An Oscillator Driven by Algebraically Decorrelating Noise

Chistophe Gomez¹, Gautam Iyer², Hai Le³, Alexei Novikov³

1 Aix Marseille Univ, CNRS, I2M, Marseille, France. E-mail: christophe.gomez@univ-amu.fr
2 Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: gautam@math.cmu.edu
3 Department of Mathematics, Pennsylvania State University, State College, Centre County, PA 16802, USA. E-mail: leviethaipsu3@gmail.com

Received: 15 July 2021 / Accepted: 4 April 2023
Published online: 5 June 2023 – © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract: We consider a stochastically forced nonlinear oscillator driven by a stationary Gaussian noise that has an algebraically decaying covariance function. It is well known that such noise processes can be renormalized to converge to fractional Brownian motion, a process that has memory. In contrast, we show that the renormalized limit of the nonlinear oscillator driven by this noise converges to diffusion driven by standard (not fractional) Brownian motion, and thus retains no memory in the scaling limit. The proof is based on the study of a fast-slow system using the perturbed test function method.

1. Introduction

Consider a stochastically forced nonlinear oscillator with 1 degree of freedom:

$$\ddot{x}_t + f(x_t) = \varepsilon v(t), \quad x_0 \in \mathbb{R}, \quad \dot{x}_0 = y_0 \in \mathbb{R}. \quad (1.1)$$

Here $f : \mathbb{R} \to \mathbb{R}$ is a given smooth function, and $v$ is a stochastic process representing the noise. Our interest is to study the asymptotic long-time behavior of $x$ when the noise $v$ has an algebraically decaying covariance function. In this case, the rescaled noise converges to fractional Brownian motion (fBm), a process that has a memory.

We aim to study how the nonlinear dynamics affects the limiting behavior, and study whether or not the rescaled oscillator also converges to a process with memory.

A similar question was studied by Komorowski et al. [KNR14] in the case of a passive tracer particle advected by a periodic shear flow. In this case, it turns out that there is a certain parameter regime where the time rescaled dynamics is Markovian and the memory effect of the noise is forgotten. However, there is also a regime (namely the “very long-time” behavior when the Hurst index of the driving noise is larger than 1/2) where the memory effect persists. In contrast, for the oscillator (1.1), we will show that the memory effect never persists, and the effective long-time behavior is always Markovian. The main reason for this is that the oscillatory nature of the deterministic dynamics cancels slowly decaying correlations, and destroys the memory effect. To
study (1.1) we cannot use the limit theorem for additive functionals fBm’s used by
[KNR14]. Instead, we recast (1.1) as a fast-slow system and use the perturbed test
function method. Even though this method was introduced in the ’76 [PSV76,Kus84],
it has applications to many problems arising today [FGPS107,DRV21,CV21], and is the
only method we presently know that can be used to prove our main result.

To describe our setup, we first introduce the Hamiltonian
\[ H(x, y) \overset{\text{def}}{=} \frac{1}{2} y^2 + \int_0^x f(s) \, ds, \]
and recast (1.1) as a stochastically perturbed Hamiltonian system
\[
\begin{align*}
\dot{x}_t &= y_t = \partial_y H(x_t, y_t), \\
\dot{y}_t &= \dot{x}_t = -f(x_t) + \varepsilon v(t) = -\partial_x H(x_t, y_t) + \varepsilon v(t).
\end{align*}
\]
(1.2a)

(1.2b)

Setting \( X_t \overset{\text{def}}{=} (x_t, y_t)^T \), we rewrite (1.2a)–(1.2b) compactly as
\[
\dot{X}_t = \nabla^\perp H(X_t) + \varepsilon v(t) e_2, \quad X_0 = (x_0, y_0) \in \mathbb{R}^2.
\]
(1.3)

where \( \nabla^\perp \overset{\text{def}}{=} (\partial_y, -\partial_x)^T \), and \( e_2 = (0, 1)^T \).

