The exit problem for diffusions with time periodic drift and stochastic resonance

Samuel Herrmann
Institut de Mathématiques Elie Cartan
Université Henri Poincaré Nancy I
B.P. 239
54506 Vandoeuvre-lès-Nancy Cedex
France

Peter Imkeller
Institut für Mathematik
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany

December 2, 2003

Abstract

Physical notions of stochastic resonance for potential diffusions in periodically changing double well potentials such as the spectral power amplification have proved to be defective. They are not robust for the passage to their effective dynamics: continuous time finite state Markov chains describing the rough features of transitions between different domains of attraction of meta-stable points. In the framework of one-dimensional diffusions moving in periodically changing double well potentials we design a new notion of stochastic resonance which refines Freidlin's concept of quasi-periodic motion. It is based on exact exponential rates for the transition probabilities between the domains of attraction which is robust with respect to the reduced Markov chains. The quality of periodic tuning is measured by the probability for transition during fixed time windows depending on a time scale parameter. Maximizing it in this parameter produces the stochastic resonance points.

Key words and phrases: stochastic resonance; diffusion; periodic potential; double well potential; noise induced transition; exit time distribution; comparison theorem; perturbed dynamical system; spectral theory; low lying eigenvalue.

2000 AMS subject classifications: primary 60 H 10, 60 J 60; secondary: 60 J 70, 86 A 10, 60 F 10, 34 D 45

Introduction

One of the simplest and earliest stochastic climate models goes back to Nicolis [17] and Benzi et al. [1]. It intends to give a qualitative explanation of glacial cycles and is based on a deterministic differential equation for the global mean temperature expressed through a balance between the albedo-driven absorbed and the black-body type emitted radiative energies. A periodic exterior forcing comes from the slowly fluctuating solar constant and is due to Milankovich cycles caused by the gravitation of big planets. Only the addition of a stochastic term as a second forcing makes spontaneous transitions between the otherwise isolated *meta-stable states* of temperature possible. The resulting stochastically and periodically perturbed differential equation was capable of describing at least one characteristic aspect of experience: the typically short and abrupt *transitions*, observed before by Kramers [14] in reaction-diffusion phenomena. The model was soon strongly disputed. Despite its lack of realistic assumptions, the concepts underlying the model brought to light the phenomenon of *stochastic resonance*. Roughly speaking, a periodic (input) system subject to random perturbations is in stochastic resonance, if the noise intensity is tuned in such a way that the random periodic output is optimal. A very lively research field developed around this concept, drawing numerous examples from a wide spectrum of areas (see Gammaitoni et al. [10], Freund et al. [9] for a survey).

The mathematically precise understanding of the phenomenon is still under discussion at the time this paper is written. The first approach is done in Freidlin [7], where the deep large deviations' theory of Freidlin and Wentzell [8] is employed to produce a notion of stochastic resonance explaining the phenomenon in the small noise limit as approximating the periodic hopping between the energetically most favorable states in the landscape provided by a periodically weakly perturbed potential with finitely many local minima. In this sense stochastic resonance can be understood as the ability of the system to undergo quasi-periodic motion in the limit of small noise intensity. Let us briefly recall this interpretation more precisely. If noise intensity is ε, in the absence of periodic exterior forcing, the exponential order of times at which successive transitions between meta-stable states occur corresponds to the work to be done against the potential gradient to leave a well. This fact, heuristically derived by Kramers and Eyring (*Kramers' time*), is shown with mathematical accuracy in Freidlin and Wentzell [8]. The attractor basins are subdivided into a hierarchy of cycles with main states corresponding to the deepest among the cycle states. In the presence of periodic forcing with period time scale $e^{\frac{\mu}{\varepsilon}}$, in the limit $\varepsilon \to 0$ transitions between (the main states of) cycles with critical hopping work close to μ will be periodically observed. Transitions with smaller critical work may happen, but are negligible in the limit. Those with larger critical work are forbidden. In the simplest case of two minima of potential depth V and v, $v < V$, the role of which switches periodically at time T, for T larger than $e^{\frac{v}{\varepsilon}}$ the diffusion will be quasi-deterministic, i.e. close to the deterministic periodic function following the location of the deepest well.

Quasi-periodicity captures an important aspect of stochastic resonance, as it provides conditions under which stochastic trajectories are able to exhibit periodic behavior. Yet, physics literature (see Gammaitoni et al. [10], Freund et al. [9]) stipulates that stochastic resonance not only explains conditions for *stochastically periodic behavior* but comprises its optimality in a sense quite similar to the resonance notions of wave dynamics. In classical optics resonance is understood as the optimal amplitude of the response of the system to periodic excitation. In the same way, a stochastic resonance point is claimed to explain *optimal periodic tuning* of the stochastic trajectories of the diffusion responding to deterministic periodic excitation. Amplitude as a measure of quality of periodic tuning is replaced by *signal-to-noise ratio* or *spectral power amplification* (see below). Numerical simulations as for example in Milstein and Tretjakov [16] clearly support the optical evidence that beyond the threshold described by Freidlin [7] at which quasi-deterministic behavior becomes possible, for different noise intensities quite different qualities of periodicity of the random trajectories can be observed. There are parameter ranges for T in which random trajectories follow quite well the deterministic shapes of excitation curves. But as T gets even bigger, many short excursions to the wrong well during one period may occur. They will not count on the exponential scale on which quasi-periodic motion is measured, but trajectories will look less and less periodic. Physicists' quality measures for tuning therefore cannot be explained on the basis of quasi-deterministic motion alone.

The thesis by I. Pavlyukevich [18] and [13] present an attempt to provide a mathematically sound underpinning of physical notions of stochastic resonance based on optimality of periodic tuning — as opposed to the trajectorial analysis of Berglund and Gentz [2] containing very fine estimates on relaxation times. The physical concepts are mostly based on comparisons of trajectories of the noisy system and the deterministic periodic curve describing the location of the relevant meta-stable states, averaged with respect to the equilibrium measure. In the simple one-dimensional situation considered above the system switches between a double well potential state $U(x)$ with two wells of unequal depths V and v, $v < V$, during the first half period, and $U(-x)$ for the second half period. The total period length is T, and stochastic perturbation comes from the coupling to a white noise of intensity ε. The most important measures of quality studied are the *spectral power amplification*, the related *signal-to-noise ratio* or the *entropy of the equilibrium distribution*. In particular the first two mentioned play an eminent role in the physical literature. They mainly contain the L^2 average in equilibrium of the spectral component of the solution trajectories corresponding to the input period T, normalized in different ways. These measures of quality are functions of ε and T, and the problem of finding the resonance point e.g. consists in optimizing them in ε for fixed (large) T.

Let us briefly explain a striking shortcoming of these concepts of optimal periodic tuning which made us look for different ones. The first step to find optimal tuning intensities $\varepsilon(T)$ for large T consists in reducing the dynamics of the diffusion to the *inter well motion*, i.e. the pure transitions between the potential minima. In the physics literature, this corresponds to the reduction given for example by Mc Namara and Wiesenfeld [15]. One ends up with continuous time two state Markov chains with transition probabilities corresponding to the inverses of the diffusions' Kramers-Eyring' times. The mathematical analysis of stochastic resonance then proceeds along the following lines. One first determines the optimal tuning parameters $\varepsilon(T)$ for large T for the approximating Markov chains, a rather simple task. To see that the Markov chain's behavior approaches the diffusion's in the small noise limit, spectral theory of the infinitesimal generator is used. Its spatial part is seen to possess a spectral gap between the second and third eigenvalues, and therefore the closeness of equilibrium distributions of Markov chain on the one hand and diffusion on the other hand can be well controlled. Surprisingly, however, the notion of spectral power amplification is not robust for the passage from the Markov chain to the diffusion. Subtle dependencies on the geometrical fine structure of the potential function U in the minima beyond the expected curvature properties lead to quite unexpected counterintuitive effects. For example, a subtle drag away from the other well caused by the sign of the third derivative of U in the deep well suffices to make the spectral power amplification curve strictly increasing in the parameter range in which the approximating Markov chain has its resonance point. This dramatic deviation from expected behavior is due to the significance the spectral power amplification attributes to small *intra well fluctuations*.

Our main motivation in writing this paper was to investigate concepts by which on the one hand the physical intuition of optimal periodic tuning of random trajectories with a simple periodic input can be interpreted in a mathematically sound way, and which on the other hand fail to have this unfortunate defect of robustness. We deal with the framework of one-dimensional potential diffusions. The notion of quality of periodic tuning we shall investigate completely excludes the effect of small intra well fluctuations and purely relies on the transition mechanism between domains of attraction given by the potential. At the same time it generalizes the previously known results to potential functions which may vary periodically in time in a continuous, but otherwise quite general way, and whose growth at $\pm\infty$ may just be linear. More precisely, we study diffusion processes driven by a Brownian motion of intensity ε given by the stochastic differential equation

$$dX_t = -\frac{\partial}{\partial x} U(\frac{t}{T}, X_t)dt + \sqrt{2\varepsilon} dW_t, \quad t \geq 0.$$

The underlying potential landscape (see Figure 1) is described by a function $U(t,x), t \geq$

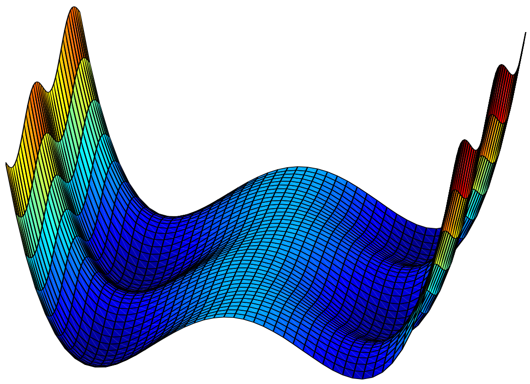

Figure 1: Potential landscape

$0, x \in \mathbb{R}$, which is periodic in time with period 1, and its temporal variation, by the rescaling with very large T, acts on the diffusion at a very small frequency. U is supposed to have exactly two wells located at ± 1, separated by a saddle at 0. The depth (measured in positive quantities) of $U(t, \cdot)$ at ± 1 is given by the 1-periodic depth functions $D_{\pm 1}(t)$ which are assumed to never fall below zero. We shall throughout look at time scales $T = \exp(\frac{\mu}{\varepsilon})$, for which the Kramers-Eyring formula indicates that transitions e.g. from the domain of attraction of -1 to the domain of attraction of 1 will occur as soon as D_{-1} gets less than μ, i.e. at time (see Figure 2)

