Barrier crossings characterize stochastic resonance

Samuel Herrmann
Technische Universität Berlin,
MA 7-5, Str. des 17. Juni 135, 10623 Berlin, Germany
herrmann@math.tu-berlin.de
Peter Imkeller
Institut für Mathematik, Humboldt-Universität zu Berlin,
Unter den Linden 6, 10099 Berlin, Germany
imkeller@mathematik.hu-berlin.de

ABSTRACT. - In a two-state Markov chain with time periodic dynamics, we study path properties such as the sojourn time in one state between two consecutive jumps or the distribution of the first jump. This is done in order to exhibit a resonance interval and an optimal tuning rate interpreting the phenomenon of stochastic resonance through quality notions related with interspike intervals. We consider two cases representing the reduced dynamics of particles diffusing in time periodic potentials: Markov chains with piecewise constant periodic infinitesimal generators and Markov chains with time-continuous periodic generators.

The physical concept of stochastic resonance was discovered around 1980 in the study of a very simple stochastic climate model: an *energy balance model* designed to give a qualitative interpretation of paleoclimatic data describing glacial cycles in the terrestrial history. While its physical understanding - based on transition laws observed very early by Eyring [5] and Kramers [13] which interpreted the prefactor in Arrhenius' law [1] - was deepened in numerous papers finding it in a big variety of areas of natural sciences (see Maier, Stein [14] and Gammaitoni et al. [8] for a review), its rigorous mathematical background is being clarified only recently. In mathematical terms, *stochastic resonance* describes the *optimal* response or tuning with respect to the noise intensity parameter of a noisy system the deterministic dynamics of which is characterized by the presence of finitely many *metastable states* between which it is bound to move through a periodic forcing of large period T. One of the simplest mathematical models of this type is encountered in the following stochastic differential equation

$$dX_t = -\left(U'(X_t) + Q\sin\frac{2\pi t}{T}\right)dt + \sqrt{\varepsilon}dB_t. \tag{1}$$

Here U is a double well potential with wells of equal depth, Q and ε are some small positive parameters, T is a large one and B is an one-dimensional Brownian motion. It describes the motion of a physical quantity X_t in a periodically changing potential landscape

$$(x,t) \mapsto U(x) + Q \sin \frac{2\pi t}{T}$$

in which Q mainly contains the amplitude of very slow periodic variation of period T of the potential wells, perturbed by a white noise $\dot{B}$ of intensity ε. In this simple framework, the problem of stochastic resonance can be paraphrased into the following terms. Fix a small periodic modulation amplitude Q, and a large period T. Can one then choose the noise intensity parameter $\varepsilon = \varepsilon(T)$ in such a way that the trajectories of the diffusion X in (1) look "as periodic as possible"?
Freidlin [7], in a more general framework, gives a first partial answer to this question. The mathematical underpinning of the Eyring-Kramers law he finds using large deviations theory (see [6]) leads him to the following conclusion: if ε is above a critical threshold determined by the typical depth a of the shallower potential well, more precisely if ε is such that $T \leq \exp(\frac{a}{\varepsilon})$, the diffusion is able to show quasi-deterministic behavior - on an exponential scale. More precisely, under this condition the amount of times t so that the diffusion at time tT is not in a small neighborhood of the discontinuous deterministic curve describing just the position of the bottom of the deeper well is negligible in the *large period limit* $T \to \infty$. In other words: Freidlin determines the lower bound of the interval $[\frac{a}{\ln T}, \infty[$ in which an "optimal tuning intensity" $\varepsilon = \varepsilon(T)$ has to be found. If, for a fixed intensity ε we paraphrase this into a statement about finding an optimal period length $T = T(\epsilon)$ above the critical threshold $\exp(\frac{a}{\varepsilon})$, it is clear that we are discussing times beyond the typical large deviations scale. The extensive pathwise investigations of periodicity and transition phenomena in Berglund, Gentz [2], [3], and the forerunner paper Eckmann, Thomas [4] for two-state Markov chains, confined to smaller time scales, therefore cannot tackle the problem of finding an optimal tuning or *resonance point*. This is done in Imkeller, Pavljukevich [10] for time discrete two-state Markov chains, and in Pavljukevich [17], and Imkeller, Pavljukevich [11] for diffusions flipping periodically between two spatially antisymmetric states of a potential with wells of unequal depth.

Of course, for the purpose of optimizing periodic tuning of diffusion trajectories, relevant quality measures have to be defined. In Pavljukevich [17], as well in the framework of diffusions in periodically changing double-well potentials as for dynamically adapted two-state Markov chains jumping in continuous time between the metastable states ± 1 of the diffusions and thus describing the reduced dynamics of the diffusions, the physicists' favorite quality measure, *spectral power amplification* (SPA), has been investigated thoroughly. SPA measures the energy of the spectral component of the averaged trajectory corresponding to the periodic forcing frequency $\frac{1}{T}$ of the potential, where the average is taken with respect to the equilibrium measure of the process. For the reduced dynamics models, Pavljukevich [17] investigates a number of different quality measures of periodic tuning including entropy notions measuring chaoticity of the averaged trajectories instead. The main result is that optimal tuning with respect to SPA in the Markov chain framework happens for $\varepsilon(T) \sim \frac{1}{\ln T}\frac{V+v}{2}$, if the different depths of the potential wells are given by $\frac{V}{2}$ and $\frac{v}{2}$. This is seen to extend to the diffusion only if small fluctuations near the potential valley bottoms are suppressed. In particular the notion of SPA to measure stochastic resonance is not robust for transitions between the diffusion model and its reduced dynamics Markov chain. This also implies that physical reduced model studies such as Mc Namara, Wiesenfeld [15] only appear to give the right picture if notions for measuring periodic tuning are used which are more robust when passing between diffusion and reduced chain.

This observation is one of the main motivations for the present study. As it happens, also applications in neurophysiological models (see [8]) and in particular rather recent applications in simple reduced climate systems strongly suggest another type of measure of quality of tuning which appears more robust. Rather than taking into account all small fluctuations during the long time periods the process spends near the valley bottoms it is based more directly on transition times between the metastable states of the system given by the valley bottoms. It measures the statistics of the *interspike intervals*, i. e. the interval distribution between consecutive barrier crossings. A particularly recent appearance of these ideas is related to the analysis of the Greenland ice core record. The statistical properties of spontaneous intermediate warmings which are commonly known as Dansgaard-Oeschger events, were found to be consistent with stochastic resonance phenomena. It is observed that besides the metastable ice and warm age temperature states with transition times around multiples of $10^4 - 10^5$ years there is another metastable state at an intermediate temperature accessible from the glacial state. Transition intervals cluster around integer multiples of 1500 years. Ganopolski and Rahmstorf (see [9]) reproduce these observations by a simulation based on the CLIMBER coupled ocean-atmosphere model of moderate complexity established by the Potsdam group. A stability analysis shows the existence of the intermediate metastable state, and suitable small periodic and random excitations of the salinity balance of the North Atlantic as one of the system variables produce temperature curves with abrupt transitions of the observed type. The empirical distribution of the interspike intervals is seen to be a function of the noise amplitude and other system parameters.

In this paper we shall tackle a mathematical underpinning of the investigation of spontaneous transitions and stochastic resonance by means of the notion of interspike intervals. This will be done in the simple framework of two-state continuous time Markov chains. One should always think of these objects as representing the effective reduction of more complex diffusion models, a paradigm which will shed light on still more complex models such as the ones resulting from the CLIMBER example. So, in the case of a time periodic potential function jumping after every half period between the two antisymmetric potential states with wells of depth $\frac{V}{2}$ at -1 and depth $\frac{v}{2}$ at 1, and $v < V$, the transition rates in the $Q-$matrix defining the infinitesimal generator will also be piecewise constant on half period intervals and will be given by the following: to exponential order, the one from -1 to 1 will be determined as $\phi = \exp(-\frac{V}{2})$ and from 1 to -1 by $\exp(-\frac{v}{2})$. In particular, we shall investigate a notion of optimal tuning (resonance point) based on the *intensity of the first peak*. More precisely, the probability that the first transition, in scale T, happens in a fixed neighborhood of 1, will be maximized in ε using uniform large deviation estimates.

The paper is organized as follows. In section 1 we will study the law of the jump times for the case of a piecewise constant time periodic infinitesimal generator. In section 2, the infinitesimal generator matrix is allowed to have a continuous time periodic variation, thus corresponding to a diffusion of a physical quantity in a one-dimensional potential landscape varying periodically continuously in time. In both cases we characterize the distribution of the first jump time after fixed times in scale T for large periods T in terms of geometric laws, and show that optimal tuning rates for the first peak in this distribution correspond to the tuning rates found in [10]. The generalization of our asymptotic results to diffusion models is left for forthcoming research.

