Energy growth for systems of coupled oscillators with partial damping

Dmitry Dolgopyat*, Bassam Fayad†, Leonid Koralov‡, Shuo Yan§

Abstract

We consider two interacting particles on the circle. The particles are subject to stochastic forcing, which is modeled by white noise. In addition, one of the particles is subject to friction, which models energy dissipation due to the interaction with the environment. We show that, in the diffusive limit, the absolute value of the velocity of the other particle converges to the reflected Brownian motion. In other words, the interaction between the particles are asymptotically negligible in the scaling limit. The proof combines averaging for large energies with large deviation estimates for small energies.

1 Introduction

Understanding energy transfer in complex systems is a fundamental problem in mathematical physics. There is vast literature on this subject and, despite rigorous results in a number of important models, there are still many challenges in our understanding of this phenomenon. Energy transfer plays a key role in several important phenomena including the following:

(a) *Fermi acceleration.* The problem is to describe the motion of a particle in random media where the particle accelerates due to random energy exchange with the environment. Originally introduced by Fermi [9] in order to explain the presence of highly energetic particles in cosmic rays, this model is a subject of intense research (see, e.g., [19, 6] and references therein).

(b) *Fourier Law.* Here the problem is to describe how the heat emitted from some source(s) spreads around a given domain. In particular, one would like to understand the heat conductivity of different materials. We refer the readers to [2, 5, 18, 21] for a review of this subject. The Fourier law remains an active area of research, see e.g. [16] and references therein.

(c) *Energy cascade in turbulence.* In mathematical terms, the problem is to understand how the energy introduced at large scales is transferred to the smaller scales in Hamiltonian PDEs. This problem has certain similarities to the previous problems as it can be reduced to a system of interacting ODEs after passing to the Fourier basis. We refer the readers to [3, 7, 20] for more information on this subject.

The difficulty of the transfer problems comes from the complex interaction network of multi-particle systems (cf. [17]). A simpler situation appears in the rarified regime where each particle interacts during long time intervals only with a fixed small collection of other particles. In an effort to understand such systems, several authors considered local equilibria for a small number of particles subjected to forcing and dissipation (see, e.g., [4, 11, 15] and references therein).

*Dept of Mathematics, University of Maryland, Email: dolgop@umd.edu
†Dept of Mathematics, University of Maryland, Email: bassam@umd.edu
‡Dept of Mathematics, University of Maryland, Email: koralov@umd.edu
§Dept of Mathematics, University of Maryland, Email: shuoyan@umd.edu

In the present paper, we also deal with a simple system of this form. Namely, we consider two one-degree-of-freedom particles interacting via a bounded potential. In addition, we suppose that each particle gains energy via stochastic forcing that is modeled by white noise. The energy dissipation is modeled by friction. We suppose that only one of the particles is subject to dissipation. The question is whether the energy exchange between the particles is strong enough to transfer the excess energy to the particle with the friction and eventually remove it from the system so as to ensure that the total energy of the system does not grow on average. This turns out not to be the case for the system we consider. The mechanism is the following. The first particle loses energy quickly due to friction, so most of the time it has much smaller energy than the second particle. When the second particle has high energy, the relative change in energy is small and so it can be studied using perturbative methods. If we neglect the interaction as well as the forcing, then the motion of the second particle is integrable, and so we can average the interaction along the orbit of the unperturbed system. The averaged interaction vanishes due to the Hamiltonian nature of the two-particle system. This ensures that the energy exchange is too slow at large energies to keep the total energy of the system finite.

Let us now describe our model more precisely. Let V be a twice continuously differentiable function on $\mathbb{S}^1$ and H be defined as

$$H = \frac{1}{2}(r_1^2 + r_2^2) + V(\theta_1 - \theta_2), \text{ for } (r_1, r_2, \theta_1, \theta_2) \in \mathbb{R}^2 \times \mathbb{T}^2.$$

Consider the two-particle Hamiltonian system in $\mathbb{R}^2 \times \mathbb{T}^2$, subject to stochastic forcing and energy dissipation:

$$\begin{cases} dr_1(t) = -V'(\theta_1(t) - \theta_2(t))dt + dW_1(t) \\ dr_2(t) = V'(\theta_1(t) - \theta_2(t))dt + dW_2(t) - r_2(t)dt \\ d\theta_1(t) = r_1(t)dt \\ d\theta_2(t) = r_2(t)dt \end{cases} \tag{1.1}$$

where W_t^1 and W_t^2 are two independent Brownian motions.

Theorem 1.1. *For each initial distribution μ of the processes in* (1.1), *the process $|r_1(t \cdot T)|/\sqrt{T}$ converges weakly, as $T \to \infty$, to a Brownian motion starting and reflected at the origin.*

The proof of Theorem 1.1 is given in Section 6.

2 Expansions

In this section, we obtain preliminary results on the typical behavior of the processes on short time intervals. In particular, given $r_1(t)$, we give an accurate expansion of $r_1(t + \sigma)$, where $\sigma = 1/|r_1(t)|$. Let us introduce some notation first. Let $\mathcal{F}_t$ be the natural filtration generated by the Brownian motions $W_1(t)$ and $W_2(t)$, and $\tilde{\mathcal{F}}_t^T$ be the natural filtration generated by the Brownian motions $W_1(t \cdot T)$ and $W_2(t \cdot T)$. For a random variable A and a function B defined on the parameter space of A, we write $A = O(B)$ (or $A \lesssim B$) if there is a constant $M > 0$ such that $|A| \leq M|B|$, $A = \Theta(B)$ if there is a constant $M > 0$ such that $\|A\|_2 \leq M|B|$, and $A = \tilde{\Theta}(B)$ if there is a constant $M > 0$ such that $\|A\|_4 \leq M|B|$. In particular, the absolute value on $\mathbb{S}^1$ is defined to be the shortest distance to 0. These are often needed to describe the asymptotic behavior of A (e.g., when the time parameter tends to zero). We start our analysis by solving explicitly for $r_1(t)$ and $r_2(t)$:

$$r_1(t) = r_1(0) + W_1(t) - \int_0^t V'(\theta_1(s) - \theta_2(s))ds = r_1(0) + \tilde{\Theta}(\sqrt{t}) + O(t), \tag{2.1}$$

$$
\begin{aligned}
r_2(t) &= e^{-t} r_2(0) + e^{-t}\int_0^t e^s V'(\theta_1(s)-\theta_2(s))ds + e^{-t}\int_0^t e^s dW_2(s) \\
&= e^{-t} r_2(0) + O(1-e^{-t}) + \tilde{\Theta}(\sqrt{1-e^{-2t}}).
\end{aligned} \tag{2.2}
$$

For the other processes, we can write the following expansions:

$$
\begin{aligned}
\theta_1(t) &= \theta_1(0) + \int_0^t [r_1(0) + \tilde{\Theta}(\sqrt{s}) + O(s)]ds = \theta_1 0) + r_1(0)t + O(t^2) + \tilde{\Theta}(t^{3/2}), \\
\theta_2(t) &= \theta_2(0) + \int_0^t \left[e^{-s} r_2(0) + O(1-e^{-s}) + \tilde{\Theta}(\sqrt{1-e^{-2s}})\right] ds \\
&= \theta_2(0) + r_2(0)(1-e^{-t}) + O(t^2) + \tilde{\Theta}(t^{3/2}).
\end{aligned} \tag{2.3}
$$

As our main result, Theorem 1.1, indicates, we aim to show that, away from the origin, the coupling term $-V'(\theta_1(t)-\theta_2(t))$ does not essentially change the behavior of $r_1(t)$ on large time intervals. Namely, if we let $Z(t) = -\int_0^t V'(\theta_1(s)-\theta_2(s))ds$, then

$$
r_1(t) = r_1(0) + W_1(t) + Z(t), \tag{2.4}
$$

and $Z(t)$ is expected to be small compared with $W_1(t)$ on large time scales. When $r_1(t)$ is large and $r_2(t)$ is relatively small, with $\sigma = 1/|r_1(0)|$, we would like to show that $Z(\sigma)$ is small compared with σ. Eventually, we'll be interested in asymptotically small values of σ, but the following result holds for all $\sigma \leq 1$.

Lemma 2.1. *If $|r_1(0)| \geq 1$ and $\sigma = 1/|r_1(0)|$, then*

$$
Z(\sigma) = r_2(0)O(\sigma^2) + O(\sigma^3) + \tilde{\Theta}(\sigma^{5/2}), \tag{2.5}
$$

$$
Z(\sigma) = \frac{r_2(0)}{r_1(0)^2} V'(\theta_1(0)-\theta_2(0)) + (r_2(0)^2+1)O(\sigma^3) + \Theta(\sigma^{5/2}). \tag{2.6}
$$

It is worth noting that, compared to (2.5), (2.6) provides a exact form the leading term. However, one of the error terms depends on $r_2(0)$.

Proof. The proof utilizes the expansions we obtain above and Taylor's expansion of $V'(\theta_1(s)-\theta_2(s))$ at $\theta_1(0)-\theta_2(0)+r_1(0)s$ for $s \in [0,\sigma]$. Below, to simplify our notation, we denote $\boldsymbol{r}_i = r_i(0)$ and $\boldsymbol{\theta}_i = \theta_i(0)$ for $i = 1, 2$. By (2.3),

$$
\begin{aligned}
Z(\sigma) &= -\int_0^\sigma V'(\theta_1(s)-\theta_2(s))ds \\
&= -\int_0^\sigma V'\left(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \boldsymbol{r}_1 s - \boldsymbol{r}_2(1-e^{-s}) + O(s^2) + \tilde{\Theta}(s^{3/2})\right) ds \\
&= -\int_0^\sigma V'(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \boldsymbol{r}_1 s)ds + \int_0^\sigma \boldsymbol{r}_2 O(1-e^{-s}) + O(s^2) + \tilde{\Theta}(s^{3/2}))ds \\
&= \boldsymbol{r}_2 O(\sigma^2) + O(\sigma^3) + \tilde{\Theta}(\sigma^{5/2}).
\end{aligned} \tag{2.7}
$$

This proves (2.5). Next, (2.7) gives

$$
\begin{aligned}
Z(\sigma) = &-\int_0^\sigma V'(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \boldsymbol{r}_1 s)ds \\
&+ \int_0^\sigma V''(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \boldsymbol{r}_1 s)[\boldsymbol{r}_2(1-e^{-s}) + O(s^2) + \tilde{\Theta}(s^{3/2})]ds + \boldsymbol{r}_2^2 O(\sigma^3) + O(\sigma^5) + \Theta(\sigma^4)
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^\sigma V''(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \boldsymbol{r}_1 s)\boldsymbol{r}_2 s ds \\
&\quad + \int_0^\sigma V''(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2 + \boldsymbol{r}_1 s)[(\boldsymbol{r}_2 O(s^2) + O(s^2) + \tilde{\Theta}(s^{3/2})]ds + \boldsymbol{r}_2^2 O(\sigma^3) + O(\sigma^5) + \Theta(\sigma^4) \\
&= \frac{\boldsymbol{r}_2}{\boldsymbol{r}_1^2} V'(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) + \boldsymbol{r}_2 O(\sigma^3) + O(\sigma^3) + \tilde{\Theta}(\sigma^{5/2}) + \boldsymbol{r}_2^2 O(\sigma^3) + O(\sigma^5) + \Theta(\sigma^4) \\
&= \frac{\boldsymbol{r}_2}{\boldsymbol{r}_1^2} V'(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) + (\boldsymbol{r}_2^2 + 1) O(\sigma^3) + \Theta(\sigma^{5/2}). \qquad \square
\end{aligned}
$$

Let us briefly examine the right-hand side of (2.6) in order to motivate our further analysis. This formula will be used iteratively to express the increment of r_1 on longer time intervals. On short time scales (of order $\sigma = 1/|r_1|$, as above), the difference between the increment of r_1 and the Brownian motion is small (of order $1/r_1^2(0)$), provided that r_2 is bounded and r_1 is large. However, at larger times, the cumulative effect of the terms containing $V'(\theta_1 - \theta_2)$ needs to be controlled. Namely, we'll use an averaging effect to show that the sum of N such terms (each corresponding to a new starting point obtained after a new rotation) is much smaller than N/r_1^2, provided that N is large (but not too large so that r_1 does not change much on the corresponding time interval).

Our plan is as follows. First, in Section 3, we obtain various estimates on r_2 and its joint distribution with θ_2. Those will help us to control the terms with the powers of r_2 on the right-hand side of (2.6) and prove averaging for the term containing $V'(\theta_1 - \theta_2)$. Next, in Section 4, we will show that on relatively large time intervals r_1 behaves as a Brownian motion plus a drift term that is small, provided that r_1 is large. The times when r_1 is near the origin need to be analyzed separately. In Section 5, we show that the exit times from a neighborhood of zero can be bounded from above. After we control the number of excursions to the origin (see Lemma 6.2), it will be seen that the time spent by the process near the origin can be ignored. Finally, the main result is proved in Section 6.

In the rest of the paper, we make statements with assumptions on the initial condition of r_1 and r_2. Unless specified otherwise, the results hold uniformly in all the initial conditions of θ_1 and θ_2. In addition, for simplicity, we assume that the function V is bounded by 1 along with its first and second derivatives. From the proofs, it will be easy to see that the results hold for all twice continuously differentiable V.

3 Behavior of r_2 and θ_2

We start this section by giving a result about the supremum of $r_2(t)$ during a large interval of time. It demonstrates that, most of the time, $r_2(t)$ can be expected to be relatively small.

