ON GUMBEL LIMIT FOR THE LENGTH OF REACTIVE PATHS

YURI BAKHTIN

Abstract. We give a new proof of the vanishing noise limit theorem for exit times of 1-dimensional diffusions conditioned on exiting through a point separated from the starting point by a potential wall. We also prove a scaling limit for exit location in a model 2-dimensional situation.

1. Introduction

A random variable $Z$ has standard Gumbel distribution if

$P\{Z \leq x\} = \Lambda(x) = e^{-e^{-x}}, \ x \in \mathbb{R}.$

The Gumbel distribution is mostly well-known as a limit law in the extreme value theory, see, e.g., [dHF06, Chapter 1].

Surprisingly, it also appears as a limiting distribution (in the limit of vanishing noise) for normalized exit times of diffusions conditioned on unlikely exit locations. The first result of this kind was obtained in [Day90], where the phenomenon of cycling for exits through a characteristic repelling boundary was discovered and explained via the asymptotics of exit times, see [BG13] for most recent progress on cycling and a bibliography.

In a recent paper [CGLM13] (see also references therein for related results) a similar result for conditional exit times was obtained for the case where the prescribed exit location for the diffusion is separated from the starting point by a potential wall.

Namely, consider a strong solution $X_\varepsilon$ of the following SDE driven by additive white noise of small intensity $\varepsilon > 0$ and a smooth vector field $b$ on an interval $[q_-, q_+]$ containing $0$:

\begin{align*}
(1) \quad dX_\varepsilon(t) &= b(X_\varepsilon(t)) dt + \varepsilon dW(t), \\
(2) \quad X_\varepsilon(0) &= x_0,
\end{align*}

Here $x_0 \in (q_-, 0)$, the drift $b$ satisfies $b(0) = 0$, $b'(0) > 0$, $b(x) < 0$ for $x \in (q_-, 0)$, and $b(x) > 0$ for $x \in (0, q_+]$, and $W$ is a standard Wiener process on a complete probability space $(\Omega, \mathcal{F}, P)$.

We are interested in the conditional distribution of the first exit time

$\tau_\varepsilon = \inf\{t \geq 0 : X_\varepsilon(t) \in \{q_-, q_+\}\},$

The author was partially supported by the NSF CAREER Award DMS-0742424.
conditioned on the event $C_\varepsilon(q_-, x_0, q_+) = \{X_\varepsilon(\tau_\varepsilon) = q_+\}$. This is an extremely improbable event for small $\varepsilon$ since for $C_\varepsilon(q_-, x_0, q_+)$ to happen, the process $X_\varepsilon$ has to travel against the drift. The resulting diffusion trajectories are often called reactive paths.

**Theorem 1 (CGLM13).** If, in addition to the conditions given above, $q_+ > 0$, then there are constants $c_1, c_2, c_3$ such that as $\varepsilon \to 0,$

$$\text{Law} \left[ \tau_\varepsilon - c_1 \ln \frac{1}{\varepsilon} \bigg| C_\varepsilon(q_-, x_0, q_+) \right] \Rightarrow \text{Law} \left[ c_2 Z + c_3 \right],$$

where $Z$ is a standard Gumbel random variable and “$\Rightarrow$” denotes weak convergence of probability measures.

The core of the argument in CGLM13 was based on asymptotic analysis of solutions of second-order differential equations describing the moment generating functions of exit times. This is a powerful general method, and in CGLM13 it was also used to study other interesting settings. Unfortunately, it does not seem to answer why it is natural for the max-stable Gumbel distribution to appear in this context. The path-based arguments for slightly different situations in Day90 and Day95 are more elucidating, and it would be natural to supplement the argument of CGLM13 by a more probabilistic one.

In Bak we observed that the connection between the extreme value theory and exit times is provided by the theory of residual life times, see BdH74. However, to make our point in that paper we chose to work with an initial condition $x_0$ depending on $\varepsilon$, namely, we set $x_0 = -\varepsilon a$ and considered two consecutive limit transitions instead of one: first $\varepsilon \to 0$, and then $a \to +\infty$.

One of the goals of the present paper is to show that purely probabilistic reasoning based on basic properties of the Wiener process is, in fact, not much harder for the original problem with fixed $x_0 \in (q_-, 0)$ and only one limiting transition $\varepsilon \to 0$.