1 Two-state Markov chain with piecewise constant infinitesimal generator

1.1 Introduction

In this first section, we describe some properties of the stochastic resonance model introduced by Pavlyukevich [17]. Let us consider a time-continuous Markov chain $\{Y_t,\ t \geq 0\}$ on the state space $\mathcal{S}^Y = \{-1, 1\}$ ($Y_0 = -1$), where the infinitesimal generator is temporally periodic with period $2T$:

$$Q_1 = \begin{pmatrix} -\varphi & \varphi \\ \psi & -\psi \end{pmatrix} \quad \text{for} \quad 0 \leq t < T,$$

and

$$Q_2 = \begin{pmatrix} -\psi & \psi \\ \varphi & -\varphi \end{pmatrix} \quad \text{for} \quad T \leq t < 2T,$$

periodically continued on $\mathbb{R}_+$.
Here $\varphi = pe^{-V/\varepsilon}$, $\psi = qe^{-v/\varepsilon}$ with $p, q > 0$ and $0 < v < V < +\infty$.
In order to understand the parameters appearing in the infinitesimal generator of the Markov chain, let us recall that, in the small noise limit $\epsilon \to 0$, for a time homogeneous diffusion in a double well potential, the mean transition time between the wells is given by Kramers' law. If the diffusion starts in the minimum of one well, the mean exit time is equivalent to $\exp -\frac{2\Delta U}{\epsilon}$, where ΔU is the height of the barrier between the respective minimum and the saddle separating the wells of the potential. Hence $\frac{V}{2}$ is the barrier height if the diffusion starts in the deep well, and $\frac{v}{2}$ corresponds to the height of the shallow one (see Figure 1).

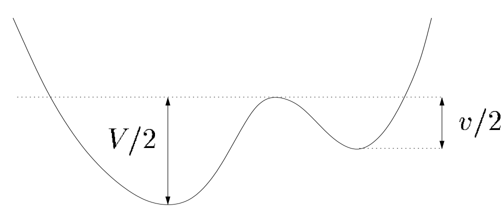

Figure 1: Double-well potential

The states ± 1 of the chain represent the two metastable states of the system, which are given by the positions of the minima of the potential. The remaining two parameters p and q play the roles of prefactors of smaller than exponential order appearing in asymptotic expansions of the laws of transition times or invariant densities. They are beyond the scope of large deviation theory. They are well known to be related to the geometry (curvature) of the potential in the minima and the saddle point of the potential landscape (see Maier, Stein [14]) and given by

$$p = \frac{1}{2\pi}\sqrt{U''(-1)|U''(0)|}, \quad q = \frac{1}{2\pi}\sqrt{U''(1)|U''(0)|}.$$

In what follows we shall let p and q be arbitrary positive numbers. The depencence of the main results on these parameters will be explicitly exhibited.

Let us denote by T_n the time of the nth jump from one state to the other (if n is odd, the Markov chain jumps from -1 to $+1$ and the opposite happens if n is even). Then we define the normalized time of jump:

$$\tau_n := \frac{T_n}{T}, \quad \text{for} \quad n \in \mathbb{N}. \tag{2}$$

Due to the periodic changes of the potential (matrix of infinitesimal probabilities) the law of the nth jump depends strongly on the position of the Markov chain and on the time of the last jump.

We define $\lfloor u \rfloor$ to be the largest integer less than or equal to the real number u. For the intuitive understanding of the formulas in the following statement, let us remark that for $n \in \mathbb{N}$, $u \geq 0$ the formal condition $\tau_{n-1} = u$ and $(-1)^{\lfloor u \rfloor + n} = +1$ just means that the number n of the following jump and the half period interval in which it occurs have "equal parity", i.e. either the jump goes to 1 and the corresponding position of the potential has the shallow well at 1 or the same statement with -1 replacing 1 holds. The condition $\tau_{n-1} = u$ and $(-1)^{\lfloor u \rfloor + n} = -1$ analogously expresses "unequal parity". The conditional law of τ_n is then given by the following

Lemma 1 *Let $u \geq 0$.*

• *Given $\tau_{n-1} = u$ and $(-1)^{\lfloor u \rfloor + n} = +1$, the conditional density $p_+(t)$ of the law of τ_n is equal to*

$$\begin{aligned} p_+(t) &= \xi_1 T e^{-\xi_1 (t-u) T} \mathbb{1}_{\{u \leq t < \lfloor u \rfloor + 1\}}(t) \\ &+ e^{-\xi_1 (1 + \lfloor u \rfloor - u) T} \sum_{k \geq 0} \xi_k T \exp - \left(\sum_{j=1}^{k} \xi_{j-1} T + \xi_k T (t - \lfloor t \rfloor) \right) \mathbb{1}_{\Gamma_k}(t) \end{aligned} \tag{3}$$

with $\Gamma_k = \{t \in \mathbb{R}_+ : \ \lfloor u \rfloor + k + 1 \leq t < \lfloor u \rfloor + k + 2\}$

and $\xi_k = \begin{cases} \varphi & \text{if } k \text{ is even,} \\ \psi & \text{otherwise.} \end{cases}$

• *Given $\tau_{n-1} = u$ and $(-1)^{\lfloor u \rfloor + n} = -1$, the conditional density $p_-(t)$ of the law of τ_n is equal to*

$$\begin{aligned} p_-(t) &= \xi_0 T e^{-\xi_0 (t-u) T} \mathbb{1}_{\{u \le t < \lfloor u \rfloor + 1\}}(t) \\ &+ e^{-\xi_0 (1 + \lfloor u \rfloor - u) T} \sum_{k \ge 0} \xi_{k+1} T \exp - \left(\sum_{j=1}^{k} \xi_j T + \xi_{k+1} T (t - \lfloor t \rfloor) \right) \mathbb{1}_{\Gamma_k}(t). \end{aligned} \tag{4}$$

Proof: Straightforward, since, for a constant matrix of infinitesimal probabilities, the first time the Markov chain, starting in one state, reaches the other one, is exponentially distributed. QED

Remark: *In fact, the conditional law depends only on u (time of the last jump) and on the expression $(-1)^{\lfloor u \rfloor + n}$ which gives the following information: the process is in the deep well at time u or not. Furthermore, since Y is a Markovian process, the density of the time of the first jump after a given normalized time $u \ge 0$ is equal to p_+ if $Y_{uT} = +1$ and p_- otherwise.*

1.2 Asymptotic behaviour

Obviously the conditional densities p_+ and p_- depend on ε. Let us therefore study more closely the behaviour of the jumps probabilitites if ε changes. Since the period is large and the variance ε small, we assume that $T = T(\varepsilon)$ eventually depends on ε to get only one scale.

The following result states that the asymptotic behaviour of the Markov chain depends on the ratio between T and ξ_i, $i \in \{0, 1\}$. Recall that ξ_i is the transition rate from -1 to 1 if $i = 0$ and from 1 to -1 if $i = 1$. In fact, we shall see that asymptotic properties of the products of $\xi_i T$ will determine the asymptotic laws for transitions in the small noise limit $\varepsilon \to 0$. For instance, the condition $\lim_{\varepsilon \to 0} \xi_i T = 0$ requires the jump rate to go to 0 exponentially, while T may or may not depend on ε. It may for example be a fixed constant of order 1. In contrast, the condition $\lim_{\varepsilon \to 0} \xi_0 T = 0, \lim_{\varepsilon \to 0} \xi_1 T = c$ for some $c > 0$ requires that $T = T(\varepsilon)$ be asymptotically equal to $cq \exp(\frac{v}{\varepsilon})$ and actually implies the weaker first statement since the exponential decay of $\exp(-\frac{V}{\varepsilon})$ is stronger due to $V > v$.

Top state our main result on the asymptotic laws of jumps let us define S_u to be the normalized time of the first jump of the process Y after time uT.

Theorem 1 *Let $n \ge 2$. Then, as $\varepsilon \to 0$,*

• *if $\lim\limits_{\varepsilon \to 0} \xi_i T = +\infty$ for $i \in \{0, 1\}$, the law of $S_u - u$ converges to the Dirac measure δ_0.*

• *if $\lim\limits_{\varepsilon \to 0} \xi_i T = 0$ for $i \in \{0, 1\}$, the measure of probability of S_u tends weakly* *to the null measure.*

• *if* $\lim_{\varepsilon\to 0}\xi_0 T = 0$ *and* $\lim_{\varepsilon\to 0}\xi_1 T = +\infty$, *the law of* $S_u - u$ *tends to* δ_0, *if* $(-1)^{\lfloor u\rfloor} Y_{uT} = +1$, *and to* $\delta_{\lfloor \tau_{n-1}\rfloor+1}$ *otherwise.*

• *if for some* $c > 0$ $\lim_{\varepsilon\to 0}\xi_0 T = c$ *and* $\lim_{\varepsilon\to 0}\xi_1 T = +\infty$, *the law of* $S_u - u$, *tends to* δ_0, *if* $(-1)^{\lfloor u\rfloor} Y_{uT} = +1$, *and, otherwise, to the law of* $\inf(G, \lfloor u\rfloor + 1 - u)$ *where* G *is an exponentially distributed random variable with parameter* c.