3.1 Growth of r_2

Proposition 3.1. *Consider a function D such that $D(T) \to \infty$ as $T \to \infty$. Then for each $t > 0$ there is $T_0(t)$ such that for $T \geq T_0(t)$ we have that for each initial distributions of $r_1(0)$ and $r_2(0)$*

$$
\mathbf{P}(\sup_{0 \leq s \leq t \cdot T} |r_2(s)| > |r_2(0)| + D(T)) \leq 18tTD(T)e^{-(D(T)-1)^2} \tag{3.1}
$$

Note that the right-hand side of (3.1) converges to 0 as $T \to \infty$ provided that $D(T) \geq 2\sqrt{\log T}$.

Proof. Let $f(x) = \int_0^x e^{y^2} dy$. By (2.2), we know that

$$
\mathbf{P}(\sup_{0 \leq s \leq t \cdot T} |r_2(s)| > |r_2(0)| + D(T)) \leq \mathbf{P}(\sup_{0 \leq s \leq t \cdot T} |r(s)| > D(T) - 1), \tag{3.2}
$$

where $r(s)$ is the Ornstein-Uhlenbeck process defined as:

$$dr(s) = -r(s)dt + dW(s), \; r(0) = 0. \tag{3.3}$$

Then it is not hard to see that $f(r(s))$ is a martingale. In order to apply Doob's martingale inequality, we compute an upper bound for $\mathbf{E}(f(r(s)) \vee 0)$. Notice that $xf(x) < e^{x^2}$ and $r(s) \sim \mathcal{N}(0, \frac{1}{2}(1-e^{-2s}))$. Hence, for all s sufficiently large,

$$\begin{aligned}
\mathbf{E}(f(r(s)) \vee 0) &= \mathbf{E}\chi_{\{0<r(s)<1\}} f(r(s)) + \mathbf{E}\chi_{\{r(s)\geq 1\}} f(r(s)) \\
&\leq e + \mathbf{E}\chi_{\{r(s)\geq 1\}} e^{r(s)^2}/r(s) \leq e + \int_1^\infty \frac{1}{x} \exp(x^2 - \frac{x^2}{1-e^{-2s}})dx \\
&= e + \int_1^\infty \frac{1}{x} \exp(-\frac{x^2}{e^{2s}-1})dx = e + \int_0^\infty \exp(-e^{2y}/(e^{2s}-1))dy \\
&\leq e + \int_0^\infty \exp(-e^{2y-2s})dy \leq e + 2s + \int_{2s}^\infty \exp(-e^{2y-2s})dy \\
&\leq e + 2s + \int_{2s}^\infty \exp(-y^2)dy \leq 3s.
\end{aligned}$$

So, by Doob's martignale inequality, for all T sufficiently large,

$$\begin{aligned}
&\mathbf{P}(\sup_{0\leq s\leq t\cdot T} r(s) > D(T) - 1) = \mathbf{P}(\sup_{0\leq s\leq t\cdot T} f(r(s)) > f(D(T)-1)) \\
&\leq \mathbf{E}(f(r(tT)) \vee 0)/f(D(T)-1) \leq 3tT/f(D(T)-1).
\end{aligned}$$

Since $r(s)$ is symmetric, the right-hand side of (3.2) is at most $6tT/f(D(T)-1)$. The inequality in (3.1) follows from the fact that $3xf(x) \geq \exp(x^2)$ for all x sufficiently large. □

Now let us focus on the situation where $|r_1(0)| = R$ is large and $|r_2(0)| \leq R^\alpha$, with $0 < \alpha < 1$, is relatively small. This situation is indeed representative, since we expect $|r_1(t\cdot T)|/\sqrt{T}$ to behave like a Brownian motion reflected at the origin, hence it is typically of order $\sqrt{T}$ during time of order T, while Proposition 3.1 indicates that $|r_2(t)|$ typically stays below $2\sqrt{\log T}$. We are interested in the behavior of the processes at time $0 \leq t(R) \leq R^{\alpha_t}$, where $\alpha_t > 0$ is a constant less than $2/3$. The subscript t is used to stress that the constant appears in the upper bound on the function $t(R)$. During this time, it is unlikely for $r_1(t)$ to get too far away from its original location. Namely, let $c(R) = R^{\alpha_c}$ where $\alpha_t/2 < \alpha_c < 1/3$. We will see that the probability for $r_1(t)$ to exit from $[R - c(R), R + c(R)]$ within time $t(R)$ is small. In order to describe the evolution of the process $r_1(t)$, we will use the expansions in Section 2 and divide the time $t(R)$ into small intervals that correspond to the full rotations of $\theta_1(t)$. Let $\tau_c = \inf\{t : |r_1(t) - r_1(0)| = c(R)\}$. Define

$$\begin{aligned}
&\sigma_0 = 0, \\
&\sigma_{k+1} = \sigma_k + \chi_{\{\sigma_k < t(R)\wedge\tau_c\}} 1/|r_1(\sigma_k)|, \; k \geq 0,
\end{aligned} \tag{3.4}$$

and $\tilde{n} = \inf\{k : \sigma_k \geq t(R) \wedge \tau_c\} \leq t(R)(R + 2c(R))$. The large parameter R, the related functions $t(R)$, $c(R)$, the stopping times τ_c, σ_k, $k \geq 0$, and the quantity $\tilde{n}$ are frequently used in the remainder of the paper. Although we will choose different $t(R)$ and $c(R)$ in different results and their proofs, unless otherwise specified, the stopping times and $\tilde{n}$ are always defined as above. Since the error terms in (2.6) involve $r_2(t)$ and $r_2(t)^2$, we need bounds on $r_2(t)$ at those full rotation time steps σ_k, $0 \leq k \leq \tilde{n}$, which explains the need for the next lemma.

Lemma 3.2. *Let $|r_1(0)| = R$, $|r_2(0)| \le R^\alpha$, where $0 \le \alpha < 1$, $0 \le t(R) \le R^{\alpha_t}$, where $\alpha_t < 2/3$, and $c(R) = R^{\alpha_c}$, where $\alpha_t/2 < \alpha_c < 1/3$. Then we have, for all R large and each $k \ge 0$,*

$$\mathbf{E}\chi_{\{k\le\tilde{n}\}}|r_2(\sigma_k)|^4 \le (2r_2(0)e^{-k/R} + 3)^4, \tag{3.5}$$

where σ_k, $k \ge 0$, and $\tilde{n}$ is defined as in (3.4).

Proof. Since we have good control over $r_2(t)$ at deterministic times, we will compare $r_2(\sigma_k)$ with $r_2(k/R)$. Note that σ_k is close to k/R if $k \le \tilde{n}$. To be more precise,

$$\begin{aligned}\left|\chi_{\{k\le\tilde{n}\}}\left(\sigma_k - \frac{k}{R}\right)\right| &\le \chi_{\{k\le\tilde{n}\}}\sum_{j=0}^{k-1}\left|\sigma_{j+1} - \sigma_j - \frac{1}{R}\right| \\ &\le \frac{t(R)}{1/(R+2c(R))}\left(\frac{1}{R-c(R)} - \frac{1}{R}\right) \le \frac{2t(R)c(R)}{R}.\end{aligned} \tag{3.6}$$

By the explicit solution in (2.2), we know that

$$\|r_2(t)\|_4 \le r_2(0)e^{-t} + 2. \tag{3.7}$$

Note that for each $k \in \mathbb{N}$, we have

$$r_2(\sigma_k) = r_2\left(\frac{k}{R}\right)e^{\frac{k}{R}-\sigma_k} + e^{-\sigma_k}\int_{k/R}^{\sigma_k} e^s V'(\theta_1(s) - \theta_2(s))ds + e^{-\sigma_k}\int_{k/R}^{\sigma_k} e^s dW_s. \tag{3.8}$$

(In the case of $\sigma_k < k/R$, $\int_{k/R}^{\sigma_k} e^s dW_s$ means $-\int_{\sigma_k}^{k/R} e^s dW_s$.) Then it follows that

$$\begin{aligned}\|\chi_{\{k\le\tilde{n}\}}(r_2(\sigma_k) - r_2(k/R))\|_4 &\le \|\chi_{\{k\le\tilde{n}\}}(r_2\left(\frac{k}{R}\right) - r_2\left(\frac{k}{R}\right)e^{\frac{k}{R}-\sigma_k})\|_4 \\ +\|\chi_{\{k\le\tilde{n}\}}e^{-\sigma_k}\int_{k/R}^{\sigma_k} e^s V'(\theta_1(s)-\theta_2(s))ds\|_4 &+ \|\chi_{\{k\le\tilde{n}\}}e^{-\sigma_k}\int_{k/R}^{\sigma_k} e^s dW_s\|_4.\end{aligned} \tag{3.9}$$

Here the first term on the right-hand side is bounded by

$$\begin{aligned}\|\chi_{\{k\le\tilde{n}\}}(r_2\left(\frac{k}{R}\right) - r_2\left(\frac{k}{R}\right)e^{\frac{k}{R}-\sigma_k})\|_4 &\le \frac{3t(R)c(R)}{R}\|\chi_{\{k\le\tilde{n}\}}r_2(k/R)\|_4 \\ &\le \frac{3t(R)c(R)}{R}\|r_2(k/R)\|_4 \le \frac{3t(R)c(R)}{R}(r_2(0)e^{-k/R} + 2).\end{aligned}$$

The second term in (3.9) is bounded by $2|\chi_{\{k\le\tilde{n}\}}(\sigma_k - k/R)| \le \frac{4t(R)c(R)}{R} \to 0$. The third term is bounded by the fourth root of:

$$\mathbf{E}|\chi_{\{k\le\tilde{n}\}}e^{-\sigma_k}\int_{k/R}^{\sigma_k} e^s dW_s|^4 = \mathbf{E}|\chi_{\{k\le\tilde{n}\}}e^{k/R-\sigma_k}\int_{k/R}^{\sigma_k} e^{s-k/R}dW_s|^4.$$

Note that the integral in the right-hand side is a time change of a Brownian motion. Therefore

$$\begin{aligned}\mathbf{E}|\chi_{\{k\le\tilde{n}\}}e^{-\sigma_k}\int_{k/R}^{\sigma_k} e^s dW_s|^4 &= \mathbf{E}|\chi_{\{k\le\tilde{n}\}}e^{k/R-\sigma_k}\widetilde{W}(\frac{1}{2}(e^{2(\sigma_k-k/R)} - e^{-2k/R})) - \widetilde{W}(\frac{1}{2}(1-e^{-2k/R}))|^4 \\ &\le \mathbf{E}\sup_{[-\frac{4t(R)c(R)}{R},\frac{4t(R)c(R)}{R}]}|\widetilde{W}(\frac{1}{2}(e^{2t} - e^{-2k/R})) - \widetilde{W}(\frac{1}{2}(1-e^{-2k/R}))|^4 \to 0,\end{aligned}$$

where $\widetilde{W}$ is another Brownian motion. So we have the estimate on the L^4 norm of difference

$$\left\|\chi_{\{k\le\tilde{n}\}}\left(r_2(\sigma_k) - r_2\left(\frac{k}{R}\right)\right)\right\|_4 = o(r_2(0)e^{-k/R} + 1), \tag{3.10}$$

proving the desired result. □

3.2 Invariant measure

The next objective is to show that the third term in (2.6) will be averaged to $o(r_1^{-2})$ if one considers an appropriate number of rotations. To this end we show that, if $|r_2|$ is much smaller than $|r_1|$, then the particles asymptotically decouple, so time averages involving r_2 could be well approximated by the invariant measure of the decoupled process.

Lemma 3.3. *Let (r, θ) be the Markov process on $\mathbb{R} \times \mathbb{S}^1$ defined by*

$$\begin{aligned} dr(t) &= -r(t)dt + dW(t), \\ d\theta(t) &= r(t)dt. \end{aligned} \tag{3.11}$$

Then $\mathbf{E}[r(t)V'(\theta(t) - \theta_0)] \to 0$ as $t \to \infty$, uniformly in $r(0)$ on a compact set, $\theta(0) \in \mathbb{S}^1$, and $\theta_0 \in \mathbb{S}^1$.

Note that in (3.11), r is an Ornstein-Uhlenbeck process. In particular, the stationary measure for this process is the following: r has Gaussian distribution on $\mathbb{R}$ and θ is uniformly distributed on the circle and independent of r.

Proof. The process $(r(t), \theta(t))$ on $\mathbb{R} \times \mathbb{S}^1$ has the generator L defined by $Lu = -ru'_r + ru'_\theta + \frac{1}{2}u''_{rr}$. The unique invariant measure has the density $p(r, \theta) = \frac{1}{\sqrt{\pi}}e^{-r^2}$, and

$$\int_{\mathbb{R}\times\mathbb{S}^1} rV'(\theta - \theta_0) \cdot p(r, \theta)drd\theta = 0. \tag{3.12}$$

We show that $(r(t), \theta(t))$ converges, in total variation, to the invariant measure uniformly in the initial data on a compact set. (Note that the convergence in total variation itself is not enough since we have an unbounded function in the expectation. However, it is easy to see that $r(t) = r(0)e^{-t} + e^{-t}\int_0^t e^s dW_s$, and hence $r(t)$ has bounded L^2-norm for all non-negative t uniformly in $r(0)$ on a compact set. Combining both facts, with the help of the Cauchy-Schwarz inequality, we can obtain the desired result.) Here we use Harris's Theorem (see [10, 12], and the original statement in [13]). The Harris Theorem implies the convergence in total variation if there is a function U such that the following conditions are satisfied (see Assumption 3.1, Exercise 3.3, and Assumption 3.4 in [10]):

(i) There exist positive constants c and K such that $LU \le K - cU$;

(ii) For every $R > 0$, there exist $\alpha > 0$ and $t > 0$ such that $\mathrm{TV}(\mu_x^t, \mu_y^t) < 2(1-\alpha)$, if $U(x)+U(y) \le R$, where μ_x^t and μ_y^t are the measures induced by $(r(t), \theta(t))$ with starting points x and y, respectively, and TV is the total variation distance.