As in Bak, we make a simplifying assumption $b(x) = \lambda x$ for a constant $\lambda > 0$ (in general, one has to approximate the nonlinear case with a linearization). We will prove the following result with purely probabilistic tools:

**Theorem 2 (CGLM13).** Let $Z$ denote a standard Gumbel random variable. If $q_+ > 0$, then

$$\text{Law} \left[ \tau_\varepsilon - \frac{2}{\lambda} \ln \frac{1}{\varepsilon} \bigg| C_\varepsilon(q_-, x_0, q_+) \right] \Rightarrow \text{Law} \left[ \frac{\ln(2\lambda q_+ | x_0|)}{\lambda} + \frac{1}{\lambda} Z \right], \quad \varepsilon \to 0.$$  

We stress that this result per se is not new. In CGLM13 it is given as Proposition 3.1 and plays the central role in the proof of Theorem 1 of which it is a specific case. However, our point is to give a proof based on properties of the Wiener process and to stress the connection to residual life times. We give our proof in Section 2. Our method is somewhat similar to the
proof of the following statement which is a version of a result from [Day90] (where, in fact, a more general nonlinear version is considered) for the case of characteristic boundary, i.e. \( q_+ = 0 \):

**Theorem 3** ([Day90]).

\[
\text{Law} \left[ \tau_c - \frac{1}{\lambda} \ln \frac{1}{\varepsilon} \left| C_{\varepsilon}(q_-, x_0, 0) \right| \right] \Rightarrow \text{Law} \left[ \frac{\ln (x_0^2 \lambda)}{2\lambda} + \frac{1}{2\lambda} Z \right], \quad \varepsilon \to 0.
\]

For completeness we give a proof of this result in Section 2 too.

We notice that Theorem 2 fits nicely with the results of [Day95, Bak08, Bak11] which for the linear 1-dimensional situation with \( q_+ > 0 \) and \( x_0 = 0 \) take the following form:

**Theorem 4** ([Day95]). Let \( q_+ > 0 \). Then

\[
\text{Law} \left[ \tau_c - \frac{1}{\lambda} \ln \frac{1}{\varepsilon} \left| C_{\varepsilon}(q_-, 0, q_+) \right| \right] \Rightarrow \text{Law} \left[ \frac{\ln q_+}{\lambda} + \frac{\ln(2\lambda)}{2\lambda} + \frac{1}{\lambda} \Theta \right],
\]

where \( \Theta = -\ln |N| \) and \( N \) is a standard Gaussian random variable.

The connection is the following: to realize the exit described in Theorem 2 the solution must first reach 0 and then reach \( q_+ \) from 0. The strong Markov property allows us to conclude that the limit in (3) is equal to the convolution of limits in (4) and (5). This can be checked independently by a straightforward computation of the Laplace transforms or characteristic functions of the right-hand sides of (3), (4), and (5), and invoking the duplication formula for the gamma function.

The second goal of this paper is to consider systems with small noise in high dimensions and use Theorem 2 to obtain a vanishing noise scaling limit for conditional distributions of exit locations. Namely, we are interested in exit from a neighborhood of a saddle point of the driving vector field under conditioning on a rare event that the solution travels through the potential wall.

In the unconditioned situation, scaling limits for exit distributions for neighborhoods of saddle points have been obtained in [Bak11], [AMB11a], and [AMB11b]. These results are crucial for in the theory of heteroclinic networks under small noise in [Bak11] (see also [Bak10] for a less technical exposition).

The only known to us results on scaling limits for exit distributions under conditioning on unlikely exit locations in multiple dimensions are those of [BS] where a case with no invariant sets (such as critical points) in the domain under several additional technical conditions is considered. The analysis in [BS] is based on using Doob’s \( h \)-transform to reduce the problem to the Levinson case studied in [AMB11b]. However, the details are involved and require obtaining a new gradient estimate for solutions of nonlinear elliptic equation.
Here we study only a model 2-dimensional situation with a linear saddle and additive isotropic noise that can be addressed using the asymptotics provided by Theorem 2, although we conjecture that the behavior is similar in the general nonlinear case of tunneling through a potential wall near a saddle point.

Let us consider the following stochastic dynamics in two dimensions:

\[ dX^1_\varepsilon(t) = \lambda X^1_\varepsilon(t) dt + \varepsilon dW^1(t), \]
\[ dX^2_\varepsilon(t) = -\mu X^2_\varepsilon(t) dt + \varepsilon dW^2(t), \]

where \( \mu \) and \( \lambda \) are positive constants, and \( W^1, W^2 \) are two independent Wiener processes. The drift \( b(x_1, x_2) = (\lambda x_1, -\mu x_2) \) of this model system is linear and the origin is a saddle critical point with first axis serving as the unstable manifold and second axis as the stable one.

Let us assume that the initial conditions \((X^1_\varepsilon(0), X^2_\varepsilon(0))\) for this system satisfy the following scaling limit: \( X^1_\varepsilon(0) = x_1, \) where \( x_1 < 0 \) is a constant, and \( X^2_\varepsilon = \varepsilon^\alpha \xi_\varepsilon \) for some \( \alpha \in \mathbb{R} \) and a family of random variables \( (\xi_\varepsilon)_{\varepsilon > 0} \) independent of the driving Wiener processes \( W^1, W^2 \) and converging in distribution as \( \varepsilon \to 0 \) to a random variable \( \xi_0 \) that is not concentrated at 0.