To study the long-time behavior, we consider the time rescaled process \( X^\varepsilon(t) \overset{\text{def}}{=} X(t/\varepsilon^2) \), and observe
\[
\dot{X}^\varepsilon_t = \frac{1}{\varepsilon^2} \nabla^\perp H(X^\varepsilon_t) + \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) e_2, \quad X^\varepsilon_0 = X_0 = (x_0, y_0) \in \mathbb{R}^2.
\]
(1.4)

When the noise \( v \) is white, the behavior of \( X^\varepsilon \) as \( \varepsilon \to 0 \) is completely described by the
averaging principle of Freidlin and Wentzell [FW93,FW94,FW98]. To briefly explain
this, we note that in the absence of noise (i.e. when \( v \equiv 0 \)), the process \( X^\varepsilon \) travels very
fast along level sets of the Hamiltonian. In the presence of noise, the process \( X^\varepsilon \) will
also diffuse slowly across these level sets. To capture the limiting behavior we factor
out the fast motion by projecting to the \textit{Reeb graph} of the Hamiltonian, which has the
effect of identifying all closed trajectories of the Hamiltonian system to points. Now, as
\( \varepsilon \to 0 \), this projection converges to a diffusion driven by a standard Brownian motion and
we refer the reader to [FW93,FW94,FW98] for details. (We also remark that when the
noise \( v \) is white, one may also study the behavior of \( X \) on time scales shorter than \( 1/\varepsilon^2 \). In
some scenarios, a robust and stable limiting behavior is also observed on these time scales
and has been studied by several authors [You88,YJ91,Bak11,HV14a,HV14b,HKPG16,
TV16,HIK+18].) In this paper, our interest is to study the case when the driving noise
in (1.4) is colored in time and has algebraically decaying correlations. It is well known
that this noise can be renormalized to converge to a fractional Brownian motion (fBm),
with Hurst parameter that is determined from the decay rate of the correlation function
(see for instance [Taq75,Taq77,Mar05]). When the Hurst parameter is not half (which
happens when the correlation function decays like \( 1/t^\gamma \) with \( \gamma \neq 1 \)), the renormalized
limit of the noise has memory and is a non-Markovian process. We aim to study whether
the memory effect is also present in the \( \varepsilon \to 0 \) limit of the nonlinear oscillator (1.4),
where the noise is combined with the Hamiltonian dynamics.
At present, we are only able to analyze the scenario where the Hamiltonian is smooth with exactly one non-degenerate critical point. In this case, it is convenient to think about the dynamics of \( X^\varepsilon \) in terms of action-angle coordinates. The angular coordinate of \( X^\varepsilon \) changes very fast, and has no meaningful limit as \( \varepsilon \to 0 \). The action coordinate of \( X^\varepsilon \), on the other hand, changes slowly as a result of the interaction between the noise and the averaged angular coordinate and has a non-trivial limit \( \varepsilon \to 0 \). To study this we use the Hamiltonian itself as a proxy for the action coordinate and state our results in terms of convergence of the process \( H(X^\varepsilon) \). (To relate this to the Freidlin–Wentzell framework, we note that when \( H \) has exactly one non-degenerate critical point, the Reeb graph has exactly one vertex and one edge, and \( H(X^\varepsilon) \) is precisely the projection of \( X^\varepsilon \) onto the Reeb graph.)

In this paper, we prove two results. They show that even though the driving noise has memory, the time correlations are destroyed by the oscillatory dynamics of (1.4), and \( H(X^\varepsilon) \) converges to a diffusion driven by standard Brownian motion (a process without memory). Roughly speaking our main results are as follows:

1. If the Hamiltonian \( H \) is quadratic, and the noise \( \nu \) is any stationary Gaussian process with an algebraically decaying covariance function, then \( H(X^\varepsilon) \) converges to a diffusion process driven by a Brownian motion. In this particular case, the limiting diffusion is a rescaling of the square of the 2-dimensional Bessel process.
2. If the Hamiltonian is not quadratic (but still smooth with exactly one non-degenerate critical point), then we can show \( H(X^\varepsilon) \) converges to a diffusion driven by Brownian motions, provided the noise \( \nu \) is chosen suitably. (The noise \( \nu \) is still a stationary Gaussian process with an algebraically decaying covariance function. However, it must be expressible as a superposition of Ornstein–Uhlenbeck processes as we make use of certain exponential mixing estimates in the proof.)