Moreover, the convergence of μ_x^t in total variation, as $t \to \infty$, is uniform in x on $\{U(x) \le R\}$ for any $R > 0$. We take $U(r, \theta) = r^2$. Part (i) is trivial since $LU = 1 - 2U$. To prove part (ii), it is enough to show that μ_x^t, $t > 0$, has a density function $p_x^t(x')$ that is positive everywhere and continuous in both x and x'. This follows from Example 3.1 in [14]. ☐

The next lemma provides a similar result for the process (r_2, θ_2), which is not a Markov one by itself. The idea is to wait for r_2 to decrease to a compact set and, from then, we consider a time scale on which (r_2, θ_2) is close to its Markov analogue defined as in (3.11), while the process in (3.11) has the desired properties.

Lemma 3.4. *Suppose that* $|r_1(0)| = R$, $|r_2(0)| \leq R^{\alpha}$, *where* $\alpha \in (0,1)$. *For each* $\beta \in (0,1)$, *we have that* $\mathbf{E}[r_2(t)V'(\theta_2(t) - \theta_0)] \to 0$ *uniformly in all* $|\log R|^2 \leq t \leq R^{\beta}$ *and* $\theta_0 \in \mathbb{S}^1$, *as* $R \to \infty$.

Proof. Consider the distribution of r_2 at time $t - \log R \geq \log R$. Recall that $r_2(u)$ can be expressed explicitly:

$$r_2(u) = e^{-u} r_2(0) + e^{-u} \int_0^u e^s V'(\theta_1(s) - \theta_2(s))ds + e^{-u} \int_0^u e^s dW_2(s). \tag{3.13}$$

Then it follows that $\|r_2(t - \log R)\|_2 \leq 4$ for R sufficiently large. On the other hand, since $t \leq R^{\beta}$, $\mathbf{P}(|r_1(t - \log R) - R| > 2R^{\beta}) \to 0$. By the boundedness of $\|r_2(t)\|_2$ and the Cauchy-Schwarz inequality, it suffices to prove that, for each $C > 0$, given that $|r_2(0)| \leq C$ and $|r_1(0) - R| \leq 2R^{\beta}$,

$$\mathbf{E} r_2(\log R) V'(\theta_2(\log R) - \theta_0) \to 0, \text{ as } R \to \infty. \tag{3.14}$$

We will apply Lemma 3.2 to prove this. We start by choosing $t(R) = 2\log R$ and $c(R) = R^{1/4}$ and recalling the definitions of τ_c, σ_k, $k \geq 0$, and $\tilde{n}$ in (3.4). Since $(r_2(u), \theta_2(u))$ is not a Markov process, we need to compare $(r_2(u), \theta_2(u))$ with $(r(u), \theta(u))$ defined in (3.11) with starting point $(r_2(0), \theta_2(0))$. Note $r(u)$ can also be expressed explicitly:

$$r(u) = e^{-u} r_2(0) + e^{-u} \int_0^u e^s dW_2(s). \tag{3.15}$$

By (2.5), we have that if $k < \tilde{n}$, then

$$\begin{aligned}
&\int_{\sigma_k}^{\sigma_{k+1}} e^s V'(\theta_1(s) - \theta_2(s))ds \\
&= e^{\sigma_k} \int_0^{\sigma_{k+1} - \sigma_k} e^s V'(\theta_1(\sigma_k + s) - \theta_2(\sigma_k + s))ds \\
&= e^{\sigma_k} \int_0^{\sigma_{k+1} - \sigma_k} V'(\theta_1(\sigma_k + s) - \theta_2(\sigma_k + s))ds \\
&\quad + e^{\sigma_k} \int_0^{\sigma_{k+1} - \sigma_k} (e^s - 1) V'(\theta_1(\sigma_k + s) - \theta_2(\sigma_k + s))ds \\
&= e^{\sigma_k} [(r_2(\sigma_k) + 1) O(R^{-2}) + O(R^{-3}) + \tilde{\Theta}(R^{-5/2})].
\end{aligned}$$

Then, for each $0 \leq n \leq t(R)R/2$, we obtain with the help of Lemma 3.2,

$$\begin{aligned}
\mathbf{E}|r_2(\sigma_n) - r(\sigma_n)| &= \mathbf{E}\left| e^{-\sigma_n} \int_0^{\sigma_n} e^s V'(\theta_1(s) - \theta_2(s))ds \right| \\
&\leq \mathbf{E} \sum_{k=0}^{n-1} \chi_{\{k < \tilde{n}\}} \left| e^{-\sigma_{k+1}} \int_{\sigma_k}^{\sigma_{k+1}} e^s V'(\theta_1(s) - \theta_2(s))ds \right| \\
&= O(t(R)/R).
\end{aligned} \tag{3.16}$$

For any $u \leq \log R$, we take $\tilde{k}(u) = \lfloor uR \rfloor \leq t(R)R/2$ and compare $r_2(u)$ with $r_2(\tilde{k}(u)/R)$.

$$\mathbf{E}|r_2(u) - r_2(\tilde{k}(u)/R)| \leq \mathbf{E}|\int_{\tilde{k}(u)/R}^{u} r_2(s)ds| + \frac{1}{R} + \mathbf{E}|\int_{\tilde{k}(u)/R}^{u} dW_2(s)| = O\left(\frac{1}{\sqrt{R}}\right). \tag{3.17}$$

On the other hand,

$$\mathbf{E}|r_2(\tilde{k}(u)/R) - r_2(\sigma_{\tilde{k}(u)})| = \mathbf{E}\chi_{\{\tilde{k}(u) < \tilde{n}\}} |r_2(\tilde{k}(u)/R) - r_2(\sigma_{\tilde{k}(u)})| \tag{3.18}$$

$$+\mathbf{E}\chi_{\{\tilde{k}(u)\geq\tilde{n}\}}|r_2(\tilde{k}(u)/R)-r_2(\sigma_{\tilde{k}(u)})|. \tag{3.19}$$

The first term (3.18) can be bounded by $O\left(\sqrt{\frac{t(R)c(R)}{R}}\right)=o\left(\frac{1}{\sqrt[3]{R}}\right)$ as in Lemma 3.2, and the second term (3.19) can be bounded by

$$\mathbf{E}\chi_{\{\tau_c<\log R\}}|r_2(\tilde{k}(u)/R)-r_2(\sigma_{\tilde{n}})|=o(1/R), \tag{3.20}$$

since $\mathbf{P}(\tau_c<\log R)$ is exponentially small w.r.t. R and L^2 norm of $r_2(\sigma_{\tilde{n}})$ can be bounded by

$$\|r_2(\sigma_{\tilde{n}})\|_2\leq C+1+\|e^{-\sigma_{\tilde{n}}}\int_0^{\sigma_{\tilde{n}}}e^s dW_s\|_2\leq C+1+\|\int_0^{\sigma_{\tilde{n}}}e^s dW_s\|_2\leq C+1+\sqrt{\mathbf{E}\exp(2\sigma_{\tilde{n}})}\leq C+1+R^3$$

since $\sigma_{\tilde{n}}<3\log R$. Thus, by combining this with (3.16) and (3.17) and noticing that the same argument applies to r, we have, for each $u\leq\log R$,

$$\mathbf{E}|r_2(u)-r(u)|=O(t(R)/R)+O(R^{-1/2})+o(R^{-1/3})+o(1/R)=o(R^{-1/4}), \tag{3.21}$$

as R tends to infinity. By integrating r_2-r over time u, we obtain

$$\mathbf{E}|\theta_2(u)-\theta(u)|=o(\log R\cdot R^{-1/4}). \tag{3.22}$$

(3.14) and, consequently, the desired result follow from Lemma 3.3 and the fact that

$$\begin{aligned}&\mathbf{E}|r_2(u)V'(\theta_2(u)-\theta_0)-r(u)V'(\theta(u)-\theta_0)|\\&\leq\mathbf{E}|(r_2(u)-r(u))V'(\theta(u)-\theta_0)|+\|r_2(u)\|_2\|V'(\theta_2(u)-\theta_0)-V'(\theta(u)-\theta_0)\|_2\\&\leq\mathbf{E}|r_2(u)-r(u)|+(C+2)\cdot\sqrt{\mathbf{E}|\theta_2(u)-\theta(u)|}\to 0.\end{aligned}$$

□

4 Behavior of r_1 away from the origin

In this section, we will deal with different temporal and spatial scales that are powers of R. Those exponents will be denoted by α with different subscripts. We will prove that, if initially $|r_1(0)|$ is large and $|r_2(0)|$ is relatively small, then, after a certain time, the increment of $r_1(t)$ is made up of a Brownian motion and a small correction term. Recall $Z(t)=-\int_0^t V'(\theta_1(s)-\theta_2(s))ds$, and

$$r_1(t)=r_1(0)+W_1(t)+Z(t). \tag{4.1}$$

Let $X_t^T=r_1(t\cdot T)/\sqrt{T}$. The main result of this section essentially states that, if we start with X_0^T of order 1 and stop it when X reaches a fixed (small) level ε, then the resulting process converges as $T\to\infty$ to the Brownian motion stopped at the same level (in fact, in order to complete the proof of convergence we need a tightness result (Lemma 6.4) proved in Section 6).

Lemma 4.1. *For $\varepsilon>0$, let $\eta=\inf\{s: X_s^T=\varepsilon\}$. Then, for each $t>0$ and bounded $f\in\mathbf{C}^3([0,+\infty))$ with bounded derivatives up to order three, we have*

$$\mathbf{E}\left[f(X_{t\wedge\eta}^T)-f(X_0^T)-\frac{1}{2}\int_0^{t\wedge\eta}f''(X_s^T)ds\right]\to 0, \tag{4.2}$$

as $T\to\infty$, uniformly in X_0^T on a compact set $K\subset[\varepsilon,+\infty)$ and $|r_2(0)|\leq 2\log T$.

The proof is quite long, so we divide it into several steps.

4.1 Short time behavior

Proposition 4.2. *Suppose that* $|r_1(0)| = R$, $|r_2(0)| \le R^\alpha$, *where* $\alpha \in (0, \frac{2}{3})$. *Suppose that* α_t *and* α_t' *satisfy* $\alpha < \alpha_t' < \alpha_t < 2/3$. *Then, uniformly in* $R^{\alpha_t'} \le t(R) \le R^{\alpha_t}$,

$$\mathbf{E}Z(t(R))^2 = O((t(R)/R)^2),\ \mathbf{E}Z(t(R)) = o(t(R)/R),\ \textit{as } R \to \infty. \tag{4.3}$$

The proof of this proposition is deferred to later after we obtain technical results that will be needed. The next lemma provides a rough estimate that shows that $r_1(t)$ cannot exit a large interval too quickly.

Lemma 4.3. *Suppose that* $|r_1(0)| = R$, $|r_2(0)| \le R^\alpha$, *where* $\alpha \in (0, \frac{2}{3})$. *For each* α_t, α_t', *and* α_c *that satisfy* $\alpha < \alpha_t' < \alpha_t < 2/3$ *and* $\alpha_t/2 < \alpha_c < 1/3$, *uniformly in* $R^{\alpha_t'} \le t(R) \le R^{\alpha_t}$,

$$\mathbf{P}(\tau_c < t(R)) = O((t(R)/Rc(R))^4),\ \textit{as } R \to \infty, \tag{4.4}$$

where $c(R) = R^{\alpha_c}$ *and* $\tau_c = \inf\{t : |r_1(t) - r_1(0)| = c(R)\}$.

