Stochastic resetting and hierarchical synchronization
Meylahn, J.M.

Citation

Meylahn, J. M. (2019, September 24). *Stochastic resetting and hierarchical synchronization*. Retrieved from https://hdl.handle.net/1887/78560

Version:	Publisher's Version
License:	[Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](https://hdl.handle.net/1887/78560)
Downloaded from:	https://hdl.handle.net/1887/78560

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/78560 holds various files of this Leiden University dissertation.

Author: Meylahn, J.M.
Title: Stochastic resetting and hierarchical synchronization
Issue Date: 2019-09-24

PART I

LARGE DEVIATIONS OF STOCHASTIC PROCESSES WITH RESETTING

CHAPTER 2

Large deviations for diffusions with resetting

This chapter is based on:[92].

Abstract

Markov processes restarted or reset at random times to a fixed state or region in space have been actively studied recently in connection with random searches, foraging, and population dynamics. Here we study the large deviations of time-additive functions or observables of Markov processes with resetting. By deriving a renewal formula linking generating functions with and without resetting, we are able to obtain the rate function of such observables, characterizing the likelihood of their fluctuations in the long-time limit. We consider as an illustration the large deviations of the area of the Ornstein-Uhlenbeck process with resetting. Other applications involving diffusions, random walks, and jump processes with resetting or catastrophes are discussed.

§2.1 Introduction

Stochastic processes with restarting or reset events, corresponding to random transitions in time to a given state or region in space, have been the subject of active studies in physics and mathematics in recent years. In physics, such processes have been studied as a mechanism for power-law distributions [89] and, more recently, as random search models suggested by common experience (e.g., losing one's keys) in which periods of diffusive exploration are interspaced with random returns to a starting point [48, 47, 50, 49, 74, 58, 87, 88]. In this context, a reset is also called a restart [67] or a teleportation [11] and can be considered as part of more general intermittent search strategies combining different exploration dynamics [97].

In mathematics, processes with reset have been studied mostly in the context of birth-death processes modelling the evolution of populations in which partial or complete extinction or emigration events happen at random times [101, 19, 18, 75, 102, 45, 43]. In this context, a reset is more often referred to as a catastrophe, disaster or decimation and can also be seen as an absorbing or "killing" state that triggers, when reached, a restart or "resurrection" of the process [102]. Similar jump processes have been studied for modelling queues where random "failures" clearing the content or occupation of a queue are followed by "repaired phases" in which the queue functions normally [31, 29, 71].

The focus of these studies, both on the physical and mathematical sides, is on determining time-dependent and stationary distributions, as well as survival and first-passage time statistics using modified Master or Fokker-Planck equations that account for the effect of resetting. Renewal representations of distributions and first-passage time statistics have also been obtained for jump processes [75, 102, 45] and diffusion equations [48, 47, 50]. First-passage times are especially important for search applications, as they provide a measure of the efficiency of adding resetting to random walks.

Here, we consider a different problem involving resetting, namely, that of deriving large deviation functions for additive observables. The study of large deviations for "normal" Markov processes is an active area of probability theory having many applications in queueing theory, estimation, and control [20, 119, 39]. Large deviation functions also play a fundamental role in statistical physics by providing rigorous versions of the notions of entropy and free energy for equilibrium systems [46], which can be generalized to nonequilibrium systems driven in steady states [130, 40, 42, 15, 61]. In this context, an additive observable is simply a quantity integrated over time for a physical system evolving stochastically due to the influence of noise, external forces, and boundary reservoirs. It can represent, for example, the work done when pulling a Brownian particle with laser tweezers [132], the stretch of a molecular motor attached to a protein [91], or the total energy or particle current exchanged between different reservoirs in a given time interval [42]. In all cases, the fluctuations of the observable studied are characterized in the long-time limit by the so-called rate function, which is the central function of large deviation theory [46, 39, 40, 130].