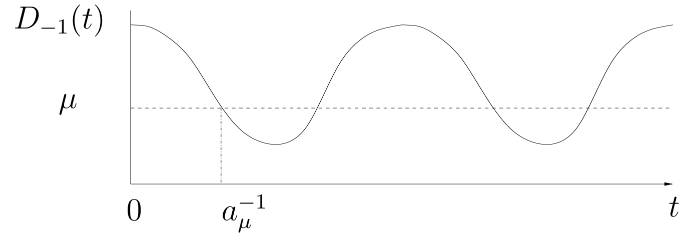

Figure 2: Definition of a_μ^{-1}

$$a_\mu^{\pm 1} = \inf\{t \geq 0 : D_{\pm 1}(t) \leq \mu\}.$$

This triggers periodic behavior of the diffusion trajectories on long time scales. The modern theory of meta-stability in *time homogeneous diffusion processes* complements the fundamental large deviations' theory presented by Freidlin and Wentzell [8] to produce the exponential decay rates of transition probabilities between different domains of attraction of a potential landscape together with very sharp multiplicative error estimates, uniformly on compacts in system parameters. Their sharpest forms are presented in some recent papers by Bovier et al. [3], [4], improving Day's previous results obtained in [5], [6]. They are derived from deep relationships of large deviations' theory with the spectral and capacity theory of the infinitesimal generator. We shall make use of this powerful machinery to obtain very precise estimates of the exponential tails of the laws of the transition times between domains of attraction. In fact, we have to extend the estimates by Bovier et al. [4] to the framework of *time inhomogeneous diffusions*. In the underlying one-dimensional situation, this is roughly achieved by freezing the time dependence of the potential on small time intervals to define lower and upper bound time homogeneous potentials not differing very much from the original one. Consequently comparison theorems are used to control the transition behavior from above and below by the behavior of the corresponding time homogeneous diffusions. This allows very precise estimates on the probabilities with which the diffusion at time scale $T = \exp(\frac{\mu}{\varepsilon})$ transits from the domain of attraction of -1 to the domain of attraction of 1 or vice versa within time windows $[(a_\mu^{\pm 1} - h)T, (a_\mu^{\pm 1} + h)T]$ for small $h > 0$. If $\tau_x(X)$ denotes the transit time to x, it is shown to be given by

$$\lim_{\varepsilon \to 0} \varepsilon \ln(1 - M(\varepsilon, \mu)) = \max_{i = \pm 1} \left\{ \mu - D_i(a_\mu^i - h) \right\},$$

with

$$M(\varepsilon, \mu) = \min_{i = \pm 1} \mathbb{P}_i(\tau_{-i}(X) \in [(a_\mu^i - h)T, (a_\mu^i + h)T]), \quad \varepsilon > 0, \mu \in I_R,$$

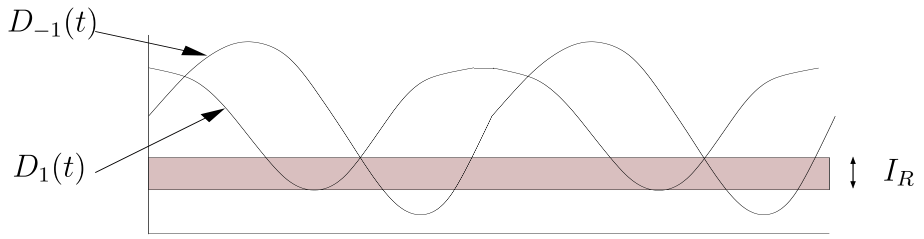

Figure 3: Resonance interval

and where I_R is the *resonance interval* (Figure 3), i.e. the set of scale parameters for which trivial or chaotic transition behavior of the trajectories is excluded.

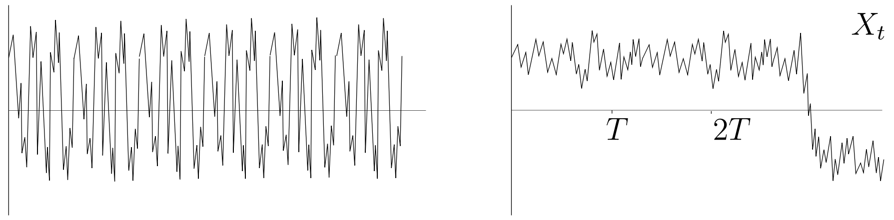

Figure 4: Chaotic and trivial transition behavior of the trajectories

The stated convergence is *uniform* in μ on compact subsets of I_R. This allows us to take $M(\varepsilon, \mu)$ as our measure of periodic tuning, compute the scale $\mu_0(h)$ for which the transition rate is optimal, and define the *stochastic resonance point* as the eventually existing limit of $\mu_0(h)$ as $h \to 0$. This notion of stochastic resonance is strongly related to the notions of periodic tuning based on *interspike intervals* (see Figure 5 and [11]), which describe the probability distribution for transitions as functions of time with exponentially decaying *spikes* near the integer multiples of the forcing periods. As

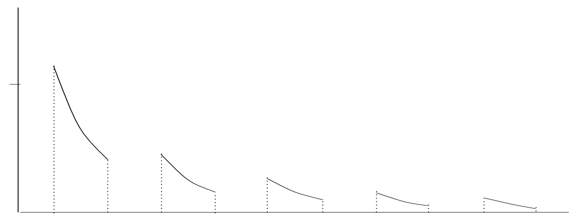

Figure 5: Length distribution of the interspike intervals in a simplified model (two state Markov chain)

opposed to the physics notions based on spectral decomposition of the statistics of the solution trajectories investigated in Pavlyukevich [18] or [13], it has the big advantage of being robust for the passage from the diffusion to the two state Markov chain reducing its behavior to the features of pure transitions between the two domains of attractions of meta-stable points.

Here is an outline of the organization of the material in the paper. Section 1 presents a review of results from the asymptotic theory of time homogeneous diffusions and their meta-stable sets needed for our purposes (Theorem 1.1). In section 2 we bring to work the tools of comparison theorems to deduce the sharp exponential transition rates for our time periodic diffusions from the time homogeneous results (Theorem 2.1). Section 3 is devoted to applying these sharp estimates to identify stochastic resonance points for diffusions (Theorem 3.2), compare them to their counterparts for the reduced Markov chains and prove robustness of our notion of resonance (Theorem 3.4).

1 Exponential distribution of transition times for time homogeneous diffusions

It will turn out to be crucial for our approach of periodic tuning to be discussed later to obtain large deviation type estimates for the exponential decay rate of the law of transition times uniformly in a time scale parameter. We shall make use of a technique of freezing time dependent potentials on small subintervals of the periodicity interval $[0, T]$ on their states taken at fixed times in the intervals, to be able to use known results for time homogeneous diffusions. In this setting, the uniformity problem translates into uniformity of the convergence to exponential decay rates in compact subsets of the domain of attraction the diffusion starts in and in time. It is clear that we are lead directly into large deviations' estimates for exit time distributions of *time homogeneous diffusions* such as presented in the pioneering book by Freidlin and Wentzell [8]. But for obtaining uniformity in space and time, one has to use sharpened versions of these estimates developed later for controlling in particular the exponential errors in the estimates. The purpose of this section is to summarize what we shall need from this fine well established theory.

We shall refer to the most recent and advanced development of sharp estimates for transition times presented in Bovier, Gayrard, Klein [3], [4]. They are valid far beyond our modest framework, both in the multidimensional case and for any finite number of local minima of the potential. Their quality comes from a detailed analysis of the relationship between transition times and low lying eigenvalues of the spectrum of the infinitesimal generator of the diffusion. We shall state them in the simple one-dimensional setting given here. A more complex multi-dimensional version can also be found in Day [5]. For this purpose, suppose that Q is a purely space dependent $\mathcal{C}^2$ potential function (see Figure 6) possessing only $-1, 1$ as local minima, separated by

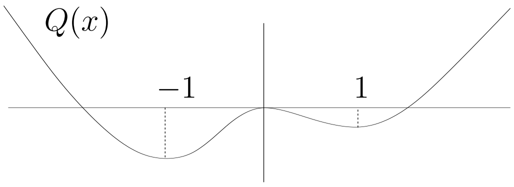

Figure 6: Potential

the saddle point 0 at which Q takes the value 0. Suppose that the curvature of Q at -1 is strictly positive, i.e. $Q''(-1) > 0$. As for ultra- or hypercontractivity type properties for Q, we shall assume that it has exponentially tight level sets, i.e. there is $a_0 > 0$ such that for any $a \geq a_0$ there exists a constant $C(a)$ such that for $\varepsilon \leq 1$

$$\int_{\{y:Q(y)\geq a\}} \exp(-\frac{Q(z)}{\varepsilon})dz < C(a)\exp(-\frac{a}{\varepsilon}). \tag{1}$$

We shall concentrate in this situation on a transition out of the domain of attraction of the stable point -1 for the diffusion associated with the SDE

$$\begin{cases} dY_t^\varepsilon = -Q'(Y_t^\varepsilon)\,dt + \sqrt{2\varepsilon}dW_t, \\ Y_0^\varepsilon = y. \end{cases}$$

Let C be a closed interval of the form $[d, \infty[$ with $d \neq 0$. To state our aim in a slightly different version, we will be interested in the asymptotics of the entrance time of Y^ε into C:

$$\tau_C^\varepsilon = \inf\{t > 0 : \ Y_t^\varepsilon \in C\}.$$

Then we obtain the following result (see [5] or [4]).

Theorem 1.1 *Let λ^ε denote the principal eigenvalue of the linear operator*

$$\mathcal{L}^\varepsilon u = \varepsilon u'' - Q'u'$$

with Dirichlet boundary conditions on ∂C. Then for every compact $K \subseteq]-\infty, 0[$ there is a constant $c > 0$ such that

$$\mathbb{P}_y(\tau_C^\varepsilon > t) = e^{-\lambda^\varepsilon t}(1 + \mathcal{O}_K(e^{-c/\varepsilon})), \tag{2}$$

where $\mathcal{O}_K$ denotes an error term which is uniform in $y \in K, t \geq 0$. Moreover, for the asymptotic behavior of the eigenvalue λ^ε the following holds

$$\lambda^\varepsilon \mathbb{E}_y[\tau_C^\varepsilon] \to 1 \quad \textit{uniformly on compacts} \quad K \subseteq]-\infty, 0[. \tag{3}$$

as $\varepsilon \to 0$.

