected Ornstein–Uhlenbeck (O–U) process plays much the same role in the context of queueing models with reneging or balking as does reflected Brownian motion (RBM) in the setting of conventional queues. Specifically, in a related paper [21], we show that the reflected O–U process serves as a good approximation for a Markovian queue with reneging when the arrival rate is either close to or exceeds the processing rate and the reneging rate is small. In another related paper [20], and for the same part of the parameter space, we show that the reflected O–U process also well approximates queues having renewal arrival and service processes in which customers have deadlines constraining total sojourn time. Customers either renege from the queue when their deadline expires or balk if the conditional expected waiting time given the queue-length exceeds their deadline.

Another setting in which the reflected O–U process is relevant is that of multiserver loss models. Both Borovkov [5] and Srikant and Whitt [19] (under different conditions) find that the reflected O–U process serves as a good approximation for the number-in-system process in a $G/M/s/0$ queueing model when both the number of servers and the arrival rate are large.

Given the central role of the reflected O–U process as an approximation to a number of different queueing models, the study of the structure of the reflected O–U process is clearly of some importance. Our goal, in this paper, is to initiate the study of the properties and structure of the reflected O–U process. A mathematical complication in this work is that many of the methods that have been successfully employed in dealing with reflected Brownian motion do not apply to the analysis of reflected O–U. This is because reflected O–U cannot be defined as the image of a "free" process under a continuous regulator mapping as RBM is in [9]. (The presence of the state-dependent drift rate in (1.1) makes the concept of a "free" process meaningless.)

Suppose $X$ is a reflected O–U process with infinitesimal drift $\alpha - \gamma x$ and infinitesimal variance $\sigma^2$, where $\gamma, \sigma > 0$. The process $X$ can be defined as the strong solution (whose existence is guaranteed by a careful extension of the results of Lions and Sznitman [15], which treat bounded domains) to the stochastic differential equation (SDE):

$$\mathrm{d}X(t) = \big(a - \gamma X(t)\big)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t) + \mathrm{d}L(t), \tag{1.1}$$

subject to $X(0) = x \geqslant 0$. Here $L = (L(t)\colon t \geqslant 0)$ is the minimal nondecreasing process which makes $X(t) \geqslant 0$ for $t \geqslant 0$. The process $L$ increases only when $X$ is zero, so that

$$\int_{[0,\infty)} I\big(X(t) > 0\big)\,\mathrm{d}L(t) = 0,$$

where $I(\cdot)$ is the indicator function.

This paper offers the first extensive study of the properties of the reflected O–U process and our body of results expands what is known about the reflected O–U process, both qualitatively and quantitatively. Specifically, our contributions are as follows:

1. To provide a rigorous proof of the formula for its steady-state distribution; see proposition 1.
2. To establish a stochastic ordering result for RBM and reflected O–U; see proposition 2.
3. To offer an approximation for its transient moments via the more tractable RBM process; see (3.4).
4. To obtain a perturbation expansion (also via the more tractable RBM process) for its transition density; see (3.6).
5. To obtain approximations for level crossing times; see theorem 2 and corollary 3.
6. To study the growth rate of its maximum process; see theorem 4 and corollary 5.

Recall that the reflected O–U is the key approximating process for reneging models, certain balking models, and multiserver loss models. Therefore, knowing the steady-state distribution of reflected O–U, and having approximations for its transient moments and transition density, enhances our understanding of how these models behave over various time scales. In particular, such formulae give insight into how these systems approach steady-state. Our motivation for studying level crossing times for reflected O–U comes from the analysis of finite buffer systems. There, a key performance characteristic is the time to buffer overflow, and the distribution of this random variable coincides with the corresponding level crossing time for the infinite buffer analog. Finally, there has been significant interest historically within the queueing community on understanding the behavior of maximum processes; see [3] for a recent survey and [18] for a study in the context of birth–death processes. This is because the behavior of the maximum process provides insight into the "worst-case" behavior of the number-in-system process for a queueing model over a given time interval.

Since a RBM process is well known to be analytically tractable (for example, see [9]), one might be tempted, at least in settings where $\gamma$ is close to zero, to replace a reflected O–U approximation with a RBM approximation by letting $\gamma = 0$ in equation (1.1). In particular, one might wish to approximate a queue with small amounts of reneging or balking with a RBM. Therefore, in stating our results for a reflected O–U process, for purposes of comparison, we also state the equivalent result for a RBM process. Our results confirm what one would suspect by examining the SDE's defining the two processes: the behavior of the reflected O–U process is significantly different than that of a RBM process. This conclusion reinforces the results of the numerical study in [21]. There, it was shown that the RBM approximation to a queue with reneging has surprisingly large error relative to that associated with reflected O–U, even for very small reneging rates.

The rest of this paper is organized as follows. In section 2, we discuss the steady-state of the reflected O–U process. In section 3, we provide an approximation for the transient moments and the transition density of a reflected O–U process in terms of those for a RBM process. In section 4, we provide an approximation for the distribution of level crossing times valid for large levels. Finally, in section 5, we show how the maximum process of a reflected O–U grows with time.

## 2. Steady-state behavior

Perhaps the single most important performance measure for a stochastic process is its steady-state distribution. Therefore, we begin our study of the reflected O–U process with a theorem that computes this distribution. Related computations of the steady-state of reflected O–U can be found in [6,7].