Proof. Let the stopping times σ_k, $k \ge 0$, be defined as in (3.4). We will use the expansion (2.5) to control the desired probability.

$$\begin{aligned} r_1(\sigma_{\tilde n}) &= r_1(0) + W_1(\sigma_{\tilde n}) + \sum_{k=0}^{\tilde n - 1} r_2(\sigma_k) \cdot O(R^{-2}) + O(t(R)R^{-2}) + \tilde\Theta(t(R)R^{-3/2}) \\ &=: r_1(0) + W_1(\sigma_{\tilde n}) + \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3. \end{aligned}$$

By Lemma 3.2 and the fact that $\tilde n < 2t(R)R$, we obtain that $\|\mathcal{A}_1\|_4 = O(t(R)/R)$. Since $\sigma_{\tilde n} < t(R) + 2/R$, the probability of $|W_1(\sigma_{\tilde n})|$ being significantly larger than $\sqrt{t(R)}$ is exponentially small. So

$$\begin{aligned} \mathbf{P}(|r_1(\sigma_{\tilde n}) - r_1(0)| > \frac{4c(R)}{5}) &\le \mathbf{P}(|W_1(\sigma_{\tilde n})| > c(R)/5) + \sum_{j=1}^{3} \mathbf{P}(|\mathcal{A}_j| > c(R)/5) \\ &\lesssim \frac{1}{R^4} + \frac{t(R)^4}{R^4 c(R)^4} + 0 + \frac{t(R)^2}{R^4 c(R)^6} \le \frac{t(R)^4}{R^4 c(R)^4}. \end{aligned}$$

On the other hand, since on the event $\{\tau_c < t(R)\}$, we have $0 \le \sigma_{\tilde n} - \tau_c < 2/R$. Then,

$$\begin{aligned} &\mathbf{P}(\tau_c < t(R), |r_1(\sigma_{\tilde n}) - r_1(0)| \le \frac{4c(R)}{5}) \\ &\le \mathbf{P}(\tau_c < t(R), |r_1(\tau_c) - r_1(\sigma_{\tilde n})| > c(R)/5) \\ &\le \mathbf{P}(\tau_c < t(R), |\int_{\tau_c}^{\sigma_{\tilde n}} V'(\theta_1(s) - \theta_2(s) ds| + |W_1(\sigma_{\tilde n}) - W_1(\tau_c)| > c(R)/5) \\ &\le \mathbf{P}(\tau_c < t(R), |W_1(\sigma_{\tilde n}) - W_1(\tau_c)| > c(R)/6) \\ &\le \mathbf{P}(\sup_{0 \le s \le 2/R} |W_1(\tau_c + s) - W_1(\tau_c)| > c(R)/6) \end{aligned}$$

which is exponentially small w.r.t. R. The proof is completed by taking the sum of two probabilities. □

Lemma 4.4. *Suppose that* $|r_1(0)| = R$, $|r_2(0)| \le R^\alpha$, *where* $\alpha \in (0, \frac{2}{3})$. *Suppose that* α_t, α_t', *and* α_c *satisfy* $\alpha < \alpha_t' < \alpha_t < 2/3$ *and* $\alpha_t/2 < \alpha_c < 1/3$. *Let* τ_c, σ_k, $k \ge 0$, *and* $\tilde n$, *be defined as in* (3.4) *with* $c(R) = R^{\alpha_c}$ *and let* $\tilde Z = Z(\sigma_{\tilde n})$. *Then, uniformly in* $R^{\alpha_t'} \le t(R) \le R^{\alpha_t}$,

$$\mathbf{E}\tilde Z^2 = O((t(R)/R)^2),\ \mathbf{E}\tilde Z = o(t(R)/R),\ \textit{as } R \to \infty. \tag{4.5}$$

Proof. The result on the second moment does not require much extra work compared to (2.5). Indeed, from (1.1) and (2.5), we know that $\tilde{Z} = \sum_{k=0}^{\tilde{n}-1} \tilde{Z}_k$, where

$$\tilde{Z}_k = \int_{\sigma_k}^{\sigma_{k+1}} V'(\theta_1(s) - \theta_2(s))ds = r_2(\sigma_k) \cdot O(R^{-2}) + O(R^{-3}) + \tilde{\Theta}(R^{-5/2}).$$

Thus, by Lemma 3.2, we obtain

$$\|\tilde{Z}_k\|_2 \lesssim \|r_2(\sigma_k)\|_2/R^2 + R^{-5/2} \le (2|r_2(0)|e^{-k/R} + 3)/R^2 + R^{-5/2} \lesssim (t(R)e^{-k/R} + 1)/R^2.$$

Note that $\tilde{n} < L(R) := t(R)(R + 2c(R))$. So $\|\tilde{Z}\|_2 \le \sum_{k=0}^{\tilde{n}-1} \|\tilde{Z}_k\|_2 = O(t(R)/R)$.

The result on the first moment requires more delicate treatment. Since $\tilde{n} < L(R)$, by (2.6),

$$\begin{aligned} \mathbf{E}\tilde{Z} &= \sum_{k=0}^{L(R)} \mathbf{E}\left(\chi_{\{k<\tilde{n}\}}\tilde{Z}_k\right) = \sum_{k=0}^{L(R)} \mathbf{E}\left(\chi_{\{k<\tilde{n}\}}\mathbf{E}(\tilde{Z}_k|\mathcal{F}_{\sigma_k})\right) \\ &= \sum_{k=0}^{L(R)} \mathbf{E}\left(\chi_{\{k<\tilde{n}\}}\left[\frac{r_2(\sigma_k)}{r_1(\sigma_k)^2} V'(\theta_1(\sigma_k) - \theta_2(\sigma_k))\right]\right) \qquad (4.6) \\ &\quad + \sum_{k=0}^{L(R)} \mathbf{E}\left(\chi_{\{k<\tilde{n}\}}(r_2(\sigma_k)^2 + 1) \cdot O(1/R^3)\right) + O(t(R)R^{-3/2}). \qquad (4.7) \end{aligned}$$

By Lemma 3.2, (4.7) is bounded by

$$O((r_2(0)^2R + Rt(R))/R^3) + O(t(R)/R^{-3/2}) = O(r_2(0)^2/R^2) + o(t(R)/R) = o(t(R)/R)$$

since $t(R) \ge R^\alpha$.

It remains to consider (4.6). We will replace $r_1(\sigma_k)$ with $r_1(0)$, $\theta_1(\sigma_k)$ with $\theta_1(0)$, $r_2(\sigma_k)$ with $r_2(k/R)$, and $\theta_2(\sigma_k)$ with $\theta_2(k/R)$, for $k < \tilde{n}$, in the following steps.

1. To start with, we substitute $r_1(\sigma_k)$ by $r_1(0)$ for $k < \tilde{n}$. Due to the bound on $r_2(\sigma_k)$ in Lemma 3.2, the difference made to (4.6) will be $o(t(R)/R)$. So it remains to show that

$$\sum_{k=0}^{L(R)} \mathbf{E}\left(\chi_{\{k<\tilde{n}\}}\left[r_2(\sigma_k)V'(\theta_1(\sigma_k) - \theta_2(\sigma_k))\right]\right) = o(t(R)R). \qquad (4.8)$$

2. We observe that $\theta_1(\sigma_k)$, $0 \le k < \tilde{n} < L(R)$, are obtained from each other by almost exact rotations. Namely, for each such k, we have that

$$\|\theta_1(\sigma_k) - \theta_1(0)\|_2 = O(t(R)/R), \qquad (4.9)$$

as $R \to \infty$. Indeed, recalling the expansion (2.3) of θ_1, we obtain:

$$\theta_1(\sigma_k) = \theta_1(0) + \sum_{j=0}^{k-1} \chi_{\{j<\tilde{n}\}} \cdot O(1/R^2) + \sum_{j=0}^{k-1} \chi_{\{j<\tilde{n}\}} \int_{\sigma_j}^{\sigma_{j+1}} \int_{\sigma_j}^{t} dW_1(u)dt. \qquad (4.10)$$

The first sum here is $O(t(R)/R)$ because $k < L(R)$. Thus (4.9) follows from

$$\mathbf{E}\left[\sum_{j=0}^{k-1} \chi_{\{j<\tilde{n}\}} \int_{\sigma_j}^{\sigma_{j+1}} \int_{\sigma_j}^{t} dW_1(u)dt\right]^2 = \mathbf{E}\left[\sum_{j=0}^{k-1} \int_{\sigma_j}^{\sigma_{j+1}} \int_{\sigma_j}^{t} dW_1(u)dt\right]^2$$

$$= \mathbf{E}\sum_{j=0}^{k-1}\left[\int_{\sigma_j}^{\sigma_{j+1}}\int_{\sigma_j}^{t} dW_1(u)dt\right]^2 \le \mathbf{E}\sum_{j=0}^{L(R)}\left[\int_{\sigma_j}^{\sigma_{j+1}}\int_{u}^{\sigma_{j+1}} dtdW_1(u)\right]^2$$

$$\le \sum_{j=0}^{L(R)} \mathbf{E}\left(\int_{\sigma_j}^{\sigma_{j+1}} (\sigma_{j+1}-u)^2 du\right) \le \frac{t(R)}{R^2}.$$

Now we can replace the left-hand side (4.8) by

$$\sum_{k=0}^{L(R)} \mathbf{E}\chi_{\{k<\tilde{n}\}}\left[r_2(\sigma_k)V'(\theta_1(0)-\theta_2(\sigma_k))\right], \tag{4.11}$$

with difference estimated by

$$\sum_{k=0}^{L(R)} \left|\mathbf{E}\chi_{\{k<\tilde{n}\}}r_2(\sigma_k)[V'(\theta_1(0)-\theta_2(\sigma_k)) - V'(\theta_1(\sigma_k)-\theta_2(\sigma_k))]\right|$$

$$\le \sum_{k=0}^{L(R)} \|\chi_{\{k<\tilde{n}\}}r_2(\sigma_k)\|_2 \cdot \|\theta_1(0)-\theta_1(\sigma_k)\|_2 \lesssim \sum_{k=0}^{L(R)} (2e^{-k/R}|r_2(0)|+3)\cdot t(R)/R$$

$$\lesssim t(R)R \cdot 2t(R)/R = o(t(R)R).$$

3. It remains to estimate (4.11). Recalling (3.10), (4.11) can be approximated by

$$\sum_{k=0}^{L(R)} \mathbf{E}\chi_{\{k<\tilde{n}\}}\left[r_2(k/R)V'(\theta_1(0)-\theta_2(\sigma_k))\right] \tag{4.12}$$

since $\mathbf{E}|\chi_{\{k\le\tilde{n}\}}(r_2(\sigma_k)-r_2(k/R))| = o(r_2(0)e^{-k/R}+1)$. Now we discard the first $\sqrt{t(R)}R$ terms, which will not make a big difference since, by (2.2) (cf. (3.7)),

$$\begin{aligned}&\left|\sum_{k=0}^{\sqrt{t(R)}R} \mathbf{E}\chi_{\{k<\tilde{n}\}}\left[r_2(k/R)V'(\theta_1(0)-\theta_2(\sigma_k))\right]\right| \\ &\le \sum_{k=0}^{\sqrt{t(R)}R} (|r_2(0)|e^{-k/R}+2) = O(r_2(0)R) + O(\sqrt{t(R)}R) = o(t(R)R).\end{aligned} \tag{4.13}$$

4. For the tail of the series, we substitute $\theta_2(\sigma_k)$ by $\theta_2(k/R)$, and the difference is

$$\sum_{k=\sqrt{t(R)}R}^{L(R)} \mathbf{E}\chi_{\{k<\tilde{n}\}}\left[r_2(k/R)[V'(\theta_1(0)-\theta_2(k/R)) - V'(\theta_1(0)-\theta_2(\sigma_k))]\right]$$

$$\lesssim \sum_{k=\sqrt{t(R)}R}^{L(R)} \|r_2(k/R)\|_2 \cdot \sqrt{\mathbf{E}\chi_{\{k<\tilde{n}\}}|\theta_2(k/R)-\theta_2(\sigma_k)|}$$

$$\lesssim \sum_{k=\sqrt{t(R)}R}^{L(R)} \|r_2(k/R)\|_2 \cdot \sqrt{\mathbf{E}\int_{k/R-2t(R)c(R)/R}^{k/R+2t(R)c(R)/R} |r_2(s)|ds}$$

$$\lesssim \sum_{k=\sqrt{t(R)}R}^{L(R)} \sqrt{t(R)c(R)/R} = o(t(R)R),$$

where we used the Cauchy-Schwarz inequality and the fact that both V' and V'' are bounded by 1 in the first inequality, (3.6) in the second inequality, and (2.2) (cf. (3.7)) in the third inequality.

Similarly, we can discard the last $4t(R)c(R)$ terms. Now the problem reduces to proving

$$\sum_{k=\sqrt{t(R)}R}^{t(R)(R-2c(R))} \mathbf{E}\chi_{\{k<\tilde{n}\}}\left[r_2(k/R)V'(\theta_1(0)-\theta_2(k/R))\right] = o(t(R)R). \tag{4.14}$$

We can get rid of the indicator function by the Cauchy-Schwarz inequality, (3.7), and Lemma 4.3, since $\mathbf{P}(k \geq \tilde{n}) \leq \mathbf{P}(\tau_c < t(R))$ for $k \leq t(R)(R-2c(R))$. We finally arrive at proving

$$\sum_{k=\sqrt{t(R)}R}^{t(R)(R-2c(R))} \mathbf{E}\left[r_2(k/R)V'(\theta_1(0)-\theta_2(k/R))\right] = o(t(R)R), \tag{4.15}$$

which is a consequence of Lemma 3.4. □

Proof of Proposition 4.2. Choose α_c such that $\alpha_t/2 < \alpha_c < 1/3$. Let $c(R) = R^{\alpha_c}$ and stopping times τ_c, σ_k, $k \geq 0$, $\sigma_{\tilde{n}}$ be defined as in (3.4). By (1.1), we have

$$\begin{aligned}
r_1(t(R)) &= r_1(\sigma_{\tilde{n}}) + W_1(t(R)) - W_1(\sigma_{\tilde{n}}) + \int_{t(R)\wedge\tau_c}^{t(R)} V'(\theta_1(s)-\theta_2(s))ds \\
&\quad - \int_{t(R)\wedge\tau_c}^{\sigma_{\tilde{n}}} V'(\theta_1(s)-\theta_2(s))ds \\
&= r_1(\sigma_{\tilde{n}}) + W_1(t(R)) - W_1(\sigma_{\tilde{n}}) + \chi_{\{\tau_c<t(R)\}}\cdot O(t(R)) + O(1/R).
\end{aligned}$$

The desired result follows from Lemma 4.4 and Lemma 4.3. □

Corollary 4.5. *Suppose that $|r_1(0)| = R$, $|r_2(0)| \leq R^\alpha$, where $\alpha \in (0, \frac{2}{3})$. For each α_t and α_t' that satisfy $\alpha < \alpha_t' < \alpha_t < 2/3$, as $R \to \infty$, uniformly in $R^{\alpha_t'} \leq t(R) \leq R^{\alpha_t}$,*

$$\mathbf{E}r_1(t(R)) = r_1(0) + o(t(R)/R), \tag{4.16}$$

$$\mathbf{E}r_1(t(R))^2 = r_1(0)^2 + t(R) + o(t(R)). \tag{4.17}$$

$$\mathbf{E}[r_1(t(R)) - r_1(0)]^2 = t(R) + o(t(R)), \tag{4.18}$$

4.2 Moderate time behavior

The results of §4.1 concern a relatively small time scale, and $r_1(t)$ typically does not change much compared to the its own magnitude. The next lemma gives time and probability estimates on a larger interval. The idea is to iterate the process on time scale $t(R)$ until it exits the interval.