Proof:

There are two small issues which deserve comments.

First, the uniformity over compacts in $]-\infty, 0[$ claimed in the main statements. Day [5] tackles it. But he considers only exits from bounded domains. Bovier, Gayrard, Klein [4] have a version for unbounded domains, but uniformity over compacts of the domain of attraction in which the diffusion starts is not explicitly proved. It is, however, hidden in their method of proof of Theorem 1.3 ([4], pp. 30, 31) which makes use of an eigenfunction expansion. But due to regularity results on the eigenfunctions (see [4], pp. 16-18) they must be bounded on compacts in the domain of attraction $]-\infty, 0[$. This implies the desired uniformity.

The second comment concerns our assumptions on C. Translated into our setting, in [4] the target set C is assumed to be closed, to contain a neighborhood of 1 if the potential is deeper there than at -1, and to have a positive distance from the saddle 0. Since we are in a one-dimensional setting, we may reduce these conditions to the simple one $d \neq 0$. If necessary, we may always cut out of C a small open neighborhood of 0 without changing the law of τ_C if starting from the domain of attraction of -1. $\square$

2 Exponential transition rates between moving domains of attraction

We shall now consider a potential diffusion given by the one-dimensional stochastic differential equation

$$dX_t = -\frac{\partial U}{\partial x}(\frac{t}{T}, X_t)\, dt + \sqrt{2\varepsilon} dW_t. \tag{4}$$

The time periodic potential U of period 1 is supposed to fulfill the following conditions. First of all, its global rough geometry is the one of a double well potential with temporally moving wells, but time independent critical points. For simplicity we suppose that its local minima are given by ± 1, and its only saddle point by 0, independently of time. So ± 1 are the only metastable states of the potential on the whole time axis. Outside of $0, \pm 1$, $\frac{\partial U}{\partial x}$ is supposed to be continuous in (t, x). Our main concern will be the asymptotics of the transition times from the domain of attraction $]-\infty, 0[$ of -1 to the domain of attraction $]0, \infty[$ associated with 1 of the *time inhomogeneous diffusion* in the small noise limit $\varepsilon \to 0$. More precisely we will be interested in describing the exponential transition rate from the domain of -1 to the domain of 1. Our potential not being time homogeneous, we shall make use of comparison arguments with diffusions possessing time independent potentials in order to find a careful reduction of the inhomogeneous exit problem to the homogeneous one, and use the asymptotic results stated in Theorem 1.1 in this framework. This will be achieved by freezing the driving force derived from the potential on small time intervals on the mimimal or maximal level it takes there. To be more precise, for each interval $I \subset \mathbb{R}_+$ let

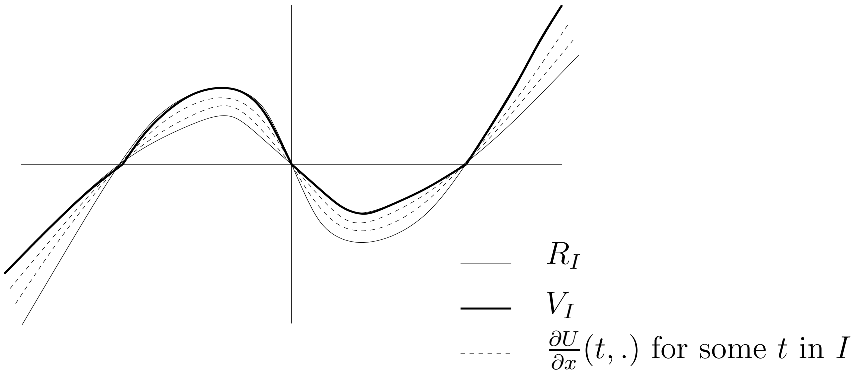

Figure 7: Definition of V_I and R_I

$$V_I(x) = \sup_{t \in I} \frac{\partial U}{\partial x}(t, x) \quad \text{and} \quad R_I(x) = \inf_{t \in I} \frac{\partial U}{\partial x}(t, x). \tag{5}$$

The regularity conditions valid for U imply that V and R are continuous functions. Moreover $V_I(-1) = R_I(-1) = 0$. If $I = [a, b]$, we denote by $\overline{X}^I$ the solution of the SDE on $\mathbb{R}_+$

$$\begin{cases} d\overline{X}^I_t = -R_I(\overline{X}^I_t)dt + \sqrt{2\varepsilon}dB_t, \\ \overline{X}^I_0 = X_{aT}. \end{cases} \tag{6}$$

$\underline{X}^I$ is defined in the same way replacing R_I by V_I. These two *time homogeneous* diffusions are used to control the *time inhomogeneous* diffusion X as long as time runs in the interval I. In fact, we have $P-$a.s.

$$\underline{X}^I_{tT} \leq X_{(t+a)T} \leq \overline{X}^I_{tT}, \quad t \in [0, b-a].$$

Of course, to make use of the asymptotic results stated in the previous section, we need ultra- or hypercontractivity properties for the frozen potentials. To formulate a hypothesis which is both not too restrictive and easy to handle for time dependent potentials, let us give the following easy sufficient criterion for exponential tightness of levels of a time independent potential Q.

Lemma 2.1 *Assume that Q is a real valued differentiable function on $\mathbb{R}$, and that there are constants $K_1, K_2 > 0$ such that*

$$\begin{aligned} Q'(x) &\leq -K_2 \quad \textit{for} \quad x \leq -K_1, \\ Q'(x) &\geq K_2 \quad \textit{for} \quad x \geq K_1. \end{aligned} \tag{7}$$

Then Q has exponentially tight level sets.

Proof:
It is obviously enough to argue on $\mathbb{R}_+$. Due to (7), we know that near ∞, Q is strictly increasing with inverse Q^{-1}. So for $a > 0$ big enough and $\varepsilon > 0$ we have

$$\begin{aligned} \int_{\{y:Q(y)\geq a\}} \exp(-\frac{Q(z)}{\varepsilon})dz &= \int_a^\infty \exp(-\frac{t}{\varepsilon})\frac{1}{Q'(Q^{-1}(t))}dt \\ &\leq \frac{1}{K_2}\int_a^\infty \exp(-\frac{t}{\varepsilon})dt \\ &= \frac{\varepsilon}{K_2}\exp(-\frac{a}{\varepsilon}). \end{aligned}$$

This clearly implies exponential tightness.□

Conditions like (7) are practical in our setting for the following reason: if we assume them to be satisfied uniformly for $t \in [0, 1]$, it is clear that all frozen potential functions will inherit the property. We will therefore assume throughout that there exist constants $K_1, K_2 > 0$ such that

$$\begin{aligned} \sup_{t\geq 0} \frac{\partial U}{\partial x}(t, x) &\leq -K_2 \quad \text{for} \quad x \leq -K_1, \\ \inf_{t\geq 0} \frac{\partial U}{\partial x}(t, x) &\geq K_2 \quad \text{for} \quad x \geq K_1. \end{aligned} \tag{8}$$

We measure periods T on the logarithmic scale μ given by $T = T_\varepsilon = \exp(\frac{\mu}{\varepsilon})$. The reason for this is hidden in the classical formula of Kramers-Eyring. It states that in the small noise limit T is the time it takes the diffusing particle to climb a height μ in a potential landscape U. This formula has the following intuitive consequences. Assume the diffusion faces an obstacle of constant potential height $U^+ > \mu$ for exiting if it diffuses on time scale T. Then asymptotically it never exits on this scale. On the other hand, if it faces an obstacle of height $U_- < \mu$ diffusing at time scale T, it has to exit immediately in the small noise limit. We introduce the depth function at $x \in \mathbb{R}$ by

$$D_x(t) = U(t,0) - U(t,x), \quad t \geq 0.$$

The maximal well depths D_{-1} and D_1 will be of particular importance. We shall assume that they satisfy the assumption (see Figure 2)

$$\text{all local extrema of } \quad D_1, D_{-1} \quad \text{are global and the} \atop \text{functions are strictly monotonous between the extrema.} \tag{9}$$

For $\mu \geq 0$ let now

$$a_\mu = \inf\{t \geq 0 : D_{-1}(t) \leq \mu\}.$$

This is the same function as a_μ^{-1} defined above. Here we omit the superscript since we always concentrate on transitions from -1 to 1. The interval

$$I_{-1} =]\inf_{t\geq 0} D_{-1}(t), \sup_{t\geq 0} D_{-1}(t)[\tag{10}$$

contains all possible depths the potential minimum located at -1 takes during one period of time. Note that $a_\mu = 0$ for μ strictly above the upper boundary of I_{-1}, and $a_\mu = \infty$ for μ strictly below the lower boundary. If $\mu \in I_{-1}$, then for times beyond a_μ, the barrier height between -1 and 0 has dropped below the critical level so that on time scale T the diffusion exits immediately. According to this heuristics, the diffusion running on time scale T should exit the domain of attraction $]-\infty, 0[$ through 0 and then transit to the other well immediately after $TD_{-1}(t)$ drops below μ. We shall be interested in the exponential rate at which this happens, uniformly in starting points taken from a compact in $]-\infty, 0[$. For this purpose, for a regular diffusion Y on $\mathbb{R}$ and $a \in \mathbb{R}$, we denote by $\tau_a(Y)$ the first hitting time of a by Y.

Our first aim is to prove the precise estimate

$$\lim_{\varepsilon\to 0} \varepsilon \ln \mathbb{P}_x(\tau_1(X) \leq (a_\mu - h)T) = \mu - D_{-1}(a_\mu - h) \tag{11}$$

for $x < 0$, uniformly in μ on compact subintervals Γ of (10) and for $0 < h < \inf_{\mu\in\Gamma} a_\mu$. Note that given such a compact Γ, due to the continuity of D_{-1} we have $\inf_{\mu\in\Gamma} a_\mu > 0$. Fix Γ, x and h from now on. Our approach proceeds in essentially two steps.