We obtain in the following large deviation functions for processes with resetting by deriving two representations for the generating function of additive observables: one

that is essentially a reset generalization of the Feynman-Kac formula and another that links, via a renewal argument, the generating function of an observable with resetting to its generating function without resetting. The derivation of rate functions follows from these results by studying, as is common in large deviation theory, the long-time asymptotics of generating functions. As an illustration of these results, we consider in Sec. 2.4 the large deviations of the integral (area) of the reset Ornstein-Uhlenbeck process, which can be considered as a simple model of molecular motor with resetting [91]. Other applications related to birth-death processes and queues are mentioned in the conclusion of the paper.

§2.2 Problem

To simplify the presentation, we consider the case of one-dimensional diffusions. Higher-dimensional diffusions and jump processes such as birth-death processes can be treated with minor changes of notation.

We thus consider an ergodic diffusion process $X_t \in \mathbb{R}$ described by the stochastic differential equation (SDE)

$$\mathrm{d}X_t = F(X_t)\mathrm{d}t + \sigma \mathrm{d}W_t, \tag{2.2.1}$$

which is reset to the fixed position x_r at random times distributed according to an exponential distribution with parameter $r \geq 0$. Considering the evolution of X_t over an infinitesimal time $\mathrm{d}t$, this means that X_t is either reset to $X_{t+\mathrm{d}t} = x_r$ with probability $r\mathrm{d}t$ or that X_t diffuses with probability $1 - r\mathrm{d}t$ according to the SDE 2.2.1, which involves the drift $F(X_t)$, the noise power $\sigma > 0$, and the Brownian motion or Wiener process W_t.

As shown in [48, 47], the resetting modifies the Fokker-Planck equation governing the evolution of the probability density $p(x, t|x_0)$ of X_t started at $X_0 = x_0$ by adding a uniform sink and a source at x_r:

$$\frac{\partial}{\partial t}p(x,t|x_0) = -\frac{\partial}{\partial x}F(x)p(x,t|x_0) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}p(x,t|x_0)$$
$$-rp(x,t|x_0) + r\delta(x - x_r). \tag{2.2.2}$$

Alternatively, $p(x, t|x_0)$ can be obtained by noting that X_t can reach $x \neq x_r$ by diffusing either from its last reset position x_r, which occurred at the random time $t - \tau$, or from its initial state x_0 without resetting, so that

$$p(x,t|x_0) = e^{-rt}p_0(x,t|x_0) + \int_0^t re^{-r\tau}\, p_0(x,\tau|x_r)\mathrm{d}\tau, \tag{2.2.3}$$

where $p_0(x, t|x_0)$ is the free propagator solving the standard Fokker-Planck equation (2.2.2) with $r = 0$ [48, 47]. Similar renewal formulae connecting time-dependent distributions with and without resetting have been obtained in the context of jump processes modelling population dynamics [19, 18, 75, 45] and queues [71, 31, 29]. Modified Fokker-Planck equations with resetting have also been obtained by studying the diffusive or Kramers-Moyal limit of reset jump processes; see [31, 29, 43, 44].

Here, we study the probability density not of the process itself but of functionals or observables of X_t having the general time-additive form

$$A_T = \frac{1}{T} \int_0^T f(X_t) \mathrm{d}t, \tag{2.2.4}$$

where f is a real function of X_t. Such observables naturally arise in manmade and physical systems, as mentioned, and are often characterized by a probability density having the form

$$P(A_T = a) = e^{-TI(a) + o(T)} \tag{2.2.5}$$

in the limit of large integration times T, with $o(T)$ denoting any correction term that grows slower than T. This scaling of probabilities is known in large deviation theory as the *large deviation principle* (LDP) [46, 39, 40, 130] and implies that fluctuations of A_T are exponentially unlikely to be observed in the long-time limit. This applies for all values $A_T = a$ such that the rate of decay or rate function $I(a)$ is positive. In general, $I(a)$ also has (at least) one zero a^* determining the *typical value* of A_T around which $P(A_T = a)$ concentrates exponentially as $T \to \infty$. The rate function is thus important as it characterizes in the long-time limit the typical value of A_T, which corresponds to its ergodic or stationary value, as well as the atypical fluctuations around this ergodic value.

