Universality of first-passage and residence-time
distributions in non-adiabatic stochastic resonance

NILS BERGLUND\(^1\)(*) and BARBARA GENTZ\(^2\)(**)  
\(^1\) Centre de Physique Théorique (***) , CNRS Luminy  
Case 907, 13288 Marseille Cedex 9, France  
\(^2\) Weierstrass Institute for Applied Analysis and Stochastics  
Mohrenstr. 39, 10117 Berlin, Germany  

PACS. 02.50.Ey – Stochastic processes.  
PACS. 05.10.Gg – Stochastic analysis methods (Fokker–Planck, Langevin, etc.).  
PACS. 05.40.-a – Fluctuation phenomena, random processes, noise, and Brownian motion.  

Abstract. – We present mathematically rigorous expressions for the first-passage-time and  
residence-time distributions of a periodically forced Brownian particle in a bistable potential.  
For a broad range of forcing frequencies and amplitudes, the distributions are close to  
periodically modulated exponential ones. Remarkably, the periodic modulations are governed  
by universal functions, depending on a single parameter related to the forcing period. The  
behaviour of the distributions and their moments is analysed, in particular in the low- and  
high-frequency limits.  

The amplification by noise of a weak periodic signal acting on a multistable system is known  
as stochastic resonance (SR). A simple example of a system showing SR is an overdamped  
Brownian particle in a symmetric double-well potential, subjected to deterministic periodic  
forcing as well as white noise. Despite of the amplitude of the forcing being too small to  
enable the particle to switch from one potential well to the other, such transitions can be  
made possible by the additive noise. For sufficiently large noise intensity, depending on the  
forcing period, the transitions between potential wells can become close to periodic. This  
mechanism was originally proposed by Benzi \textit{et al.} and Nicolis and Nicolis [1–3] in order to  
offer an explanation for the close-to-periodic occurrence of the major Ice Ages. Since then,  
it has been observed in a large variety of physical and biological systems (for reviews see,  
\textit{e.g.}, [4–7]).  

Although much progress has been made in the quantitative description of the phenomenon  
of SR, many of its aspects are not yet fully understood. Mathematically rigorous results have  
so far been limited to the regimes of exponentially slow forcing [8,9], or moderately slow  
forcing of close-to-threshold amplitude [10,11].

\(*\) E-mail: berglund@cpt.univ-mrs.fr  
\(**\) E-mail: gentz@wias-berlin.de  
\(\text{***}\) Unité Mixte de Recherche (UMR 6207) du CNRS et des Universités d’Aix Marseille 1, Aix Marseille 2  
et Sud Toulon–Var, affiliée à la FRUMAM.

\(\text{© EDP Sciences}\)
One of the measures introduced in order to quantify SR is the residence-time distribution, that is, the distribution of the random time spans the Brownian particle spends in each potential well between transitions. SR is characterized by the fact that residence times are more likely to be close to odd multiples of half the forcing period than not. The residence-time distribution was first studied by Eckmann and Thomas for a two-level system [12]. For continuous systems, it has been estimated, in the case of adiabatic forcing, by averaging the escape rate for the frozen potential over the distribution of jump phases [13,14].

For larger forcing frequencies, however, the adiabatic approximation can no longer be used. An alternative approach is to consider time as an additional dynamic variable, which yields a two-dimensional (2D) problem (with degenerate noise). In the absence of noise, the system has two stable periodic orbits, one oscillating around each potential well, and one unstable periodic orbit, which oscillates around the saddle and separates the basins of attraction of the two stable orbits. The residence-time distribution is closely related to the distribution of first passages of the stochastic process through the unstable orbit. 2D systems with periodic orbits are much harder to study than 2D gradient systems, in particular their invariant density is not known in general, and their action functional [15,16] may not be smooth [17]. Graham and Tél studied consequences of this lack of smoothness on the dynamics near unstable periodic orbits [18,19]. Distributions of first-passage locations through unstable periodic orbits were first analysed by Day [20–22], who found them to exhibit a nontrivial dependence on the noise intensity, called cycling. Other related studies relevant to our approach include those by Maier and Stein [23,24], Lehmann et al. [25,26], and Dykman et al. [27].