2.1 Upper bound for the exponential exit rate

In the first step we shall find upper bounds for the exponential rates. For this purpose we shall partition the relevant time intervals $[0, a_\mu - h]$. Fix some $\delta > 0$ such that

$\delta < |x|$. Since $\frac{\partial U}{\partial x}$ is continuous we may choose an equidistant partition $0 = r_0 < r_1 < \cdots < r_n = a_\mu - h$ of $[0, a_\mu - h]$ with mesh γ small enough to ensure

$$\sup_{s,t\in[r_j,r_{j+1}]} \sup_{x\in[-1,0]} \left| \frac{\partial U}{\partial x}(t,x) - \frac{\partial U}{\partial x}(s,x) \right| \leq \delta. \tag{12}$$

Denote $I_j = [r_{j-1}, r_j], 1 \leq j \leq n$. Though the choice of the intervals depends on h and Γ, their number will be bounded by a universal constant. Using this partition, we may start our search for an upper bound by freezing the time dependent potential on I_j at its value R_{I_j}, $1 \leq j \leq n$, and then comparing the diffusion X there with $\overline{X}^{I_j}$. There is a little difficulty with this procedure. The drift coefficients R_{I_j} which govern the sde for $\overline{X}^{I_j}$ were defined by taking infima over time intervals. This operation may destroy their differentiability properties in the spatial variable. Therefore it may be necessary to compare X on the intervals I_j with smoother diffusions still dominating it. But this can be done at no extra cost. For each $1 \leq j \leq n$ we may choose $R_j \in \mathcal{C}^1(\mathbb{R})$ satisfying

$$\begin{cases} R_j \leq R_{I_j} \leq R_j + \delta, \\ \text{there are } \underline{m}_j \in]-1-\delta, -1+\delta[, s_j \in]-\delta, \delta[, \overline{m}_j \in]1-\delta, 1+\delta[\\ \quad \text{such that } R_j|_{]-\infty,\underline{m}_j[\cup]s_j,\overline{m}_j[} < 0, R_j|_{]\underline{m}_j,s_j[\cup]\overline{m}_j,\infty[} > 0, \\ R_j'(\underline{m}_j) > 0, \\ \text{there are constants } K_1, K_2 > 0 \text{ such that } R_j(x) \leq -K_2 \text{ for } x \leq -K_1, \\ R_j(x) \geq K_2 \text{ for } x \geq K_1. \end{cases} \tag{13}$$

Let X^j be the diffusion associated with R_j, $1 \leq j \leq n$. Let us choose a partition $x = x_0 < x_1 < \cdots < x_n = -\delta$ of the interval $[x, -\delta]$ which will typically not be supposed to be equidistant. In the following inequality the diffusions X and X^j on I_j are compared and the Markov property is employed. We have

$$\begin{aligned} \mathbb{P}_x(\tau_1(X) \leq (a_\mu - h)T) &\leq \sum_{j=1}^{n-1} \mathbb{P}_x(\tau_{x_{j-1}}(X) \geq r_{j-1}T, \tau_{x_j}(X) \leq r_j T) \\ &\quad + \mathbb{P}_x(\tau_{x_{n-1}}(X) \geq r_{n-1}T, \tau_1(X) \leq r_n T) \\ &\leq \sum_{j=1}^{n-1} \mathbb{P}_x\left[\{\tau_{x_j}(X_{\cdot+r_{j-1}T}) \leq \gamma T\} \cap \{\tau_{x_{j-1}}(X) \geq r_{j-1}T\}\right] \\ &\quad + \mathbb{P}_x\left[\{\tau_1(X_{\cdot+r_{n-1}T}) \leq \gamma T\} \cap \{\tau_{x_{n-1}}(X) \geq r_{n-1}T\}\right] \\ &\leq \sum_{j=1}^{n-1} \mathbb{E}_x\left[\mathbb{P}_{X_{r_{j-1}T}}(\tau_{x_n}(X^j) \leq \gamma T) 1_{\{\tau_{x_{j-1}}(X) \geq r_{j-1}T\}}\right] \\ &\quad + \mathbb{E}_x\left[\mathbb{P}_{X_{r_{n-1}T}}(\tau_1(X^n) \leq \gamma T) 1_{\{\tau_{x_{n-1}}(X) \geq r_{n-1}T\}}\right] \\ &\leq \sum_{j=1}^{n-1} \mathbb{P}_{x_{j-1}}(\tau_{x_j}(X^j) \leq \gamma T) \\ &\quad + \mathbb{P}_{x_{n-1}}(\tau_1(X^n) \leq \gamma T). \end{aligned} \tag{14}$$

Let us now fix $1 \leq j \leq n$ and continue estimating the term $\mathbb{P}_{x_{j-1}}(\tau_{x_j}(X^j) \leq \gamma T)$ and $\mathbb{P}_{x_{n-1}}(\tau_1(X^n) \leq \gamma T)$ individually. For this purpose we apply Theorem 1.1 for

$Q = R_j, d = x_j$ to obtain that

$$\begin{aligned} \mathbb{P}_{x_{j-1}}(\tau_{x_j}(X^j) \leq \gamma T) &\leq 1 - e^{-\lambda_j^\varepsilon \gamma T}(1 - e^{-c/\varepsilon}) \leq 1 - e^{-\lambda_j^\varepsilon \gamma T} + e^{-c/\varepsilon}, \\ \mathbb{P}_{x_{n-1}}(\tau_1(X^n) \leq \gamma T) &\leq 1 - e^{-\lambda_n^\varepsilon \gamma T}(1 - e^{-c/\varepsilon}) \leq 1 - e^{-\lambda_n^\varepsilon \gamma T} + e^{-c/\varepsilon} \end{aligned} \tag{15}$$

uniformly in T, hence uniformly in μ. Here λ_j^ε denotes the principal eigenvalue of the operator $\mathcal{L}_j^\varepsilon$ defined by

$$\mathcal{L}_j^\varepsilon u = \varepsilon u'' - R_j \, u'$$

with Dirichlet boundary conditions at x_j. We now come to the crucial part of the derivation of an upper estimate. We shall use precise asymptotics of the eigenvalues λ_j^ε.

Lemma 2.2 *There exists $C > 0$ such that for $1 \leq j \leq n$*

$$\left| \lim_{\varepsilon \to 0} \varepsilon \ln \lambda_j^\varepsilon - [U(r_j, x_j) - U(r_j, -1)] \right| \leq C\delta. \tag{16}$$

Proof:
Fix $1 \leq j \leq n$. Define the pseudopotential corresponding to the drift R_j by

$$V_j(x, z) = \inf \left\{ \frac{1}{2} \int_0^t (\phi_s' + \frac{R_j}{2}(\phi_s))^2 \, ds, \ \phi_0 = x, \ \phi_t = z, \ t > 0 \right\},$$

where ϕ stands for absolutely continuous functions defined on the time interval $[0, t]$. Since due to our assumptions $\underline{m}_j$ is the only local minimum of the potential corresponding to R_j on $]-\infty, 0[$, the sharpened form of the exit time Theorem of Freidlin-Wentzell (see Bovier, Gayrard, Klein [4], Theorem 1.1) implies that

$$\lim_{\varepsilon \to 0} \varepsilon \ln \lambda_j^\varepsilon = -V_j(\underline{m}_j, x_j).$$

Let us estimate the pseudopotential. We have

$$\begin{aligned} V_j(\underline{m}_j, x_j) &= \int_{\underline{m}_j}^{x_j} R_j(\theta) d\theta \\ &= \int_{-1}^{x_j} \frac{\partial U}{\partial x}(r_j, \theta) d\theta + \int_{\underline{m}_j}^{x_j} \left(R_j(\theta) - \frac{\partial U}{\partial x}(r_j, \theta) \right) d\theta \\ &\quad - \int_{[-1, \underline{m}_j]} \frac{\partial U}{\partial x}(r_j, \theta) d\theta. \end{aligned}$$

Continuity of $\frac{\partial U}{\partial x}$ in (t, x) entails the existence of $C_1 < 0$ such that

$$\left| \int_{[-1, \underline{m}_j]} \frac{\partial U}{\partial x}(r_j, \theta) d\theta \right| \leq C_1 \delta.$$

To estimate the second remainder term, recall that the mesh γ was chosen to produce at most δ as modulus of continuity of $\frac{\partial U}{\partial x}$ (see (12)), and that R_j is also at most a distance δ away (see (13)). We therefore obtain

$$\left| \int_{\underline{m}_j}^{x_j} \left(R_j(\theta) - \frac{\partial U}{\partial x}(r_j, \theta) \right) d\theta \right| \leq 2\delta.$$

Hence

$$\left|\lim_{\varepsilon\to 0}\varepsilon\ln\lambda_j^\varepsilon - [U(r_j,x_j) - U(r_j,-1)]\right| \le (2+C_1)\delta. \tag{17}$$

The asserted asymptotic result follows. □

As a consequence we obtain an upper bound for the exponential convergence rate for the exit time from the domain of attraction of the potential well at -1.

Proposition 2.1 *Let $x < 0$, Γ a compact subset of (10). Then there exists $0 < h_0 < \inf_{\mu\in\Gamma} a_\mu$ such that for $h \le h_0$*

$$\limsup_{\varepsilon\to 0}\varepsilon\ln \mathbb{P}_x(\tau_1(X) \le (a_\mu - h)T) \le \mu - D_{-1}(a_\mu - h) \tag{18}$$

uniformly for $\mu \in \Gamma$.