For processes with no resetting, the rate function is generally obtained by calculating the *scaled cumulant generating function* (SCGF) of A_T defined by the limit

$$\lambda_0(k) = \lim_{T \to \infty} \frac{1}{T} \ln E_x^0 \big[e^{TkA_T} \big], \tag{2.2.6}$$

where $k \in \mathbb{R}$ and $E_x^0[\cdot]$ denotes the expectation with respect to the non-reset process X_t started at $X_0 = x$. For Markov processes, it is known that this function coincides under general conditions with the dominant eigenvalue of the so-called *tilted generator* [46, 39, 40, 130], which for the SDE (2.2.1) has the form

$$\mathcal{L}_k = L + kf, \tag{2.2.7}$$

where

$$L = F \frac{\partial}{\partial x} + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} \tag{2.2.8}$$

is the generator of the diffusion X_t without resetting. In this case, the calculation of large deviations is therefore essentially a spectral problem. Assuming that $\lambda_0(k)$ can be obtained and is differentiable, we then have from an important result of large deviation theory, known as the Gärtner-Ellis Theorem [46, 39, 40, 130], that A_T satisfies an LDP with rate function $I_0(a)$ given by the Legendre-Fenchel transform of the SCGF:

$$I_0(a) = \sup_k \big\{ ka - \lambda_0(k) \big\}. \tag{2.2.9}$$

This method can be applied in principle to processes with resetting, but the generator of X_t in this case is not a pure differential operator: it is a mixed operator involving the pure part (2.2.8) and a singular integral kernel accounting for the delta

source in the Fokker-Planck equation. Finding the SCGF by spectral method then becomes a complicated and singular problem, so that other methods must be used. We propose one in the next section based on the renewal representation of reset processes.

§2.3 Results

We obtain the large deviations of A_T for the process X_t with resetting by studying, following the limit (2.2.6), the time evolution of the generating function:

$$G_r(x, k, t) = E_x\big[e^{tkA_t}\big] = E_x\big[e^{k\int_0^t f(X_s)\mathrm{d}s}\big], \tag{2.3.1}$$

where $E_x[\cdot]$ denotes the expectation with respect to the process X_t with resetting started at $X_0 = x$. Without resetting ($r = 0$), this function is known to evolve according to the Feynman-Kac (FK) formula

$$\frac{\partial}{\partial t}G_0 = \mathcal{L}_k G_0, \tag{2.3.2}$$

which is a parabolic linear partial differential equation for $G_0 = G_{r=0}$ with initial condition $G_0(x, k, 0) = 1$ [84].

A modified FK formula that includes resetting can be derived similarly to the reset-free case by considering an additional time step $\mathrm{d}t$ in the generating function, so as to write

$$
\begin{aligned}
G_r(x, k, t + \mathrm{d}t) &= E_x\big[e^{\int_0^{\mathrm{d}t} f(X_s)\mathrm{d}s}e^{\int_{\mathrm{d}t}^{t+\mathrm{d}t} f(X_s)\mathrm{d}s}\big] \\
&= e^{f(x)\mathrm{d}t}E_x\big[e^{\int_{\mathrm{d}t}^{t+\mathrm{d}t} f(X_s)\mathrm{d}s}\big], \tag{2.3.3}
\end{aligned}
$$

using $X_0 = x$. From this initial state, the process can either reset to $X_{\mathrm{d}t} = x_r$ with probability $r\mathrm{d}t$ or diffuse to $X_{\mathrm{d}t}$ according to the SDE (2.2.1) with the complementary probability $1 - r\mathrm{d}t$, so that

$$
G_r(x, k, t + \mathrm{d}t) = e^{f(x)\mathrm{d}t}\bigg\{r\mathrm{d}t\, G_r(x_r, k, t) \\
+ (1 - r\mathrm{d}t)\int_{-\infty}^{\infty} \mathrm{d}\xi\, K(\xi)G_r(x + \xi, k, t)\bigg\}, \tag{2.3.4}
$$

where $K(\xi)$ is the probability distribution of the increment $X_{\mathrm{d}t} - X_0 = \xi$ as determined from (2.2.1). In this way, we separate the resetting from the pure diffusion (2.2.1). Expanding $G_r(x + \xi, k, t)$ up to second order in ξ and performing the integral then yields

$$\frac{\partial}{\partial t}G_r = (\mathcal{L}_k - r)G_r + rG_r(x_r, k, t) \tag{2.3.5}$$

with the initial condition $G_r(x, k, 0) = 1$.