At first glance, however, the 2D approach seems to produce a paradoxical result. Indeed, it is known from the classical Wentzell–Freidlin theory [15–17] that the distribution of first-passage locations through a periodic orbit looks uniform on the level of exponential asymptotics [20]. This is due to the fact that translations along the periodic orbit do not contribute to the cost in terms of action functional. How can this fact be conciled with the quasistatic picture, which yields residence times concentrated near odd multiples of half the forcing period? Obviously, the answer has to lie in the subexponential behaviour of the distribution of transitions.

In this Letter, we extend the results of [22–26] to a mathematically rigorous expression for the first-passage-time distribution up to multiplicative errors in the subexponential prefactor, valid for a broad range of forcing periods [28,29], from which we then deduce the residence-time distribution. A particularly interesting aspect of the result is that both distributions are governed by universal periodic functions, depending only on the period of the unstable periodic orbit times its Lyapunov exponent. All the model-dependent properties of the distributions can be eliminated by a deterministic time change.

Model. – We consider one-dimensional stochastic differential equations of the form

\[ \mathrm{d}x_t = -\frac{\partial}{\partial x} V(x_t, t) \mathrm{d}t + \sigma \mathrm{d}W_t, \]

where \( W_t \) is a standard Wiener process, describing white noise, and the small parameter \( \sigma \) measures the noise intensity (the diffusion constant being \( D = \sigma^2/2 \)). The double-well potential \( V(x, t) \) depends periodically on time, with period \( T \). The simplest example is

\[ V(x, t) = \frac{1}{4} x^4 - \frac{1}{2} x^2 - A \sin(\omega t) x, \]

where the forcing has angular frequency \( \omega = 2\pi/T \) and amplitude \( |A| < \sqrt{4/27} \).
Our results apply to a general class of $T$-periodic double-well potentials. We assume that for each fixed $t$, $V(x,t)$ has two minima at $X_{1,2}^t$ and a saddle at $X^u(t)$, such that, for all times, $X_1^t(t) < c_1 < X^u(t) < c_2 < X_2^t(t)$ for two constants $c_1, c_2$ (in the particular case of the potential (2), one can take $c_2 = -c_1 = 1/\sqrt{3}$). Using Poincaré maps, it is then straightforward to show that in the absence of noise, the system (1) has exactly three periodic orbits, one of them unstable and staying between $c_1$ and $c_2$, which we denote by $x^{\text{per}}(t)$. We denote by $a(t) = -\partial_x^2 V(x^{\text{per}}(t), t)$ the curvature of the potential at $x^{\text{per}}(t)$, and by

$$
\lambda = \frac{1}{T} \int_0^T a(t) \, dt
$$

(3)

the Lyapunov exponent of the unstable orbit. We assume that $\lambda$ is of order 1, but $T$ can become comparable to Kramers’ time.

Finally, we need a non-degeneracy assumption for the system, which assures that the action functional is minimized on a discrete set of paths, and excludes symmetries other than time-periodicity [28]. In particular, it should not be possible to transform the equation into an autonomous one by a time-periodic change of variables. In the special case of the potential (2), this condition is met whenever $A \neq 0$. In addition, we will assume that $|A|$ is of order 1, while $\sigma^2 \ll |A|$. 

**First-passage-time distribution.** Assume the system starts at time $t_0$ in a given initial point in the left-hand potential well. We call first-passage time the random first time $\tau$ at which $x_t$ crosses the unstable periodic orbit $x^{\text{per}}(t)$. In [28], we derived the distribution of $\tau$ in the simplified setting of a piecewise quadratic potential. The general case will be treated in [29].

Consider sample paths first reaching $x^{\text{per}}(t)$ in a time interval $[t, t + \Delta] \subset [nT, (n+1)T]$. Most of these paths are close to a minimizer of the action functional [16], called most probable exit path (MPEP). It turns out that there are $n$ such MPEPs, all of them spending considerable time either near the bottom of the potential well, or near the unstable orbit. The $k$th MPEP remains inside the left-hand well for $n - k$ periods, and then idles along $x^{\text{per}}(t)$ during the remaining $k$ periods [23–27]. The contribution of paths tracking the $k$th MPEP to the probability of reaching $x^{\text{per}}(t)$ during $[t, t + \Delta]$ is of the form $c_k(t, \sigma) e^{-\sqrt{V}/\sigma^2}$, where $\sqrt{V}$ is the constant value of the quasipotential on $x^{\text{per}}(t)$, which can be computed by a variational method (see [16]). In the limit of small forcing amplitude, $\sqrt{V}$ reduces to twice the potential barrier height.