**Proposition 1.** Let $X$ be the reflected O–U process specified in (1.1). Then, $X$ has a unique stationary distribution $\pi$ with density

$$\begin{aligned} p(x) &= P\left[N\left(\frac{\alpha}{\gamma}, \frac{\sigma^2}{2\gamma}\right) \in \mathrm{d}x \,\middle|\, N\left(\frac{\alpha}{\gamma}, \frac{\sigma^2}{2\gamma}\right) \geqslant 0\right] \\ &= \sqrt{\frac{2\gamma}{\sigma^2}} \frac{\phi(\sqrt{2\gamma/\sigma^2}(x-\alpha/\gamma))}{1-\Phi(-\sqrt{2\alpha^2/\gamma\sigma^2})} \end{aligned}$$

for $x \geqslant 0$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution of a $N(0,1)$ random variable, respectively. Furthermore, $X(t) \Rightarrow X(\infty)$ as $t \to \infty$, where $X(\infty)$ has distribution $\pi$.

*Proof.* Suppose $\pi$ is a probability measure with density $p$. Choose $f$ twice continuously differentiable on $[0,\infty)$ with compact support and so that $f'(0)=0$. It is straightforward to verify (using integration by parts) that

$$\int_{[0,\infty)} (Af)(x)\pi(\mathrm{d}x) = 0 \tag{2.1}$$

for all such $f$. On the other hand, Itô calculus implies that

$$E_x f\big(X(t)\big) - f(x) = \int_0^t E_x(Af)\big(X(s)\big)\,\mathrm{d}s. \tag{2.2}$$

It follows from (2.1) and (2.2) that

$$E_\pi f\big(X(t)\big) = E_\pi f\big(X(0)\big)$$

for all such $f$. This establishes that $P_\pi(X(t) \in \cdot) = P_\pi(X(0) \in \cdot)$ and hence $\pi$ is a stationary distribution.

We now prove that $X(t) \Rightarrow X(\infty)$ as $t \to \infty$. Suppose first that $X(0)=0$, and let $X^* = (X^*(t)\colon t \geqslant 0)$ be an independent version of $X$ initiated under $\pi$. It follows from proposition 4 in section 5, that $T^*(0) < \infty$ a.s. where $T^*(y) = \inf\{t \geqslant 0\colon X^*(t) = y\}$. Since $X(0) \leqslant X^*(0)$, and both $X$ and $X^*$ have continuous paths, evidently $T^* = \inf\{t \geqslant 0\colon X(t) = X^*(t)\} \leqslant T^*(0)$. Thus, $T^*$ is an a.s. finite coupling time for $X$. Consequently (see [13,14]), $X(t) \Rightarrow X(\infty)$ as $t \to \infty$. The extension to $X(0) > 0$ is straightforward.

To prove that $\pi$ is the unique stationary distribution, suppose that $\eta$ is another such distribution. Then, $P_\eta(X(t) \in \cdot) = P_\eta(X(0) \in \cdot) \Rightarrow \pi(\cdot)$, showing that $\eta = \pi$. □

Let $X^{\mathrm{R}}$ be a reflected Brownian motion with infinitesimal drift $\alpha$ and infinitesimal variance $\sigma^2$ so that $X^{\mathrm{R}}$ is the unique strong solution to the SDE given in (1.1) under the assumption that $\gamma = 0$. Here and throughout the rest of this paper, we use $X^{\mathrm{R}}$ to compare the properties of RBM with those of reflected O–U. When $\alpha < 0$, the steady-state distribution of $X^{\mathrm{R}}$ is exponential with mean $\sigma^2/2|\alpha|$. (If $\alpha \geqslant 0$, the process is not positive recurrent and so does not have a steady-state.) Notice the strikingly different tail behavior between the steady-state distributions of $X$ and $X^{\mathrm{R}}$. As one would expect, the probability of the reflected Brownian motion $X^{\mathrm{R}}$ being at high levels in steady state is much greater than that for the reflected O–U $X$.

## 3. Transient analysis

In this section, we study the computation of transient expectations of the form $E_x f(X(t))$, where $E_x(\cdot) \triangleq E(\cdot \mid X(0) = x)$. This further allows us to write a perturbation expansion for the transition density of $X$. We start our analysis with a an intuitive result showing that reflected O–U is always stochastically smaller than RBM. Recall that $Y_1 \overset{\text{st}}{\leqslant} Y_2$ if $Ef(Y_1) \leqslant Ef(Y_2)$ for all non-negative, nondecreasing functions $f$; see, for example, [17].

**Proposition 2.** Suppose that $X$ is a reflected O–U process with infinitesimal drift $\alpha-\gamma x$ and infinitesimal variance $\sigma^2$, starting from initial position $x_0$. Let $X^{\mathrm{R}}$ be a reflected Brownian motion process with infinitesimal drift $\alpha$ and infinitesimal variance $\sigma^2$, starting from initial position $x_1 \geqslant x_0$. Then, for each $t \geqslant 0$, $X(t) \overset{\text{st}}{\leqslant} X^{\mathrm{R}}(t)$.

*Proof.* Let $B = (B(t)\colon\ t \geqslant 0)$ be a standard Brownian motion and let

$$\begin{aligned} \mathrm{d}X(t) &= (\alpha - \gamma X(t))\,\mathrm{d}t + \sigma\,\mathrm{d}B(t) + \mathrm{d}L_1(t),\\ \mathrm{d}X^{\mathrm{R}}(t) &= \alpha\,\mathrm{d}t + \sigma\,\mathrm{d}B(t) + \mathrm{d}L_2(t), \end{aligned}$$

subject to $X(0) = x_0$, $X^{\mathrm{R}}(0) = x_1$, where $L_1$ and $L_2$ are the local time processes that increase only when $X$ and $X^{\mathrm{R}}$ are at the origin, respectively. We will show that $X(t) \leqslant X^{\mathrm{R}}(t)$ for $t \geqslant 0$, from which $X(t) \overset{\text{st}}{\leqslant} X^{\mathrm{R}}(t)$ follows immediately.