This modified FK formula with uniform sink and source at x_r is similar to equations obtained for the first-passage problem with resetting [48, 47, 50] and must be solved, as for this problem, by considering the source term $G_r(x_r, k, t)$ as a constant

and by matching the solution $G_r(x, k, t)$ self-consistently for $x = x_r$. This is a difficult task in general, which does not suggest in our experience an efficient way to obtain large deviations, especially since we need the generating function for large times in order to obtain the limit

$$\lambda_r(k) = \lim_{T \to \infty} \frac{1}{T} \ln G_r(x, k, T), \qquad (2.3.6)$$

which is the reset version of (2.2.6).

For the purpose of calculating this limit, a more useful renewal representation of $G_r(x, k, t)$ similar to (2.2.3) can be derived. To this end, assume that the time interval $[0, T]$ witnesses n resettings with periods $\tau_1, \tau_2, \dots, \tau_n$ such that

$$T = \sum_{i=1}^{n+1} \tau_i \qquad (2.3.7)$$

and

$$TA_T = \sum_{i=1}^{n+1} \int_{\sum_{j=1}^{i} \tau_{j-1}}^{\tau_i} f(X_s)\mathrm{d}s, \qquad (2.3.8)$$

where τ_{n+1} is the last period without resetting leading to T. Because of the additive form of A_T, it is clear that G_r can be decomposed, when conditioned on these n resettings, into a product of generating functions G_0 involving only pure diffusion between resettings. To write the full G_r, we then have to sum over all possible reset number and reset times. Since the probability of having a reset at time τ is $re^{-r\tau}$ and the probability of no reset until the time τ is $e^{-r\tau}$, we thus obtain

$$G_r(x, k, T) = \sum_{n=0}^{\infty} \int_0^T \mathrm{d}\tau_1 \, re^{-r\tau_1} G_0(x, k, \tau_1) \int_0^T \mathrm{d}\tau_2 \, re^{-r\tau_2} G_0(x_r, k, \tau_2) \cdots$$

$$\times \int_0^T \mathrm{d}\tau_{n+1} \, e^{-r\tau_{n+1}} G_0(x_r, k, \tau_{n+1}) \, \delta\left(T - \sum_{i=1}^{n+1} \tau_i\right). \qquad (2.3.9)$$

Notice that the first G_0 term starts at the initial condition $X_0 = x$, while the others start after resetting at x_r. The probability of the last period τ_{n+1} is also different from the other periods, since it is determined by the prior n reset periods and the constraint (2.3.7), included in (2.3.9) with the delta function.

To deal with this constraint, it is natural to consider the Laplace transform in time of the generating function

$$\tilde{G}_r(x, k, s) = \int_0^{\infty} G_r(x, k, T)e^{-sT} \, \mathrm{d}T, \qquad (2.3.10)$$

which yields, after integration over the τ_i's,

$$\tilde{G}_r(x, k, s) = \tilde{G}_0(x, k, s+r) \sum_{n=0}^{\infty} r^n \tilde{G}_0(x_r, k, s+r)^n, \qquad (2.3.11)$$

where $\tilde{G}_0$ denotes the Laplace transform of G_0. Assuming that

$$r\tilde{G}_0(x_r, k, s+r) < 1, \qquad (2.3.12)$$

we therefore obtain

$$\tilde{G}_r(x, k, s) = \frac{\tilde{G}_0(x, k, s+r)}{1 - r\tilde{G}_0(x_r, k, s+r)}. \qquad (2.3.13)$$

This is our main result connecting the generating function of A_T with resetting to its generating function without resetting. It can be verified that this formula is equivalent (by Laplace transform) to the modified FK equation (2.3.5), though (2.3.13) is simpler, as it expresses G_r explicitly in terms of the free generating function G_0.