Proof:
According to what has been proved, there are constants $\varepsilon_0 > 0$ and $K > 0$ such that for $\varepsilon \le \varepsilon_0, \mu \in \Gamma$

$$\begin{aligned} \mathbb{P}_x(\tau_1(X) \le (a_\mu - h)T) &\le \sum_{j=1}^{n-1}\mathbb{P}_{x_{j-1}}(\tau_{x_j}(X^j) \le \gamma T) + \mathbb{P}_{x_{n-1}}(\tau_1(X^n) \le \gamma T) \\ &\le K\gamma T\sum_{j=1}^{n}\lambda_j^\varepsilon + ne^{-\frac{c}{\varepsilon}}. \end{aligned} \tag{19}$$

Taking logarithms on both sides, multiplying by ε and using the equation

$$\lim_{\varepsilon\to 0}\varepsilon\ln(f(\varepsilon)+g(\varepsilon)) = \max[\lim_{\varepsilon\to 0}\varepsilon\ln(f(\varepsilon)), \lim_{\varepsilon\to 0}\varepsilon\ln(g(\varepsilon))]$$

for two positive functions f and g, we may apply Lemma 2.2 to get

$$\begin{aligned} &\limsup_{\varepsilon\to 0}\varepsilon\ln\mathbb{P}_x(\tau_1(X) \le (a_\mu - h)T) \\ &\quad \le [\max[\mu - [U(r_j,x_j) - U(r_j,-1)] : 1 \le j \le n] + C\delta)] \vee (-c) \\ &\quad \le [\max[\mu - [U(r_j,x_1) - U(r_j,-1)] : 1 \le j \le n] + C\delta)] \vee (-c). \end{aligned}$$

Recalling the definition of a_μ and that $x_1 < 0$ is arbitrary, we may conclude

$$\begin{aligned} &\limsup_{\varepsilon\to 0}\varepsilon\ln\mathbb{P}_x(\tau_1(X) \le (a_\mu - h)T) \\ &\quad \le [\max[\mu - D_{-1}(r_j) : 1 \le j \le n] + C\delta)] \vee (-c) \\ &\quad \le [(\mu - D_{-1}(a_\mu - h)) + C\delta)] \vee (-c) \end{aligned}$$

uniformly for $\mu \in \Gamma$. Now choose $h_0 > 0$ small enough so that for $h \le h_0$ we have

$$\inf_{\mu\in\Gamma}(\mu - D_{-1}(a_\mu - h)) > -c.$$

Finally, since δ is arbitrary, we may let δ tend to zero. This way we obtain the desired upper bound for the exponential rate. □

2.2 Lower bound for the exponential exit rate

In the second step of our approach, we shall establish lower bounds for the exponential rates at which the diffusion exits from the basin of attraction of -1. Let us first prove an auxiliary result. It states that the probability of exiting the interval $[l, 0]$ via l is exponentially small with exponential order increasing in $|l|$, due to hypercontractivity. Recall the constants K_1 and K_2 from (8).

Lemma 2.3 *There exist positive constants C and ε_0 such that for $\varepsilon \leq \varepsilon_0$ and $l < x \wedge -K_1$ and $\mu > 0$ we have*

$$\mathbb{P}_x(\tau_l(X) \leq T) \leq \frac{C}{\varepsilon} \exp(\frac{2K_2(l - x \wedge (-K_1)) + \mu}{\varepsilon}) \tag{20}$$

Proof:

We give arguments for the case $x < -K_1$, the other case being easier. Recalling that by (8) the gradient of U is bounded below by $-K_2$ on $\mathbf{R}_+ \times]-\infty, -K_1]$, we may compare the diffusion X with the diffusion Z on the interval $]-\infty, x]$ reflected at x with constant drift equal to K_2. It can be given by the SDE

$$dZ_t = K_2 dt + dL_t + \sqrt{2\varepsilon} dW_t,$$

where $Z_0 = x$, and L is an increasing process satisfying $\int_0^t (Z_s - x) dL_s = 0, t \geq 0$. See for example Revuz, Yor [19]. The comparison clearly yields

$$\mathbb{P}_x(\tau_l(X) \leq T) \leq \mathbb{P}_x(\tau_l(Z) \leq T). \tag{21}$$

By Tchebyshev's inequality

$$\mathbb{P}_x(\tau_l(Z) \leq T) \leq e\mathbb{E}_x[\exp(-\frac{1}{T}\tau_l(Z))]. \tag{22}$$

Let $\varphi(y) = \mathbb{E}_y[\exp(-\frac{1}{T}\,\tau_l(Z))]$, $l \leq y \leq x$. Our task consists in an estimation of $\varphi(x)$. According to the Feynman-Kac and Dynkin formulae φ solves the boundary value problem

$$\begin{cases} \varepsilon\,\varphi'' + K_2\varphi' - \frac{1}{T}\,\varphi = 0 \quad \text{on} \quad]l, x[, \\ \varphi'(x) = 0, \quad \varphi(l) = 1. \end{cases}$$

The eigenvalues of the differential equation are determined by the equation

$$\varepsilon\lambda^2 + K_2\lambda - \frac{1}{T} = 0,$$

hence by $\lambda^{\pm} = \frac{1}{2\varepsilon}[-K_2 \pm \sqrt{K_2^2 + \frac{4\varepsilon}{T}}]$. Taking the boundary conditions into account leads to the equation

$$\varphi(y) = \frac{\lambda^+ e^{\lambda^+ x + \lambda^- y} - \lambda^- e^{\lambda^- x + \lambda^+ y}}{\lambda^+ e^{\lambda^+ x + \lambda^- l} - \lambda^- e^{\lambda^- x + \lambda^+ l}}, \quad y \in [l, x].$$

Neglecting the second term in the denominator of the fraction we obtain

$$\varphi(x) \leq \frac{\sqrt{K_2^2 + 4\varepsilon/T}}{\varepsilon\lambda^+ e^{\lambda^-(l-x)}}.$$

Now for ε small enough, $\lambda^+ \geq \frac{1}{K_2 T}$ and $\lambda^- \leq \frac{-K_2}{\varepsilon}$. Therefore, for ε small enough there exists a constant $C_0 > 0$ independent of ε and T such that

$$\varphi(x) \leq C_0 K_2^2 \frac{T}{\varepsilon} \exp[\frac{K_2(l-x)}{\varepsilon}] = \frac{C_0 K_2^2}{\varepsilon} \exp[\frac{K_2(l-x)+\mu}{\varepsilon}].$$

This implies the desired inequality (20). □

We shall continue to use the partition $(I_j : 1 \leq j \leq n)$ of the interval $[0, a_\mu - h]$ of the preceding subsection. This time, we shall compare with homogeneous diffusions by freezing the potential derivative at an upper level, which results in working with the drifts V^{I_j}, and the diffusions $\underline{X}^{I_j}$. Again, these drifts may fail to possess the regularity properties required to apply Theorem 1.1. For this reason we may choose smoothed versions $V_j \in \mathcal{C}^1(\mathbb{R})$ of the potentials with corresponding diffusion processes Y^j satisfying

$$\begin{cases} V_j \geq V_{I_j} \geq V_j + \delta, \\ \text{there are } \underline{m}_j \in]-1-\delta, -1+\delta[, s_j \in]-\delta, \delta[, \overline{m}_j \in]1-\delta, 1+\delta[\\ \quad \text{such that } V_j|_{]-\infty, \underline{m}_j[\cup]s_j, \overline{m}_j[} < 0, V_j|_{]\underline{m}_j, s_j[\cup]\overline{m}_j, \infty[} > 0, \\ V_j'(\underline{m}_j) > 0, \\ \text{there are constants } K_1, K_2 > 0 \text{ such that } V_j(x) \leq -K_2 \text{ for } x \leq -K_1, \\ V_j(x) \geq K_2 \text{ for } x \geq K_1. \end{cases} \tag{23}$$

To deduce a lower estimate, we shall compare X via X^{I_n} with Y^n on the interval I_n in the scale T. We may write for $l < x$

$$\begin{aligned} \mathbb{P}_x(\tau_1(X) \leq (a_\mu - h)T) &\geq \mathbb{P}_x(\tau_1(X) \leq r_n T, \tau_1(X) \wedge \tau_l(X) \geq r_{n-1}T) \qquad (24) \\ &\geq \mathbb{E}_x(1_{\{\tau_1(X)\wedge\tau_l(X) \geq r_{n-1}T\}} P_{X_{r_{n-1}T}}(\tau_1(X_{\cdot+r_{n-1}T}) \leq \gamma T)) \\ &\geq \mathbb{P}_l(\tau_1(Y^n) \leq \gamma T) \times \mathbb{P}_x(\tau_1(X) \wedge \tau_l(X) \geq r_{n-1}T). \end{aligned}$$

As a consequence of $\mu - D_{-1}((a_\mu - h)T) < 0$ and the arguments presented in the preceding subsection, we note that uniformly on our compact set Γ

$$\lim_{\varepsilon \to 0} \mathbb{P}_x(\tau_1(X) \leq r_{n-1}T) = 0.$$

This clearly implies that there is $\varepsilon_0 > 0$ and a constant $C > \frac{1}{2}$ such that for $\varepsilon \leq \varepsilon_0$

$$\mathbb{P}_x(\tau_1(X) \geq r_{n-1}T) \geq C. \tag{25}$$

Moreover, by Lemma 2.3, for l small enough, there exists $\varepsilon_1 > 0$ such that for $\varepsilon \leq \varepsilon_1$

$$\mathbb{P}_x(\tau_1(X) \geq r_{n-1}T) \geq C.$$

Hence for ε small enough we have

$$\mathbb{P}_x(\tau_1(X) \wedge \tau_l(X) \geq r_{n-1}T) \geq C - \frac{1}{2} > 0.$$

It therefore remains to find lower bounds for $\mathbb{P}_l(\tau_1(Y^n) \leq \gamma T)$.

We may now apply the same arguments as those developed in the preceding subsection. We just have to use Lemma 2.2 for the eigenvalues of the operator

$$\mathcal{L}^\varepsilon u = \varepsilon u'' - V_n\, u'.$$

with Dirichlet boundary conditions at 0 in the sense of lower bounds uniformly on the compact Γ. As a consequence we obviously obtain, with a constant $C > 0$ independent of Γ

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_l(\tau_1(Y^n) \leq \gamma T) \geq (\mu - D_{-1}(a_\mu - h) - C\delta) \vee (-c),$$

uniformly for $\mu \in \Gamma$. Let us now choose $h_0 > 0$ and $\delta_0 > 0$ small enough such that for $h \leq h_0, \delta \leq \delta_0$ we have $\mu - D_{-1}((a_\mu - h)T) - C\delta \geq -c$. Since δ is arbitrary, we obtain

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_l(\tau_1(Y^n) \leq \gamma T) \geq (\mu - D_{-1}((a_\mu - h)T)),$$

uniformly for $\mu \in \Gamma$. Recalling (24) and (25), we finally obtain

$$\liminf_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_x(\tau_1(X) \leq (a_\mu - h)T) \geq \mu - D_{-1}((a_\mu - h)T),$$

uniformly on Γ. With this result we have established the desired lower bound for the exponential exit rate.

Proposition 2.2 *Let $x < 0$, Γ a compact subset of (10). Then there exists $0 < h_0 < \inf_{\mu \in \Gamma} a_\mu$ such that for $h \leq h_0$*

$$\lim_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_x(\tau_1(X) \leq (a_\mu - h)T) \geq \mu - D_{-1}(a_\mu - h) \tag{26}$$

uniformly for $\mu \in \Gamma$.