Suppose that there exists $t > 0$ for which $X(t) > X^{\mathrm{R}}(t)$. Since $X(0) \leqslant X^{\mathrm{R}}(0)$ and $X - X^{\mathrm{R}}$ has continuous sample paths, there exists $s \in [0, t)$ such that $X(s) = X^{\mathrm{R}}(s)$ with $X(u) > X^{\mathrm{R}}(u)$ for $s < u \leqslant t$. Note that

$$\begin{aligned} X(t) - X^{\mathrm{R}}(t) &= X(s) - X^{\mathrm{R}}(s) - \gamma \int_s^t X(u)\,\mathrm{d}u\\ &\quad + \big(L_1(t) - L_1(s)\big) - \big(L_2(t) - L_2(s)\big)\\ &\leqslant L_1(t) - L_1(s) \end{aligned}$$

since $X$ is non-negative and $\mathrm{L}_2$ is nondecreasing. But $X(u) > X^{\mathrm{R}}(u) \geqslant 0$ for $u \in (s, t]$, so $L_1(t) = L_1(s)$ (since $L_1$ is a continuous process that increases only when $X$ is at the origin). We conclude that $X(t) \leqslant X^{\mathrm{R}}(t)$, which is a contradiction. □

Note that proposition 2 provides an upper bound on $E_x f(X(t))$ in terms of the highly tractable process $X^{\mathrm{R}}$ (that is valid for every non-negative nondecreasing function $f$).

We turn next to computing $E_x f(X(t))$ itself. The computation of $u(t,x) = E_x f(X(t))$ involves solving the Kolmogorov backwards partial differential equation (PDE)

$$\frac{\partial}{\partial t}u(t,x) = (\alpha - \gamma x)\frac{\partial}{\partial x}u(t,x) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}u(t,x),$$

subject to $u(0,x) = f(x)$ and $(\partial/\partial x)u(t,0) = 0$. Unfortunately, this PDE appears to be generally intractable.

However, we note that if $\gamma = 0$, then the above PDE is precisely the backwards Kolmogorov PDE for RBM. As noted above, $X^{\mathrm{R}}$ has a quite tractable associated transient theory. In particular, its transition density can be explicitly computed in terms of the normal cumulative distribution function, and excellent approximations for its transient moments exist; refer to [9] and the series of papers by Abate and Whitt [1,2]. This suggests that if $\gamma$ is small, we ought to view the reflected O–U process as a small perturbation of an RBM process and thereby attempt to compute the first term of a perturbation expansion of reflected O–U in terms of RBM. This approach can be made mathematically rigorous by appealing to Girsanov's formula; see [10] for a discussion of this important result.

**Theorem 1.** Suppose that under the probability $P$, $B = (B(t)\colon\ t \geqslant 0)$ is a standard Brownian motion and

$$\mathrm{d}X^{\mathrm{R}}(t) = \alpha\,\mathrm{d}t + \sigma\,\mathrm{d}B(t) + \mathrm{d}L(t)$$

subject to $X^{\mathrm{R}}(0) = x \geqslant 0$, where $L$ is the local time process that increases only when $X^{\mathrm{R}}$ is zero. Put

$$M(t,\gamma) = \exp\biggl(-\int_0^t \frac{\gamma}{\sigma} X^{\mathrm{R}}(s)\,\mathrm{d}B(s) - \frac{\gamma^2}{2\sigma^2}\int_0^t X^{\mathrm{R}}(s)^2\,\mathrm{d}s\biggr).$$

Then:

(i) For each $\gamma \geqslant 0$, $(M(t,\gamma)\colon\ t \geqslant 0)$ is a martingale adapted to the filtration generated by $B$, so that for each $t \geqslant 0$, $P_{\gamma,t}$ is a probability, where $\mathrm{d}P_{\gamma,t} = M(t,\gamma)\,\mathrm{d}P$.

(ii) Under $P_{\gamma,t}$, $(X^{\mathrm{R}}(s)\colon\ 0 \leqslant s \leqslant t)$ is a reflected O–U process having drift $\alpha - \gamma x$, infinitesimal variance $\sigma^2$, and starting from initial position $x$.

*Proof.* Conclusion (ii) is an immediate consequence of (i); see [10, theorem 5.1, p. 191]. For an (unreflected) O–U process, conclusion (i) is covered by standard SDE Novikov type conditions. For the reflected O–U process considered here, the key is to prove that a local martingale is a martingale. To do this, we "localize" the martingale and then justify the necessary limit interchange.

If we apply Itô's formula to compute the stochastic differential of $X^{\mathrm{R}}(t)^2$, we get the identity

$$\int_0^t X^{\mathrm{R}}(s)\,\mathrm{d}B(s) = \frac{1}{2\sigma}\bigl(X^{\mathrm{R}}(t)^2 - X^{\mathrm{R}}(0)^2\bigr) - \int_0^t \biggl[\frac{\alpha}{\gamma}X^{\mathrm{R}}(s) + \frac{\sigma}{2}\biggr]\mathrm{d}s.$$

If we substitute this identity into the exponent of $M(t,\gamma)$, we obtain

$$M(t,\gamma) = \exp\biggl(\frac{\gamma}{2\sigma^2}\bigl(X^{\mathrm{R}}(0)^2 - X^{\mathrm{R}}(t)^2\bigr) + \frac{\gamma\alpha}{\sigma^2}\int_0^t X^{\mathrm{R}}(s)\,\mathrm{d}s + \frac{\gamma t}{2} - \frac{\gamma^2}{2\sigma^2}\int_0^t X^{\mathrm{R}}(s)^2\,\mathrm{d}s\biggr) \tag{3.1}$$