This is more convenient for obtaining large deviations. Assuming that the limit (2.3.6) defining the SCGF $\lambda_r(k)$ of A_T exists implies the following scaling of the generating function:

$$G_r(x, k, T) \sim e^{\lambda_r(k)T} \qquad (2.3.14)$$

as $T \to \infty$, which translates in Laplace space into

$$\tilde{G}_r(x, k, s) \sim \frac{1}{s - \lambda_r(k)}. \qquad (2.3.15)$$

As a result, we see that the SCGF of A_T for the resetting process can be determined by locating the largest (simple and real) pole of the right-hand side of (2.3.13), which is also a zero (in s) of the denominator $1 - r\tilde{G}_0$ when $\tilde{G}_0$ is finite. If $\lambda_r(k)$ is differentiable as a function of k, we then obtain the rate function $I_r(a)$ of A_T similarly to (2.2.9) by taking the Legendre-Fenchel transform of $\lambda_r(k)$.

These calculations are based only on the knowledge of the generating function G_0 of A_T without resetting. In some cases, the large-time asymptotics of that generating function proves to be sufficient to obtain the desired pole $\lambda_r(k)$, which means that the large deviations of A_T for the process with resetting can be obtained directly from the large deviations of A_T without resetting. This important result is illustrated next.

§2.4 Example

We consider in this section the reset Langevin equation (or reset Ornstein-Uhlenbeck process) obtained by adding resettings at x_r with rate r to the diffusion

$$dX_t = -\gamma X_t dt + \sigma dW_t, \qquad (2.4.1)$$

where γ is the friction coefficient, σ is the noise strength, and W_t is the Wiener process. The stationary distribution of this model was studied recently in [103]. The observable that we consider is the integral of the state,

$$A_T = \frac{1}{T} \int_0^T X_t dt. \qquad (2.4.2)$$

This reset process can be considered physically as a simple model of filament dynamics in motility assays [116, 115, 126], wherein filaments are pulled by spring-like motor proteins attached to a substrate at one end and moving on filaments at the other [9]. In this context, A_T represents the mean force exerted on one filament over a time T, which is proportional to the stretch X_t of the motor protein attached to it, while resetting happens when the motor randomly detaches from the filament and a new motor attaches itself with zero stretch [91].

The generating function G_0 of A_T for the reset-free Langevin equation is known in closed form, but its Laplace transform is relatively complicated to work with. For our purpose, it is more convenient to expand G_0, following the FK formula (2.3.2), in spectral form as

$$G_0(x, k, T) = \sum_{i=0}^{\infty} \psi_{k,i}(x) e^{\lambda_{0,i}(k)T}, \tag{2.4.3}$$

where $\lambda_{0,i}(k)$ are the eigenvalues of the tilted generator $\mathcal{L}_k$ without resetting and $\psi_{k,i}$ are the corresponding eigenfunctions. Such a spectral decomposition can be obtained in principle for any Markov process. By symmetrization to the quantum oscillator (see the Appendix), we explicitly find here

$$\lambda_{0,i}(k) = \frac{k^2 \sigma^2}{2\gamma^2} - i\gamma, \quad i = 0, 1, \ldots \tag{2.4.4}$$

and

$$\psi_{k,i}(x) = \frac{(-1)^i \gamma^{-3i/2} k^i \sigma^i e^{\frac{kx}{\gamma} - \frac{3k^2 \sigma^2}{4\gamma^3}} H_i\left(\frac{\sqrt{\gamma}x}{\sigma} - \frac{k\sigma}{\gamma^{3/2}}\right)}{\sqrt{2^i i!}\sqrt{(2i)!!}} \tag{2.4.5}$$

where H_i is ith Hermite polynomial. The SCGF $\lambda_0(k)$ of A_T corresponds to the largest eigenvalue:

$$\lambda_0(k) = \max_i \lambda_{0,i}(k) = \frac{\sigma^2 k^2}{2\gamma^2}. \tag{2.4.6}$$

From the Legendre-Fenchel transform (2.2.9), we thus find the rate function of A_T without resetting to be

$$I_0(a) = \frac{\gamma^2 a^2}{2\sigma^2}, \tag{2.4.7}$$

which implies that the fluctuations of A_T are Gaussian-distributed around $A_T = 0$ [1].