Let us take two numbers \( q_-, q_+ \) satisfying \( q_- < x_1 < 0 < q_+ \) and define the time \( \tau_\varepsilon \) as the first time the solution exits from the strip \((q_-, q_+) \times \mathbb{R})\), i.e., \( \tau_\varepsilon = \inf\{t \geq 0 : X^1_\varepsilon(t) \in \{q_-, q_+\}\}\). We would like to study the exit distribution of the process \( X_\varepsilon = (X^1_\varepsilon, X^2_\varepsilon) \) conditioned on \( C_\varepsilon = \{X^1_\varepsilon(\tau_\varepsilon) = q_+\} \), i.e., on the rare event that the solution makes it over the potential wall to the other side of the saddle.

**Theorem 5.** Let \( V, N, \xi \) be independent random variables such that \( V \) is standard exponential, i.e., \( P\{V \geq x\} = e^{-x} \) for \( x \geq 0 \), \( N \) is standard Gaussian, and \( \xi \) is distributed as \( \xi_0 \). Let \( \beta = 1 \wedge (2\mu/\lambda + \alpha) \). Then

\[
\text{Law} \left[ \frac{X^2_\varepsilon(\tau_\varepsilon)}{\varepsilon^\beta} \middle| C_\varepsilon \right] \Rightarrow \text{Law} \left[ 1 - 2^{\mu/\lambda} + \alpha \frac{V^{\mu/\lambda} \xi}{(2\lambda q_+ |x_1|)^{\mu/\lambda}} + 1 \beta = 1 \frac{N}{\sqrt{2\mu}} \right], \quad \varepsilon \to 0.
\]

In particular, we observe effects similar to those described in [Bak11] and [Bak10]. The character of the limiting exit distribution under rescaling depends not only on the scaling exponent \( \alpha \) of the distribution of the initial condition, but also on the ratio of contraction and expansion rates.

In particular, if contraction is strong enough \( (2\mu/\lambda + \alpha > 1) \), then at the exit the system asymptotically forgets the initial condition, and the exit distribution is asymptotically centered Gaussian and scales as \( \varepsilon \), the magnitude of the noise.

However, if the contraction is not strong enough \( (2\mu/\lambda + \alpha < 1) \), then the Gaussian term disappears and the dependence of the limiting exit distribution on the initial distribution is nontrivial. In particular, this can lead to the following memory effect: if the distribution of \( \xi \) is supported by the positive (or negative) semi-axis, then the limiting exit distribution will also retain this strong asymmetry.
There is also an intermediate case \((2\mu/\lambda + \alpha = 1)\) where both Gaussian and non-Gaussian terms contribute to the limiting distribution.

We give the proof of Theorem 5 in Section 3. It is clear from the proof that it is also easy to obtain a generalization of Theorem 5 for higher dimensions.

### 2. Proof of Theorems 2 and 3

#### 2.1. Asymptotic equivalence of families of events.

More than once in this section we will need to replace conditioning on one event by conditioning on another event. To that end it is convenient to introduce a notion that will allow to control the change in the conditional probability.

We will say that families of events \((A_\varepsilon)_{\varepsilon>0}\) and \((B_\varepsilon)_{\varepsilon>0}\) are asymptotically equivalent as \(\varepsilon \to 0\) if

\[
\operatorname{Law}\left[ Y_\varepsilon \mid A_\varepsilon \right] \Rightarrow \mu \text{ as } \varepsilon \to 0
\]

for a measure \(\mu\).

If \((B_\varepsilon)_{\varepsilon>0}\) is a family of events asymptotically equivalent to \((A_\varepsilon)_{\varepsilon>0}\), then

\[
\operatorname{Law}\left[ Y_\varepsilon \mid A_\varepsilon \right] \Rightarrow \mu \text{ as } \varepsilon \to 0.
\]

The roles of \(A_\varepsilon\) and \(B_\varepsilon\) are not symmetric in \((6)\). However, if \((6)\) holds, then

\[
P(B_\varepsilon) = P(A_\varepsilon)(1 + o(1)) \text{ as } \varepsilon \to 0,
\]

and we see that \(P(B_\varepsilon) > 0\) for sufficiently small \(\varepsilon\) and

\[
\lim_{\varepsilon \to 0} \frac{P(A_\varepsilon \triangle B_\varepsilon)}{P(A_\varepsilon)} = 0.
\]