Lemma 4.6. *Suppose that $\alpha \in (0, \frac{2}{3})$. For every $\beta > 1$, $\kappa > 0$, and for all R sufficiently large (depending on β and κ), if $|r_1(0)| = R$ and $|r_2(0)| \leq R^\alpha$, we have*

$$\mathbf{E}\tau \leq (2^{1/\beta}-1)(1-(\frac{1}{2})^{1/\beta})R^2 + \kappa R^2, \tag{4.19}$$

$$\mathbf{P}(|r_1(\tau)|^\beta = \frac{1}{2}R^\beta) \leq \frac{2}{3}, \tag{4.20}$$

where $\tau = \inf\{t : |r_1(t)|^\beta = \frac{1}{2}R^\beta \text{ or } 2R^\beta\} \wedge \inf\{t : |r_2(t)| = R^\alpha\}$.

Proof. Without loss of generality, we assume that $r_1(0) = R$. Fix $\alpha < \alpha_t < 2/3$ and choose $t(R) = R^{\alpha_t}$. Define $\eta_k = k \cdot t(R)$ and $\tilde{k} = \inf\{k : \eta_k \geq \tau\}$. Note that the difference between τ and $\eta_{\tilde{k}}$ is bounded deterministically by $t(R)$. So, by the equation of $r_1(t)$, we have

$$\|r_1(\tau) - r_1(\eta_{\tilde{k}})\|_1 \leq \|r_1(\tau) - r_1(\eta_{\tilde{k}})\|_2 \lesssim t(R). \tag{4.21}$$

By (4.17) in Corollary 4.5, using the Markov property, we have that

$$\mathbf{E}\chi_{\{k<\tilde{k}\}}[r_1(\eta_{k+1})^2 - r_1(\eta_k)^2] = \mathbf{E}\chi_{\{k<\tilde{k}\}}t(R) + o(\mathbf{E}\chi_{\{k<\tilde{k}\}}t(R)),$$

hence

$$\mathbf{E}[r_1(\eta_{\tilde{k}})^2 - R^2] = \sum_{k=0}^{\infty} \mathbf{E}\chi_{\{k<\tilde{k}\}}[r_1(\eta_{k+1})^2 - r_1(\eta_k)^2] = \mathbf{E}\eta_{\tilde{k}} + o(\mathbf{E}\eta_{\tilde{k}}). \tag{4.22}$$

Similarly, by (4.16), we have

$$\mathbf{E}[r_1(\eta_{\tilde{k}}) - R] = o(\mathbf{E}\eta_{\tilde{k}}/R). \tag{4.23}$$

We compute the difference in the second moment by the Cauchy-Schwarz inequality and (4.21):

$$\mathbf{E}[r_1(\eta_{\tilde{k}})^2 - r_1(\tau)^2] = \mathbf{E}[2r_1(\tau) + (r_1(\eta_{\tilde{k}}) - r_1(\tau))][(r_1(\eta_{\tilde{k}}) - r_1(\tau)] \lesssim Rt(R) = o(R^2). \tag{4.24}$$

Then, we obtain

$$\begin{aligned}
\mathbf{E}[r_1(\tau) - R]^2 &= \mathbf{E}r_1(\tau)^2 - 2R\mathbf{E}r_1(\tau) + R^2 \\
&= \mathbf{E}r_1(\eta_{\tilde{k}})^2 + o(R^2) - 2R(R + O(t(R)) + o(\mathbf{E}\eta_{\tilde{k}}/R)) + R^2 \\
&= \mathbf{E}\eta_{\tilde{k}} + o(\mathbf{E}\eta_{\tilde{k}}) + o(R^2) \\
&= \mathbf{E}\tau + o(\mathbf{E}\tau) + o(R^2),
\end{aligned} \tag{4.25}$$

where the second line follows from (4.24), (4.21), and (4.23); the third line follows from (4.22); and the last line follows from the fact that $|\tau - \eta_{\tilde{k}}| \leq t(R)$. By the definition of τ, we see that

$$\mathbf{E}[r_1(\tau) - R - (2^{1/\beta} - 1)R][r_1(\tau) - R + (1 - (\frac{1}{2})^{1/\beta})R] \leq 0,$$

since the function inside the expectation in non-positive on $[2^{-1/\beta}R, 2^{1/\beta}R]$. Opening the bracket, we have

$$\mathbf{E}[r_1(\tau) - R]^2 - \left(2^{1/\beta} + (\frac{1}{2})^{1/\beta} - 2\right) R \cdot \mathbf{E}[r_1(\tau) - R] \leq (2^{1/\beta} - 1)(1 - (\frac{1}{2})^{1/\beta})R^2. \tag{4.26}$$

As follows from (4.21) and (4.23),

$$|\mathbf{E}[r_1(\tau) - R]| \leq |\mathbf{E}[r_1(\eta_{\tilde{k}}) - R]| + \mathbf{E}|r_1(\tau) - r_1(\eta_{\tilde{k}})| = o(\mathbf{E}\tau/R) + O(t(R)), \tag{4.27}$$

By combining this with (4.25) and (4.26), we obtain, for $\beta > 1$, $\kappa > 0$ small enough, and all R sufficiently large,

$$\mathbf{E}\tau \leq (2^{1/\beta} - 1)(1 - (\frac{1}{2})^{1/\beta})R^2 + \kappa R^2 < R^2.$$

Returning to (4.27), we obtain that, for each $\beta > 1$,

$$\mathbf{E}r_1(\tau) = R + o(R) \geq \left[\frac{1}{3} \cdot 2^{1/\beta} + \frac{2}{3}(\frac{1}{2})^{1/\beta}\right] R,$$

for all R sufficiently large. The desired result follows. □

4.3 Long time behavior

Proof of Lemma 4.1. We divide the time into smaller intervals and iteratively apply Corollary 4.5. By Proposition 3.1, with overwhelming probability, $r_2(t)$ stays relatively small hence satisfies the assumption in Corollary 4.5.

Let $\sigma = \inf\{s : r_1(s\cdot T) = 2t\cdot T\} \wedge \inf\{s : |r_2(s\cdot T)| = 3\log T\}$. We introduce this stopping time for technical reasons only, and $\{\sigma \le t\}$ is not likely to happen. Fix $t(x) = x^{1/6}$. Define $\tau_0 = 0$ and inductively define $\tau_{k+1} = \tau_k + t(r_1(\tau_k))$, and $\hat{n} = \inf\{k : \tau_k \ge (t\wedge\eta\wedge\sigma)\cdot T\}$. By the definitions, we immediately see that $\tau_{\hat{n}} \le t\cdot T + (2t\cdot T)^{1/6}$. Before carrying out our plan of applying Corollary 4.5 repeatedly, we first prove a claim that reduces the proof.

Claim. It is enough to prove (4.2) with $t\wedge\eta$ replaced by $\tau_{\hat{n}}/T$.

Proof of the claim. Since $\mathbf{P}(\sigma < t\wedge\eta) \le \mathbf{P}(\sigma < t) \to 0$ as $T\to\infty$, by (1.1), the assumption that V' is bounded by 1, and Proposition 3.1, it is enough to show (4.2) with $t\wedge\eta$ replaced by $t\wedge\eta\wedge\sigma$. Now it remains to consider the difference induced during the interval $[t\wedge\eta\wedge\sigma, \tau_{\hat{n}}/T]$ which has length no more than $t(r_1(\tau_{\hat{n}-1}))/T = O(T^{-5/6})$. Then, uniformly in X_0^T on K and $|r_2(0)| \le 2\log T$,

$$
\begin{aligned}
&\mathbf{E}\left[f(X^T_{\tau_{\hat{n}}/T}) - f(X^T_{t\wedge\eta\wedge\sigma}) - \frac{1}{2}\int_{t\wedge\eta\wedge\sigma}^{\tau_{\hat{n}}/T} f''(X_s^T)ds\right]\\
&\lesssim \left|\mathbf{E}\left[f(X^T_{\tau_{\hat{n}}/T}) - f(X^T_{t\wedge\eta\wedge\sigma})\right]\right| + \mathbf{E}|\tau_{\hat{n}}/T - (t\wedge\eta\wedge\sigma)|\\
&\lesssim \mathbf{E}\left[\frac{1}{\sqrt{T}}|r_1(\tau_{\hat{n}}) - r_1((t\wedge\eta\wedge\sigma)T)|\right] + O(T^{-5/6}) = O(T^{-1/3}) \to 0,
\end{aligned}
$$

where the last line follows from (1.1) and the fact that $|\tau_{\hat{n}} - (t\wedge\eta\wedge\sigma)T| = O(T^{1/6})$. The claim is proved. $\square$

Now let us prove (4.2) with $t\wedge\eta$ replaced by $\tau_{\hat{n}}/T$. We divide the interval $[0,T]$ into smaller subintervals using τ_k, $k\ge 0$, and obtain

$$
\begin{aligned}
&\mathbf{E}\left[f(X^T_{\tau_{\hat{n}}/T}) - f(X_0^T) - \frac{1}{2}\int_0^{\tau_{\hat{n}}/T} f''(X_s^T)ds\right]\\
&= \mathbf{E}\left[f(r_1(\tau_{\hat{n}})/\sqrt{T}) - f(r_1(0)/\sqrt{T}) - \frac{1}{2T}\int_0^{\tau_{\hat{n}}} f''(r_1(s)/\sqrt{T})ds\right]\\
&= \sum_{k=0}^{\infty}\mathbf{E}\left(\chi_{\{k<\hat{n}\}}\left[f(r_1(\tau_{k+1})/\sqrt{T}) - f(r_1(\tau_k)/\sqrt{T}) - \frac{1}{2T}\int_{\tau_k}^{\tau_{k+1}} f''(r_1(s)/\sqrt{T})ds\right]\right)\\
&= \sum_{k=0}^{\infty}\mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[f(r_1(\tau_{k+1})/\sqrt{T}) - f(r_1(\tau_k)/\sqrt{T}) - \frac{1}{2T}\int_{\tau_k}^{\tau_{k+1}} f''(r_1(s)/\sqrt{T})ds\middle|\mathcal{F}_{\tau_k}\right]\right)\\
&= \sum_{k=0}^{\infty}\mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[\frac{1}{\sqrt{T}}f'(r_1(\tau_k)/\sqrt{T})(r_1(\tau_{k+1}) - r_1(\tau_k))\middle|\mathcal{F}_{\tau_k}\right]\right)\\
&\quad + \sum_{k=0}^{\infty}\mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[\frac{1}{2T}f''(r_1(\tau_k)/\sqrt{T})(r_1(\tau_{k+1}) - r_1(\tau_k))^2 - \frac{1}{2T}f''(\tilde{\xi}_k)(\tau_{k+1} - \tau_k)\middle|\mathcal{F}_{\tau_k}\right]\right)\\
&\quad + \sum_{k=0}^{\infty}\mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[\frac{1}{6T\sqrt{T}}f'''(\xi_k)(r_1(\tau_{k+1}) - r_1(\tau_k))^3\middle|\mathcal{F}_{\tau_k}\right]\right)\\
&=: \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3,
\end{aligned}
$$

where ξ_k and $\tilde{\xi}_k$ are (random) numbers between $r_1(\tau_k)/\sqrt{T}$ and $r_1(\tau_{k+1})/\sqrt{T}$. Recall (4.16) in Corollary 4.5. Let us deal with the first term $\mathcal{T}_1$:

$$\begin{aligned}\mathcal{T}_1 &= \sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[\frac{1}{\sqrt{T}}f'(\frac{r_1(\tau_k)}{\sqrt{T}})(r_1(\tau_{k+1})-r_1(\tau_k))\Big|\mathcal{F}_{\tau_k}\right]\right)\\ &\lesssim \sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}o(\frac{t(r_1(\tau_k))}{\varepsilon T})\right) = o(\mathbf{E}\tau_{\hat{n}}/T).\end{aligned}$$

Now we recall (4.18) in Corollary 4.5 and deal with the second term $\mathcal{T}_2$:

$$\begin{aligned}\mathcal{T}_2 &= \sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[\frac{1}{2T}f''(\frac{r_1(\tau_k)}{\sqrt{T}})(r_1(\tau_{k+1})-r_1(\tau_k))^2 - \frac{1}{2T}f''(\tilde{\xi}_k)(\tau_{k+1}-\tau_k)\Big|\mathcal{F}_{\tau_k}\right]\right)\\ &= \frac{1}{2T}\sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}f''(\frac{r_1(\tau_k)}{\sqrt{T}})\mathbf{E}\left[(r_1(\tau_{k+1})-r_1(\tau_k))^2 - t(r_1(\tau_k))\Big|\mathcal{F}_{\tau_k}\right]\right)\\ &\quad + \frac{1}{2T}\sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}t(r_1(\tau_k))\mathbf{E}\left[f''(\frac{r_1(\tau_k)}{\sqrt{T}}) - f''(\tilde{\xi}_k)\Big|\mathcal{F}_{\tau_k}\right]\right)\\ &\lesssim \frac{1}{2T}\sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}o(t(r_1(\tau_k)))\right) + \frac{1}{2T}\sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}t(r_1(\tau_k))\mathbf{E}\left[\frac{1}{\sqrt{T}}|r_1(\tau_{k+1})-r_1(\tau_k)|\Big|\mathcal{F}_{\tau_k}\right]\right)\\ &= o(\mathbf{E}\tau_{\hat{n}}/T)\end{aligned}$$

since $\tilde{\xi}_k$ is a number between $r_1(\tau_k)/\sqrt{T}$ and $r_1(\tau_{k+1})/\sqrt{T}$, and (4.18) is used in the penultimate step. Finally,

$$\begin{aligned}\mathcal{T}_3 &= \sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}\mathbf{E}\left[\frac{1}{6T\sqrt{T}}f'''(\xi_k)(r_1(\tau_{k+1})-r_1(\tau_k))^3\Big|\mathcal{F}_{\tau_k}\right]\right)\\ &\lesssim \sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}t(r_1(\tau_k))^3/(T\sqrt{T})\right)\\ &\lesssim \sum_{k=0}^{\infty} \mathbf{E}\left(\chi_{\{k<\hat{n}\}}t(r_1(\tau_k))/T\cdot t(r_1(\tau_k))^2/\sqrt{T}\right) = o(\mathbf{E}\tau_{\hat{n}}/T),\end{aligned}$$

since $\chi_{\{k<\hat{n}\}}t(r_1(\tau_k)) = O(\sqrt[6]{T})$ for all k. Thus, the desired result follows from that $\tau_{\hat{n}} = O(T)$. □

The same argument can be applied to prove the following extension of Lemma 4.1:

Lemma 4.7. *For $\varepsilon>0$, let $\eta = \inf\{s : X_s^T = \varepsilon\}$. Then, for each $t>0$ and bounded $f \in \mathbf{C}^3([0,+\infty))$ with bounded derivatives up to order three, we have*

$$\mathbf{E}\left[f(X_{\sigma\wedge\eta}^T) - f(X_0^T) - \frac{1}{2}\int_0^{\sigma\wedge\eta} f''(X_s^T)ds\right] \to 0,$$

as $T\to\infty$, uniformly in X_0^T on a compact set $K\subset[\varepsilon,+\infty)$, $|r_2(0)|\le 2\log T$, and stopping time σ (w.r.t. $\tilde{\mathcal{F}}_\cdot^T$) bounded by t.