2.3 The exponential smallness of the rate of too long transitions

Having proved (11) in the preceding two propositions, the second aim of this section is to show that the exponential rate at which $\tau_1(X)$ exceeds $(a_\mu + h)T$ is arbitrarily small. In fact, we shall make precise that the rate at which transitions happen which take at least as long as $(a_\mu + h)T$ vanishes to all exponential orders, for $h > 0$ arbitrary.

Proposition 2.3 *Let $x < 0$, Γ a compact subset of (10) not containing $D_{-1}(0)$. Then there exists $h_0 > 0$ such that for all $0 < h \leq h_0$ and $\mu \in \Gamma$ we have*

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_x(\tau_1(X) \geq (a_\mu + h)T) = -\infty.$$

Proof:
Let $\delta > 0$ and $h > 0$ be given. Let Γ be a compact subset of (10). First, for $l < x$ we may write

$$\begin{aligned} \mathbb{P}_x(\tau_1(X) \geq (a_\mu + h)T) &\leq \mathbb{P}_x(\tau_1(X) \wedge \tau_l(X) \geq (a_\mu + h)T) && (27)\\ &\quad + \mathbb{P}_x(\tau_l(X) \leq (a_\mu + h)T). && (28)\end{aligned}$$

To estimate the second term on the right hand side of (27), we employ Lemma 2.3. In fact, for $l < x \wedge (-K_1)$ we have

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_x(\tau_l(X) \leq (a_\mu + h)T) \leq 2K_2(l - x \wedge (-K_1)) + \sup_{\mu \in \Gamma} \mu.$$

Therefore

$$\lim_{l \to -\infty} [\limsup_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_x(\tau_l(X) \leq (a_\mu + h)T)] = -\infty.$$

It therefore remains to estimate the first term on the right hand side of (27) for l small but fixed. Let $0 = r_0 < r_1 < \cdots < r_n = a_\mu + h$ be an equidistant partition of the interval $[0, a_\mu + h]$ of mesh $\gamma < \frac{h}{2}$ and denote $I_j = [r_{j-1}, r_j], 1 \leq j \leq n$. Then we have

$$\begin{aligned} &\mathbb{P}_x(\tau_1(X) \wedge \tau_l(X) \geq (a_\mu + h)T) && (29)\\ &= \mathbb{P}_x(\tau_0(X) \wedge \tau_l(X) \geq r_{n-1}T, \tau_1(X_{\cdot + r_{n-1}T}) \wedge \tau_l(X_{\cdot + r_{n-1}T}) \geq \gamma T)\\ &\leq \mathbb{E}_x(1_{\{\tau_0(X) \wedge \tau_l(X) \geq r_{n-1}T\}} \mathbb{P}_{X_{r_{n-1}T}}(\tau_1(\underline{X}^{I_n}) \geq \gamma T)\\ &\leq \max_{y \in]l,0[} \mathbb{P}_y(\tau_1(\underline{X}^{I_n}) \geq \gamma T)\\ &= \mathbb{P}_l(\tau_1(\underline{X}^{I_n}) \geq \gamma T)\\ &\leq \mathbb{P}_l(\tau_1(X^n) \geq \gamma T). \end{aligned}$$

Here, we compare the inhomogeneous diffusion X on I_n with the time homogeneous one X^{I_n} corresponding to drift R_{I_n} and finally with X^n subject to drift R_n to be described below, and we use monotonicity of

$$y \mapsto \mathbb{P}_y(\tau_1(\underline{X}^{I_n}) \geq \gamma T).$$

We assume γ to be small enough to ensure

$$\sup_{s,t \in [r_{n-1}, r_n]} \sup_{x \in [l,0]} \left| \frac{\partial U}{\partial x}(t,x) - \frac{\partial U}{\partial x}(s,x) \right| \leq \delta. \tag{30}$$

We may choose the drift R_n to satisfy

$$\begin{cases} R_n \leq R_{I_n} \leq R_n + \delta, \\ \text{there are } \underline{m}_n \in]-1-\delta, -1+\delta[, s_n \in]-\delta, \delta[, \overline{m}_n \in]1-\delta, 1+\delta[\\ \quad \text{such that } R_n|_{]-\infty, \underline{m}_n[\cup]s_n, \overline{m}_n[} < 0, R_n|_{]\underline{m}_n, s_n[\cup]\overline{m}_n, \infty[} > 0, \\ R_n'(\underline{m}_n) > 0, \\ \text{there are constants } K_1, K_2 > 0 \text{ such that } R_n(x) \leq -K_2 \text{ for } x \leq -K_1, \\ R_n(x) \geq K_2 \text{ for } x \geq K_1. \end{cases} \tag{31}$$

We are ready to apply Theorem 1.1, this time for $Q = R_n, d = -\delta$ to obtain that

$$\mathbb{P}_l(\tau_1(X^n) \geq \gamma T) \quad \leq \quad e^{-\lambda_n^\varepsilon \gamma T}(1 + e^{-c/\varepsilon}) \tag{32}$$

uniformly on Γ. Here λ_n^ε stands for the principal eigenvalue of the operator $\mathcal{L}_n^\varepsilon$ defined by

$$\mathcal{L}_n^\varepsilon u = \varepsilon u'' - R_n\, u'$$

with Dirichlet boundary conditions at $-\delta$. The asymptotic properties of λ_n^ε can be deduced in a similar way to Lemma 2.2. We estimate the pseudopotential corresponding to R_n taking (30) into account. We obtain that there exists $C > 0$ such that

$$|\lim_{\varepsilon \to 0} \varepsilon \lambda_n^\varepsilon T - (\mu - D_{-1}(r_n))| \leq C\delta, \tag{33}$$

uniformly on Γ. Now recall that due to the choice of γ, we have $a_\mu + \frac{h}{2} < r_n$, hence

$$\mu - D_{-1}(r_n) > 0, \quad \mu \in \Gamma,$$

and by compactness of Γ even

$$\inf_{\mu \in \Gamma} [\mu - D_{-1}(r_n)] > 0.$$

This in turn implies that

$$\lim_{\varepsilon \to 0} \lambda_n^\varepsilon\, T = \infty$$

uniformly on Γ. But due to (32) we are allowed to conclude

$$\lim_{\varepsilon \to 0} \mathbb{P}_l(\tau_1(X^n) \geq \gamma T) = -\infty,$$

uniformly in Γ. According to (29), this completes the proof.□

Finally, we may summarize the results of the preceding subsections, and state the main result on asymptotic exponential decay rates of transition probabilities.

Theorem 2.1 *Let $x < 0$, Γ be a compact subset of (10). For $\varepsilon > 0, \mu \geq 0$ let $T = \exp(\frac{\mu}{\varepsilon}$. Then there exists $h_0 > 0$ such that for $h \leq h_0$*

$$\lim_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}_x(\tau_1(X) \notin [(a_\mu - h)T, (a_\mu + h)T]) = \mu - D_{-1}(a_\mu - h) \tag{34}$$

uniformly for $\mu \in \Gamma$.

Proof:
This is an immediate consequence of Propositions 2.1, 2.2, and 2.3. □

3 Stochastic resonance in a double-well potential

Let us now turn to the main subject of this paper, a characterization of the notion of *stochastic resonance*. Let us recall that we look for a characterization of the concept of *optimal periodic tuning* which is extensively studied in the physics literature by notions such as the signal-to-noise ratio or the spectral power amplification (see Pavlyukevich [18]). Let us also remark that this concept implicitly uses and refines the concept of stochastic resonance studied by Freidlin [7] which paraphrases the ability of periodically perturbed stochastic systems to follow the periodic excitation in the small noise limit, and exhibit *quasi-periodic motion*. In more mathematical terms and the notation introduced before we aim at choosing the noise intensity parameter ε such that in the *large period limit* $T \to \infty$ the diffusion trajectories follow the periodic excitation of the system hidden in U in an optimal way to be made precise. In the first subsection we shall show that a quality measure of goodness of periodic tuning is given by the exponential rate at which the first transition to the other well happens within a fixed interval around $a_\mu T$. In the second subsection we establish robustness of this notion of quality: we show that in the small noise limit the diffusion and its reduced model, a Markov chain living on a two point state space, have the same resonance pattern.

3.1 Transition probabilities as a measure of quality

The local extrema of the depth functions $D_{\pm 1}$ of U are supposed to be global, and $D_{\pm 1}$ is strictly increasing between its extrema. Recall that we work with exponential time scales μ related to the natural time T by the equation $T = \exp(\frac{\mu}{\varepsilon})$. In this section, we have to work with scale functions depending on the starting well and eventually on arbitrary starting times. So we let

$$a^i_\mu(s) = \inf\{t \geq s : D_i(t) \leq \mu\}, \quad i = \pm 1, \quad \mu \geq 0.$$

The relevant time scales μ will be chosen from the intervals

$$I_i =]\inf_{t\geq 0} D_i(t), \sup_{t\geq 0} D_i(t)[, \quad i = \pm 1.$$

Our aim is to observe periodic behavior of the diffusion. This will in principle mean that the process can travel from one well to the other and back on the time scales in which we let the diffusion run, but not instantaneously. So, on the one hand, we have to work on time scales on which it *does not get stuck in one of the wells* of the potential. On the other hand, the time scales we are concentrating on should also *not allow for chaotic behavior*, i.e. immediate re-bouncing after changing the well.