1.1. Main results for a quadratic Hamiltonian. We now state our results precisely when \( H \) is quadratic. Hereafter we assume that the noise \( \nu \) is a stationary Gaussian process with covariance function

\[
R(t) \overset{\text{def}}{=} EV(t)\nu(0) = EV(t+s)\nu(s).
\] (1.5)

Our main result when the Hamiltonian is quadratic is as follows:

**Proposition 1.1.** Let \( H \) be the quadratic Hamiltonian

\[
H(x) \overset{\text{def}}{=} \frac{|x|^2}{2}, \quad x = (x_1, x_2) \in \mathbb{R}^2,
\] (1.6)

and suppose the noise \( \nu \) is a stationary Gaussian process whose covariance function, \( R \), is of the form

\[
R(t) = \frac{L(t)}{t^\gamma}
\] (1.7)

for some \( \gamma \in (0, 2) \) and a function \( L \) that is slowly varying at infinity.\(^1\) If \( \gamma \in (0, 1] \), we further assume that \( L \) has the slow increase property

\[
\lim_{t \to \infty} \frac{L'(t)}{L(t)/t} = 0.
\] (1.8)

---

\(^1\) Recall a function \( L \) is said to be slowly varying at infinity if for every \( s > 0 \) we have \( L(st)/t \to 1 \) as \( t \to \infty \).
Then the family of processes $H(X^\varepsilon)_{\varepsilon>0}$ converges in distribution to $H(W^D)$, where $W^D$ is a 2D Brownian motion with $W_0 = X_0$ and covariance matrix $D$ given by

$$D_{11} = D_{22} \equiv \lim_{x \to \infty} \int_0^x R(z) \cos(z) \, dz \quad \text{and} \quad D_{12} = D_{21} = 0. \quad (1.9)$$

**Remark 1.1.** For $\gamma \in (0, 1]$ we note that (1.8) implies that $R' < 0$ near infinity. Since $\cos(z)$ is periodic and mean zero, the limit in (1.9) exists and is finite. For $\gamma \in (1, 2)$, we note $R \in L^1(\mathbb{R})$ and so the limit in (1.9) also exists and is finite.

As mentioned earlier, it is well known [Taq75, Taq77, Mar05] that the renormalized noise converges to a fractional Brownian motion. For the reader’s convenience, we state a result to that effect next.

**Proposition 1.2.** Let $v$ be a stationary Gaussian process with covariance function $R$ given by (1.7) for some $\gamma \in (0, 2)$ and a slowly varying function $L$.

If $\gamma \in (1, 2)$, we additionally suppose

$$\int_0^\infty R(t) \, dt = 0. \quad (1.10)$$

Let

$$\sigma(\varepsilon) = \begin{cases} L(\varepsilon^{-2})^{1/2} \varepsilon^{\gamma} & \gamma \neq 1, \\ L(\varepsilon^{-2})^{1/2} \ln \varepsilon^{1/2} & \gamma = 1, \end{cases} \quad \text{and} \quad u^\varepsilon(t) \equiv \frac{1}{\sigma(\varepsilon)} \int_0^t v\left(\frac{s}{\varepsilon^2}\right) \, ds. \quad (1.11)$$

Then, as $\varepsilon \to 0$, the family of processes $(u^\varepsilon)_{\varepsilon>0}$ converges in distribution to $\sigma^{\mathcal{H}} B^H$, where $B^H$ is a standard fractional Brownian motion with Hurst index $\mathcal{H} = 1 - \gamma/2$, and

$$\sigma^{2\mathcal{H}} \equiv \begin{cases} \frac{1}{\mathcal{H}(2\mathcal{H}-1)} & \mathcal{H} \neq \frac{1}{2}, \\ 1 & \mathcal{H} = \frac{1}{2}. \end{cases} \quad (1.12)$$