Put $T_n = \inf\{t \geqslant 0\colon X^{\mathrm{R}}(t) \geqslant n\}$. Since $|X^{\mathrm{R}}(t \wedge T_n)| \leqslant n$ for all $t \geqslant 0$, it is easily seen that $M(t \wedge T_n, \gamma)$ is integrable. A straightforward calculation then establishes that $(M(t \wedge T_n, \gamma)\colon t \geqslant 0)$ is a martingale adapted to the filtration of $B$. Consequently,

$$E\big(M\big((t+s) \wedge T_n, \gamma\big) \mid B(u)\colon 0 \leqslant u \leqslant t\big) = M(t \wedge T_n, \gamma) \tag{3.2}$$

for $t, s \geqslant 0$. We can send $n \to \infty$ in (3.2) to conclude that $(M(t, \gamma)\colon t > 0)$ is a martingale, provided that we verify that the limit in $n$ can be interchanged with the (conditional) expectation.

Note that for $\gamma \geqslant 0$,

$$M\big((t+s) \wedge T_n, \gamma\big) \leqslant \exp\left(\frac{\gamma}{2\sigma^2} x + \frac{\gamma(t+s)}{2}\right) \exp\left(\frac{(s+t)\gamma\alpha}{\sigma^2} \max_{0 \leqslant u \leqslant t+s} X^{\mathrm{R}}(u)\right).$$

But

$$\begin{aligned} X^{\mathrm{R}}(u) &= x + \alpha u + \sigma B(u) + \sup_{0 \leqslant r \leqslant u} \big[x_0 + \alpha r + \sigma B(r)\big]^- \\ &\leqslant 2 \sup_{0 \leqslant r \leqslant u} \big|x + \alpha r + \sigma B(r)\big| \\ &\leqslant 2x + 2|\alpha| u + 2\sigma \sup_{0 \leqslant r \leqslant u} \big|B(r)\big| \end{aligned}$$

so

$$\begin{aligned} M\big((t+s) \wedge T_n, \gamma\big) \leqslant {} & \exp\left(x\left(\frac{\gamma}{2\sigma^2} + 2(s+t)\frac{\gamma|\alpha|}{\sigma^2}\right) + \frac{\gamma(t+s)}{2} + 2(s+t)^2 \frac{\alpha^2\gamma}{\sigma^2}\right) \\ & \times \exp\left(2(s+t)\gamma \frac{|\alpha|}{\sigma} \max_{0 \leqslant r \leqslant t+s} \big|B(r)\big|\right). \end{aligned}$$

However, $E \exp(\theta \max_{0 \leqslant r \leqslant t+s} |B(r)|) < \infty$ for all $\theta \in \Re$, proving that $M((t + s) \wedge T_n, \gamma)$ is uniformly (in $n$) dominated by an integrable random variable. The Dominated Convergence Theorem for conditional expectations then establishes that the limit/expectation interchange is valid. □

Suppose that $X_\gamma$ is a reflected O–U process having drift $\alpha - \gamma x$ and infinitesimal variance $\sigma^2$, starting from $x$. According to theorem 1, part (ii),

$$Ef\big(X_\gamma(t)\big) = E_P f\big(X^{\mathrm{R}}(t)\big) M(t, \gamma),$$

provided $f$ is non-negative, where $E_P(\cdot)$ is the expectation operator associated with the probability $P$ of theorem 1. If $f$ grows at most exponentially fast, it is easy to prove that the $k$th derivative can be interchanged with $E_P(\cdot)$, yielding

$$\frac{\mathrm{d}^k}{\mathrm{d}\gamma^k} E_P f\big(X^{\mathrm{R}}(t)\big) M(t, \gamma)\big|_{\gamma=\gamma_0} = E_P f\big(x^{\mathrm{R}}(t)\big) \frac{\mathrm{d}^k}{\mathrm{d}\gamma^k} M(t, \gamma)\Big|_{\gamma=\gamma_0}$$

for $\gamma_0 \geqslant 0$, where the $k$th derivative at $\gamma_0 = 0$ is interpreted as a right-handed derivative. It follows that

$$Ef\big(X_\gamma(t)\big) = Ef\big(X^{\mathrm{R}}(t)\big) + \gamma E_P f\big(X^{\mathrm{R}}(t)\big)M^{(1)}(t,0) + \mathrm{O}\big(\gamma^2\big)$$

as $\gamma \downarrow 0$. Consequently, for $f$ having growth at most exponential,

$$\begin{aligned} Ef\big(X_\gamma(t)\big) = {} & E_P f\big(X^{\mathrm{R}}(t)\big) + \frac{\gamma}{2\sigma^2} E_P f\big(X^{\mathrm{R}}(t)\big)\big(X^{\mathrm{R}}(0)^2 - X^{\mathrm{R}}(t)^2\big) \\ & + \frac{\gamma\alpha}{\sigma^2}\int_0^t E_P f\big(X^{\mathrm{R}}(t)\big)\bigg(X^{\mathrm{R}}(s) + \frac{\sigma^2}{2\alpha}\bigg)\,\mathrm{d}s + \mathrm{O}\big(\gamma^2\big) \end{aligned} \tag{3.3}$$

Observe that if $\alpha < 0$ so that $X^{\mathrm{R}}$ has a steady-state, $-\sigma^2/(2\alpha)$ is the steady-state mean of $X^{\mathrm{R}}$. Thus, the integral on the right-hand side of (3.3) converges, as $t \to \infty$, to

$$\frac{\gamma\alpha}{\sigma^2}\int_0^\infty \mathrm{cov}\big(f\big(X^{\mathrm{R}}(0)\big), X^{\mathrm{R}}(s)\big)\,\mathrm{d}s,$$

where $\mathrm{cov}(\cdot)$ is the covariance operator associated with a stationary version of $X^{\mathrm{R}}$. (We use here the fact that $X^{\mathrm{R}}$, in stationarity, is a reversible process.) Consequently, the approximation on the right-hand side of (3.3) is bounded in $t$ when $X^{\mathrm{R}}$ has a steady-state. This is reassuring, given that the left-hand side converges as $t \to \infty$.