Similarly, \((7)\) is implied by \((6)\), so the roles of \(A_\varepsilon\) and \(B_\varepsilon\) are, in fact, interchangeable.

\[\text{Lemma 1.} \quad \text{Let } (Y_\varepsilon)_{\varepsilon>0} \text{ be a family of random variables and assume that for a family of events } (A_\varepsilon)_{\varepsilon>0}, \ \operatorname{Law}[Y_\varepsilon | A_\varepsilon] \Rightarrow \mu \text{ as } \varepsilon \to 0 \text{ for a measure } \mu. \text{ If } (B_\varepsilon)_{\varepsilon>0} \text{ is a family of events asymptotically equivalent to } (A_\varepsilon)_{\varepsilon>0}, \ \text{then } \operatorname{Law}[Y_\varepsilon | A_\varepsilon] \Rightarrow \mu \text{ as } \varepsilon \to 0.\]

\[\text{PROOF:} \quad \text{We are given that for any continuous bounded function } f : \mathbb{R} \to \mathbb{R},
\]

\[
\frac{\mathbb{E}[f(Y_\varepsilon)1_{A_\varepsilon}]}{P(A_\varepsilon)} \to \int_{\mathbb{R}} f(x) \mu(dx), \quad \varepsilon \to 0.
\]

Therefore, for any continuous bounded function \(f : \mathbb{R} \to \mathbb{R},\)

\[
\frac{\mathbb{E}[f(Y_\varepsilon)1_{B_\varepsilon}]}{P(B_\varepsilon)} = \frac{\mathbb{E}[f(Y_\varepsilon)1_{A_\varepsilon}] + \mathbb{E}[f(Y_\varepsilon)1_{B_\varepsilon \setminus A_\varepsilon}] - \mathbb{E}[f(Y_\varepsilon)1_{A_\varepsilon \setminus B_\varepsilon}]}{P(A_\varepsilon)(1 + o(1))} = \frac{\mathbb{E}[f(Y_\varepsilon)1_{A_\varepsilon}] + o(P(A_\varepsilon))}{P(A_\varepsilon)(1 + o(1))} \to \int_{\mathbb{R}} f(x) \mu(dx), \quad \varepsilon \to 0.
\]

and the proof is completed. \qed
2.2. The representation of solution and reflection principle for an auxiliary process. The strong solution of \((1)\)–\((2)\) is given by variation of constants:
\[
X_\varepsilon(t) = e^{\lambda t}(x_0 + \varepsilon U(t)), \quad t \geq 0,
\]
where
\[
U(t) = \int_0^t e^{-\lambda s} dW(s), \quad t \geq 0.
\]
It will be useful for us to introduce
\[
D_\varepsilon(z, r) = \{ \varepsilon U(t) = z \text{ for some } t \in [0, r) \}, \quad z \in \mathbb{R}, \quad r \in [0, \infty],
\]
(in particular, \(D_\varepsilon(|x_0|, \infty)\) is the event that \(X_\varepsilon\) ever reaches 0) and prove a representation for probabilities of these events. Let
\[
\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy, \quad x \in \mathbb{R},
\]
be the standard Gaussian distribution function.

**Lemma 2.** For any \(z \in \mathbb{R}, \ r \in [0, \infty]\)
\[
P(D_\varepsilon(z, r)) = 2 \left(1 - \Phi \left( \frac{|z| \sqrt{2\lambda}}{\varepsilon \sqrt{1 - e^{-2\lambda r}}} \right) \right).
\]
**Proof:** Notice that the process \(U\) admits a representation
\[
U(t) = B \left( \frac{1 - e^{-2\lambda t}}{2\lambda} \right), \quad t \geq 0,
\]
for a standard Wiener process \(B\). Therefore,
\[
D_\varepsilon(z, r) = \left\{ \varepsilon B(s) = z \text{ for some } s \in \left[0, \frac{1 - e^{-2\lambda r}}{2\lambda} \right) \right\}, \quad r \in [0, \infty],
\]
so the reflection principle for the Wiener process implies
\[
P(D_\varepsilon(z, r)) = P \left\{ \sup_{s \in \left[0, \frac{1 - e^{-2\lambda r}}{2\lambda} \right]} B(s) \geq \frac{|z|}{\varepsilon} \right\} = 2P \left\{ B \left( \frac{1 - e^{-2\lambda r}}{2\lambda} \right) \geq \frac{|z|}{\varepsilon} \right\},
\]
and the lemma follows. \(\square\)

2.3. **Proof of Theorem** \([3]\) This is essentially the same proof as in \([\text{Day90}]\) given here for completeness. Let us define
\[
\tau_{0, \varepsilon} = \inf \{ t \geq 0 : X_\varepsilon(t) = 0 \} = \inf \{ t \geq 0 : x_0 + \varepsilon U(t) = 0 \} \in [0, \infty].
\]
We have \(\tau_{0, \varepsilon} \geq \tau_\varepsilon\) and \(C_\varepsilon := C_\varepsilon(q_-, x_0, 0) = \{ \tau_\varepsilon = \tau_{0, \varepsilon} \}\). Let us denote
\[
D_\varepsilon = D_\varepsilon(|x_0|, \infty) = \{ \tau_{0, \varepsilon} < \infty \}.
\]
Our goal is to replace conditioning on \(C_\varepsilon\) by conditioning on \(D_\varepsilon\).