To determine the effect of resetting on these fluctuations, we insert the Laplace transform of the spectral representation (2.4.3),

$$\tilde{G}_0(x, k, s) = \sum_{i=0}^{\infty} \frac{\psi_{k,i}(x)}{s - \lambda_{0,i}(k)}, \tag{2.4.8}$$

into the Laplace formula (2.3.13) and locate the largest pole of the resulting expression for a given truncation $0 \leq i \leq m$. The result is shown for $x_r = 1$, $r = 2$, and various truncation orders m in Fig. 2.1 and compared with the reset-free SCGF $\lambda_0(k)$. As

---

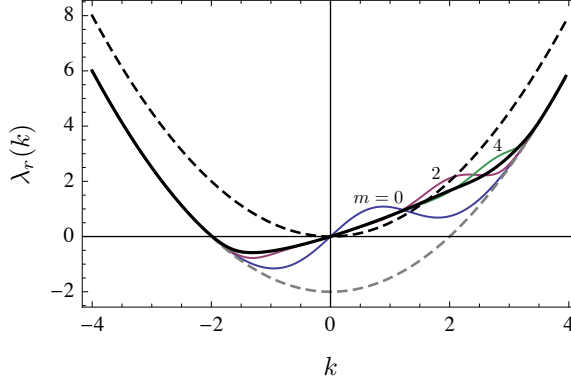[1]This is also evident from the fact that A_T is a linear integral of a linear Gaussian process.

Figure 2.1: (Color online) Dominant pole of $\tilde{G}_r(x, k, s)$ for increasing truncation orders: $m = 0$ (blue), $m = 2$ (purple), $m = 4$ (green). Black curve: Convex $\lambda_r(k)$ obtained for $m \geq 6$. Dashed black curve: Non-reset $\lambda_0(k)$. Dashed gray curve: Tail approximation of $\lambda_r(k)$ shown in (2.4.11). Parameters: $x_r = 1$, $r = 2$, $\gamma = 1$, $\sigma = 1$.

can be seen, the dominant pole is nonconvex in k for low truncation orders, which means that it does not represent a valid SCGF, since SCGFs are always convex by definition [46, 39, 40, 130]. By increasing however the truncation order, the pole does converge to a convex function, identified from (2.3.15) as $\lambda_r(k)$. For the parameter values used in Fig. 2.1, convergence is attained essentially for $m \gtrsim 6$; for larger values of $|x_r|$ or r, more modes are generally required.

This applies to the part of $\lambda_r(k)$ close to $k = 0$, which describes the small fluctuations of A_T. For the large fluctuations associated with the tails of $\lambda_r(k)$, convergence appears immediately for one mode, as can be seen in Fig. 2.1, which implies the following approximation:

$$\lambda_r(k) \approx \lambda_0(k) - r + r\psi_{k,0}(x_r). \tag{2.4.9}$$

Here, we have explicitly

$$\psi_{k,0}(x) = e^{kx/\gamma - 3k^2\sigma^2/(4\gamma^3)}, \tag{2.4.10}$$

so that (2.4.9) can be simplified in fact to

$$\lambda_r(k) \approx \lambda_0(k) - r \tag{2.4.11}$$

for $|k| \to \infty$.