When $f(x) = x^k$, (3.3) yields

$$\begin{aligned} EX_\gamma(t)^k = {} & E_P X^{\mathrm{R}}(t)^k + \frac{\gamma}{2\sigma^2} E_P X^{\mathrm{R}}(t)^k\big(X^{\mathrm{R}}(0)^2 - X^{\mathrm{R}}(t)^2\big) \\ & + \frac{\gamma\alpha}{\sigma^2}\int_0^t E_P X^{\mathrm{R}}(t)^k\bigg(X^{\mathrm{R}}(s) + \frac{\sigma^2}{2\alpha}\bigg)\,\mathrm{d}s + \mathrm{O}\big(\gamma^2\big). \end{aligned} \tag{3.4}$$

Now that we have a perturbation expansion for $Ef(X_\gamma(t))$, we can use this to formally derive a perturbation expansion for the transition density of $X_\gamma$. Let $p_\gamma(t,x,y)$ be the transition density for $X_\gamma$ and let $p(t,x,y)$ be the transition density for the RBM $X^{\mathrm{R}}$. Assume $p_\gamma(t,x,y)$ is differentiable in $\gamma$ and suppose

$$p_\gamma(t,x,y) = p(t,x,y) + \gamma p'(t,x,y) + \mathrm{o}(\gamma). \tag{3.5}$$

Then, by theorem 1, part (ii), we expect

$$\begin{aligned} & \int_0^\infty f(y)p'(t,x,y)\,\mathrm{d}y \\ & = E_x f\big(X^{\mathrm{R}}(t)\big)\frac{\mathrm{d}}{\mathrm{d}\gamma}M(t,\gamma)\Big|_{\gamma=0} \\ & = E_x f\big(X^{\mathrm{R}}(t)\big)\bigg(\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}X^{\mathrm{R}}(t)^2 + \frac{1}{2}t + \frac{\alpha}{\sigma^2}\int_0^t X^{\mathrm{R}}(s)\,\mathrm{d}s\bigg) \\ & = \int_0^\infty f(y)\bigg[\bigg(\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}y^2 + \frac{1}{2}t\bigg)p(t,x,y) \\ & \quad + \frac{\alpha}{\sigma^2}\int_0^t\int_0^\infty zp(s,x,z)p(t-s,z,y)\,\mathrm{d}z\,\mathrm{d}s\bigg]\,\mathrm{d}y \end{aligned}$$

which suggests that

$$p'(t,x,y) = \left(\frac{1}{2\sigma^2}x^2 - \frac{1}{2\sigma^2}y^2 + \frac{1}{2}t\right)p(t,x,y) + \frac{\alpha}{\sigma^2}\int_0^t\int_0^\infty zp(s,x,z)p(t-s,z,y)\,\mathrm{d}z\,\mathrm{d}s.$$

Substituting the above into (3.5) we find

$$p_\gamma(t,x,y) = \left(1 + \frac{\gamma}{2\sigma^2}x^2 - \frac{\gamma}{2\sigma^2}y^2 + \frac{\gamma}{2}t\right)p(t,x,y) + \frac{\alpha\gamma}{\sigma^2}\int_0^t\int_0^\infty zp(s,x,z)p(t-s,z,y)\,\mathrm{d}z\,\mathrm{d}s + \mathrm{o}(\gamma).$$

This provides an approximation to the transition density of $X_\gamma$ (valid for small $\gamma$), in terms of the known transition density for RBM. As indicated earlier, the transition density for a RBM having drift $\alpha$ and infinitesimal variance $\sigma^2$ can be found in [9].

## 4. Level crossing times

Let $T(b) = \inf\{t \geqslant 0\colon X(t) \geqslant b\}$ be the first time the process $X$ crosses the level $b$. Computing the exact distribution of $T(b)$ appears difficult. However, for large $b$, our next result offers an approximation. As earlier, for $x \geqslant 0$, let $P_x(\cdot) \triangleq P(\cdot \mid X(0) = x)$ and $E_x(\cdot) \triangleq E(\cdot \mid X(0) = x)$.

**Theorem 2.** For $b > x$,

$$E_xT(b) = \sqrt{\frac{4\pi}{\gamma\sigma^2}}\int_x^b \exp\left(\frac{(\gamma v-\alpha)^2}{\gamma\sigma^2}\right)P\left(0 \leqslant N\left(\frac{\alpha}{\gamma}, \frac{\sigma^2}{2\gamma}\right) \leqslant v\right)\mathrm{d}v. \tag{4.1}$$

Furthermore,

$$\frac{T(b)}{E_xT(b)} \Rightarrow \exp(1)$$

as $b \to \infty$.