**Lemma 3.** Families \((C_\varepsilon)_{\varepsilon > 0}\) and \((D_\varepsilon)_{\varepsilon > 0}\) are asymptotically equivalent as \(\varepsilon \to 0\).
Using Lemma 2 and Gaussian tail asymptotics (11), we obtain for any $E$ and $\theta$:

$$1 - \Phi(x) \sim \frac{1}{x\sqrt{2\pi}} e^{-x^2/2}, \quad x \to \infty,$$

Now applying Lemma 2 and the standard estimate

(11)

to $P(D_\varepsilon(|q_1|, \infty))$ and $P(D_\varepsilon(|x_0|, \infty))$, we complete the proof.

We turn now to computing the conditional distribution of $\tau_{0,\varepsilon}$ given $D_\varepsilon$. Using Lemma 2 and Gaussian tail asymptotics (11), we obtain for any $r \in \mathbb{R}$:

$$P\left( \tau_{0,\varepsilon} < \frac{1}{\lambda} \ln \frac{1}{\varepsilon} + r \mid D_\varepsilon \right) = P \left( \tau_{0,\varepsilon} < \frac{1}{\lambda} \ln \frac{1}{\varepsilon} + r \mid D_\varepsilon \right)$$

$$= P \left( D_\varepsilon \left( |x_0|, \frac{1}{\lambda} \ln \frac{1}{\varepsilon} + r \right) \right)$$

$$= \frac{1 - \Phi \left( \frac{|x_0|\sqrt{2\lambda}}{\varepsilon\sqrt{1 - e^{-2\lambda\varepsilon^2}}} \right)}{1 - \Phi \left( \frac{|x_0|\sqrt{2\lambda}}{\varepsilon} \right)} \sim e^{-x_0^2\lambda e^{-2\lambda r}}, \quad \varepsilon \to 0.$$

The right-hand side is the distribution function of $\frac{\ln(x_0^2\lambda)}{2\lambda} + \frac{1}{2\lambda} Z$. Lemma 3 allows us to apply Lemma 1 to $Y_\varepsilon = \tau_{0,\varepsilon} - \frac{1}{\lambda} \ln \frac{1}{\varepsilon}$ and families $(D_\varepsilon)_{\varepsilon \geq 0}$ and $(C_\varepsilon)_{\varepsilon \geq 0}$. Since $\tau_\varepsilon = \tau_{0,\varepsilon}$ on $C_\varepsilon$, (11) follows.

2.4. Proof of Theorem 2 Let us recall that the solution $X_\varepsilon$ is given by (9). We need to study the exit time $\tau_{\varepsilon}$ conditioned on the event $C_\varepsilon(q_-, x_0, q_+) = \{X_\varepsilon(\tau_\varepsilon) = q_+\}$. From now on we use $C_\varepsilon$ as a shorthand for $C_\varepsilon(q_-, x_0, q_+)$. In this proof we will change the conditioning twice.

Let us define

$$\theta_\varepsilon = \inf \left\{ t \geq 0 : X_\varepsilon(t) = e^{\lambda t} (x_0 + \varepsilon U(t)) = q_+ \right\} \in (0, \infty],$$

and $E_\varepsilon = \{ \theta_\varepsilon < \infty \}$. Notice that $C_\varepsilon = \{ \theta_\varepsilon = \tau_\varepsilon \}$.

Lemma 4. Families $(C_\varepsilon)_{\varepsilon > 0}$ and $(E_\varepsilon)_{\varepsilon > 0}$ are asymptotically equivalent as $\varepsilon \to 0$.

Proof: The proof repeats the proof of Lemma 3.

We will need one more asymptotic equivalence statement since it is most convenient to work with events $F_\varepsilon = \{ x_0 + \varepsilon U(\infty) > 0 \}$. 

□
Lemma 5. Families \((E_\varepsilon)_{\varepsilon > 0}\) and \((F_\varepsilon)_{\varepsilon > 0}\) are asymptotically equivalent as \(\varepsilon \to 0\).