• *if* $\lim_{\varepsilon\to 0}\xi_0 T = 0$ *and if for some* $c > 0$ $\lim_{\varepsilon\to 0}\xi_1 T = c$, *the law of* S_u *tends to the following law:*

- *if* $(-1)^{\lfloor u\rfloor} Y_{uT} = -1$: *the law of* $\mathcal{A} + \lfloor u\rfloor + 1$ *where* $\mathcal{A}$ *is exponentially distributed on the interval* $[2\mathcal{E}, 2\mathcal{E}+1]$ *with parameter* c. *Here* $\mathcal{E}$ *is a geometrical random variable on* $\mathbb{N}$ *with parameter* e^{-c} *independent of* $\mathcal{A}$,

- *if* $(-1)^{\lfloor u\rfloor} Y_{uT} = +1$: *the law of* $\mathcal{B}$ *defined as follows:*

$$\mathcal{B} := \begin{cases} \lfloor u\rfloor + \mathcal{E}_2 & \text{if} \quad \mathcal{E}_1 \in [1 - u + \lfloor u\rfloor, \infty[\\ u + \mathcal{E}_1 & \text{otherwise.} \end{cases}$$

Here $\mathcal{E}_1$ *is exponentially distributed on* $\mathbb{R}_+$, $\mathcal{E}_2$ *is exponentially distributed on* $[2\mathcal{G}, 2\mathcal{G}+1]$, *both with parameter* c, *and* $\mathcal{G}$ *is a geometrical random variable on* $\mathbb{N}^*$ *with parameter* e^{-c}, *and* $\mathcal{E}_1$,$\mathcal{E}_2$,$\mathcal{G}$ *independent.*

Interpretation: *1)* In the first case described in the statement of the proposition, the asymptotic behaviour of the Markov chain is characterized by instantaneous jumps on an exponential scale. This just means that a clock ticking in units of T will record all jumps of the process as instantaneous, since they occur on a smaller scale.

2) In the second case, the time scale T is too small compared to the transition rates. Consequently no transitions will be observed, and the process never jumps on this scale.

3) In the third case we encounter a mixture of the preceding two. A clock ticking in units of T under the conditions of this case, i.e. $\lim_{\varepsilon\to 0}\xi_0 T = 0$ and $\lim_{\varepsilon\to 0}\xi_1 T = +\infty$, will be still too slow for observing transitions from the shallow well and too fast for transitions from the deep one. Therefore the former will be recorded as instantaneous jumps, while for the latter we just have to wait until the potential state switches, before recording an instantaneous jump. This happens as soon as the half period has passed.

4) If $\lim_{\varepsilon\to 0}\xi_0 T = c$ and $\lim_{\varepsilon\to 0}\xi_1 T = +\infty$, the behaviour is similar to the preceding case, except that the clock runs on the right scale for jumps from the deep well. So if the process starts in the deep well, it does not have to wait for jumping: the jump is exponentially distributed on the interval

$[0, \lfloor u \rfloor + 1 - u]$, and then, if the process has not jumped before the half period has elapsed, it jumps at that time.
5) Let us now interpret the last case. Here the clock on scale T runs too slowly for seeing transitions from the deep well, and just has the right speed for transitions from the shallow one. So if we start in the deep position, i.e. $Y_{uT} = -1$, in units of T we will have to wait until time $\lfloor u \rfloor + 1$, after which we choose one of the intervals during which the process is in the shallow well, i.e. every second half period interval, according to an exponential law (represented by $\mathcal{E}$), to be the one in which the jump occurs. Once this interval is chosen, the jump from the shallow well inside it is again exponentially distributed (represented by $\mathcal{A}$). So in particular $\mathcal{E}$ and $\mathcal{A}$ are independent. If we start in the shallow position, the process may jump with exponential law before the next half period interval during which it is in the deep position (described by $\mathcal{E}_1$ during $[u, \lfloor u \rfloor + 1]$). If it does not jump during this period, the preceding statements apply.

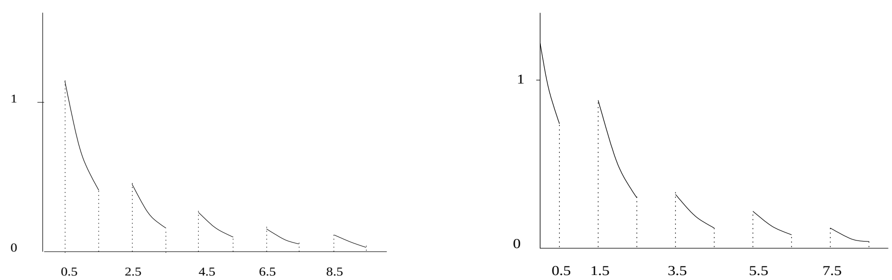

Figure 2: Asymptotic distributions of $S_u - u$ for $c = 1$ and $\lfloor u \rfloor = 0.5$

The following definition singles out the window of scales for which the Markov chain can be said to show periodic behaviour. Given our stepwise constant transition dynamics which clearly exhibits the two scales given by the Kramers times to leave a shallow well of depth $\frac{v}{2}$ and a deep well of depth $\frac{V}{2}$ as extreme scales, the definition might seem complicated at first sight. For analogy with a similar notion in more complex models in the subsequent section we prefer to leave it in this form.

Definition: Let us assume that $T(\varepsilon) = \exp \mu/\varepsilon$, $\mu \in \mathbb{R}$, and let $\mathcal{L}_{\pm}^{(u)}(\lambda)$ be the Laplace transform of the law of S_u given $(-1)^{\lfloor u \rfloor} Y_{uT} = \pm 1$, where the dependence of this transform on μ resides in the scaling of time. Then the interval

$$I_R = \{\mu : \text{for all } u \in \mathbb{R}_+,\ \text{all } \lambda \in \mathbb{R}_+ \lim_{\varepsilon \to 0} \mathcal{L}_-^{(u)}(\lambda) \neq 0,\ \lim_{\varepsilon \to 0} \mathcal{L}_-^{(u)}(\lambda) \neq e^{-\lambda u}\}$$

is called *interval of resonance.*
In this interval, characterized by the last three cases developed in the statement of the Theorem, the asymptotic behaviour of the chain is almost periodic: with probability one, during each half period, the process spends positive time in the deep well. Moreover in these scales, its asymptotic jumping does not happen instantaneously. Here the interval is clearly given by $[v, V]$ (see Figure 1). The scales corresponding to the interval boundaries just correspond to the Kramers times to exit from the shallow resp. deep well, if as usual our Markov chain is considered as the reduced dynamics of a potential diffusion.
Proof of Theorem 1: i) Let us first compute the Laplace transform of the law of S_u, given the position of Y_{uT}, using the densities (3) and (4) and the Markov property of Y. We obtain, for $u \geq 0$,

$$\begin{aligned} \mathcal{L}_-^{(u)}(\lambda) &:= \mathbb{E}\left[e^{-S_u}\Big|(-1)^{\lfloor u\rfloor}Y_{uT}=-1\right] \qquad (5)\\ &= \frac{\xi_0T}{\xi_0T+\lambda}(e^{-\lambda u}-e^{-\xi_0T(1+\lfloor u\rfloor-u)-\lambda(\lfloor u\rfloor+1)})+e^{-\xi_0T(1+\lfloor u\rfloor-u)}\\ &\times \sum_{k\geq 0}\frac{\xi_{k+1}T(1-e^{-\xi_{k+1}T-\lambda})}{\xi_{k+1}T+\lambda}\exp-\left(\sum_{j=1}^{k}\xi_jT+\lambda(\lfloor u\rfloor+k+1)\right), \end{aligned}$$

and we obtain a similar result for

$$\mathcal{L}_+^{(u)}(\lambda) := \mathbb{E}\left[e^{-S_u}\Big|(-1)^{\lfloor u\rfloor}Y_{uT}=+1\right] \qquad (6)$$

replacing in formula (5) all even indices by odd ones and vice versa.
We shall let ε tend to zero and describe the limit of the Laplace transform, according to the limit of ξ_iT.
ii) If $\lim_{\varepsilon\to 0}\xi_iT=+\infty$ for $i\in\{0,1\}$, then, obviously $\lim_{\varepsilon\to 0}\mathcal{L}_\pm^{(u)}(\lambda)=e^{-\lambda u}$. We deduce that the law of the normalized time of the first jump after time uT tends to the Dirac measure δ_u.
iii) If $\lim_{\varepsilon\to 0}\xi_iT=0$ for $i\in\{0,1\}$, then $\lim_{\varepsilon\to 0}\mathcal{L}_\pm^{(u)}(\lambda)=0$, which implies the weak convergence to the null measure.
iv) If $\lim_{\varepsilon\to 0}\xi_0T=0$ and $\lim_{\varepsilon\to 0}\xi_1T=+\infty$, we get $\lim_{\varepsilon\to 0}\mathcal{L}_+^{(u)}(\lambda)=e^{-\lambda u}$. This means that the law of S_u-u tends to the Dirac measure in the origin if $(-1)^{\lfloor u\rfloor}Y_{uT}=+1$. Otherwise $\lim_{\varepsilon\to 0}\mathcal{L}_-^{(u)}(\lambda)=e^{-\lambda\lfloor u\rfloor+1}$ so that the law of S_u-u tends to the Dirac measure in $\lfloor u\rfloor+1$ if $(-1)^{\lfloor u\rfloor}Y_{uT}=-1$.
v) If $\lim_{\varepsilon\to 0}\xi_0T=c$ (and then $\lim_{\varepsilon\to 0}\xi_1T=+\infty$), we obtain, on one hand, $\lim_{\varepsilon\to 0}\mathcal{L}_+^{(u)}(\lambda)=e^{-\lambda u}$, so the same conclusion holds as in **iv)**. On the other

hand, we get

$$\lim_{\varepsilon\to 0}\mathcal{L}_-^{(u)}(\lambda)=\frac{c}{c+\lambda}\,e^{-\lambda u}+\frac{\lambda}{c+\lambda}\,e^{-c(\lfloor u\rfloor+1-u)-\lambda(\lfloor u\rfloor+1)}. \tag{7}$$