5 Excursions to small values of r_1

5.1 Length of one excursion

Fix $\alpha_1 = 6/7$, $\alpha_2 = 5/9$, and let β in Lemma 4.6 be chosen as 3/2. Note that Lemma 4.6 will be used in this section with $\alpha = \alpha_2/\alpha_1 < 2/3$. Let $\tau(x) = \inf\{t : |r_1(t)| = x\}$. In this section, we obtain the upper bound on time spent by $r_1(t)$ near 0. More precisely, we consider the time $|r_1(t)|$ spends on the interval $[0, \varepsilon\sqrt{T}]$ before exiting, where ε is a small parameter that will be specified later. For time spent below $|\log T|^{\alpha_1}$, we use the boundedness of the drift term to get a crude estimate. Once $|r_1(t)|$ gets large enough, we can apply the result in Lemma 4.6.

Lemma 5.1. *As $T \to \infty$, $\mathbf{E}\tau(|\log T|^{\alpha_1}) = o(\sqrt[10]{T})$, uniformly in $|r_1(0)| \le |\log T|^{\alpha_1}$ and $r_2(0) \in \mathbb{R}$.*

Proof. Since the drift term in the equation for $r_1(t)$ in (1.1) is bounded by 1 in absolute value, $\mathbf{E}(\tau(|\log T|^{\alpha_1})) \le u(0)$, where $u(x) = -\frac{1}{2}e^{2x} + x + \frac{1}{2}e^{2|\log T|^{\alpha_1}} - |\log T|^{\alpha_1}$ is the solution to

$$\begin{cases} \frac{1}{2}u'' - u' = -1 \\ u'(0) = u(|\log T|^{\alpha_1}) = 0 \end{cases}.$$

It remains to see that $u(0) = -\dfrac{1}{2} + \dfrac{1}{2}e^{2|\log T|^{\alpha_1}} - |\log T|^{\alpha_1} = o(\sqrt[10]{T})$. □

Proposition 5.2. *Let $\zeta = \inf\{t : |r_2(t)| = |\log T|^{\alpha_2}\}$. For each $\varepsilon > 0$ small enough, if $|r_1(0)| \le \varepsilon\sqrt{T}$, $|r_2(0)| \le |\log T|^{\alpha_2}$, then $\mathbf{E}(\tau(\varepsilon\sqrt{T}) \wedge \zeta) \lesssim \varepsilon^2 T$, as $T \to \infty$.*

Proof. For all $k \in \mathbb{N}$ such that $\lfloor \beta \log_2(|\log T|^{\alpha_1}) \rfloor - 1 \le k \le \lfloor \beta \log_2(\varepsilon\sqrt{T}) \rfloor$, let

$$t_k = \mathbf{E}(\text{time spent by } |r_1(t)| \text{ in } [2^{k/\beta}, 2^{(k+1)/\beta}] \text{ before } \tau(\varepsilon\sqrt{T}) \wedge \zeta).$$

Be aware that, in the steps below, we use the same notation $\mathbf{P}$ and $\mathbf{E}$ in the formulas, however with different initial conditions, specified before the formulas. Lemma 4.6 and the strong Markov property imply that, if $|r_1(0)| \in [2^{k/\beta}, 2^{(k+1)/\beta}]$ and $|r_2(0)| \le |\log T|^{\alpha_2}$, then

$$\mathbf{E}\tau(2^{(k-1)/\beta}) \wedge \tau(2^{(k+2)/\beta}) \wedge \zeta \le 10 \cdot 2^{2k/\beta}. \tag{5.1}$$

Let $K > 0$ be specified later, $\tilde{\tau}_1 = \inf\{t > K \cdot 2^{2k/\beta} : |r_1(t)| \in [2^{k/\beta}, 2^{(k+1)/\beta}]\}$, and $\tilde{\tau}_{j+1} = \inf\{t > K \cdot 2^{2k/\beta} + \tilde{\tau}_j : |r_1(t)| \in [2^{k/\beta}, 2^{(k+1)/\beta}]\}$. Here the idea is that the process $|r_1(t)|$ starting on $[2^{k/\beta}, 2^{(k+1)/\beta}]$ can make an attempt to reach $2^{(k+2)/\beta}$ within time $K \cdot 2^{2k/\beta}$, and then reaches $\varepsilon\sqrt{T}$ before coming back to the interval $[2^{k/\beta}, 2^{(k+1)/\beta}]$. The stopping times $\tilde{\tau}_j$, $j \ge 1$, are used to describe the "failed" attempts. Take $K = 100$. Then uniformly in $|r_1(0)| \in [2^{k/\beta}, 2^{(k+1)/\beta}]$,

$$\begin{aligned} &\mathbf{P}(\tau(2^{(k+2)/\beta}) \wedge \zeta < K \cdot 2^{2k/\beta}) \\ &\ge \mathbf{P}(\tau(2^{(k-1)/\beta}) \wedge \tau(2^{(k+2)/\beta}) \wedge \zeta < K \cdot 2^{2k/\beta}, \tau(2^{(k+2)/\beta}) \wedge \zeta < \tau(2^{(k-1)/\beta})) \\ &\ge \mathbf{P}(\tau(2^{(k+2)/\beta}) \wedge \zeta < \tau(2^{(k-1)/\beta})) - \mathbf{P}(\tau(2^{(k-1)/\beta}) \wedge \tau(2^{(k+2)/\beta}) \wedge \zeta \ge K \cdot 2^{2k/\beta}) \\ &\ge \frac{1}{9} - \frac{1}{10} \ge \frac{1}{100}, \end{aligned} \tag{5.2}$$

by Lemma 4.6 and (5.1). Note that, by (4.20), for $|r_1(0)| = 2^{(k+2)/\beta}$ and $|r_2(0)| \le |\log T|^{\alpha_2}$, $\mathbf{P}(\tau(\varepsilon\sqrt{T}) \wedge \zeta < \tau(2^{(k+1)/\beta}))$ has a lower bound. Namely,

$$\mathbf{P}(\tau(\varepsilon\sqrt{T}) \wedge \zeta < \tau(2^{(k+1)/\beta})) \ge \frac{2-1}{2^{\lfloor \beta \log_2(\varepsilon\sqrt{T}) \rfloor - (k+1)} - 1} \cdot \frac{1}{3} \ge \frac{2^{k+1}}{(\varepsilon\sqrt{T})^\beta} \cdot \frac{1}{3}. \tag{5.3}$$

Then, for each $|r_1(0)| \in [2^{k/\beta}, 2^{(k+1)/\beta}]$,

$$\begin{aligned}
\mathbf{P}(\tilde{\tau}_1 > \tau(\varepsilon\sqrt{T}) \wedge \zeta) &\geq \mathbf{P}(\tau(2^{(k+2)/\beta}) < \zeta, \tau(2^{(k+2)/\beta}) < K \cdot 2^{2k/\beta}, \tau(\varepsilon\sqrt{T}) \wedge \zeta < \tilde{\tau}_1) \\
&\quad + \mathbf{P}(\zeta \leq \tau(2^{(k+2)/\beta}), \zeta < K \cdot 2^{2k/\beta}) \\
&\geq \mathbf{P}(\tau(2^{(k+2)/\beta}) \wedge \zeta < K \cdot 2^{2k/\beta}) \cdot \inf_{\substack{|r_1(0)|=2^{(k+2)/\beta} \\ |r_2(0)| \leq |\log T|^{\alpha_2}}} \mathbf{P}(\tau(\varepsilon\sqrt{T}) \wedge \zeta < \tau(2^{(k+1)/\beta})) \\
&\geq \frac{1}{100} \cdot \frac{2^{k+1}}{(\varepsilon\sqrt{T})^\beta} \cdot \frac{1}{3},
\end{aligned}$$

by the strong Markov property applied at $\tau(2^{(k+2)/\beta}) \wedge \zeta$, (4.20), and (5.2). To see the first inequality, note that the events on the right-hand side are disjoint subsets of the event on the left-hand side. Thus,

$$\begin{aligned}
t_k &\leq K \cdot 2^{2k/\beta} \sum_{j=1}^{\infty} \mathbf{P}(\tilde{\tau}_j < \tau(\varepsilon\sqrt{T}) \wedge \zeta) \\
&\lesssim 2^{2k/\beta} \sum_{j=1}^{\infty} \sup_{\substack{|r_1(0)| \in [2^{k/\beta}, 2^{(k+1)/\beta}] \\ |r_2(0)| \leq \log T^{\alpha_2}}} \mathbf{P}(\tilde{\tau}_1 < \tau(\varepsilon\sqrt{T}) \wedge \zeta)^j \lesssim 2^{2k/\beta - k}(\varepsilon\sqrt{T})^\beta.
\end{aligned}$$

Denote $\hat{k} = \lfloor \beta \log_2(|\log T|^{\alpha_1}) \rfloor - 1$. Then it follows that

$$\sum_{k=\hat{k}}^{\lfloor \beta \log_2(\varepsilon\sqrt{T}) \rfloor} t_k \lesssim \varepsilon^2 T.$$

It remains to consider $\hat{t} = \mathbf{E}(\text{time spent by } |r_1(t)| \text{ in } [0, 2^{\hat{k}/\beta}] \text{ before } \tau(\varepsilon\sqrt{T}) \wedge \zeta)$.

Similarly to the way it was done before, we define following stopping times:

$$\hat{\tau}_1 = \inf\{t > \sqrt[9]{T} : |r_1(t)| \in [0, 2^{\hat{k}/\beta}]\}, \quad \text{and} \quad \hat{\tau}_{j+1} = \inf\{t > \sqrt[9]{T} + \hat{\tau}_j : |r_1(t)| \in [0, 2^{\hat{k}/\beta}]\}.$$

By Lemma 5.1 for each $|r_1(0)| \in [0, 2^{\hat{k}/\beta}]$ we have

$$\mathbf{P}(\tau(2^{(\hat{k}+1)/\beta}) < \sqrt[9]{T}) \geq \frac{1}{2}. \tag{5.4}$$

As in (5.3), by (4.20), for $|r_1(0)| = 2^{(\hat{k}+1)/\beta}$ and $|r_2(0)| \leq |\log T|^{\alpha_2}$,

$$\mathbf{P}(\tau(\varepsilon\sqrt{T}) \wedge \zeta < \tau(2^{\hat{k}/\beta})) \geq \frac{2-1}{2^{\lfloor \beta \log_2(\varepsilon\sqrt{T}) \rfloor - \hat{k}} - 1} \cdot \frac{1}{3} \geq \frac{2^{\hat{k}}}{(\varepsilon\sqrt{T})^\beta} \cdot \frac{1}{3}. \tag{5.5}$$

By (5.4), (5.5), and the strong Markov property applied at $\tau(2^{(\hat{k}+1)/\beta}) \wedge \zeta$, for each $|r_1(0)| \in [0, 2^{\hat{k}/\beta}]$,

$$\begin{aligned}
&\mathbf{P}(\hat{\tau}_1 > \tau(\varepsilon\sqrt{T}) \wedge \zeta) \\
&\geq \mathbf{P}\left(\zeta \leq \tau(2^{(\hat{k}+1)/\beta}) < \sqrt[9]{T}\right) + \mathbf{P}\left(\tau(2^{(\hat{k}+1)/\beta}) < \sqrt[9]{T}, \tau(2^{(\hat{k}+1)/\beta}) < \zeta, \tau(\varepsilon\sqrt{T}) \wedge \zeta < \hat{\tau}_1\right) \\
&\geq \mathbf{P}(\tau(2^{(\hat{k}+1)/\beta}) \wedge \zeta < \sqrt[9]{T}) \cdot \inf_{\substack{|r_1(0)|=2^{(\hat{k}+1)/\beta} \\ |r_2(0)| \leq |\log T|^{\alpha_2}}} \mathbf{P}(\tau(\varepsilon\sqrt{T}) \wedge \zeta < \tau(2^{\hat{k}/\beta}))
\end{aligned}$$

$$\geq \frac{1}{2} \cdot \frac{2^{\hat{k}}}{(\varepsilon\sqrt{T})^\beta} \cdot \frac{1}{3}.$$

Thus, by the strong Markov property, $\hat{t} \leq \sqrt[9]{T} \sum_{j=1}^{\infty} (1 - \frac{1}{6} \frac{2^{\hat{k}}}{(\varepsilon\sqrt{T})^\beta})^j \lesssim \sqrt[9]{T} (\varepsilon\sqrt{T})^\beta = o(T)$. ☐

5.2 Number of excursions

Recall that $X_t^T = r_1(t \cdot T)/\sqrt{T}$. Our next task is to prove a result (Corollary 5.4) that will eventually be helpful in showing that the number of excursions from 2ε to ε of this process is not too large (see (6.7)).