*Proof.* Fix $b > 0$ and let $u(x)$ be defined via the right-hand side of (4.1). Then, $u$ satisfies the differential equation

$$(Au)(x) = -1, \tag{4.2}$$

subject to $u(b) = 0$ and $u'(0) = 0$, where $A$ is the second-order differential operator

$$A = (\alpha - \gamma x)\frac{\mathrm{d}}{\mathrm{d}x} + \frac{\sigma^2}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}.$$

Let $\tilde{u}(\cdot)$ be a twice continuously differentiable function with compact support that agrees with $u$ on $[0, b]$. Clearly, $\tilde{u}$ and its derivatives are bounded. Applying Itô's formula, we find that

$$\begin{aligned} \mathrm{d}\tilde{u}\big(X(t)\big) &= (A\tilde{u})\big(X(t)\big) + \sigma\tilde{u}'\big(X(t)\big)\,\mathrm{d}B(t) + \tilde{u}'\big(X(t)\big)\,\mathrm{d}L(t) \\ &= (A\tilde{u})\big(X(t)\big) + \sigma\tilde{u}'\big(X(t)\big)\,\mathrm{d}B(t), \end{aligned}$$

because $L$ increases only when $X$ is at the origin and $\tilde{u}'(0) = 0$. The boundedness of $\tilde{u}'$ implies that

$$\tilde{u}\big(X(t)\big) - \int_0^t (A\tilde{u})\big(X(s)\big)\,\mathrm{d}s$$

is a martingale adapted to $B$. The Optional Sampling Theorem yields

$$E_x\tilde{u}\big(X\big(T(b) \wedge t\big)\big) - \tilde{u}(x) = E_x \int_0^{T(b)\wedge t} (A\tilde{u})\big(X(s)\big)\,\mathrm{d}s$$

for $t \geqslant 0$. But $\tilde{u}(X(t)) = u(X(t))$ for $t \leqslant T(b)$. Since $Au = -1$ on $[0, b]$, we conclude that

$$u(x) - E_x u\big(X\big(T(b) \wedge t\big)\big) = E_x T(b) \wedge t. \tag{4.3}$$

Sending $t \to \infty$, the Monotone Convergence Theorem and the non-negativity of $u$ permit us to conclude that

$$u(x) \geqslant E_x T(b).$$

Consequently, $T(b) < \infty$, $P_x$ a.s. Returning to (4.3) and applying the Bounded Convergence Theorem yields the identity

$$u(x) - E_x u\big(X\big(T(b)\big)\big) = E_x T(b).$$

But $u(X(T(b))) = u(b) = 0$, establishing the desired identity for $E_x T(b)$.

Theorem 4 establishes that $X$ visits the origin infinitely often, so that $X$ is a regenerative process. Applying the theorem in [12, p. 867] then proves the second assertion of the theorem. □

**Corollary 3.**

$$b \exp\left(-\frac{(\gamma b - \alpha)^2}{\gamma\sigma^2}\right) T(b) \Rightarrow c\ \exp(1)$$

as $b \to \infty$, where $c = \sqrt{\pi\sigma^2/\gamma^3}\,P(N(\alpha/\gamma, \sigma^2/2\gamma) \geqslant 0)$.

*Proof.* We need only establish that

$$E_x T(b) \sim \frac{c\ \exp((\gamma b - \alpha)^2/\gamma\sigma^2)}{b}$$

as $b \to \infty$. This can be proved easily via 1'Hopital's rule. □

According to corollary 3, the time to hit level $b$ for a reflected O–U process grows exponentially in the square of $b$. Continuing our comparison of the reflected O–U process $X$ with the RBM $X^{\mathrm{R}}$, we state the counterpart to theorem 2 for $X^{\mathrm{R}}$. As one would expect, an RBM process crosses level $b$ in a much shorter time span, namely, in a time which grows exponentially in $b$.

**Proposition 3.** For $b > 0$, let $T^{\mathrm{R}}(b) = \inf\{t \geqslant 0\colon\ X^{\mathrm{R}}(t) \geqslant b\}$, where $X^{\mathrm{R}}$ is RBM with drift $-\mu$ and variance $\sigma^2$. Then, for $b > x$,

$$E_x T^{\mathrm{R}}(b) = \frac{\sigma^2}{2\mu^2}\left(\exp\left(\frac{2\mu b}{\sigma^2}\right) - \exp\left(\frac{2\mu x}{\sigma^2}\right)\right) - \frac{b-x}{\mu}$$

Furthermore,

$$\frac{T^{\mathrm{R}}(b)}{E_x T^{\mathrm{R}}(b)} \Rightarrow \exp(1)$$

as $b \to \infty$.

The proof of this proposition follows the arguments in theorem 2 with $-\mu(\mathrm{d}/\mathrm{d}x) + (\sigma^2/2)(\mathrm{d}^2/\mathrm{d}x^2)$ replacing $A$, and is therefore omitted. A related calculation can be found in [8].

## 5. Behavior of the maximum

Let $X^*(t) = \max\{X(s)\colon\ 0 \leqslant s \leqslant t\}$ be the maximum level the reflected O–U process $X$ attains in the time span $[0, t]$. To analyze $X^*(t)$, we exploit the regenerative structure of $X$. Specifically, we exploit the regenerative cycles formed by consecutive visits to the origin that are "interlaced" with visits to level 1. Thus, a typical cycle involves a "first passage" from 0 to 1, followed by another "first passage" from 1 to 0. The asymptotic behavior of $X^*(t)$ for large $t$ will depend on $E\tau$, where $\tau$ is the duration of a typical cycle. Note that $E\tau = E_0T(1) + E_1T(0)$. We computed $E_0T(1)$ in theorem 2; we now compute $E_1T(0)$.