To reiterate our main point, we note that Proposition 1.1 implies

$$H \left(X_0 + \frac{1}{\varepsilon^2} \int_0^t \nabla^\perp H(X^\varepsilon_s) \, ds + \frac{1}{\varepsilon} \int_0^t v\left(\frac{s}{\varepsilon^2}\right) e_2 \, ds\right) \xrightarrow{\varepsilon \to 0} H(W^D_t), \quad (1.13)$$

for a Brownian motion $W^D$ with covariance matrix $D$. However,

$$\frac{1}{\sigma(\varepsilon)} \int_0^t v\left(\frac{s}{\varepsilon^2}\right) \, ds \xrightarrow{\varepsilon \to 0} \sigma^{\mathcal{H}} B^H_t,$$

for a fBm $B^H$ with Hurst index $\mathcal{H}$. Of course, $H(W^D)$ is a Markov process with no memory, but $B^H$ is a non-Markovian process with memory.

Note further that when $\gamma \in (0, 1)$, we see $\varepsilon \ll \sigma(\varepsilon)$ and so the term $\frac{1}{\varepsilon} \int_0^t v(s/\varepsilon^2) \, ds$ appearing in (1.13) diverges as $\varepsilon \to 0$. On the other hand when $\gamma \in (1, 2)$ (and Eq. (1.10) holds), we see $\sigma(\varepsilon) \ll \varepsilon$, and so this term vanishes. In both cases, the oscillatory Hamiltonian term contributes non-trivially and as a result, the time correlations are forgotten and $H(X^\varepsilon)$ converges to a memoryless Markov process.

We now explain the reason for the assumption (1.10), which is required in Proposition 1.2 when $\gamma \in (1, 2)$, but not in Proposition 1.1. If $\gamma \in (1, 2)$ by the central limit
theorem one can show that \( \frac{1}{\epsilon} \int_0^t v(s/\epsilon^2) \, ds \) converges to a Brownian motion. The covariance of this Brownian motion is proportional to \( \int_0^\infty R(t) \, dt \). If this is 0 (as required by assumption (1.10)), then one can divide \( \int_0^t v(s/\epsilon^2) \, ds \) by the smaller factor \( \epsilon^\gamma = \sigma(\epsilon) \), and look for a non-trivial limit. This is precisely what is given by Proposition 1.2. The reason the assumption (1.10) is not needed for Proposition 1.1 is because in (1.13) the noise is only scaled by a factor of \( 1/\epsilon \) and not \( 1/\sigma(\epsilon) \).

Finally, we briefly summarize the main idea of the proofs of Propositions 1.1 and 1.2. First, Proposition 1.2 only requires convergence of Gaussian processes and can be proved by directly computing covariances. Similar results are well known \[ Taq75, Taq77, Mar05 \]. Since the exact result we need isn’t readily available, we prove it in Appendix B.

The proof of Proposition 1.1 requires a little more work. The main simplification obtained from the assumption that \( H \) is quadratic is that the evolution of the angle coordinate is independent of the action coordinate. As a result one can perform a (time dependent) rotation in space and prove convergence of \( H(X^\epsilon) \) by studying integrals of the form
\[
\frac{1}{\epsilon} \int_0^t v \left( \frac{s}{\epsilon^2} \right) \cos \left( \frac{s}{\epsilon^2} \right) \, ds \quad \text{and} \quad \frac{1}{\epsilon} \int_0^t v \left( \frac{s}{\epsilon^2} \right) \sin \left( \frac{s}{\epsilon^2} \right) \, ds.
\]
These are now Gaussian and one can study convergence by computing covariances directly. We note that the study of similar integrals arises in the study of random media where several authors \[ Mar05, FGPS07, Gar97 \] have used oscillatory dynamics to decorrelate the medium.