This simple tail behavior of $\lambda_r(k)$ can be understood by noting that very large fluctuations of A_T are brought about, for relatively small reset positions x_r, by long excursions of the process far away from x_r having very few or no reset events. As a result, the renewal representation (2.3.9) is dominated by purely diffusive trajectories whose large deviations are determined by the dominant mode of G_0 as $T \to \infty$. The r factor in (2.4.11) only accounts for the probability of seeing such trajectories without resetting. Conversely, more modes of G_0 are needed to describe the small fluctuations
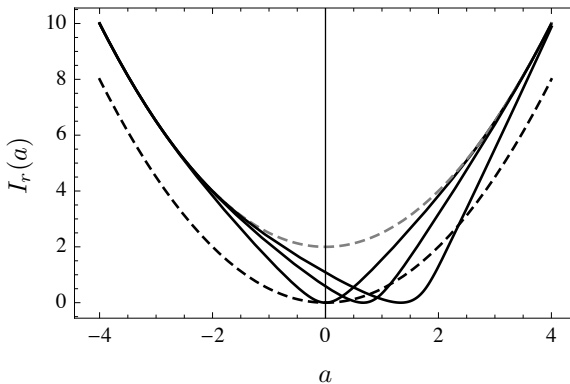
Figure 2.2: Black curves: $I_r(a)$ for $x_r = 0, 1, 2$ (from left to right). The first two curves were obtained for $m = 10$, while the last for $x_r = 2$ was obtained for $m = 20$. Dashed black curve: Non-reset rate function $I_0(a)$. Dashed gray curve: Tail approximation of $I_r(a)$ shown in (2.4.12). Parameters: $r = 2$, $\gamma = 1$, $\sigma = 1$.

of A_T close to x_r because such fluctuations are brought about by trajectories that have many resettings and, therefore, many short diffusive trajectories for which the large deviation limit is not effective. The number of modes m that must be used to recover the correct $\lambda_r(k)$ depends on the parameters used: generally, the larger r or $|x_r|$ is, the higher m should be since resetting takes place more often.

Once that number is set, the rate function $I_r(a)$ can be computed as the Legendre-Fenchel transform of $\lambda_r(k)$. The result is shown in Fig. 2.2 for $r = 2$ and different resetting positions x_r. As expected, the rate function I_r is narrower than I_0 and shifts towards the resetting position x_r, since X_t is more likely with resetting to stay near x_r. Note, however, that the minimum a^* of the rate function, corresponding to the most probable value of A_T in the ergodic limit $T \to \infty$, is not exactly x_r because the friction in the Langevin equation brings X_t near $x = 0$. It is difficult to study this competing effect analytically, since it is strongly linked to resetting, and so cannot be treated perturbatively using a mode expansion of G_0. Numerically, we find that a^* varies linearly with x_r with a slope $c(r)$ shown in Fig. 2.3. As $r \to \infty$, $c(r) \to 1$, and thus $a^* \to x_r$, as expected.

Looking back at Fig. 2.2, we can also see that the tails of $I_r(a)$ are mostly unaffected by resetting, except for a constant shift. This comes again from the large fluctuations of A_T being the result of large diffusive excursions that have very few or no resetting events, so that (2.4.11) holds. Inserting this tail approximation into the Legendre-Fenchel transform leads to the dual approximation

$$I_r(a) \approx I_0(a) + r \tag{2.4.12}$$

as $|a| \to \infty$. This gives a good approximation of the rate function, as can be seen in Fig. 2.2.

This tail result implies with (2.4.7) that A_T has large Gaussian fluctuations, reflecting with a shift its Gaussian fluctuations (2.4.7) seen without resetting. The small fluctuations of A_T around its typical value and mean a^* are also Gaussian, as can
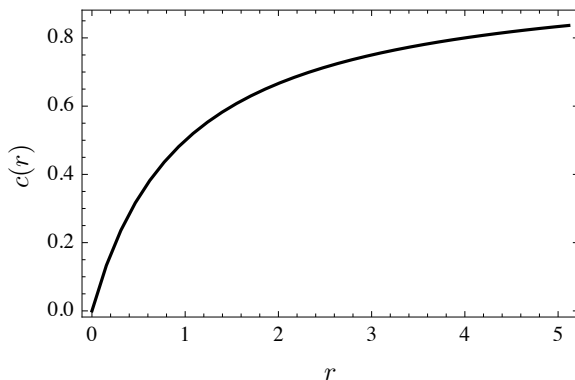
Figure 2.3: Proportionality coefficient $c(r)$ between the minimum a^* of $I_r(a)$ and the resetting position x_r. Parameters: $\gamma = 1$, $\sigma = 1$.

be seen by expanding $I_r(a)$ around its minimum a^*, but with a reset-modified variance determined by $\lambda_r''(0)$ or $I_r''(a^*)^{-1}$ [130]. Finally, in the intermediate region away from a^*, where (2.4.12) is not an accurate approximation of $I_r(a)$, the competition between resetting and diffusion leads to non-Gaussian fluctuations, characterized by the non-parabolic rate function seen in Fig. 2.2.