Proof: This proof is also very similar to that of Lemma 3. First we notice that \(P(F_\varepsilon \setminus E_\varepsilon) = 0\), and then we use the strong Markov property to compute
\[
\frac{P(E_\varepsilon \setminus F_\varepsilon)}{P(E_\varepsilon)} = P(F_\varepsilon^c | E_\varepsilon) = P(D_\varepsilon(q_+, \infty)) \to 0, \quad \varepsilon \to 0.
\]

We have \(F_\varepsilon \subset E_\varepsilon\), i.e., on \(F_\varepsilon\) we have
\[
e^{\lambda \theta_\varepsilon}(x_0 + \varepsilon U(\theta_\varepsilon)) = q_+.
\]
Therefore, on \(F_\varepsilon\) we have
\[
\theta_\varepsilon = \frac{1}{\lambda} \ln \frac{q_+}{x_0 + \varepsilon U(\theta_\varepsilon)} = \frac{1}{\lambda} \ln \frac{q_+}{x_0 + \varepsilon U(\infty)} + \frac{1}{\lambda} \ln \frac{x_0 + \varepsilon U(\infty)}{x_0 + \varepsilon U(\theta_\varepsilon)}
= \frac{1}{\lambda} \ln q_+ + \frac{1}{\lambda} \ln \frac{1}{\varepsilon} - \frac{1}{\lambda} \ln \left(\frac{x_0}{\varepsilon} + U(\infty)\right) + \frac{1}{\lambda} \ln \frac{x_0 + \varepsilon U(\infty)}{x_0 + \varepsilon U(\theta_\varepsilon)}.
\]
so, introducing a standard Gaussian random variable \(N = U(\infty)\sqrt{2\lambda}\), we can write
\[
\theta_\varepsilon - \frac{2}{\lambda} \ln \frac{1}{\varepsilon} = \frac{1}{\lambda} \ln (2\lambda q_+ |x_0|) + \frac{1}{\lambda} R_\varepsilon + Q_\varepsilon,
\]
where
\[
R_\varepsilon = - \ln \left(\frac{N - |x_0| \sqrt{2\lambda}}{\varepsilon}\right) - \ln \left(\frac{|x_0| \sqrt{2\lambda}}{\varepsilon}\right),
\]
and
\[
Q_\varepsilon = \frac{1}{\lambda} \ln \frac{x_0 + \varepsilon U(\infty)}{x_0 + \varepsilon U(\theta_\varepsilon)}.
\]

Let us recall that for each \(\varepsilon > 0\), we are studying the above random variables on the event \(F_\varepsilon\) that can be expressed as \(F_\varepsilon = \{N > |x_0| \sqrt{2\lambda}/\varepsilon\}\).

Lemma 6. As \(\varepsilon \to 0\), \(\text{Law}[R_\varepsilon | F_\varepsilon] \Rightarrow \Lambda\).

Proof: There are several ways to prove this statement. One way is to introduce \(r(\varepsilon) = |x_0| \sqrt{2\lambda}/\varepsilon\) and directly write, as in the proof of Corollary 1 of [Bak],
\[
\frac{P(-\ln(N - r) - \ln r < x)}{P(N > r)} = \frac{P(N > r + e^{-x}/r)}{P(N > r)}
= \frac{1}{\sqrt{2\pi (r + e^{-x}/r)^2}} e^{-(r + e^{-x}/r)^2/2} \sim \frac{1}{\sqrt{2\pi r}} e^{-r^2/2} \sim \Lambda(x), \quad r \to \infty.
\]
Remark 1. The statement of Lemma 6 is not specific to the Gaussian distribution of the random variable $N$ as the proof may suggest. Theorem 5 of [Bak] shows that convergence of distributions of logarithms of residual life times to the Gumbel distribution holds if $N$ is replaced by any other distribution belonging to the domain of attraction of the Gumbel distribution as a max-stable law. In fact, the theory of residual life times allows to describe the entire small pool of distributions that can appear as a result of such a limiting procedure. We refer to [Bak] for further explanation of connection between extreme values, residual life times, and conditional exit times. We also notice that the proof of Theorem 3 in Section 2.3 is based on the same kind of asymptotics.

Lemma 7. As $\varepsilon \to 0$, $\text{Law}[Q_\varepsilon | F_\varepsilon] \Rightarrow \delta_0$ (Dirac measure at 0).