1.2. Main results for Hamiltonians with only one critical point. We now study the case where the Hamiltonian has exactly one non-degenerate critical point but is not necessarily quadratic. In this case we will need to work with a stationary Gaussian noise process \( v \) that can be obtained by super-imposing Ornstein–Uhlenbeck processes and written in the form
\[
v(t) \equiv \int_S e^{-\mu |p|^{2\beta}(t-u)} B(du, dp).
\]
Here \( S \subseteq \mathbb{R} \) is a bounded symmetric open interval, \( \mu, \beta \) are constants, and \( B \) is a Gaussian random measure that is white in time and colored in space and will be constructed explicitly in Sect. 3 below to ensure that the covariance of \( v \) decays algebraically. Our main theorem can now be stated as follows.

**Theorem 1.3.** Let \( H : \mathbb{R}^2 \to \mathbb{R} \) be a smooth Hamiltonian with exactly one non-degenerate critical point at \((0,0)\), which is also the global minimum, and let \( v \) be the stationary Gaussian noise described above.

Then
\[
H(X^\epsilon) \xrightarrow{\epsilon \to 0} \mathcal{X},
\]
where \( \mathcal{X} \) is a diffusion driven by standard Brownian motion, with infinitesimal generator
\[
L = \frac{1}{2\Lambda(\mathcal{X})} \partial_{\mathcal{X}}(\Sigma(\mathcal{X}) \partial_{\mathcal{X}}).
\]

Here the coefficients $\Lambda$ and $\Sigma$ are defined by

$$\Lambda(\mathcal{X}) = \oint_{\{H = \mathcal{X}\}} \frac{dl}{|\nabla H|},$$

and

$$\Sigma(\mathcal{X}) = 2 \int_{u=0}^{\infty} du \, R(u) \oint_{\{H = \mathcal{X}\}} \partial_y H(\tilde{X}_u^{x,y}) \partial_y H(x, y) \frac{dl(x, y)}{|\nabla H(x, y)|},$$

where $\tilde{X}_{t}^{x,y}$ is the solution to

$$\partial_t \tilde{X}_t^{x,y} = \nabla \perp H(\tilde{X}_t^{x,y}), \quad \text{with} \quad \tilde{X}_{t=0}^{x,y} = (x, y).$$

Theorem 1.3 follows immediately from the more general Theorem 4.2, below, concerning fast-slow systems, and the proof of Theorem 1.3 is presented shortly after the statement of Theorem 4.2. Moreover, Theorem 4.2 also explicitly characterizes the limiting system and can be used to explicitly characterize the diffusion $\mathcal{X}$ in Theorem 1.3.

Note that when $H$ is not quadratic the convergence of $H(X^\varepsilon)$ can’t be reduced to Gaussian integrals as in Proposition 1.1, and so the proof of Theorem 1.3 is more involved. We prove Theorem 1.3 by first switching to action-angle coordinates and converting (1.4) to a fast-slow system, where the slow variables are mean-zero in the fast variable. We then use the perturbed test function (PTF) method [Kus84] to prove convergence of this system. This approach provides more information than just the convergence of $H(X^\varepsilon)$, as stated in Theorem 1.3, and we refer the reader to Theorem 4.2, below, for the precise statement. Of course, the convergence of action coordinate is equivalent to convergence of the processes $H(X^\varepsilon)$. However, our proof also identifies an asymptotic diffusive behavior of the angle coordinate on the slow scale.

We remark that our fast-slow system is different from that recently considered by Hairer and Li [HL20]. Indeed, in [HL20], the authors proved an averaging principle for a coupled fast-slow system driven by fBm, with Hurst index $\mathcal{H} > \frac{1}{2}$, and they obtained convergence in probability to the naively averaged system. However, they require the fast variable to be driven by an independent standard Brownian motion. In the scenario that arises in the present paper, the fast variable is driven by noise that converges to a fractional Brownian motion when renormalized, and the same noise is also used to drive the slow variable. Moreover the limiting system we obtain is not the naively averaged one, but a diffusion with an averaged generator akin to the limiting behavior of fast-slow systems driven by standard Brownian noise (see for instance [PS08]).