Let us consider an exponentially distributed random variable G with parameter c (we write $G\sim\mathcal{E}(c)$) and define $Y=\inf(G,l)$ for $l>0$. Then

$$\begin{aligned}\mathbb{E}[e^{-\lambda Y}] &= \int_0^l ce^{-\lambda t-ct}dt+e^{-\lambda l}\mathbb{P}(G\geq l)\\ &= \frac{c}{c+\lambda}(1-e^{-\lambda l-cl})+e^{-cl-\lambda l}. \end{aligned}\tag{8}$$

Dividing $\mathcal{L}_-^{(u)}(\lambda)$ in (7) by $e^{-\lambda u}$ and identifying the result with (8) we obtain that the limit of S_u-u has the same law as $\inf(G,\lfloor u\rfloor+1-u)$ where $G\sim\mathcal{E}(c)$.
vi) If $\lim_{\varepsilon\to 0}\xi_0 T=0$ and $\lim_{\varepsilon\to 0}\xi_1 T=c$ then, by (5), we obtain the following limit

$$\begin{aligned}\lim_{\varepsilon\to 0}\mathcal{L}_-^{(u)}(\lambda) &= \frac{c}{c+\lambda}\,e^{-\lambda(\lfloor u\rfloor+1)}(1-e^{-c-\lambda})\sum_{k\geq 0}e^{-ck-2\lambda k}\\ &= \frac{c}{c+\lambda}\frac{1-e^{-c-\lambda}}{1-e^{-c-2\lambda}}\,e^{-\lambda(\lfloor u\rfloor+1)}. \end{aligned}\tag{9}$$

Let us now define a random variable $\mathcal{A}$ which is exponentially distributed on the interval $[2\mathcal{E},2\mathcal{E}+1]$ with parameter ν and $\mathcal{E}$ is a geometric random variable with parameter μ. Then we get

$$\begin{aligned}\mathbb{E}[e^{-\lambda\mathcal{A}}] &= \sum_{k\geq 0}\int_{2k}^{2k+1}\frac{\nu e^{-\nu t-\lambda t}}{e^{-2k\nu}-e^{-(2k+1)\nu}}dt\,\mathbb{P}(\mathcal{E}=k)\\ &= \frac{\nu}{\nu+\lambda}(1-\mu)\frac{1-e^{-\lambda-\nu}}{1-e^{-\nu}}\sum_{k\geq 0}\mu^k e^{-2k\lambda}\\ &= \frac{\nu}{\nu+\lambda}\frac{1-\mu}{1-\mu e^{-2\lambda}}\frac{1-e^{-\lambda-\nu}}{1-e^{-\nu}}. \end{aligned}$$

By identification with equation (9) we obtain $\nu=c$ and $\mu=e^{-c}$ which leads to the announced result. Moreover we also obtain a limit for the Laplace transform $\mathcal{L}_+$:

$$\begin{aligned}\lim_{\varepsilon\to 0}\mathcal{L}_+^{(u)}(\lambda) &= \frac{c}{c+\lambda}\left(e^{-\lambda u}-e^{-c(\lfloor u\rfloor+1-u)-\lambda(\lfloor u\rfloor+1)}\right)\\ &+ \frac{c}{c+\lambda}\frac{1-e^{-c-\lambda}}{1-e^{-c-2\lambda}}\,e^{-c(1+\lfloor u\rfloor-u)-\lambda\lfloor u\rfloor}. \end{aligned}\tag{10}$$

Let us define the random variable $\mathcal{B}$ as in the statement of Proposition 1 and let us compute its Laplace transform. We get

$$\begin{aligned} \mathbb{E}[e^{-\lambda\mathcal{B}}] &= \int_0^{1-u+\lfloor u\rfloor} ce^{-ct-\lambda t-\lambda u}dt \\ &+ e^{-c(1-u+\lfloor u\rfloor)}\sum_{k\geq 1}\int_{2k}^{2k+1}\frac{ce^{-\lambda t-ct-\lambda\lfloor u\rfloor}}{e^{-2kc}-e^{-(2k+1)c}}\,dt. \end{aligned}$$

Then, by straightforward computation, we obtain the expression (10). This implies the announced convergence result. QED

1.3 Optimal tuning using escape-time distribution

As mentioned in the introduction, stochastic resonance may be based on measures of quality of periodic tuning given by the entropy $H(\varepsilon, T)$ of the equilibrium measure (see Pavljukevich [17]) of the system considered. In this context, optimal tuning is expressed by minimality of $\varepsilon \mapsto H(\varepsilon, T)$, so that the resonance point corresponds to the noise intensity value for which the system in equilibrium has minimal entropy. Nature is believed to eventually choose parameter values such as this one, see Neiman et al. [16]. If not necessarily in the example of Dansgaard-Oeschger events mentioned in the following subsection, where we actually are confronted with a particular choice of a tuning parameter, this fact stresses the significance of resonance points in physical applications quite generally.

In this subsection, we shall concentrate on a measure of quality of periodic tuning which is based on the interspike distributions investigated above, and find the time scales for optimal tuning, in the small noise limit $\varepsilon \to 0$. Following [12], in physical jargon, we consider the intensity of the first peak of the *escape time distribution*. Mathematically, in our setting this intensity may be described by the probability to jump for the first time from the initial state -1 at time 0 corresponding to the deep well to the other state 1 during the normalized time interval $[1-\eta, 1+\eta]$, for $\eta > 0$ small enough. In the original scale, it means tuning the noise intensity to a value which maximizes the probability to observe the first jump in an exponentially wide time interval. As we shall see, the resonance point corresponds to findings in [17]. It is known from diffusion models that the relaxation time for the jump descreases with decreasing noise intensity (see Berglund, Gentz [2]). But at the moment we do not know if for our optimization problem we can do with shrinking intervals centered at 1 in scale T. We believe that our notion of tuning quality will prove to be robust when passing from reduced Markov chains back to potential diffusions. Of course, the subsequent peaks in the escape time distribution are able to provide analogous notions of quality of tuning. We remark at this place that the optimal time scales found for later peaks differ from the one determined in the following Proposition. We recall that τ_1 is the normalized time of the first jump of the Markov chain Y.

Proposition 1 *Set* $\Lambda(T) := \mathbb{P}(\tau_1 \in [1-\eta, 1+\eta])$. *Then* $\Lambda(T)$ *reaches a maximum for*

$$T_0 = \frac{1}{\eta(\varphi+\psi)} \ln \frac{\varphi+\psi\eta}{\varphi(1-\eta)}. \tag{11}$$

Hence, as $\varepsilon \to 0$, T_0 *is equivalent to* $\dfrac{V-v}{\eta\varepsilon q} \exp \dfrac{v}{\varepsilon}$. *Moreover,*

$$1 - \Lambda(T) = \frac{(V-v)p}{\varepsilon\eta q} e^{\frac{v-V}{\varepsilon}}(1 + \mathcal{O}(1)).$$

Proof: By the expression of the density (4), we get

$$\begin{aligned} \Lambda(T) &= e^{-\varphi T(1-\eta)} - e^{-\varphi T} + e^{-\varphi T}(1 - e^{-\psi\eta T}) \\ &= e^{-\varphi T(1-\eta)} - e^{-\varphi T - \psi\eta T}. \end{aligned}$$

Hence

$$\Lambda'(T) = -\varphi(1-\eta)e^{-\varphi T(1-\eta)} + (\varphi + \psi\eta)e^{-\varphi T - \psi\eta T}.$$

We deduce that $\Lambda'(T) = 0$ if, and only if, T_0 satisfies (11). Since $\Lambda(T) \geq 0$, $\lim_{T\to 0} \Lambda(T) = 0$ and $\lim_{T\to\infty} \Lambda(T) = 0$, Λ is maximal in T_0. QED

We can also study the variation of the kth peak of the density, considering the following probability

$$\Lambda_k(T) := \mathbb{P}(\tau_1 \in [2k-1-\eta, 2k-1+\eta]).$$

By similar computations as those used to prove Proposition 1, we obtain an equivalence, as $\varepsilon \to 0$, of the tuning maximizing the intensity of the kth peak:

$$T^{(k)} \sim \frac{e^{v/\varepsilon}}{\eta q} \ln(1 + \frac{\eta}{k-1}).$$

We deduce that, for k large enough, the optimal frequency grows like a constant times $(k-1)$. Let us note that this property was already pointed out in the context of *residence-time*: average of escape times for a large number of periods (see [8]).