§2.5 Conclusion

We have derived in this paper a general renewal formula (2.3.13) that can be used to obtain the large deviation functions of additive observables of Markov processes with resetting, and have illustrated this result for the Langevin equation with resetting. Other applications should follow this example either via the exact calculation of the generating function or via the general spectral expansion (2.4.3), keeping in mind for this expansion to include enough modes, as demonstrated, to obtain properly scaled convex cumulant generating functions in the long-time limit.

Although we have considered reset diffusions, it is clear that our main results expressed in terms of generating functions also hold for birth-death and jump processes in general, in addition to Markov chains with resetting or catastrophes, thus opening the way for many other applications. In birth-death processes, for example, one could consider as an observable the total number of births over a given time period or any birth-related cost (e.g., insurances) accumulated in that period. Similarly, for queueing models with resetting, the observable may represent the number of clients entering a queue or any cost associated with clients which is additive in time.

For these examples, we expect the main results that we have obtained for the reset Langevin equation to hold. In particular, it is clear that as long as large fluctuations of A_T are the results of long trajectories involving few resetting, as is the case for the Langevin equation, then the large deviations functions obtained with reset are a shift of the large deviations obtained without reset, following the approximations (2.4.11) and (2.4.12) that we have derived, with the shift coming from the probability

of having few or no resettings over the time T.

For future work, it would be interesting to study whether observables that do not have a large deviation principle without resetting acquire that principle when resetting is introduced. It is known that resetting adds an effective confinement that can transform a non-stationary process (e.g., Brownian motion [48]) into a stationary one, but this might not be enough on its own to force a large deviation principle. Another interesting problem is to generalize our results to observables involving an integral of the increments of the process considered (in the case of pure diffusions) or a sum over its jumps (in the case of pure jump processes); see [25] for more detail. These observables represent physically quantities, such as particle currents and entropy production, playing an important role in nonequilibrium statistical physics.

Appendix

§2.A   Spectral decomposition of the generating function

The generating function $G_0(x, k, T)$ evolves without resetting according to the linear partial differential (2.3.2) and can therefore be decomposed in the eigenbasis of the tilted generator $\mathcal{L}_k$, shown in (2.2.7). For the Ornstein-Uhlenbeck process, $\mathcal{L}_k$ is not hermitian, but can be mapped via a unitary transformation to a hermitian, Schrödinger-type operator, so its spectrum is real. This transformation or symmetrization is the same as the one used for the Fokker-Planck equation; see, e.g., Sec. 5.4 of [109].

Denote by $\rho(x) = e^{-U(x)}$ the stationary distribution of X_t satisfying $L^\dagger \rho = 0$. The symmetrization of $\mathcal{L}_k$ is given by

$$\mathcal{H}_k = \rho^{1/2} \mathcal{L}_k \rho^{-1/2} = e^{-U/2} \mathcal{L}_k e^{U/2}. \tag{2.A.1}$$

For the Ornstein-Uhlenbeck process, we have $U(x) = \gamma x^2/\sigma^2$ up to a constant, which leads to

$$\mathcal{H}_k = \frac{\sigma^2}{2} \frac{d^2}{dx^2} - \frac{\gamma^2 x^2}{2\sigma^2} + \frac{\gamma}{2} + kx. \tag{2.A.2}$$

This is the Schrödinger operator of a shifted and inverted quantum harmonic oscillator with mass $m = 1$ and $\hbar = \sigma$ [85]. From the known spectrum of the harmonic oscillator, we therefore arrive at the eigenvalues (2.4.4). As for the eigenfunctions $\psi_{k,i}$, they are obtained by

$$\psi_{k,i}(x) = \rho(x)^{-1/2} \Psi_{k,i}(x) = e^{U(x)/2} \Psi_{k,i}(x), \tag{2.A.3}$$

where $\Psi_{k,i}$ are the eigenfunctions of $\mathcal{H}_k$, normalized in the usual quantum way.