Lemma 5.3. *Assume that $X_0^T = 2\varepsilon$ and $|r_2(0)| < \log T$. Let $\eta = \inf\{t : X_t^T = \varepsilon\}$. Then for all ε sufficiently small and T sufficiently large,* $\mathbf{P}(\eta \geq 1/2) \geq \varepsilon/5$.

Proof. We first prove two useful claims that will be used with the strong Markov property later.

Claim 1. Suppose that $r_1(0) = \sqrt{T}$ and $|r_2(0)| \leq 2 \log T$. Let $\sigma = \inf\{s : X_s^T = \varepsilon \text{ or } 2\}$. Then $\mathbf{P}(\sigma > 1/2) > 1/4$.

Proof of the claim. Let $f(x) = -x^2 + (2+\varepsilon)x - 2\varepsilon$ on $[\varepsilon, 2]$. Note that $f(\varepsilon) = f(2) = 0$ and $f'' = -2$. Hence Lemma 4.7 gives that

$$\mathbf{E}\left[f(X^T_{\sigma\wedge\frac{1}{2}}) - f(1) - \frac{1}{2}\int_0^{\sigma\wedge\frac{1}{2}} f''(X_s^T) ds \right] \to 0, \tag{5.6}$$

as $T \to \infty$. Then it follows that, for T sufficiently large so that the left-hand side of (5.6) is less than 1/8 in absolute value, and $\varepsilon < 1/8$,

$$\mathbf{E} f(X^T_{\sigma\wedge\frac{1}{2}}) \geq f(1) - \mathbf{E}\left(\sigma \wedge \frac{1}{2}\right) - \frac{1}{8} \geq \frac{1}{4},$$

Hence $\mathbf{P}(\eta \geq \frac{1}{2}) \geq \mathbf{P}(\sigma \geq \frac{1}{2}) > \mathbf{E} f(X^T_{\sigma\wedge\frac{1}{2}}) / \sup_{[\varepsilon,2]} f \geq \frac{1}{(2-\varepsilon)^2} \geq \frac{1}{4}$. ☐

Claim 2. Suppose that $X_0^T = 2\varepsilon$ and $|r_2(0)| < \log T$. Let $\zeta_T = \inf\{s : |r_2(s \cdot T)| = 2 \log T\}$ and $\sigma' = \inf\{t : X_t^T = 1\} \wedge \zeta_T \wedge \frac{1}{2}$. Then, for sufficiently large T, $\mathbf{P}(\sigma' < \eta) \geq \varepsilon$.

Proof of the claim. Let $g(x) = \dfrac{x-\varepsilon}{1-\varepsilon}$ on $[\varepsilon, 1]$. By Lemma 4.7, $\mathbf{E}\left[g(X^T_{\sigma'\wedge\eta}) - g(X_0^T)\right] \to 0$, as $T \to \infty$. So, for T sufficiently large, $\mathbf{E} g(X^T_{\sigma'\wedge\eta}) \geq g(2\varepsilon) - \varepsilon^2/(1-\varepsilon) = \varepsilon$. Since $g(\varepsilon) = 0$ it follows that $\mathbf{P}(\sigma' < \eta) \geq \mathbf{E} g(X^T_{\sigma'\wedge\eta}) / \sup_{[\varepsilon,1]} g \geq \varepsilon$. ☐

We can identify two disjoint subsets of $\{\eta \geq \frac{1}{2}\}$: (i). $\sigma' = 1/2 < \eta$; (ii). X_t^T reaches 1 before $\eta \wedge \zeta_T \wedge 1/2$, and then spends more time than 1/2 before reaching ε. We obtain by the strong Markov property

$$\begin{aligned}
\mathbf{P}\left(\eta \geq \frac{1}{2}\right) &\geq \mathbf{P}(\sigma' = \frac{1}{2}, \sigma' < \eta) + \mathbf{P}(\sigma' < \zeta_T, \sigma' < \frac{1}{2}, \eta \geq \frac{1}{2}) \\
&\geq \mathbf{P}(\sigma' = \frac{1}{2}, \sigma' < \eta) + \mathbf{P}(\sigma' < \zeta_T, \sigma' < \frac{1}{2}, \sigma' < \eta) \cdot \inf_{r_1(0)=\sqrt{T}, |r_2(0)| < 2\log T} \mathbf{P}(\eta \geq \frac{1}{2}) \\
&\geq \frac{1}{4}(\mathbf{P}(\sigma' < \eta) - \mathbf{P}(\sigma' = \zeta_T)) \geq \varepsilon/5,
\end{aligned}$$

by **Claim 1** and **Claim 2**, and the fact that $\mathbf{P}(\sigma' = \zeta_T)$ is small due to Proposition 3.1. ☐

From Lemma 5.3 we see that the following corollary holds with $\kappa = (1 - 1/\sqrt{e})/5$.

Corollary 5.4. *Assume that* $X_0^T = 2\varepsilon$ *and* $|r_2(0)| < \log T$*. Let* $\eta = \inf\{t : X_t^T = \varepsilon\}$*. Then there exists* $\kappa > 0$ *such that, for each* ε *sufficiently small,*

$$\mathbf{E}e^{-\eta} \leq 1 - \kappa\varepsilon,$$

for all T *sufficiently large.*

6 Proof of the main result

The main idea of the proof is to show that $|X_t^T|$ asymptotically solves the martingale problem for the generator of a Brownian motion reflected at the origin. The times where X_t^T is far from the origin are handled by Lemma 4.7, while the times where X_t^T is small are controlled using the results of Section 5, namely, Proposition 5.2 and Corollary 5.4.

Lemma 4.7 and Corollary 5.4 deal with the case where $r_1 > 0$. However, following the same steps, we can obtain similar results in the case where r_1 is negative. Therefore, we have

Lemma 6.1. *For* $\varepsilon > 0$*, let* $\eta = \inf\{s : |X_s^T| = \varepsilon\}$*. Then, for each* $t > 0$ *and bounded* $f \in \mathbf{C}^3([0, +\infty))$ *with bounded derivatives up to order three, we have*

$$\mathbf{E}\left[f(|X_{\sigma\wedge\eta}^T|) - f(|X_0^T|) - \frac{1}{2}\int_0^{\sigma\wedge\eta} f''(|X_s^T|)ds\right] \to 0, \tag{6.1}$$

as $T \to \infty$*, uniformly in* $|X_0^T|$ *on a compact set* $K \subset [\varepsilon, +\infty)$*,* $|r_2(0)| \leq 2\log T$*, and stopping time* σ *(w.r.t.* $\tilde{\mathcal{F}}_\cdot^T$*) bounded by* t*.*

Lemma 6.2. *Assume that* $|X_0^T| = 2\varepsilon$ *and* $|r_2(0)| \leq \log T$*. Let* $\eta = \inf\{t : |X_t^T| = \varepsilon\}$*. Then there exists* $\kappa > 0$ *such that, for all* ε *sufficiently small,*

$$\mathbf{E}e^{-\eta} \leq 1 - \kappa\varepsilon, \tag{6.2}$$

for all T *sufficiently large.*

Let us again fix $\alpha_2 = 5/9$. The key step in the proof of Theorem 1.1 is the following estimate.

Lemma 6.3. *For each* $t > 0$ *and bounded* $f \in \mathbf{C}^3([0, +\infty))$ *with bounded derivatives up to order three, such that* $f'_+(0) = 0$*,*

$$\mathbf{E}[f(|X_t^T|) - f(|X_0^T|) - \frac{1}{2}\int_0^t f''(|X_s^T|)ds] \to 0, \quad \text{as } T \to \infty, \tag{6.3}$$

uniformly in X_0^T *on a compact set and* $|r_2(0)| \leq \frac{1}{2}\log T^{\alpha_2}$*.*

Proof. Fix $\delta > 0$. Define $\zeta_T = \inf\{s : |r_2(s \cdot T)| = \log T^{\alpha_2}\}$, and a sequence of stopping times: $\eta_0 \leq \sigma_1 \leq \eta_1 \leq \cdots$, by $\eta_0 = 0$, $\sigma_k = \inf\{s \geq \eta_{k-1} : |X_s^T| = 2\varepsilon\}$, $\eta_k = \inf\{s \geq \sigma_k : |X_s^T| = \varepsilon\}$, $k \geq 1$. Since f is bounded together with its derivatives and $\mathbf{P}(\zeta_T < t) \to 0$ as $T \to \infty$, it is enough to prove

$$\mathbf{E}[f(|X_{t\wedge\zeta_T}^T|) - f(|X_0^T|) - \frac{1}{2}\int_0^{t\wedge\zeta_T} f''(|X_s^T|)ds] \to 0. \tag{6.4}$$

Then

$$\mathbf{E}[f(|X^T_{t\wedge\zeta_T}|) - f(|X^T_0|) - \frac{1}{2}\int_0^{t\wedge\zeta_T} f''(|X^T_s|)ds]$$

$$= \mathbf{E}\sum_{k=0}^{\infty}\chi_{\{\eta_k<t\wedge\zeta_T\}}\left[f(|X^T_{\sigma_{k+1}\wedge t\wedge\zeta_T}|) - f(|X^T_{\eta_k}|) - \frac{1}{2}\int_{\eta_k}^{\sigma_{k+1}\wedge t\wedge\zeta_T} f''(|X^T_s|)ds\right] \tag{6.5}$$

$$+ \mathbf{E}\sum_{k=1}^{\infty}\chi_{\{\sigma_k<t\wedge\zeta_T\}}\left[f(|X^T_{\eta_k\wedge t\wedge\zeta_T}|) - f(|X^T_{\sigma_k}|) - \frac{1}{2}\int_{\sigma_k}^{\eta_k\wedge t\wedge\zeta_T} f''(|X^T_s|)ds\right]. \tag{6.6}$$

By Lemma 6.2, we have that for each $k \geq 0$, by the strong Markov property,

$$\begin{aligned}\mathbf{P}(\eta_k < t\wedge\zeta_T) &\leq e^t\mathbf{E}(\chi_{\{\eta_k<\zeta_T\}}e^{-\eta_k}) \leq e^t(1-\kappa\varepsilon)^k,\\ \mathbf{P}(\sigma_k < t\wedge\zeta_T) &\leq \mathbf{P}(\sigma_{k-1} < t\wedge\zeta_T) \leq e^t(1-\kappa\varepsilon)^{k-1}.\end{aligned} \tag{6.7}$$

It follows that

$$\sum_{k=0}^{\infty}\mathbf{P}(\eta_k < t\wedge\zeta_T) \leq \frac{e^t}{\kappa\varepsilon}, \tag{6.8}$$

$$\sum_{k=1}^{\infty}\mathbf{P}(\sigma_k < t\wedge\zeta_T) \leq \frac{e^t}{\kappa\varepsilon}. \tag{6.9}$$

Then we can estimate the first term (6.5) using the strong Markov property:

$$\left|\mathbf{E}\sum_{k=0}^{\infty}\chi_{\{\eta_k<t\wedge\zeta_T\}}[f(|X^T_{\sigma_{k+1}\wedge t\wedge\zeta_T}|) - f(|X^T_{\eta_k}|) - \frac{1}{2}\int_{\eta_k}^{\sigma_{k+1}\wedge t\wedge\zeta_T} f''(|X^T_s|)ds]\right|$$

$$\leq |\mathbf{E}[f(|X^T_{\sigma_1\wedge t\wedge\zeta_T}|) - f(|X^T_0|) - \frac{1}{2}\int_{\eta_k}^{\sigma_1\wedge t\wedge\zeta_T} f''(|X^T_s|)ds]| + \sum_{k=1}^{\infty}\mathbf{P}(\eta_k < t\wedge\zeta_T)\cdot\mathcal{S}, \tag{6.10}$$

where

$$\mathcal{S} = \sup_{\substack{|r_1(0)|=\varepsilon\sqrt{T}\\ |r_2(0)|\leq\log T^{\alpha_2}}}\mathbf{E}\left|[f(|X^T_{\sigma_1\wedge t\wedge\zeta_T}|) - f(|X^T_0|) - \frac{1}{2}\int_0^{\sigma_1\wedge t\wedge\zeta_T} f''(|X^T_s|)ds]\right|.$$

If $|X^T_0| \geq 2\varepsilon$, then, by Lemma 6.1, for each $\varepsilon > 0$, we can choose T large enough such that the first term in (6.10) is small. On the other hand, if $|X^T_0| < 2\varepsilon$, we can choose ε small enough such that $|f(|X^T_{\sigma_1\wedge t\wedge\zeta_T}|) - f(|X^T_0|)|$ is small enough due to the boundedness of f' and the integral is also small due to the boundedness of f'' and Proposition 5.2 when T is sufficiently large. Thus, the first term in (6.10) can be bounded by $\delta/3$ for small ε and all large T. Similarly, we can choose ε small enough and then T large enough such that $\mathcal{S}/\varepsilon$ is small since $f'_+(0) = 0$, $||X^T_{\sigma_1\wedge t\wedge\zeta_T}| - |X^T_0|| \leq \varepsilon$, and $\mathbf{E}\sigma_1\wedge t\wedge\zeta_T = O(\varepsilon^2)$, as $T\to\infty$. Therefore, the second term in (6.10) can be made smaller than $\delta/3$ for small ε and all large T, by (6.8).