The main difficulty lies in the computation of the prefactor $c_k(t, \sigma)$. The time-averaged value of $c_k$, which has a double-exponential dependence on $k$, has been determined in [21, 23, 25, 26]. The major new result in [28] is the explicit time-dependence of $c_k$. The desired probability is then obtained by summing the contributions of all $k$ from 1 to $n$. Extending the sum to all $k \in \mathbb{Z}$ yields a more compact expression, and only produces an error of order $\sigma$. We thus arrive at the following result.

**Theorem 1.** For any initial time $t_0$, any $\Delta \geq \sqrt{\sigma}$, and all times $t \geq t_0$,

$$
P\{\tau \in [t, t + \Delta]\} = \int_t^{t+\Delta} p_{\text{fpt}}(s, t_0) \, ds \left[1 + r(\sigma)\right],
$$

(4)

where $r(\sigma) = O(\sqrt{\sigma})$ and

$$
p_{\text{fpt}}(t, t_0) = \frac{1}{N} Q_{\lambda T}(\theta(t) - |\ln \sigma|) \frac{\theta'(t)}{\lambda T_k(\sigma)} e^{-[\theta(t) - \theta(t_0)]/\lambda T_k(\sigma)} f_{\text{trans}}(t, t_0).
$$

(5)
The following notations are used in (5):

- \(T_K(\sigma)\) is the analogue of Kramers’ time in the autonomous case; it has the form
  \[ T_K(\sigma) = \frac{C}{\sigma} e^{V/\sigma^2}. \]  
  The prefactor has order \(\sigma^{-1}\) rather than 1 [25, 26], due to the fact that most paths reach \(x_{\text{per}}(t)\) through a bottleneck of width \(\sigma\) (the width would be larger if \(|A|\) were not of order 1 [24]).

- \(Q_{\lambda T}(y)\) is the announced universal \(\lambda T\)-periodic function; it has the explicit expression
  \[ Q_{\lambda T}(y) = 2\lambda T \sum_{k=-\infty}^{\infty} A(y - k\lambda T) \quad \text{with} \quad A(z) = \frac{1}{2} e^{-2z} \exp\left\{-\frac{1}{2} e^{-2z}\right\}, \]  
  and thus consists of a superposition of identical asymmetric peaks, shifted by a distance \(\lambda T\). The \(k\)th peak is the contribution of the \(k\)th MPEP to the crossing probability. The average of \(Q_{\lambda T}(y)\) over one period is equal to 1.

- \(\theta(t)\) contains the model-dependent part of the distribution; it is an increasing function of \(t\), satisfying \(\theta(t + T) = \theta(t) + \lambda T\), and is given by
  \[ \theta(t) = \text{const} + \int_0^t a(s) \, ds - \frac{1}{2} \ln \frac{v(t)}{v(0)}, \]  
  where \(v(t)\) is the unique periodic solution of the differential equation \(\dot{v}(t) = 2a(t)v(t)+1\). It is related to the variance of Eq. (1) linearized around \(x_{\text{per}}(t)\), and has the expression
  \[ v(t) = \frac{1}{e^{2\lambda T} - 1} \int_t^{t+T} \exp\left\{\int_s^{t+T} 2a(u) \, du\right\} ds. \]  

- \(f_{\text{trans}}(t, t_0)\) accounts for the initial transient behaviour of the system; it is an increasing function satisfying
  \[ f_{\text{trans}}(t, t_0) = \begin{cases} \mathcal{O}\left(\exp\left\{-\frac{L}{\sigma^2} \frac{e^{-\lambda(t-t_0)}}{1 - e^{-2\lambda(t-t_0)}}\right\}\right) & \text{for } \lambda(t-t_0) < 2|\ln \sigma| \\ 1 - \mathcal{O}\left(\frac{e^{-\lambda(t-t_0)}}{\sigma^2}\right) & \text{for } \lambda(t-t_0) \geq 2|\ln \sigma| \end{cases} \]  
  where \(L\) is a constant, describing the rate at which the distribution in the left-hand well approaches metastable equilibrium. The transient term thus behaves roughly like \(\exp\{-L e^{-[\theta(t)-\theta(t_0)]}/\sigma^2(1-e^{-2[\theta(t)-\theta(t_0)]})\}\). However, \(f_{\text{trans}}(t, t_0)\) can be different when starting with an initial distribution that is not concentrated in a single point.

- \(N\) is the normalization, which we compute below.