To make these conditions mathematically precise, recall that transitions become possible as soon as the potential barrier $D_{\pm 1}$ becomes smaller than the time scale parameter μ. Hence if $\mu > \inf_{t\geq 0} D_i(t)$, there is a time range during which the diffusion can leave the well centered at i. Not to get stuck in one of them, the diffusion has to be able to leave both. This is guaranteed if

$$\mu > \max_{i=\pm 1} \inf_{t\geq 0} D_i(t). \tag{35}$$

To avoid immediate re-bouncing, we have to assure that the diffusion cannot leave the domain of attraction of $-i$ at the moment it reaches it, coming from i. Suppose we consider the dynamics after time $s \geq 0$, and the diffusion is near i at that time. Its first transition to the well at $-i$ occurs at time $a_\mu^i(s)T$, and it stays there for at least a little while if $D_{-i}(a_\mu^i(s))$ is bigger than μ. This is equivalent to stating that for all $s \geq 0$ there exists $\delta > 0$ such that on $[a_\mu^i(s), a_\mu^i(s) + \delta]$ we have $\mu < D_{-i}$. But for t shortly after $a_\mu^i(s)$, we always have $D_i(t) \leq \mu$ by the very definition of a_μ^i. Hence our condition becomes equivalent to the following: for all $s \geq 0$ there exists $\delta > 0$ such that on $[a_\mu^i(s), a_\mu^i(s) + \delta]$ we have $\mu < \max_{i=\pm 1} D_i$. This in turn is more elegantly expressed by

$$\mu < \inf_{t \geq 0} \max_{i=\pm 1} D_i(t). \tag{36}$$

We may summarize our search for an appropriate set of scale parameters μ for which periodicity in the diffusion behavior will occur. We call this set "resonance interval" to indicate that we have to look for the scale of optimal periodicity, the *resonance scale*, in this interval. See [11] for the definition of the corresponding interval in the case of two state Markov chains. The interval

$$I_R =]\max_{i=\pm 1} \inf_{t \geq 0} D_i(t), \inf_{t \geq 0} \max_{i=\pm 1} D_i(t)[$$

is called *resonance interval* (see Figure 3). Let us pause for a moment at this point to compare our approach with Freidlin's [7] understanding of stochastic resonance by quasi-deterministic motion. In Freidlin's terms, stochastic resonance is given if the parameter μ exceeds the lower boundary of our resonance interval. Our concept of resonance stipulates to look for an optimal μ in the resonance interval at which in a sense to be made precise the quality of periodic tuning is optimal.

Let us now come to the discussion of the quality of periodic response of the stochastic system given by the diffusion, in dependence of the noise parameter ε and the time scale parameter μ which according to the remarks made above has to be chosen in the resonance interval. To simplify things a little, let us assume that the depth functions are related by a phase $\phi \in]0, 1[$, i.e.

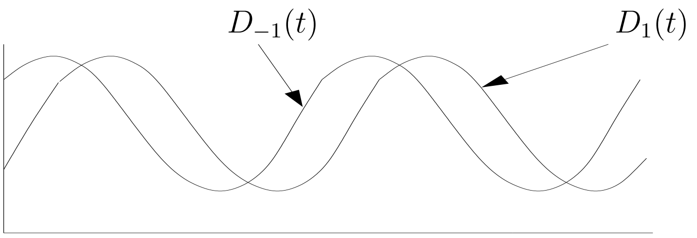

Figure 8: Depth functions in phase

$$D_{-1}(t) = D_1(t + \phi), \quad t \geq 0.$$

Moreover we assume that the diffusion starts in -1. There are many ways to describe optimality of periodic tuning. Pavlyukevich and Imkeller ([12], [13]) consider different measures of quality such as the *spectral power amplification*, the *energy*, the *energy-to-noise ratio*, the *out-of-phase measure*, the *entropy* and the *relative entropy*. The detailed study of the physicists' favorite measure, the spectral power amplification, based on the energy of the spectral component of the mean trajectory in equilibrium corresponding to the forcing frequency $\frac{2\pi}{T}$, shows one surprising defect: it is not robust as one passes from the diffusion to a reduced model described by a two state Markov chain jumping with rates corresponding to the transition rates between the meta-stable states ± 1 of the diffusion given by the potential minima. In fact, while the Markov chain's spectral power coefficient shows a pronounced peak for μ near an average well depth, the overwhelming influence of the diffusion's fluctuations in small neighborhoods of the potential wells, discovering very subtle details of the potential's geometry there, destroys this picture completely. Here we propose a notion of quality of periodic tuning which is based on the pure transition mechanism of the system between the domains of attraction of the double well potential. Generalizing the approach of a study of optimal tuning for two state Markov chain models (see [11]), we measure the quality of tuning by computing for varying time scale parameters μ the probability that, starting in i, the diffusion is transferred to $-i$ within the time window $[(a^i_\mu - h)T, (a^i_\mu + h)T]$ of width $2hT$. To find the *stochastic resonance point* for large T we have to maximize this measure of quality in $\mu \in I_R$. The probability for transition within this window will be computed by the estimates of the preceding section. Uniformity of convergence to the exponential rates will enable us to maximize in μ. Let us now make these ideas precise.

To make sure that the transition window makes sense at least for small h, we have to suppose that $a^i_\mu > 0, i = \pm 1$ for $\mu \in I_R$. This will be guaranteed if

$$D_i(0) > \inf_{t\geq 0} \max_{i=\pm 1} D_i(t), i = \pm 1. \tag{37}$$

If this is not granted from the beginning, it suffices to start the diffusion a little later, in order to be sure that (37) is satisfied. Under (37), we call

$$M(\varepsilon, \mu) = \min_{i=\pm 1} \mathbb{P}_i(\tau_{-i}(X) \in [(a^i_\mu - h)T, (a^i_\mu + h)T]), \quad \varepsilon > 0, \mu \in I_R, \tag{38}$$

transition probability for a time window of width h.

We are prepared to state our main resonance result.

Theorem 3.1 *Let Γ be a compact subset of I_R, $h_0 > 0$ be given according to Theorem 2.1. Then*

$$\lim_{\varepsilon \to 0} \varepsilon \ln(1 - M(\varepsilon, \mu)) = \max_{i=\pm 1} \left\{ \mu - D_i(a^i_\mu - h) \right\} \tag{39}$$

uniformly for $\mu \in \Gamma$.

Proof:
This proposition is an obvious consequence of Propositions 2.1, 2.2 and 2.3.□

It is clear that for h small the eventually existing global minimizer $\mu_R(h)$ of

$$I_R \ni \mu \mapsto \max_{i=\pm 1} \left\{ \mu - D_i(a^i_\mu - h) \right\}$$

is a good candidate for our resonance point. But it still depends on h. To get rid of this dependence, we shall consider the limit of $\mu_R(h)$ as $h \to 0$.

Definition 3.1 *Suppose that*

$$I_R \ni \mu \mapsto \max_{i=\pm 1} \left\{ \mu - D_i(a_\mu^i - h) \right\}$$

possesses a global minimum $\mu_R(h)$. Suppose further that

$$\mu_R = \lim_{h \to 0} \mu_R(h)$$

exists in I_R. We call μ_R the stochastic resonance point *of the diffusion X with time periodic potential U.*

We shall now show that the stochastic resonance point exists if one of the depth functions, and thus both, due to the phase lag, has a unique point of maximal decrease on the interval where it is strictly decreasing.

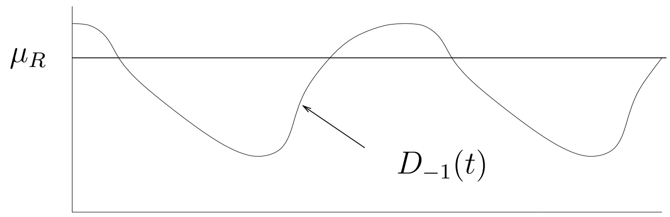

Figure 9: Point of maximal decrease

Theorem 3.2 *Suppose that D_1 is twice continuously differentiable and has its global maximum at t_1, and its global minimum at t_2, where $t_1 < t_2$. Suppose further that there is a unique point $t_1 < s < t_2$ such that $D_1|_{]t_1,s[}$ is strictly concave, and $D_1|_{]s,t_2[}$ is strictly convex. Then $\mu_R = D_1(s)$ is the stochastic resonance point.*

Proof:
First of all, note that there is $\psi \in]0,1[$ such that $D_1 = D_{-1}(\cdot + \psi)$. As a consequence of this,

$$\max_{i=\pm 1} \left\{ \mu - D_i(a_\mu^i - h) \right\} = \left\{ \mu - D_1(a_\mu^1 - h) \right\}.$$

Write $a_\mu = a_\mu^1$ and recall that on the interval of decrease of D_1, $a_\mu = D_1^{-1}(\mu)$. Therefore, the differentiability assumption yields

$$1 = D_1'(a_\mu - h) \cdot a_\mu' = D_1'(a_\mu - h) \cdot \frac{1}{D_1'(a_\mu)}.$$

Our hypothesis concerning convexity and concavity of D_1 essentially means that $D_1''(s) = 0$, and $D_1''|_{]t_1,s[} < 0, D_1''|_{]s,t_2[} > 0$, in other words that $\mu \mapsto D_1'(a_\mu)$ has a local maximum at $a_\mu = s$. Hence for h small there exists a unique point $a_\mu(h)$ such that

$$D_1'(a_\mu(h) - h) = D_1'(a_\mu(h))$$

and

$$\lim_{h\to 0} a_\mu(h) = s.$$

To show that $a_\mu(h)$ corresponds to a minimum of the function

$$\mu \mapsto [\mu - D_1(a_\mu - h)],$$

we take the second derivative of this function at $a_\mu(h)$, which is given by

$$\frac{D_1'(a_\mu(h)-h)D_1''(a_\mu(h)) - D_1''(a_\mu(h)-h)D_1'(a_\mu(h))}{D_1'(a_\mu(h))}.$$

But $D_1'(a_\mu(h)), D_1'(a_\mu(h)-h) < 0$, whereas $D_1''(a_\mu(h)-h) > 0, D_1''(a_\mu(h)) < 0$. This clearly implies that $a_\mu(h)$ corresponds to a minimum of the function. But by definition, as $h \to 0$, $a_\mu(h) \to s$. Therefore, finally, $D_1(s)$ is the stochastic resonance point.$\square$

To illustrate our results, we next discuss an example.
Example: Let us consider the double-well potential

$$U(t,x) = \frac{x^6}{6} - \cos\left\{2\pi\left(t - \frac{1}{4} + \psi\ \mathrm{sgn}(x)\right)\right\}\left(\frac{x^5}{5} - \frac{x^3}{3}\right) - \frac{x^2}{2},$$

with $T = \exp(\frac{\mu}{\varepsilon})$ and $\psi \in [0, \frac{1}{4}[$.