Proof: We need to prove that $(x_0 + \varepsilon U(\infty))/(x_0 + \varepsilon U(\theta_\varepsilon))$ given $F_\varepsilon$, converges in probability to 1, i.e., $\varepsilon(U(\infty) - U(\theta_\varepsilon))/(x_0 + \varepsilon U(\theta_\varepsilon))$ converges to 0. Due to (12), it is sufficient to prove that, given $F_\varepsilon$,

$$\Delta_\varepsilon = e^{\lambda \theta_\varepsilon} \varepsilon(U(\infty) - U(\theta_\varepsilon)) = e^{\lambda \theta_\varepsilon} \int_{\theta_\varepsilon}^{\infty} e^{-\lambda s} dW(s)$$

converges to 0. For that it is sufficient to check the following $L^2$-convergence:

$$\lim_{\varepsilon \to 0} \frac{E[\Delta_\varepsilon^2 1_{D_\varepsilon}]}{P(F_\varepsilon)} = 0,$$

which, in turn, follows from

(14) $$\lim_{\varepsilon \to 0} \frac{E[\Delta_\varepsilon^2 1_{D_\varepsilon}]}{P(F_\varepsilon)} = 0,$$

where $D_\varepsilon$ was introduced in (10), so $F_\varepsilon \subset D_\varepsilon$ since $x_0 + \varepsilon U(\infty) > 0$ implies $x_0 + \varepsilon U(t) = 0$ for some $t$.

Let us now introduce $(F_t)_{t \geq 0}$, the completion of the natural filtration of the driving Wiener process $W$ (we implicitly have used it above when dealing with the strong Markov property). Since $D_\varepsilon \in F_{\tau_\varepsilon}$, we can use the Itô isometry to derive

$$E[\Delta_\varepsilon^2 1_{D_\varepsilon}] = \varepsilon^2 E\left[1_{D_\varepsilon} e^{2\lambda \theta_\varepsilon} E\left[\left(\int_{\theta_\varepsilon}^{\infty} e^{-\lambda s} dW(s)\right)^2 \mid F_{\theta_\varepsilon}\right]\right]$$

$$= \varepsilon^2 E\left[1_{D_\varepsilon} e^{2\lambda \theta_\varepsilon} e^{-2\lambda \theta_\varepsilon} \lambda \right] = \frac{\varepsilon^2}{2\lambda} P(D_\varepsilon),$$

so (14) follows since the reflection principle gives $P(D_\varepsilon) = 2P(F_\varepsilon)$. □

Now, combining (13) with Lemmas 6 and 7, we obtain that

(15) $$\text{Law}\left[\theta_\varepsilon - \frac{2}{\lambda} \ln \frac{1}{\varepsilon} \mid F_\varepsilon\right] \Rightarrow \text{Law}\left[\frac{\ln(2\lambda q)}{\lambda} + \frac{1}{\lambda} Z\right], \quad \varepsilon \to 0.$$
Law

Let us study the limiting behavior of time in the one-dimensional model studied above. Where \( V \theta \) to compute the characteristic function \( \phi \)

F true if conditioning on Lemmas 4 and 5 allow to apply Lemma 1 and conclude that (15) remains

\( \varepsilon \) as \( \varepsilon \rightarrow 0 \).

Therefore, \( \phi_\varepsilon (r) \rightarrow e^{-r^2/(4\mu)} \) by dominated convergence, i.e.,

(17) \[ \text{Law} \left[ J_\varepsilon / \varepsilon \mid C_\varepsilon \right] \Rightarrow \text{Law} \left[ N / \sqrt{2\mu} \right], \quad \varepsilon \rightarrow 0, \]

where \( N \) is a standard Gaussian random variable.

In fact, joint convergence also holds true, and Theorem 5 is an immediate corollary of the following result:

**Lemma 8.**

\[ \text{Law} \left[ \left( \frac{I_\varepsilon}{\varepsilon^{2\mu/\lambda + \alpha}}, \frac{J_\varepsilon}{\varepsilon} \right) \mid C_\varepsilon \right] \Rightarrow \text{Law} \left[ \left( \frac{V^{\mu/\lambda}}{2\lambda q^{|x_1|} \mu/\lambda}, \frac{N}{\sqrt{2\mu}} \right) \right], \quad \varepsilon \rightarrow 0, \]

where the joint distribution of \( V, \xi \), and \( N \) is described in the statement of Theorem 2.

**Proof:** We just have to slightly modify the proofs above. Let us compute \( \phi_\varepsilon (r_1, r_2) \), the characteristic function of \( (I_\varepsilon / \varepsilon^{2\mu/\lambda + \alpha}, J_\varepsilon / \varepsilon) \) conditioned on \( C_\varepsilon \). Denoting conditional expectation given \( C_\varepsilon \) by \( E_{C_\varepsilon} \), and representing \( I_\varepsilon / \varepsilon^{2\mu/\lambda + \alpha} = F_\varepsilon (\xi_\varepsilon, \tau_\varepsilon) \) for some Borel function \( F_\varepsilon \), we can write

\[ \phi_\varepsilon (r_1, r_2) = E_{C_\varepsilon} e^{ir_1 F_\varepsilon (\xi_\varepsilon, \tau_\varepsilon)} e^{ir_2 \varepsilon^{-\mu/\lambda} \int_0^{\tau_\varepsilon} e^{\mu s} dW^2 (s)} 