If it were not for the limitation on \(\Delta\), which is due to technical reasons, this result would show that the probability density of \(\tau\) is given by \(p_{\text{fpt}}(t, t_0)[1+r(\sigma)]\). We expect the remainder to be of order \(\sigma\) rather than \(\sqrt{\sigma}\). In the simplified setting considered in [28], there is no restriction on \(\Delta\), \(r(\sigma) = \sigma\), and we obtain explicit values for \(\overline{V}, C\) and \(L\).
Fig. 1 – First-passage-time distribution $p_{\text{fpt}}(t, 0)$ (full curve) for two different parameter values. The broken curve is proportional to the average density (12), but scaled to match the peak height in order to guide the eye. The x-axis comprises 10 periods on each plot; the vertical scale is not respected between plots. Parameter values are $V = 0.5$, $\lambda = 1$, and (a) $\sigma = 0.2$ (i.e., $D = \sigma^2/2 = 0.02$), $T = 2$, (b) $\sigma = 0.4$ (i.e., $D = \sigma^2/2 = 0.08$), $T = 10$. Smaller noise intensities $\sigma$ would prolongate the transient phase and yield a slower decay of the peaks, without changing their shape.

Taking $\theta(t)/\lambda$ as new time variable in (5) eliminates the factor $\theta'(t)/\lambda$ in the density. Thus $\theta(t)/\lambda$ can be considered as a natural parametrization of time, in which one has to measure the first-passage-time distribution in order to reveal its universal character. We may thus henceforth assume that $\theta(t) = \lambda t$.

The universal periodic function $Q_{\lambda T}$ depends only on the single parameter $\lambda T$. For large $\lambda T$, it consists of well-separated asymmetric peaks, while for decreasing $\lambda T$ these peaks overlap more and more and $Q_{\lambda T}(y)$ becomes flatter. The Fourier series of $Q_{\lambda T}$ reads

$$Q_{\lambda T}(y) = \sum_{q \in \mathbb{Z}} 2\pi i q/\lambda T \Gamma \left(1 + \frac{\pi i q}{\lambda T}\right) e^{2\pi i q y/\lambda T}.$$  \hspace{1cm} (11)

Since the Euler Gamma function $\Gamma$ decreases exponentially fast as a function of the imaginary part of its argument, $Q_{\lambda T}(y)$ is close, for small $\lambda T$, to a sinusoid of mean value 1 and amplitude exponentially small in $1/2\lambda T$. The remarkable fact that $|\ln \sigma|$ enters in the argument of $Q_{\lambda T}$ has been discovered, to our best knowledge, by Day, who termed it cycling [21,22]. It means that as $\sigma$ decreases, the peaks of the first-passage-time distribution are translated along the time-axis, proportionally to $|\ln \sigma|$. See also [23] for an interpretation of this phenomenon in terms of MPEPs. The remaining, non-periodic time dependence of (5) corresponds to an averaged density, and behaves roughly like

$$\exp\left\{-\frac{L}{\sigma^2} \frac{e^{-\lambda(t-t_0)}}{1 - e^{-2\lambda(t-t_0)}} - \frac{t-t_0}{T_K(\sigma)}\right\}. \hspace{1cm} (12)$$

This function grows from 0 to almost 1 in a time of order $2|\ln \sigma|/\lambda$, and then slowly decays on the scale of the Kramers time $T_K(\sigma)$. It is maximal for $\lambda(t-t_0) \approx V/\sigma^2$.

The first-passage-time distribution is thus in effect controlled by two parameters: the quantity $\lambda T$, measuring the instability of the saddle, which determines the shape of the distribution within a period; and the Kramers time, which governs the decay of the average density (12). Fig.1 shows typical examples of distributions. Though our result requires $\sigma^2 \ll |A|$, we have chosen relatively large noise intensities for illustration: for weaker noise, the decay is slower, unless one takes larger periods, which narrows the peaks. For small $T/T_K(\sigma)$, the distribution has many peaks, of comparable height once transients have died out (Fig. 1 (a)). Increasing the period for constant noise intensity yields narrower peaks, while increasing the noise intensity shortens the transient phase, and accelerates the asymptotic decay (Fig. 1 (b)). As soon
as $T$ reaches the order of $T_K(\sigma)$, the distribution becomes dominated by a single peak, and one enters the synchronization regime, with the particle switching wells twice per period.