**Proposition 4.** For $0 \leqslant b < x$,

$$E_x T(b) = \sqrt{\frac{4\pi}{\gamma\sigma^2}} \int_b^x \exp\left(\frac{(\gamma v - \alpha)^2}{\gamma\sigma^2}\right) g(v)\,\mathrm{d}v$$

where

$$g(v) = P(0 \leqslant N \leqslant v)\left(\frac{P(0 \leqslant N \leqslant \infty)}{P(0 \leqslant N \leqslant v)} - 1\right).$$

Furthermore,

$$E\tau = \sqrt{\frac{4\pi}{\gamma\sigma^2}} P\left(N\left(\frac{\alpha}{\gamma}, \frac{\sigma^2}{2\gamma}\right) \geqslant 0\right) \int_0^1 \exp\left(\frac{(\gamma v - \alpha)^2}{\gamma\sigma^2}\right) \mathrm{d}v.$$

The key to the proof of this proposition is to calculate $E_x[T(b, z)]$, where $T(b, z) = \inf\{t\colon\ X(t) \leqslant b \text{ or } X(t) \geqslant z\}$ is the first time the process $X$ exits the interval $(b, z)$. To compute $u(x) \triangleq E_x[T(b, z)]$, one must solve the differential equation given in (4.2), subject to $u(b) = u(z) = 0$. To find $E_x T(b)$ when $b < x$, one need then only take the limit of $E_x[T(b, z)]$ as $z$ approaches $\infty$. Since many of the details of this argument are similar to those found in theorem 2, we omit the proof.

Let $\Gamma(n)$ be the time at which the $n$th regenerative cycle is completed, and let $N(t)$ be the number of cycles completed in $[0, t]$. The key to analyzing $X^*(t)$ is to observe that

$$X^*(t) \overset{D}{\approx} X^*\bigl(\Gamma\bigl(N(t)\bigr)\bigr) = \max_{1 \leqslant j \leqslant N(t)} X_j^*,$$

where $X_j^* = \max\{X(t)\colon\ \Gamma(j - 1) \leqslant t < \Gamma(j)\}$ is the maximum of $X$ over the $j$th cycle. But the regenerative structure implies that $M_n^* = \max\{X_j^*\colon\ 1 \leqslant j \leqslant n\}$ is the maximum of $n$ iid random variables, so that the classical extreme value theory for iid random variables may be invoked. This then yields a limit law for $X^*(t)$ as $t \to \infty$; the machinery for making this program rigorous can be found in [3].

Classical extreme value theory makes clear that the behavior of $M_n^*$ will be governed by the tail behavior of the $X_j^*$'s. Note that $P(X_j^* > b) = P_1(T(b) < T(0))$.

**Proposition 5.** For $0 < x < b$,

$$P_x\bigl(T(b) < T(0)\bigr) = \frac{\int_0^x \exp(2((\gamma/2)v^2 - \alpha v)/\sigma^2)\,\mathrm{d}v}{\int_0^b \exp(2((\gamma/2)v^2 - \alpha v)/\sigma^2)\,\mathrm{d}v}. \tag{5.3}$$

*Proof.* Fix $b > 0$ and let $u(x)$ be given by the right-hand side of (5.3). Note that $u$ satisfies $(Au)(x) = 0$ for $0 < x < b$, subject to $u(0) = 0$ and $u(b) = 1$. Noting that $X$ spends no time at the origin up to $T(b) \wedge T(0)$ (so that $L(T(b) \wedge T(0)) = 0$), Itô's formula can be applied as in the proof of theorem 2 to obtain the desired conclusion. □

We are now ready to state the main result of this section, namely a limit theorem for $X^*(t)$.

**Theorem 4.**

$$\sqrt{\log t}\left(X^*(t) - \frac{\alpha}{\gamma} - \sqrt{\left(\frac{\sigma^2}{\gamma}\right)\log t + \left(\frac{\sigma^2}{2\gamma}\right)\log\log t}\right) \Rightarrow \sqrt{\frac{\sigma^2}{4\gamma}}\Lambda + d$$

as $t \to \infty$, where $\Lambda$ is a Gumbel r.v. having distribution $P(\Lambda \leqslant x) = \exp(-\mathrm{e}^{-x})$ and

$$d = \frac{1}{2}\frac{\gamma^{3/2}}{\sigma^3} - \frac{\alpha\gamma^{1/2}}{\sigma^3} + \frac{1}{2}\sqrt{\frac{\gamma}{\sigma^2}}\log\left(\frac{\sigma^3}{2\gamma^{3/2}}\frac{1}{E\tau\int_0^1 \exp((\gamma/\sigma^2)(v-\alpha/\gamma)^2)\,\mathrm{d}v}\right)$$

*Proof.* By lemma 1.1 of Asmussen [3] (which follows the formulation given in [16]), the proof is complete once one shows that

$$\sqrt{\log t}\left(M^*_{\lfloor t/E\tau\rfloor} - \frac{\alpha}{\gamma} - \sqrt{\left(\frac{\sigma^2}{\gamma}\right)\log t + \left(\frac{\sigma^2}{2\gamma}\right)\log\log t}\right) \Rightarrow \sqrt{\frac{\sigma^2}{4\gamma}}\Lambda + d \qquad (5.4)$$

as $t \to \infty$. Set $a(t) = (\log t)^{-1/2}$ and $b(t) = \alpha + ((\sigma^2/\gamma)\log t + (\sigma^2/2\gamma)\log\log t)^{1/2}$. The weak convergence relation (5.4) is equivalent to

$$\frac{t}{E\tau}\log\bigl(1 - P\bigl(X_1^* > xa(t) + b(t)\bigr)\bigr) \to -\exp\left(-\sqrt{\frac{4\gamma}{\sigma^2}}(x-d)\right) \qquad (5.5)$$

as $t \to \infty$. For this choice of $a(t)$ and $b(t)$, $P(X_1^* > xa(t)+b(t)) = \mathrm{O}(1/t)$ as $t \to \infty$, so the left-handed side of (5.5) is $-tP(X_1^* > xa(t)+b(t))/E\tau + \mathrm{o}(1)$ as $t \to \infty$. Given the explicit formula (5.3) for $P(X_1^* > b)$, l'Hôpital's rule easily shows that

$$P(X_1^* > b) \sim b\tilde{d}\exp\left(-\frac{\gamma}{\sigma^2}(b-\alpha)^2\right) \qquad (5.6)$$

as $b \to \infty$, where

$$\tilde{d} = \frac{(\sigma^2/2\gamma)\exp(\gamma/\sigma^2 - 2\alpha/\sigma^2)}{\int_0^1 \exp((\gamma/\sigma^2)(v-\alpha/\gamma)^2)\,\mathrm{d}v}.$$