= \int_{\mathbb{R}} P_{\tau_\varepsilon} (dt) E_{C_\varepsilon} e^{ir_1 F_\varepsilon (\xi_\varepsilon, t)} e^{ir_2 e^{-\mu t} \int_0^t e^{\mu s} dW^2 (s)}, \]
where $P_{\tau_\varepsilon}(dt)$ denotes the distribution of $\tau_\varepsilon$. We can use independence of $W^2$ from $C_\varepsilon$ and $\xi_\varepsilon$, and continue the computation:

$$
\phi_\varepsilon(r_1, r_2) = \int_\mathbb{R} P_{\tau_\varepsilon}(dt) E_{C_\varepsilon} e^{ir_1 F_\varepsilon(\xi_\varepsilon, t)} E_{C_\varepsilon} e^{ir_2 e^{-\mu t} \int_0^t e^{\mu s} dW^2(s)}
$$

$$
= \int_\mathbb{R} P_{\tau_\varepsilon}(dt) E_{C_\varepsilon} e^{ir_1 F_\varepsilon(\xi_\varepsilon, t)} e^{-r_2^2 \frac{1-e^{-2\mu t}}{4\mu}}
$$

$$
= E_{C_\varepsilon} e^{ir_1 F_\varepsilon(\xi_\varepsilon, \tau_\varepsilon)} e^{-r_2^2 / (4\mu)} + E_{C_\varepsilon} e^{ir_1 F_\varepsilon(\xi_\varepsilon, \tau_\varepsilon)} (e^{r_2^2 e^{-2\mu \varepsilon} / (4\mu)} - 1).
$$

The second term converges to 0 as the integral of a bounded random variable converging to zero in probability. Convergence of the first term follows from (16). □

**References**

[AMB11a] Sergio Angel Almada Monter and Yuri Bakhtin. Normal forms approach to diffusion near hyperbolic equilibria. *Nonlinearity*, 24(6):1883–1907, 2011.

[AMB11b] Sergio Angel Almada Monter and Yuri Bakhtin. Scaling limit for the diffusion exit problem in the Levinson case. *Stochastic Process. Appl.*, 121(1):24–37, 2011.

[BŚ] Yuri Bakhtin and Andrzej Święch. Scaling limits for conditional diffusion exit problems, doob's $h$-transform, and asymptotics for nonlinear elliptic equations. [http://arxiv.org/abs/1310.6023](http://arxiv.org/abs/1310.6023).

[Bak] Yuri Bakhtin. Gumbel distribution in exit problems. [http://arxiv.org/abs/1307.7060](http://arxiv.org/abs/1307.7060).

[Bak08] Yuri Bakhtin. Exit asymptotics for small diffusion about an unstable equilibrium. *Stochastic Process. Appl.*, 118(5):839–851, 2008.

[Bak10] Yuri Bakhtin. Small noise limit for diffusions near heteroclinic networks. *Dyn. Syst.*, 25(3):413–431, 2010.

[Bak11] Yuri Bakhtin. Noisy heteroclinic networks. *Probab. Theory Related Fields*, 150(1-2):1–42, 2011.

[BdH74] A. A. Balkema and L. de Haan. Residual life time at great age. *Ann. Probability*, 2:792–804, 1974.

[BG13] Nils Berglund and Barbara Gentz. On the noise-induced passage through an unstable periodic orbit II: The general case. Preprint [arXiv/1208.2557](http://arxiv.org/abs/1208.2557), to appear in SIAM J. Math. Anal., 2013.

[CGLM13] Frédéric Cérou, Arnaud Guyader, Tony Leliévre, and Florent Malrieu. On the length of one-dimensional reactive paths. *ALEA, Lat. Am. J. Probab. Math. Stat.*, 10(1):359–389, 2013.

[Day90] Martin V. Day. Some phenomena of the characteristic boundary exit problem. In *Diffusion processes and related problems in analysis, Vol. I (Evanston, IL, 1989)*, volume 22 of *Progr. Probab.*, pages 55–71. Birkhäuser Boston, Boston, MA, 1990.

[Day95] Martin V. Day. On the exit law from saddle points. *Stochastic Process. Appl.*, 60(2):287–311, 1995.

[dHF06] Laurens de Haan and Ana Ferreira. *Extreme value theory*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2006. An introduction.
686 Cherry Street, School of Mathematics, Georgia Institute of Technology, Atlanta, GA, 30332-0160

E-mail address: yuri.bakhtin@gmail.com