1.4 Distribution of sojourn times

This subsection is devoted to the study of the law of the entire sojourn time of our Markov chain in one state. Our main result in particular explicitly describes the dependence of the law of sojourn times on the parameters T, V and v of the model, and, linked through the passage to the effective dynamics in the metastable states, to the few parameters related to the geometry of a potential in the diffusion setting. In physical systems modeling for example the encoding of acoustic information on the primary auditory nerve of mammalians (see [8]) this law resembles the distribution of *interspike intervals*, i. e. the length of intervals between subsequent spikes in a long spike train, for instance for sinusoidally stimulated auditory nerves. It also reminds of histograms found in the treatment of reduced climate systems describing paleoclimatic phenomena such as Dansgaard-Oeschger events (see Ganopolski and Rahmstorf [9]) describing measured or simulated sojourn times in intermediate warm states during the last ice age. The knowledge of the law may thus open a way to rigorously interpret stochastic resonance related phenomena expressed through interspike histograms - though in this context the term *noise induced transitions* seems to fit better than *stochastic resonance*, since the parameter dependence of the effect is not optimized. In fact, if we choose T large, but finite, and ε small so that our limiting law is a good approximation for the law of the first transition, we obtain a principal agreement of our density curve for example with the shapes of histograms of duration of Dansgaard-Oeschger events clustering at multiples of 1500 years. The parameter c controls the overall decay of the limiting law. Its value could be statistically inferred from the real paleoclimatic data presented by Ganopolski, Rahmstorf [9]. This may give a first idea about the parameters steering spontaneous transitions in a simple model for these events. But at this stage the conclusion might seem still premature: data and simulations rather indicate a tri-stable model than a two-stable one; at this stage we also do not have error estimates available for the goodness of approximation of our limiting distribution at given finite parameters T and ε.

So let us study a large number $n \in \mathbb{N}$ of jumps. We define $W_n = \left\lfloor \frac{\tau_n - \tau_{n-1}}{2} \right\rfloor$ the to be the number of entire periods the process sojourns in one state. Let N be the point process defined by

$$N := \sum_{k \geq 1} \delta_{W_k} \mathbb{1}_{\{W_k > 0\}}.$$

Proposition 2 *W_n is independent of τ_{n-1} and is a geometric random variable with parameter $(\varphi+\psi)T$.*

Proof: Using the density (3), we obtain, for $r \in \mathbb{N}$,

$$\begin{aligned} F(r) &:= \mathbb{P}\left(\tau_n - \tau_{n-1} \geq r \mid \tau_{n-1} = u \text{ and } (-1)^{\lfloor u \rfloor + n} = +1\right) \\ &= e^{-\xi_1(1+\lfloor u \rfloor - u)T} \sum_{j=r}^{\infty} \xi_j T \exp - \left(\sum_{i=1}^{j} \xi_{i-1} T \right) \int_0^1 e^{-\xi_j T w} dw \\ &\quad + \xi_{r-1} T e^{-\xi_1(1+\lfloor u \rfloor - u)T} \exp - \left(\sum_{i=1}^{r-1} \xi_{i-1} T \right) \int_{u-\lfloor u \rfloor}^1 e^{-\xi_{r-1} T w} dw \\ &= \exp\left(-\xi_1 T - \sum_{i=1}^{r-1} \xi_{i-1} T - (\xi_1 - \xi_{r-1}) T (\lfloor u \rfloor - u) \right) \end{aligned}$$

If r is even, then

$$F(r) = \exp -\frac{r}{2}(\xi_0 + \xi_1)T.$$

We observe that this probability neither depends on n nor on τ_{n-1}, and an analogous computation applied to the density (4) leads to the same result. QED

Corollary 1 *N has the same law as $\tilde{N}$ defined by*

$$\tilde{N} := \sum_{j=0}^{S_n} \delta_{\mathcal{E}_j},$$

where $(\mathcal{E}_j)_{0 \leq j \leq n}$ is a sequence of i.i.d geometrically distributed random variables with parameter $(\varphi+\psi)T$, independent of the binomial random variable $S_n \sim \mathcal{B}(n, \exp -(\varphi+\psi)T)$.

2 Two-state Markov chain with continuous generator

2.1 Introduction

The aim of this section is to generalize results obtained for piecewise constant time periodic infinitesimal generators to the time continuous case. But at the same time we widen the scope of diffusion models the effective dynamics of which is described by the two-state Markov chains we consider. Let us return for a moment to their parent models, diffusions moving in periodically

changing potential landscapes. In the last section the potential was assumed to switch discontinuously between two states which are spatially symmetric with respect to the origin. To put it differently, the transition rates φ, ψ from -1 to 1 resp. from 1 to -1, considered as functions of time are periodic of period $2T$, and the second one has a time lag $\phi = T$ with respect to the first one, i.e. $\psi(t) = \varphi(t+\phi), t \geq 0$. In this section, besides continuous variation, we generalize the periodic motion of the potential to a swinging with a time lag $\phi \in [0, 2T]$. So, as opposed to the previous section, we now suppose, keeping in mind that transition rates of the Markov chain come from Kramers' times of the diffusion, that while the minimum depth and thus the barrier height in 1 is proportional to the rate $\varphi(t)$, the depth of the minimum -1 corresponds to $\varphi(t+\phi), t \geq 0$. φ is a continuous $2T-$periodic function, the shape of which will be specified further below.
So, let us consider a time-continuous Markov chain $\{Y_t,\, t \geq 0\}$ in the state space $\mathcal{S}^Y = \{-1, 1\}$ with initial data $Y_0 = -1$. The infinitesimal generator is given by

$$Q = \begin{pmatrix} -\varphi(t) & \varphi(t) \\ \varphi(t+\phi) & -\varphi(t+\phi) \end{pmatrix}.$$

Its nondiagonal entries describe slowly continuously varying transition rates between the metastable states ± 1 as before. For instance, for small $h > 0$, the quantity $\phi(t)h(\phi(t+\varphi))$ asymptotically corresponds to the probability to jump from -1 to 1 (from 1 to -1) during the time interval $[t, t+h]$. In order to stay close to the bistable diffusion paradigm of stochastic resonance in [8], we choose a particular rate function φ according to the rules:

$$\varphi(t) = \exp -\frac{K(t/T)}{\varepsilon}, \tag{12}$$

with

$$K(t) = \frac{v+V}{2} + \frac{V-v}{2}\cos(\pi t). \tag{13}$$

As in the first section, we define τ_n to be the normalized time of the nth jump.

In the following lemma, for $0 \leq u < t$ we describe the ($\varepsilon-$dependent) probability densities of the law of τ_n given that $\tau_{n-1} = u$ and $(-1)^n = 1$ resp. $(-1)^n = -1$ by $p_+(t)$ resp. $p_-(t)$. So for instance $p_+(t)dt$ will give the probability that on time scale $2T$ jump number n from 1 to -1 happens in $[t, t+dt]$, given that the previous jump happened at time u.

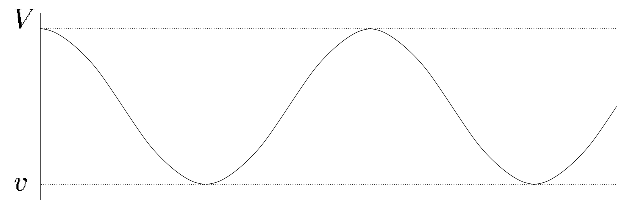

Figure 3: Definition of K

Lemma 2 *The density $p_\pm$ of τ_n, given $\tau_{n-1} = u$ and $(-1)^n = \pm 1$ is equal to*

$$p_+(t) = T\varphi(Tt+\phi)\,\mathbb{1}_{\{t\geq u\}}(t)\exp -T\int_u^t \varphi(Ts+\phi)ds, \tag{14}$$

$$p_-(t) = T\varphi(Tt)\,\mathbb{1}_{\{t\geq u\}}(t)\exp -T\int_u^t \varphi(Ts)ds. \tag{15}$$

Based on the preceding result, in the following lemma we shall give explicit descriptions of the Laplace transforms of the transition times conditioned on the times at which the previous ones occured.