In view of (5.5), the proof of the theorem is therefore complete once one shows that

$$\log\left(\frac{t}{E\tau}\right) + \log\bigl(xa(t)+b(t)\bigr) + \log\tilde{d} - \frac{\gamma}{\sigma^2}\bigl(xa(t)-b(t)-\alpha\bigr)^2 \to -\sqrt{\frac{4\gamma}{\sigma^2}}(x-d) \quad (5.7)$$

as $t \to \infty$. However, proving (5.7) is reasonably straightforward, and so we omit the details, thereby completing the proof. □

**Corollary 5.**

$$\frac{X^*(t)}{\sqrt{(\sigma^2/\gamma)\log t}} \Rightarrow 1$$

as $t \to \infty$.

*Remark 1.* The behavior of the maximum process looks roughly like that associated with Gaussian extrema. The reason for this is that the tail behavior described by (5.6) is roughly that associated with a Gaussian distribution. (It would be precisely that of a Gaussian distribution if the pre-multiplier $b$ in (5.6) were instead $b^{-1}$.)

As in the previous section, it is of interest to compare the behavior of $X^*(t)$ with the maximum process associated with RBM. Let $X^*_{\mathrm{R}}(t) = \max\{X^{\mathrm{R}}(s)\colon 0 \leqslant s \leqslant t\}$. Berger and Whitt [4] prove the following result.

**Proposition 6.** If $X^{\mathrm{R}}$ is RBM with drift $\alpha$ ($\alpha < 0$) and variance $\sigma^2$, then

$$2\frac{|\alpha|}{\sigma^2}X^*_{\mathrm{R}}(t) - \log 2\left(\frac{\alpha}{\sigma}\right)^2 t \Rightarrow \Lambda.$$

Contrasting proposition 6 to corollary 5, we see that $X^*$ grows approximately at rate $\sqrt{\log t}$, whereas $X^*_{\mathrm{R}}$ grows more like $\log t$. Thus, we see the enormous impact the linear term in the drift of the reflected O–U has on reducing the magnitude of the extreme fluctuations associated with the number-in-system process. (Note these growth rates are consistent with our results on level crossing times in section 4. Recall that there we saw the expected time for a reflected O–U to cross a large level $b$ was approximately $\mathrm{e}^{b^2}$, whereas this expected time was only $\mathrm{e}^b$ for the corresponding RBM.)

**Acknowledgements**

We thank Tom Kurtz, Ward Whitt, the two anonymous referees, and the associate editor for helpful comments relating to the issues addressed in this paper.

**References**

[1] J. Abate and W. Whitt, Transient behavior of regulated Brownian motion, I: Starting at the origin, Adv. in Appl. Probab. 19 (1987) 560–598.
[2] J. Abate and W. Whitt, Transient behavior of regulated Brownian motion, II: Non-zero initial conditions, Adv. in Appl. Probab. 19 (1987) 599–631.
[3] S. Asmussen, Extreme value theory for queues via cycle maxima, Extremes 1(2) (1998) 137–168.
[4] A. Berger and W. Whitt, Maximum values in queueing processes, Probab. Engrg. Inform. Sci. 9 (1995) 375–409.
[5] A. Borovkov, *Asymptotic Methods in Queueing Theory* (Wiley, New York, 1984).
[6] S. Browne and W. Whitt, Piecewise-linear diffusion processes, in: *Advances in Queueing: Theory, Methods, and Open Problems*, ed. J. Dshalalow (CRC Press, Boca Raton, FL, 1995) pp. 463–480.
[7] E. Coffman, A. Puhalskii, M. Reiman and P. Wright, Processor-shared buffers with reneging, Performance Evaluation 19 (1994) 25–46.
[8] P.W. Glynn and D. Iglehart, Trading securities using trailing stops, Managm. Sci. 41 (1995) 1096–1106.
[9] J.M. Harrison, *Brownian Motion and Stochastic Flow Systems* (Wiley, New York, 1985).
[10] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus* (Springer, New York, 1991).
[11] S. Karlin and H. Taylor, *A Second Course in Stochastic Processes* (Academic Press, New York, 1976).
[12] J. Keilson, A limit theorem for passage times in ergodic regenerative processes, Ann. Math. Statist. 37 (1966) 866–870.
[13] T. Lindvall, On coupling of diffusion processes, J. Appl. Probab. 20 (1983) 82–93.
[14] T. Lindvall, Lectures on the coupling method, in: *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics* (Wiley, New York, 1992).

[15] P. Lions and A. Sznitman, Stochastic differential equations with reflecting boundary conditions, Comm. Pure Appl. Math. 37 (1984) 511–537.

[16] H. Rootzen, Maxima and exceedances of stationary Markov chains, Adv. in Appl. Probab. 20 (1989) 371–390.

[17] S. Ross, *Stochastic Processes* (Wiley, New York, 1996).

[18] R. Serfozo, Extreme values of birth and death processes and queues, Stochastic Process. Appl. 27 (1988) 291–306.

[19] R. Srikant and W. Whitt, Simulation run lengths to estimate blocking probabilities, ACM Trans. Modeling Comput. Simulation 6 (1996) 7–52.

[20] A. Ward and P. Glynn, A diffusion approximation for a GI/G/1 queue with reneging, Working paper (2002).

[21] A. Ward and P. Glynn, A diffusion approximation for a Markovian queue with reneging, under review (2002).