Moments of the first-passage-time distribution can easily be computed up to a correction stemming from $r(\sigma)$ (the correction due to $f_{\text{trans}}(t)$ is of smaller order). Using the Fourier series (11), one finds (for $\theta(t) = \lambda t$)

$$E\{\tau^n\} = \frac{1}{N} n! T_K(\sigma)^n \left[ 1 + 2 \Re \sum_{q \geq 1} \frac{(2\sigma^2)^{\pi i q/\lambda T}}{(1 - 2\pi i q T K(\sigma)/T)^{n+1}} \Gamma \left( 1 + \frac{\pi i q}{\lambda T} \right) \right] [1 + O(r(\sigma))].$$

(13)

In particular, taking $n = 0$ yields the normalization $N$. Note that $\lim_{\sigma \to 0} \sigma^2 \log E\{\tau\} = \overline{V}$, in accordance with the classical Wentzell–Freidlin theory [15–17]. If either $\lambda T \ll 1$ or $T \ll T_K(\sigma)$, $E\{\tau^n\}$ is close to $n! T_K(\sigma)^n$, as for an exponential distribution.

**Residence-time distribution.** – The precise knowledge of the first-passage-time distribution allows to derive the system’s residence-time distribution. Assume that the Brownian particle makes a transition from the right-hand to the left-hand potential well, crossing the unstable orbit at time $s$, and, having visited the left-hand well, crosses the unstable orbit again at time $s + t$. The residence-time distribution $p_{\text{rt}}(t)$ is obtained [12] by integrating $p_{\text{fpt}}(s + t, s) \psi(s)$ over one period, where $\psi(s)$ is the asymptotic distribution of arrival phases (i.e., times modulo $T$). Assuming that, as for the potential (2), the two wells move half a period out of phase, $\psi(s) = Q \lambda T (\lambda(s - T/2) - |\ln \sigma|)/T$. From (5) and (7) one obtains, for $\theta(t) = \lambda t$ and up to a multiplicative error $1 + O(r(\sigma))$, the residence-time distribution

$$p_{\text{rt}}(t) = \frac{1}{N} \frac{1}{T_K(\sigma)} e^{-t/T_K(\sigma)} \tilde{f}_{\text{trans}}(t) \frac{\lambda T}{2} \sum_{k=-\infty}^{\infty} \frac{1}{\cosh^2(\lambda t + T/2 - kT)},$$

(14)

where $\tilde{f}_{\text{trans}}(t)$ has the same behaviour as $f_{\text{trans}}(t, 0)$. The average behaviour is the same as for the first-passage-time density. The periodic part of the distribution consists of identical symmetric peaks, located in odd multiples of $T/2$, which are well-separated when $\lambda T \gg 1$ (Fig. 2). Using (11) instead of (7) yields the Fourier series

$$p_{\text{rt}}(t) = \frac{1}{N} \frac{1}{T_K(\sigma)} e^{-t/T_K(\sigma)} \tilde{f}_{\text{trans}}(t) \left[ 1 + 2 \sum_{q \geq 1} (-1)^q \left( 1 + \frac{\pi i q}{\lambda T} \right)^2 \cos \left( \frac{2\pi q t}{T} \right) \right],$$

(15)

which converges quickly for small $\lambda T$. This series also allows to compute moments of the residence-time distribution, which behave similarly as (13). In particular, the $n$th moment is close to $n! T_K(\sigma)^n$ if either $\lambda T \ll 1$, or $T \ll T_K(\sigma)$.
Conclusion. – The most important aspect of our rigorous expression for the residence-time distribution is the fact that it is governed essentially by two dimensionless parameters, $\lambda T$ and $T/T_K(\sigma)$, which can be modified independently. The ratio $T/T_K(\sigma)$ between period and Kramers time appears in most quantitative measures of SR, which indicate an optimal amplification when $T$ is close to $2T_K(\sigma)$. In this regime, the probability of transitions between potential wells becomes significant during each period. The parameter $\lambda T$, by contrast, controls the concentration of residence times within each period. Large values of $\lambda T$ yield a sharply peaked residence-time distribution, regardless of the peak’s relative height.

***

It’s a pleasure to thank Anton Bovier, Arkady Pikovsky and Dan Stein for valuable discussions. N.B. thanks the WIAS for kind hospitality and financial support. B.G. thanks the Université du Sud Toulon–Var and the CPT–CNRS Luminy for kind hospitality. Financial support by the ESF Programme Phase Transitions and Fluctuation Phenomena for Random Dynamics in Spatially Extended Systems (RDSES) is gratefully acknowledged.

REFERENCES