Lemma 3 *The Laplace transform of τ_n, given $\tau_{n-1} = u$ and $(-1)^n = -1$, is equal to*

$$\begin{aligned} \mathcal{L}_-^{(u)}(\lambda) &= \left(1-\exp\left\{-T\int_0^2\varphi(Ts)ds - 2\lambda\right\}\right)^{-1} \\ &\quad\times\int_u^{u+2} T\varphi(tT)\exp\left\{-T\int_u^t\varphi(Ts)ds-\lambda t\right\}dt. \end{aligned} \tag{16}$$

A similar expression can be obtain for $(-1)^n = 1$, it suffices to replace φ by $\varphi(\cdot+\phi)$ in expression (16).
Proof: By (15),

$$\begin{aligned} \mathcal{L}_-^{(u)}(\lambda) &= \mathbb{E}\left[e^{-\lambda\tau_n}\middle|\ \tau_{n-1}=u,\quad (-1)^n=-1\right] \\ &= \int_u^\infty T\varphi(Tt)\exp\left\{-T\int_u^t\varphi(Ts)ds-\lambda t\right\}dt \\ &= \sum_{k\geq 0}\int_u^{u+2}T\varphi(Tt)e^{-\lambda t-2\lambda k} \\ &\quad\times\exp\left\{-T\int_u^t\varphi(Ts)ds - kT\int_0^2\varphi(Ts)ds\right\}dt \end{aligned}$$

Using the equality

$$\sum_{k\geq 0} \exp\left\{-kT\int_0^2 \varphi(Ts)ds - 2\lambda k\right\} = \left(1-\exp\left\{-T\int_0^2 \varphi(Ts)ds - 2\lambda\right\}\right)^{-1}$$

we obtain the expression (16). QED

2.2 Asymptotic behaviour

Since the density of τ_n given τ_{n-1} depends on ε and T, following the example of the preceding section, we shall consider the asymptotics of transition densities as $\varepsilon \to 0$ and eventually $T = T(\varepsilon) \to \infty$. So, let us introduce analogous scales by defining

$$T = T(\varepsilon) = \exp\frac{\mu}{\varepsilon} \quad \text{with} \quad \mu > 0. \tag{17}$$

Then, by the definition (12) of φ the Laplace transform $\mathcal{L}_-$ becomes

$$\begin{aligned} \mathcal{L}_-^{(u)}(\lambda) &= \left(1-\exp\left\{-\int_0^2 \exp\left\{\frac{\mu-K(s)}{\varepsilon}\right\}ds - 2\lambda\right\}\right)^{-1} \\ &\times \int_u^{u+2} \exp\left\{\frac{\mu-K(t)}{\varepsilon} - \int_u^t \exp\left\{\frac{\mu-K(s)}{\varepsilon}\right\}ds - \lambda t\right\} dt \end{aligned} \tag{18}$$

As in the preceding section, we define S_u to be the normalized time of the first jump after the normalized time u. Since the process is Markovian, the Laplace transform of S_u given $Y_{uT} = \pm 1$ is equal to $\mathcal{L}_{\pm}^{(u)}$. To motivate the main asymptotic results of the following Theorem, let us briefly return to the flip potential model of the preceding section. In the framework of Theorem 1 the choice of $T(\varepsilon)$ made above results in the following simpler relevant conditions:

$$\mu < v, \quad \mu = v, \quad v < \mu < V, \quad \mu = V, \quad V < \mu.$$

For instance, $v < \mu < V$ corresponds to the former condition $\lim_{\varepsilon\to 0} \xi_0 T = 0, \lim_{\varepsilon\to 0}\xi_1 T = \infty$, $\mu = V$ to $\lim_{\varepsilon\to 0}\xi_0 T = q, \lim_{\varepsilon\to 0}\xi_1 T = \infty$. So case 2 of the following Theorem comprises cases 3 and 4 of Theorem 1. As opposed to the piecewise constant case, however, there are now other relevant scales for the Markov chain. For $u \geq 0$, define

$$a_\mu(u) = \inf\{t > u \ : \ \mu - K(t) \geq 0\}. \tag{19}$$

Remember that we are arguing on scale T. Suppose we freeze the potential in its state it has at time $a_\mu(u)$. Then the Kramers escape time for the resulting barrier height $\frac{K(t)}{2}$ is precisely 1. Jumps are consequently likely to occur on this scale. Choosing a scale just a little below $a_\mu(u)$ would make the time scale too small to lead to recording escapes over a potential barrier which just too high according to Kramers' law. So $a_\mu(u)$ provides a critical scale at which transitions across a barrier of height $\frac{K(t)}{2}$ become noticeable.

Theorem 2 *Let ε tend to zero. Then*
- *if $\mu > V$ the law of $S_u - u$ tends to the Dirac measure in the origin.*
- *if $\mu \in]v, V]$ the conditional law of S_u given $Y_{uT} = -1$ tends to the Dirac measure in the point $a_\mu(u)$. If $Y_{uT} = 1$ then the same result holds with the Dirac measure in the point $\tilde{a}_\mu(u)$. $\tilde{a}$ is defined as a, but for the function $K(\cdot + \phi/T)$ instead of K.*
- *if $\mu \leq v$, then the probability measure of S_u tends weakly to the null measure.*

Proof: i) Let us first study the case $\mu > V$. By (18), the numerator of the Laplace transform is equal to

$$N_\varepsilon(\lambda) = \int_u^{u+2} W_\varepsilon(t,\lambda) \left(e^{(\mu - K(t))/\varepsilon} + \lambda\right) \exp\left\{-\int_u^t e^{(\mu-K(s))/\varepsilon} ds - \lambda t\right\} dt,$$

where

$$W_\varepsilon(t,\lambda) = \frac{e^{(\mu-K(t))/\varepsilon}}{e^{(\mu-K(t))/\varepsilon} + \lambda}.$$

Let us fix $\lambda \geq 0$. Since $\mu > \sup_{t \geq 0} K(t)$, we get, as $\varepsilon \to 0$, uniformly with respect to the variable t

$$1 - o(\varepsilon) \leq W_\varepsilon(t,\lambda) \leq 1.$$

Hence, asymptotically as $\varepsilon \to 0$, the numerator is equivalent to

$$\exp(-\lambda u) - \exp\left\{-\int_u^{u+2} e^{(\mu-K(s))/\varepsilon} ds - \lambda(u+2)\right\}$$

which tends to $e^{-\lambda u}$ as ε tends to zero. Moreover the denominator of the Laplace transform obviously tends to 1. So $\lim_{\varepsilon \to 0} \mathcal{L}_-^{(u)}(\lambda) = e^{-\lambda u}$. The same limit can be obtained for the Laplace transform $\mathcal{L}_+$. We deduce that the law of $S_u - u$ tends to the Dirac measure in the origin. This proof can be extended to the case $\mu = V$.
ii) In the second case $\mu \in [v, V[$, decomposing the integral $\int_0^2 e^{(\mu-K(s))/\varepsilon} ds$ into integrals on disjoint intervals, we obtain that the denominator of the

Laplace transform tends to 1, exponentially fast if $\mu > v$ and, otherwise, using the Laplace method, with a speed of order $\sqrt{\varepsilon}$ since

$$K(\eta) - K(1) \sim (V - v)\pi^2(\eta - 1)^2/2 \quad \text{for} \quad |\eta - 1| \quad \text{small.}$$

The numerator will be decomposed into three parts:

$$N_\varepsilon(\lambda) = R_1(\varepsilon, \lambda) + R_2(\varepsilon, \lambda) + R_3(\varepsilon, \lambda),$$

$$R_i(\varepsilon, \lambda) = \int_{\Delta_i} \mathcal{A}(\mu, t, \varepsilon)dt = \int_{\Delta_i} \exp\left\{\frac{\mu - K(t)}{\varepsilon} - \int_u^t e^{(\mu - K(s))/\varepsilon}ds - \lambda t\right\} dt$$

where $\Delta_1 = [u, u+2] \cap \{t \geq 0 \ : \ \mu - K(t) < -\sqrt{\varepsilon}\}$,
$\Delta_2 = [u, u+2] \cap \{t \geq 0 \ : \ \mu - K(t) > 0\}$
and $\Delta_3 = [u, u+2] \cap \{t \geq 0 \ : \ -\sqrt{\varepsilon} \leq \mu - K(t) \leq 0\}$.
By the definition of Δ_1, we get $R_1(\varepsilon, \lambda) \leq 2e^{-1/\sqrt{\varepsilon}}$ which tends to zero as ε decreases. Otherwise $R_3(\varepsilon, \lambda) \leq |\Delta_3| = \mathcal{O}(\sqrt{\varepsilon})$ where $|.|$ is the Euclidean length. It remains to determine the limit of the expression $R_2(\varepsilon, \lambda)$.
We notice that if $\mu = v$ then $\Delta_2 = \emptyset$. This implies that, in this case, the Laplace transform tends to 0: the measure of τ_n given τ_{n-1} tends to the null measure. Let us now assume that $\mu > v$ and that $a_\mu(u) > u$ (the function a_μ is defined by (19)) ; we postpone the study of the case: $a_\mu(u) = u$. There exists $\delta > 0$ small enough, independent of ε, such that

$$R_2(\varepsilon, \lambda) = \int_{[a_\mu(u), a_\mu(u)+\delta]} \mathcal{A}(\mu, t, \varepsilon)dt + \int_{\Delta_2 \cap [a_\mu(u), a_\mu(u)+\delta]^c} \mathcal{A}(\mu, t, \varepsilon)dt.$$

The second integral can be bounded above by

$$2e^{-\lambda(a_\mu(u)+\delta)} \exp - \int_u^{a_\mu(u)+\delta} e^{(\mu - K(s))/\varepsilon}ds.$$

Using Laplace's method, we get that this expression tends to zero, as $\varepsilon \to 0$. Furthermore, the first integral satisfies the inequalities

$$e^{-\lambda(a_\mu(u)+\delta)} I \leq \int_{[a_\mu(u), a_\mu(u)+\delta]} \mathcal{A}(\mu, t, \varepsilon)dt \leq e^{-\lambda a_\mu(u)} I,$$

with

$$I = \exp\left(-\int_u^{a_\mu(u)} e^{(\mu - K(s))/\varepsilon}ds\right) - \exp\left(-\int_u^{a_\mu(u)+\delta} e^{(\mu - K(s))/\varepsilon}ds\right).$$

I tends to 1 as $\varepsilon \to 0$ (δ fixed). Hence $\lim_{\varepsilon \to 0} \mathcal{L}_-^{(u)}(\lambda) = e^{-\lambda a_\mu(u)}$ for $Y_{uT} = -1$.