We can estimate the second term (6.6) similarly:

$$\left|\mathbf{E}\sum_{k=1}^{\infty}\chi_{\{\sigma_k<t\wedge\zeta_T\}}[f(|X^T_{\eta_k\wedge t\wedge\zeta_T}|) - f(|X^T_{\sigma_k}|) - \frac{1}{2}\int_{\sigma_k}^{\eta_k\wedge t\wedge\zeta_T} f''(|X^T_s|)ds]\right|$$

$$\leq \sum_{k=1}^{\infty}\left|\mathbf{E}\left(\chi_{\{\sigma_k<t\wedge\zeta_T\}}[f(|X^T_{\eta_k\wedge t\wedge\zeta_T}|) - f(|X^T_{\sigma_k}|) - \frac{1}{2}\int_{\sigma_k}^{\eta_k\wedge t\wedge\zeta_T} f''(|X^T_s|)ds]\right)\right| \leq \delta/3,$$

for all T sufficiently large, similarly by the strong Markov property, (6.9), and Lemma 6.1. □

Lemma 6.4. *The family of measures on* $\mathbf{C}([0,+\infty))$ *induced by the processes* $|X_t^T|$, $T \geq 1$, *with initial distribution* μ *of* $(r_1, r_2, \theta_1, \theta_2)$, *is tight.*

Proof. We verify the conditions of Theorem 7.3 (see also the Corollary after Theorem 7.4) in [1]:

(i) For each $\varepsilon > 0$ and $t > 0$, there exist $M > 0$ and $T_0 > 0$ such that

$$\mathbf{P}(|X_0^T| > M) < \varepsilon, \text{ for all } T > T_0;$$

(ii) For each $\varepsilon > 0$, $t > 0$, and $\kappa > 0$, there exist $0 < \delta < 1$ and $T_0 > 0$ such that

$$\frac{1}{\delta}\mathbf{P}(\sup_{s \leq u \leq s+\delta} \big||X_u^T| - |X_s^T|\big| > \kappa) < \varepsilon, \text{ for all } 0 \leq s \leq t-\delta,\ T > T_0.$$

Part (i) is trivial since we have a fixed initial distribution of r_1. To prove part (ii), we fix $\varepsilon > 0$, $t > 0$, and $\kappa > 0$, and notice that for each δ and each $0 \leq s \leq t - \delta$, we have

$$\mathbf{P}(\sup_{s \leq u \leq s+\delta} \big||X_u^T| - |X_s^T|\big| > \kappa) < \mathbf{P}(\tau < s+\delta, \sigma < s+\delta), \tag{6.11}$$

where $\tau = \inf\{u \geq s : |X_u^T| \geq 2\kappa/3\}$ and $\sigma = \inf\{u \geq \tau : |X_u^T - X_\tau^T| \geq \kappa/3\}$. (Here we introduce the stopping times to start the process X^T away from the origin in order to apply our previous estimates. The definition of the stopping time τ also implies the last inequality in the proof.) Moreover, by (1.1) and Proposition 3.1, we see that $\mathbf{P}(\sup_{0 \leq s \leq t \cdot T} |r_1(s)| > T^2) + \mathbf{P}(\sup_{0 \leq s \leq t \cdot T} |r_2(s)| > \log T) < 1/T$ for all T large enough. Thus, by the strong Markov property, it suffices to prove that for any given $2\kappa/3 \leq |X_0^\varepsilon| \leq T^{3/2}$ (or equivalently $2\kappa\sqrt{T}/3 \leq |r_1(0)| \leq T^2$) and $|r_2(0)| \leq \log T$,

$$\mathbf{P}(\tilde{\sigma} < \delta T) < \varepsilon\delta/2, \tag{6.12}$$

where $\tilde{\sigma} = \inf\{u : |r_1(\tilde{\sigma}) - r_1(0)| \geq \kappa\sqrt{T}/3\}$. Recall the definition of the process $Z(\cdot)$ in (2.4). Then

$$\mathbf{P}(\tilde{\sigma} < \delta T) < \mathbf{P}(\{\sup_{0 \leq s \leq \delta T} |W_1(s)| > \kappa\sqrt{T}/6\} \cup \{\sup_{0 \leq s \leq \delta T} |Z(s)| > \kappa\sqrt{T}/6\}), \tag{6.13}$$

Let us define $\hat{\sigma} = \inf\{s : |W_1(s)| \geq \kappa\sqrt{T}/6\} \wedge \inf\{s : |Z(s)| \geq \kappa\sqrt{T}/6\} \wedge \inf\{s : |r_2(s)| \geq \log T\} \wedge \delta T$. Then $\mathbf{P}(\tilde{\sigma} < \delta T) < \mathbf{P}(\hat{\sigma} < \delta T)$. Since $\mathbf{P}(|W_1(\hat{\sigma})| \geq \kappa\sqrt{T}/6)$ and $\mathbf{P}(|r_2(\hat{\sigma})| \geq \log T)$ are exponentially small as $\delta \downarrow 0$ or $T \to \infty$, it remains to consider $\mathbf{P}(|Z(\hat{\sigma})| \geq \kappa\sqrt{T}/6)$. We define $\eta_0 = 0$, $\eta_{k+1} = \eta_k + t(r_1(\eta_k))$, where the function $t(x)$ can be chosen as $x^{1/6}$, and $\hat{k} = \inf\{k : \eta_k \geq \hat{\sigma}\}$. Then, by Proposition 4.2 and noting that $t(\eta_{\hat{k}-1}) < \sqrt[3]{T}$, we have for all T sufficiently large,

$$\|Z(\hat{\sigma})\|_2 \leq \|Z(\eta_{\hat{k}}))\|_2 + \sqrt[3]{T} \lesssim \frac{2\delta T}{\sqrt[6]{|r_1(0)| - \kappa\sqrt{T}/3}} \cdot \frac{\sqrt[6]{|r_1(0)| + \kappa\sqrt{T}/3}}{|r_1(0)| - \kappa\sqrt{T}/3} + \sqrt[3]{T} \lesssim \frac{\delta}{\kappa}\sqrt{T}.$$

Hence, by Chebyshev's inequality, $\mathbf{P}(|Z(\hat{\sigma})| \geq \kappa\sqrt{T}/6) \lesssim \delta^2/\kappa^4 < \varepsilon\delta/4$ if we choose δ sufficiently small and then T sufficiently large. □

Proof of Theorem 1.1. By the tightness of the family of processes $|X_t^T|$ established in Lemma 6.4, for every sequence $\hat{T}_n \to \infty$, there is a subsequence $T_n \to \infty$ such that $X_t^{T_n}$ converges weakly to a continuous process Y_t. The desired weak convergence follows if we prove that, no matter which sequence $\{\hat{T}_n\}$ and $\{T_n\}$ we choose, Y_t always coincides in distribution with the Brownian motion starting and reflected at the origin. Let $S = [0, +\infty)$, $C_0(S)$ be the set of continuous functions that converge to 0 at $+\infty$, $A = \frac{1}{2}\Delta$ with the domain $\mathcal{D}(A) = \{f \in C_0(S) : f'_+(0) = 0, f'' \in C_0(S)\}$, and $\mathcal{D}$ is the set of all functions in $\mathcal{D}(A)$ that have bounded derivatives up to order three. We will show that Y_t is a solution to the martingale problem for $(A|_{\mathcal{D}}, 0)$, i.e., for each $f \in \mathcal{D}$ and $0 \le t_1 \le t_2$,

$$\mathbf{E}[f(Y_{t_2}) - f(Y_{t_1}) - \frac{1}{2}\int_{t_1}^{t_2} f''(Y_t)dt \big| \mathcal{F}_{t_1}^{Y_\cdot}] = 0, \quad Y_0 = 0. \tag{6.14}$$

It is easy to see that the pair $(C_0(S), \mathcal{D})$ satisfies the following conditions:

(i) $\mathcal{D}$ is dense in $C_0(S)$;

(ii) There exists $\lambda > 0$ such that $\mathrm{Range}(\lambda - A|_{\mathcal{D}})$ is dense in $C_0(S)$;

(iii) For each pair of measures ν_1, ν_2 on S, the equality $\int_S f d\nu_1 = \int_S f d\nu_2$ for all $f \in C_0(S)$ implies $\nu_1 = \nu_2$.

Therefore, Theorem 4.1 in Chapter 4 of [8] implies that the solution to the martingale problem (6.14) is unique.

It remains to prove every limiting process satisfies (6.14). It is easy to see that $|X_0^T| \to 0$ as $T \to \infty$. Therefore, it is sufficient to prove that, for each $f \in \mathcal{D}$, $0 \le s_1 < ... < s_k \le t_1$, and $g_1, ..., g_k \in C_0(S)$,

$$\mathbf{E}\left[\prod_{i=1}^k g_i(Y_{s_i})(f(Y_{t_2}) - f(Y_{t_1}) - \frac{1}{2}\int_{t_1}^{t_2} f''(Y_t)dt)\right] = 0.$$

Since $X_t^{T_n}$ converges to Y_t weakly, it is enough to prove

$$\mathbf{E}\left[\prod_{i=1}^k g_i(|X_{s_i}^{T_n}|)(f(|X_{t_2}^{T_n}|) - f(|X_{t_1}^{T_n}|) - \frac{1}{2}\int_{t_1}^{t_2} f''(|X_t^{T_n}|)dt)\right] \to 0,$$

which follows from Lemma 6.4, Proposition 3.1, Lemma 6.3, and the strong Markov property. □

Acknowledgements

We are grateful to Jacob Bedrossian for introducing us to this problem.

References

[1] P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968.

[2] F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet. "Fourier's law: a challenge to theorists". *Mathematical physics 2000*. Imp. Coll. Press, London, 2000, pp. 128–150.

[3] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. "Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation". *Invent. Math.* 181.1 (2010), pp. 39–113.

[4] N. Cuneo, J. Eckmann, M. Hairer, and L. Rey-Bellet. "Non-equilibrium steady states for networks of oscillators". *Electron. J. Probab.* 23 (2018), Paper No. 55, 28.

[5] A. Dhar. "Heat transport in low-dimensional systems". *Advances in Physics* 57.5 (2008), pp. 457–537.

[6] D. Dolgopyat. “Fermi acceleration”. *Geometric and probabilistic structures in dynamics.* Vol. 469. Contemp. Math. Amer. Math. Soc., Providence, RI, 2008, pp. 149–166.

[7] L. H. Eliasson, B. Grébert, and S. B. Kuksin. “KAM for the nonlinear beam equation”. *Geom. Funct. Anal.* 26.6 (2016), pp. 1588–1715.

[8] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence.* New York: John Wiley & Sons, 1986.

[9] E. Fermi. “On the Origin of the Cosmic Radiation”. *Phys. Rev.* 75 (8 Apr. 1949), pp. 1169–1174.

[10] M. Hairer. *Convergence of Markov Processes.* Lecture notes. University of Warwick. 2010.

[11] M. Hairer and J. C. Mattingly. “Slow energy dissipation in anharmonic oscillator chains”. *Comm. Pure Appl. Math.* 62.8 (2009), pp. 999–1032.

[12] M. Hairer and J. C. Mattingly. “Yet Another Look at Harris’ Ergodic Theorem for Markov Chains”. *Seminar on Stochastic Analysis, Random Fields and Applications VI.* Ed. by Robert Dalang, Marco Dozzi, and Francesco Russo. Basel: Springer Basel, 2011, pp. 109–117.

[13] T. E. Harris. “The existence of stationary measures for certain Markov processes”. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II.* Univ. California Press, Berkeley-Los Angeles, Calif., 1956, pp. 113–124.

[14] D. P. Herzog and J. C. Mattingly. “A practical criterion for positivity of transition densities”. *Nonlinearity* 28.8 (2015), pp. 2823–2845.

[15] D. P. Herzog and J. C. Mattingly. “Ergodicity and Lyapunov functions for Langevin dynamics with singular potentials”. *Comm. Pure Appl. Math.* 72.10 (2019), pp. 2231–2255.

[16] T. Komorowski, J. L. Lebowitz, and S. Olla. “Heat flow in a periodically forced, thermostatted chain”. *Comm. Math. Phys.* 400.3 (2023), pp. 2181–2225.

[17] O. E. Lanford. “Time evolution of large classical systems”. *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974).* Vol. Vol. 38. Lecture Notes in Phys. Springer, Berlin-New York, 1975, pp. 1–111.

[18] S. Lepri, R. Livi, and A. Politi. “Thermal conduction in classical low-dimensional lattices”. *Phys. Rep.* 377.1 (2003), pp. 1–80.

[19] A. J. Lichtenberg, M. A. Lieberman, and R. H. Cohen. “Fermi acceleration revisited”. *Phys. D* 1.3 (1980), pp. 291–305.

[20] Y. Pomeau and M. Tran. *Statistical physics of non equilibrium quantum phenomena.* Vol. 967. Lecture Notes in Physics. With a foreword by Linda Reichl. Springer, Cham, 2019, pp. xv+227.

[21] H. Spohn. *Large Scale Dynamics of Interacting Particles.* New York, NY: Springer, 2012.