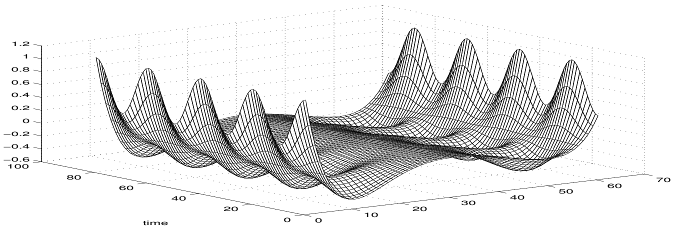

Figure 10: Double-well potential (case: $\psi = 0$)

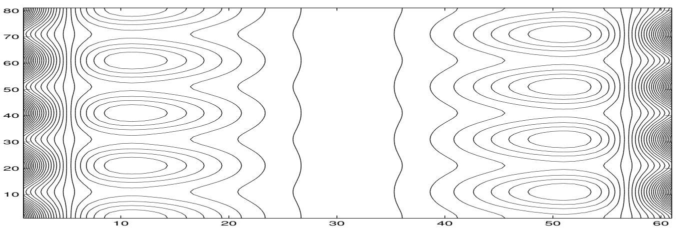

Figure 11: Level sets of the potential

U satisfies all the assumptions required for potentials above, in particular

$$\frac{\partial U}{\partial x}(t,x) = 0 \text{ iff } x \in \{-1, 0, 1\}.$$

-1 and 1 are the meta-stable states of the potential, and 0 the saddle point. We can then compute the barrier height of the two wells. For $i \in \{-1, 1\}$,

$$D_i(t) = \frac{2}{3} - i\frac{4}{15}\cos\left(2\pi t + i2\pi\psi - \frac{\pi}{2}\right).$$

Let us note that $D_1(t) = D_{-1}(t + 2\psi + \frac{1}{2})$. Since $\psi \in [0, \frac{1}{4}[$, we are in the phase case with $\phi = 2\psi + 1/2$. The resonance interval is then given by

$$I_R = \left]\frac{2}{5}, \frac{2}{3} - \frac{4}{15}\cos\left(\frac{\pi}{2} - 2\pi\psi\right)\right[.$$

In the symmetric case, that is if $\psi = 0$ we obtain $I_R =]2/5, 2/3[$. Let us now compute the optimal tuning scale applying Theorem 2.1. We obtain

$$a_\mu^{-1} = \frac{1}{2\pi}\arccos\left(\frac{15}{2}\left(\frac{\mu}{2} - \frac{1}{3}\right)\right) + \psi + \frac{1}{4}.$$

Hence, for $h > 0$ small enough,

$$\begin{aligned} F(\mu) &= \mu - D_{-1}(a_\mu^{-1} - h) \\ &= \mu - \frac{2}{3} - \frac{4}{15}\cos\left(\arccos\left\{\frac{15}{2}\left(\frac{\mu}{2} - \frac{1}{3}\right)\right\} - 2\pi h\right) \\ &= \left(\mu - \frac{2}{3}\right)(1 - \cos 2\pi h) - \frac{4}{15}\sin(2\pi h)\sqrt{1 - \left(\frac{15}{2}\left(\frac{\mu}{2} - \frac{1}{3}\right)\right)^2}. \end{aligned}$$

Let us recall that F does not depend on the phase which implies

$$\mu - D_{-1}((a_\mu^{-1} - h)T) = \mu - D_1((a_\mu^{1} - h)T).$$

Hence, to obtain optimal tuning, it suffices to compute the minimum of F for $\mu \in I_R$. Differentiating F we obtain

$$F'(\mu) = 1 - \cos(2\pi h) + \frac{15}{2}\sin(2\pi h)\frac{\frac{\mu}{2} - \frac{1}{3}}{\sqrt{1 - \left(\frac{15\mu}{4} - \frac{15}{6}\right)^2}}.$$

Hence F attains its minimum for

$$\mu_R(h) = \frac{2}{3} - \frac{2\sqrt{2}}{15}\sqrt{1 - \cos(2\pi h)}$$

and

$$\mu_R = \lim_{h \to 0}\mu_R(h) = \frac{2}{3}.$$

Thus we obtain that μ_R is the stochastic resonance point if $\mu_R \in I_R$, that is, if the phase is near to $\frac{1}{2}$, i.e. if ψ is close to 0. In the other case, the optimal tuning rate on every interval $[a, b] \subset I_R$ is given by the upper bound b.$\square$

3.2 The robustness of stochastic resonance based on transition windows

In the small noise limit $\varepsilon \to 0$, it seems reasonable to assume that the periodicity properties of the diffusion trajectories caused by the periodic forcing due to the potential, are essentially captured by a simpler, reduced stochastic process: a continuous time Markov chain which just jumps between two states representing the bottoms of the wells of the double well potential at rates corresponding to the transition mechanism of the diffusion. This is just the reduction idea ubiquitous in the physics literature, and explained for example in McNamara, Wiesenfeld [15]. In [13] it is found that this idea may conflict with the intra well fluctuations of the diffusion if the quality of periodic tuning is measured by concepts using spectral decompositions of the trajectories. We shall now show that in the small noise limit both models, diffusion and Markov chain, produce the same resonance picture, if quality of periodic tuning is measured by transition rates as discussed in the previous subsection.

We first have to describe the reduced model. Let U be a time dependent potential function generating the potential diffusions of the preceding section. Recall that the depth functions of the potential minima satisfy $D_1(t) = D_{-1}(t+\phi), t \geq 0$, with phase shift $\phi \in]0,1[$. So, let us consider a time-continuous Markov chain $\{Y_t,\, t \geq 0\}$ taking values in the state space $\{-1, 1\}$ with initial data $Y_0 = -1$. Suppose the infinitesimal generator is given by

$$G = \begin{pmatrix} -\varphi(\frac{t}{T}) & \varphi(\frac{t}{T}) \\ \psi(\frac{t}{T}) & -\psi(\frac{t}{T}) \end{pmatrix},$$

where $\psi(t) = \varphi(t+\phi), t \geq 0$, and φ is a 1-periodic function describing a rate which just produces the transition dynamics of the diffusion between the potential minima ± 1, i.e.

$$\varphi(t) = \exp(-\frac{D_{-1}(t)}{\varepsilon}), \quad t \geq 0. \tag{40}$$

Note that by choice of ϕ,

$$\psi(t) = \exp(-\frac{D_1(t)}{\varepsilon}), \quad t \geq 0. \tag{41}$$

Transition probabilities for the Markov chain thus defined are easily computed. See [11], section 2. For example, the probability density of the first transition time $\sigma_i(Y)$ is given by

$$\begin{aligned} p(t) &= \varphi(t)\exp(-\int_0^t \varphi(s)ds), \quad \text{if} \quad i=-1, \\ q(t) &= \varphi(t+\phi)\exp(-\int_0^t \varphi(s+\phi)ds), \quad \text{if} \quad i=1, \end{aligned} \tag{42}$$

$t \geq 0$. Equation (42) can be used to obtain results on exponential rates of the transition times $\sigma_i(Y)$ if starting from $-i$, $i = \pm 1$. We summarize them and apply them to the measure of quality of periodic tuning in case (37)

$$N(\varepsilon, \mu) = \min_{i=\pm 1} \mathbb{P}_i(\sigma_{-i}(Y) \in [(a^i_\mu - h)T, (a^i_\mu + h)T]), \quad \varepsilon > 0, \mu \in I_R, \tag{43}$$

which is called *transition probability for a time window of width* h for the Markov chain.

Here is the asymptotic result obtained from a slight modification of Theorems 3 and 4 of [11] which consists of allowing more general depth functions than the sinusoidal ones used there and requires just the same proof.

Theorem 3.3 *Let* Γ *be a compact subset of* I_R*,* $h_0 < \sup(a_\mu^{-1}, T/2 - a_\mu^{-1})$*. Then for* $0 < h \leq h_0$

$$\lim_{\varepsilon \to 0} \varepsilon \ln(1 - N(\varepsilon, \mu)) = \max_{i=\pm 1} \left\{ \mu - D_i(a_\mu^i - h) \right\} \tag{44}$$

uniformly for $\mu \in \Gamma$.

It is clear from Theorem 3.3 that the reduced Markov chain Y and the diffusion process X have exactly the same resonance behavior. Of course, we may define the *stochastic resonance point* for Y just as we did for X. So the following final robustness result holds true.

Theorem 3.4 *The resonance points of* X *with periodic potential* U *and of* Y *with exponential transition rate functions* $D_{\pm 1}$ *coincide.*

References

[1] R. Benzi, G. Parisi, A. Sutera, and A. Vulpiani. A theory of stochastic resonance in climatic change. *SIAM J. Appl. Math.*, 43:565–578, 1983.

[2] N. Berglund and B. Gentz. Metastability in simple climate models: pathwise analysis of slowly driven Langevin equations. *Stochastics and Dynamics*, 2(3):327–356, 2002.

[3] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes i. sharp asymptotics for capacities and exit times. *submitted*, 2002.

[4] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes ii: precise asymptotics for small eigenvalues. *submitted*, 2002.

[5] M. V. Day. On the exponential exit law in the small parameter exit problem. *Stochastics*, 8(4):297–323, 1983.

[6] M. V. Day. Recent progress on the small parameter exit problem. *Stochastics*, 20(2):121–150, 1987.

[7] M. Freidlin. Quasi-deterministic approximation, metastability and stochastic resonance. *Physica D*, 137:333–352, 2000.

[8] M. Freidlin and Wentzell A. *Random perturbations of dynamical systems*. Springer-Verlag, New York, second edition, 1998.

[9] J. A. Freund, A. Neiman, and L. Schimansky-Geier. Stochastic resonance and noise-induced phase coherence. In Imkeller, P., J.-S. von Storch, editor, *Stochastic Climate Models*. Birkhäuser, 2001.

[10] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Reviews of Modern Physics*, 70:223–287, 1998.

[11] S. Herrmann and P. Imkeller. Barrier crossings characterize stochastic resonance. *Stochastics and Dynamics*, 2(3):413–436, 2002.

[12] P. Imkeller and I. Pavlyukevich. Stochastic resonance in two-state Markov chains. *Arch. Math. (Basel)*, 77(1):107–115, 2001.

[13] P. Imkeller and I. Pavlyukevich. Model reduction and stochastic resonance. *Stochastics and Dynamics*, 2(4):463–506, 2002.

[14] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica VII*, 4:284–304, 1940.

[15] B. McNamara and K. Wiesenfeld. Theory of stochastic resonance. *Physical Review A*, 39:4854–4869, 1989.

[16] G.N. Milstein and M.V. Tretyakov. Numerical analysis of noise-induced regular oscillations. *Physica D*, 140(3-4):244–256, 2000.

[17] C. Nicolis. Stochastic aspects of climatic transitions – responses to periodic forcing. *Tellus*, 34:1–9, 1982.

[18] I. Pavlyukevich. *Stochastic resonance.* PhD thesis, Humboldt-Universität zu Berlin, 2002.

[19] D. Revuz and M. Yor. *Continuous martingales and Brownian motion. 3rd ed.* Springer, Berlin, 1999.