If $\mu > v$ and $a_\mu(u) = u$, we denote by δ the first time greater than u such that $K(\delta) = \mu$. By the definition (19), we get $\delta > u$. Let us decompose the expression $R_2(\varepsilon, \lambda)$ as follows:

$$R_2(\varepsilon, \lambda) = \int_{[u,\delta]} \mathcal{A}(\mu, t, \varepsilon)dt + \int_{\Delta_2 \cap [u,\delta]^c} \mathcal{A}(\mu, t, \varepsilon)dt.$$

The second term is bounded above by

$$2e^{-\lambda\delta} \exp - \int_u^{a_\mu(\delta)} e^{(\mu - K(t))/\varepsilon} dt.$$

This term tends to zero as $\varepsilon \to 0$. Using the arguments presented in **i)**, we obtain that the first expression tends to $e^{-\lambda u}$.
All the results of **ii)** can be proved for $Y_{uT} = 1$ using the function $K(\cdot + \phi/T)$ instead of K.
iii) In the case $\mu < v$, it is straightforward to prove that $\lim_{\varepsilon \to 0} \mathcal{L}_\pm^{(u)}(\lambda) = 0$. We deduce that the measure of probability of S_u converges weakly to the null measure. QED

As in the case of a piecewise constant infinitesimal generator, we can define in the continuous case an interval of resonance.
Definition: Let us assume that $T(\varepsilon) = e^{\mu/\varepsilon}$, recall that the dependence of the Laplace transforms on the parameter μ resides in the scaling of time, and let

$$I_0 = \{\mu \in \mathbb{R} : \lim_{\varepsilon \to 0} \mathcal{L}_-^{(u)}(\lambda) \neq 0, \text{ for all } \lambda \in \mathbb{R}_+,\ u \in \mathbb{R}_+\},$$

$$I_1 = \{\mu \in \mathbb{R} : \quad \{u \in [0, 2[, \lim_{\varepsilon \to 0} \mathcal{L}_-^{(u)}(\lambda) = \lim_{\varepsilon \to 0} \mathcal{L}_+^{(u)}(\lambda) = e^{-\lambda u} \text{ for all } \lambda \in \mathbb{R}_+\} = \emptyset\}.$$

Then $I_R := I_0 \cap I_1$ is called *interval of resonance.*
The abstract definition given above translates the following intuitive facts. The interval of resonance is set to contain those exponential scales in which the process on the one hand asymptotically cannot stay always in the same state with positive probability, and on the other hand cannot jump instantaneously from one state to the other. In the continuous case, the interval of resonance depends on the phase between the infinitesimal probabilities. Indeed, on the set of scales given by

$$\Omega := \{t \in [0, 2[:\ \mu - K(t) < 0\} \cap \{t \in [0, 2[:\ \mu - K(t + \phi/T) < 0\}$$

the process will asymptotically exhibit instantaneous jumping back and forth between ± 1. So the interval of resonance should be equal to the set of all $\mu \in]v, V]$ for which the set Ω of instantaneous jumping in both directions is empty. This is exactly what led us to the definition of I_1 above. Figure 4 presents a $\mu \in]v, V]$ such that Ω has a positive Lebesgue measure.

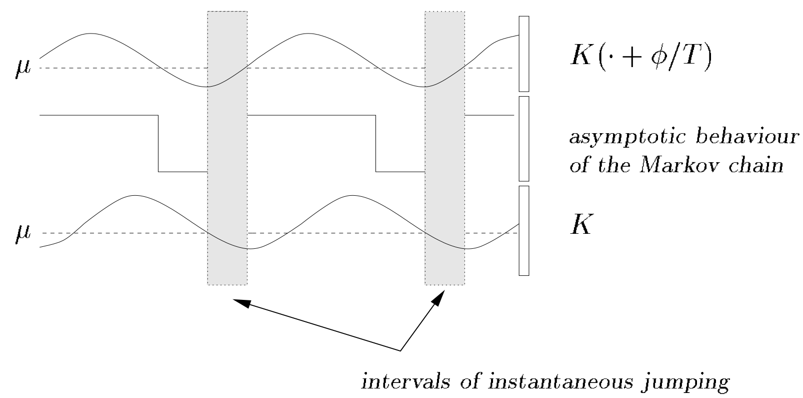

Figure 4: Instantaneous jumping

By (12), the set of all μ belonging to this particular interval satisfies the conditions

$$\frac{2T(\pi - \arccos(\frac{2\mu - V - v}{V - v}))}{\pi} \leq \phi \leq \frac{2T \arccos(\frac{2\mu - V - v}{V - v})}{\pi} \quad \text{and} \quad \mu > v. \quad (20)$$

Let us now assume that $\phi = T$. The behaviour of the two-state Markov chain then imitates the barrier crossings of a diffusion in a double-well potential. Here the potential is like $U(x) + x\cos(\pi t/T)$, where U is a symmetric double-well potential (this case has a physical meaning, see [8]). Indeed, whether the diffusion starts in the deepest position of the left well at time t or starts in the right well at time $t + T$, it has to cross a barrier of the same height to reach the other well. So the phase between the infinitesimal rate of jumps of the associated Markov chain has to be equal to T. In this particular case, we obtain the following interval of resonance:

$$I_R = \left] v, \frac{v + V}{2} \right].$$

The time scale corresponding to the upper bound is of the same order as the mean exit time of one well (Kramers' rate), by the diffusion starting in the deepest position of this well, for the symmetric double-well potential U.

This potential is, in fact, the average potential of $U(x) + x\cos(\pi t/T)$ over one period.
Let us note that $\mu_0 = \frac{v+V}{2}$ is a bifurcation point: for $\mu \leq \mu_0$ the asymptotic behaviour has no instantaneous jump part, for $\mu > \mu_0$ instantaneous jumping occurs.

2.3 Distribution of sojourn times

Let us consider a large number of jumps $n \in \mathbb{N}$. We define as in section 1.4 $W_n = \left\lfloor \frac{\tau_n - \tau_{n-1}}{2} \right\rfloor$ (τ_n is the normalized time of the nth jump).

Proposition 3 *W_n is independent of τ_{n-1} and is a geometric random variable with parameter $T \int_0^2 \varphi(Ts)ds$.*

The proof of this proposition is similar to the one in the previous section. Moreover we can also obtain an equivalence in law with a particular point process: it suffices to replace $\varphi + \psi$ by $\int_0^2 \varphi(Ts)ds$ in the statement of Corollary 1.

2.4 Optimal tuning using escape-time distribution and large deviations

Again following the preceding section, we next determine an optimal tuning rate for stochastic resonance. It will again be based on the density of the first jump, in particular the intensity of its first peak, which we propose as a new measure of quality of tuning. The optimal time scale will be determined by a combination of a large deviations result concerning the first jump of the Markov chain parametrized by the logarithmic scale μ of time, and a maximization problem for the uniformly obtained large deviation rates.
By (19), a_μ is well defined on the interval $[v, V]$. We extend this function continuously to $]-\infty, V]$: $a_\mu = a_v$ for $\mu \leq v$.
For $\mu \leq V$ and $\delta > 0$, $\mathcal{A}^\delta$ will be the event defined by

$$|\tau_1 - (a_\mu(0) + 2k)| \geq \delta, \quad \text{for all} \quad k \in \mathbb{N}. \tag{21}$$

The following Theorem, proved by large deviations methods, essentially determines exponential rates of probabilities for $\mathcal{A}_\delta$, in particular if the exponential scale μ is in the resonance interval. Its primary result (formula (22)) states that the main contribution to this probability of jumping comes from the instant during the time interval under consideration at which just the minimal barrier height is obtained. Its exponential rate is given by the *remaining barrier height* to overcome.

Theorem 3 • *If* $\mu \in]v, V[$ *and* $\delta < \sup(a_\mu(0), T/2 - a_\mu(0))$, *then*

$$\lim_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}(\mathcal{A}^\delta) = \mu - K(a_\mu(0) - \delta), \tag{22}$$

• *If* $\mu < v$ *then, for* δ *small enough,*

$$\lim_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}(\mathcal{A}^\delta) = K(a_\mu(0)) - K(a_\mu(0) - \delta). \tag{23}$$

Proof: Using the expression (15) for the density, we get

$$\mathbb{P}(\mathcal{A}^\delta) = \int_{\mathcal{D}_\delta} T\varphi(Tt) \exp -T \int_0^t \varphi(Ts)ds\, dt,$$

where $\mathcal{D}_\delta = \{t \geq 0, \ \ |t - (a_\mu(0) + 2k)| \geq \delta, \ \ \forall k \in \mathbb{N}\}$. Then, for δ small,

$$\begin{aligned} \mathbb{P}(\mathcal{A}^\delta) \;\; &= \;\; \int_0^{a_\mu(0)-\delta} T\varphi(Tt) \exp -T \int_0^t \varphi(Ts)ds \\ &+ \;\; \frac{\exp - \int_0^{a_\mu(0)} T\varphi(Ts)ds}{1 - \exp -T \int_0^2 \varphi(Ts)ds} \\ &\qquad \times \int_\delta^{2-\delta} T\varphi(T(a_\mu(0) + t)) \exp -T \int_0^t \varphi(T(a_\mu(0) + s))ds\, dt. \end{aligned}$$

Then $\mathbb{P}(\mathcal{A}^\delta) = A + B \times C$, where

$$A = 1 - \exp - \int_0^{a_\mu(0)-\delta} e^{(\mu - K(s))/\varepsilon} ds,$$

$$B = \frac{\exp - \int_0^{a_\mu(0)+\delta} e^{(\mu - K(s))/\varepsilon} ds}{1 - \exp - \int_0^2 e^{(\mu - K(s))/\varepsilon} ds},$$

$$C = 1 - \exp - \int_\delta^{2-\delta} e^{(\mu - K(s + a_\mu(0)))/\varepsilon} ds.$$

• If $\mu \in]v, V[$, then, using Laplace's method to get equivalents of the integrals as ε tends to zero, we obtain

$$\begin{aligned} \mathbb{P}(\mathcal{A}^\delta) \;\; &\sim \;\; \left(1 - \exp - \int_0^{a_\mu(0)-\delta} e^{(\mu - K(s))/\varepsilon} ds\right) \sim \int_0^{a_\mu(0)-\delta} e^{(\mu - K(s))/\varepsilon} ds \\ &\sim \;\; \frac{2\varepsilon}{\pi(V - v)\sin(\pi(a_\mu(0) - \delta))} e^{(\mu - K(a_\mu(0) - \delta))/\varepsilon}. \end{aligned} \tag{24}$$

• If $\mu < v$, then, again by Laplace's method, as $\varepsilon \to 0$, we get

$$\varepsilon \ln(A) \sim (\mu - K(a_\mu(0) - \delta)), \quad \varepsilon \ln(B) \sim (K(a_\mu(0)) - \mu),$$

and $\varepsilon \ln(C) \sim (\mu - K(a_\mu(0) - \delta))$. Hence

$$\lim_{\varepsilon \to 0} \varepsilon \ln \mathbb{P}(\mathcal{A}^\delta) = \sup\left(\mu - K(a_\mu(0) - \delta),\, K(a_\mu(0)) - K(a_\mu(0) - \delta)\right).$$

We deduce (23) for δ small enough. QED

Let us now describe the optimal tuning rate corresponding to the stronger first peak of the density of the first jump. Let us also recall our remark that we optimize the probability for the first jump to occur in an exponentially wide interval of rescaled width δ. The optimal rate will be seen to have a dependence on δ which is not desirable, but rather weak. At the moment we do not know how (by eventual reduction of the interval) this dependence can be ruled out.

Theorem 4 *Let $[a, b] \subset]v, V[$. For δ small, we define $T_0^\varepsilon := \exp \frac{\mu_0^\varepsilon}{\varepsilon}$ such that*

$$\mu_0^\varepsilon = \inf\{a \le \lambda \le b,\ \mathbb{P}(\tau_1 \in [a_\lambda(0) - \delta, a_\lambda(0) + \delta]) \ \textit{ is maximal } \}.$$

Then, $\lim_{\varepsilon \to 0} \mu_0^\varepsilon = \frac{v+V}{2} - \frac{V-v}{2} \sin \frac{\pi\delta}{2}$.

Remark: The optimal tuning for this particular observable belongs to the resonance interval if the phase ϕ equals T. If $\frac{v+V}{2} - \frac{V-v}{2} \sin \frac{\pi\delta}{2}$ does not belong to the resonance interval, then the optimal tuning in this interval is equal to the right boundary (see the following proof).

Proof: Let us study the family of functions $\varepsilon \ln F_\varepsilon(\mu)$ where $F_\varepsilon(\mu)$ is defined as an integral like $\mathbb{P}(\mathcal{A}^\delta)$ but on the domain $\{t \ge 0 \ :\ |t - a_\mu(0)| \ge \delta\}$. We get

$$\begin{aligned} F_\varepsilon(\mu) &= 1 - \exp\left(-\int_0^{a_\mu(0)-\delta} e^{(\mu - K(s))/\varepsilon} ds\right) \\ &+ \exp\left(-\int_0^{a_\mu(0)+\delta} e^{(\mu - K(s))/\varepsilon} ds\right). \end{aligned}$$

In order to study maximality of $\mathbb{P}(\tau_1 \in [a_\lambda(0) - \delta, a_\lambda(0) + \delta])$ in ε, we determine the minimum of $\varepsilon \ln F_\varepsilon(\mu)$ on the interval $[a, b]$. By Theorem 3 and due to the result concerning the interspike distribution (Proposition 3), we get the simple convergence of the sequence of functions

$$\lim_{\varepsilon \to 0} \varepsilon \ln F_\varepsilon(\mu) = \mu - K(a_\mu(0) - \delta).$$

Let us now prove uniform convergence. By Ascoli's theorem, it suffices to prove that the derivative with respect to the variable μ is bounded as $\varepsilon \to 0$. There exists a constant $C > 0$, such that

$$\begin{aligned}\left|\frac{\varepsilon F_\varepsilon'(\mu)}{F_\varepsilon(\mu)}\right| &\leq C\, F_\varepsilon^{-1}(\mu)\left(\int_0^{a_\mu(0)-\delta} e^{(\mu-K(s))/\varepsilon} ds + \right.\\ &+ \left. \varepsilon e^{(\mu-K(a_\mu(0)-\delta)/\varepsilon}\frac{\partial a_\mu(0)}{\partial\mu} + o\left(\varepsilon e^{(\mu-K(a_\mu(0)-\delta)/\varepsilon}\right)\right).\end{aligned}$$

We deduce that this derivative is bounded using the arguments related to (24). In order to finish this proof, we shall point out that the limit function $\mu - K(a_\mu(0) - \delta)$ reaches its minimum in only one point. Let us recall that

$$K(x) = \frac{v+V}{2} + \frac{V-v}{2}\cos(\pi x).$$

We deduce that

$$a_\mu(0) = \frac{1}{\pi}\arccos\left(\frac{2\mu - v - V}{V-v}\right).$$

Hence the minimum of $\mu - K(a_\mu(0) - \delta)$ is reached iff

$$\arccos\left(\frac{2\mu_0 - v - V}{V-v}\right) = \frac{\pi}{2}(1+\delta),$$

$$\text{i.e. iff}\ \ \mu_0 = \frac{V-v}{2}\cos\left(\frac{\pi}{2} + \frac{\pi\delta}{2}\right) + \frac{v+V}{2}.$$

Finally let us note that the limit function $\mu - K(a_\mu(0) - \delta)$ is decreasing for $a \leq \mu \leq \mu_0$. Hence, if μ_0 does not belong to the interval of resonance i.e. if there exists $r < \mu_0$ such that $I_R =]v, r]$, the optimal tuning in this interval is r. QED

Extensions of the results proved here for the reduced dynamics of two-state Markov chains to the full dynamics of one-dimensional potential diffusions can be obtained, and will be dealt with elsewhere.

Acknowledgement: We are much indebted to an anonymous referee for very constructive and knowledgeable criticism.

References

[1] S. Arrhenius. *J. Phys. Chem.*, 4:226, 1889.

[2] N. Berglund and B. Gentz. A sample-paths approach to noise-induced synchronization: Stochastic resonance in a double-well potential. 2000. to appear in *Ann. Appl. Prob.*

[3] N. Berglund and B. Gentz. Metastability in simple climate models: Pathwise analysis of slowly driven langevin equations. 2001. to appear in *Stochastics and Dynamics.*

[4] J.-P. Eckmann and L. E. Thomas. Remarks on stochastic resonance. *J. Phys. A*, 15:L261–L266, 1982.

[5] H. Eyring. The activated complex in chemical reactions. *Journal of Chemical Physics*, 3:107–115, 1935.

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

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

[8] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Reviews of Modern Physics*, 70(1):223–287, 1998.

[9] A. Ganopolski and S. Rahmstorf. Abrupt glacial climate changes due to stochastic resonance. *Phy. Rev. Letters*, 88(3), 2002.

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

[11] P. Imkeller and I. Pavlyukevich. Model reduction and stochastic resonance. *to appear in Stochastics and Dynamics*, 2002.

[12] P. Jung, F. Moss, and T. Zhou. Escape-time distributions of a periodically modulated bistable system with noise. *Phy. Rev. A*, 42(6):3161–3169, 1990.

[13] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[14] R. S. Maier and D. L. Stein. Escape problem for irreversible systems. *Physical Review E*, 48(2):931–938, 1993.

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

[16] A. Neiman, B. Shulgin, V. Anishchenko, W. Ebeling, L. Schimansky-Geier, and J. Freund. Dynamical Entropies Applied to Stochastic Resonance. *Physical Review Letters*, 77:4851+, December 1996.

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