Theorem 1.3 is stated for Hamiltonians which have only one (non-degenerate) critical point. The Reeb graph of such Hamiltonians is simply the half line, and the limiting process in Theorem 1.3 is a diffusion on the half line. For general Hamiltonians, Reeb graph may have multiple vertices, and the limiting process should be a diffusion on this Reeb graph. When the driving noise is a standard Brownian motion (as in the Freidlin–Wentzell theory), this diffusion is characterized through its generator on each edge of the Reeb graph, and certain gluing conditions on the vertices. While the proofs in the Freidlin–Wentzell theory rely heavily on the Markov property, and do not apply to our situation, the form of the infinitesimal generator (1.15) is similar to that in [FW98, Eq. 1.8]. As a result, we expect similar gluing conditions to hold in our context.

Explicitly, for a sufficiently general class of Hamiltonians, we expect that the process $(H(X^\varepsilon))_{\varepsilon > 0}$ converges in distribution to a diffusion process on the Reeb graph. The
generator of this diffusion should be

\[ L_j h_j(\mathcal{X}) = \frac{1}{2\Lambda_j(\mathcal{X})} \partial_\mathcal{X}(\Sigma_j(\mathcal{X}) \partial_\mathcal{X} h_j) \]

on each edge \( I_j \). The coefficients \( \Lambda_j \) and \( \Sigma_j \) are obtained by taking line integrals as follows. Let \( \mathcal{C}_j(\mathcal{X}) \) the connected component of the level set \( \{ H = \mathcal{X} \} \) corresponding to a point \( \mathcal{X} \in I_j \). Now \( \Sigma_j \) and \( \Lambda_j \) are given by

\[ \Sigma_j(\mathcal{X}) = 2 \int_{u=0}^{\infty} du R(u) \oint_{\mathcal{C}_j(\mathcal{X})} \partial_y H(\check{X}^{x,y}_u) \partial_y H(\check{x}, \check{y}) \frac{dl(x, y)}{|\nabla H(x, y)|}, \]

\[ \Lambda_j(\mathcal{X}) = \oint_{\mathcal{C}_j(\mathcal{X})} \frac{dl(x, y)}{|\nabla H(x, y)|}. \]

To describe the gluing condition, consider an interior vertex \( O_k \) corresponding to the saddle point \((x_k, y_k)\) and the level set \( \{ H = \mathcal{X}_k \} \), at which the collection of edges \( \{ I_j \} \) meet. (The typical situation is a “figure eight” scenario, which has three edges meeting at the vertex \( O_k \).) Now define

\[ \beta_{kj} = \lim_{\mathcal{X} \to \mathcal{X}_k} C_j(\mathcal{X}), \]

where the limit is taken along the edge \( I_j \). The gluing condition at the vertex \( O_k \) can then be written as

\[ \sum_{j : I_j \sim O_k} \pm \beta_{kj} \left[ h_j'(H(x, y)) \right] = 0, \quad (1.16) \]

where sign before \( \beta_{kj} \) is positive if \( \mathcal{X} > \mathcal{X}_k \) along the edge \( I_j \), and negative if \( \mathcal{X} < \mathcal{X}_k \) along the edge \( I_j \).

The main difference between our main result, Theorem 1.3, and the standard Freidlin–Wentzell theory [FW98] is in the coefficient \( \Sigma \). The definition of \( \Sigma \) in Theorem 1.3 involves an average of \( \partial_y \check{H}(\check{X}^{x,y}_u) \partial_y H(\check{x}, \check{y}) \), with a shift of one factor to the point \( \check{X}^{x,y}_u \). In [FW98] the coefficient \( \Sigma \) depends only on \((\partial_y H(\check{x}, \check{y}))^2\). The reason for this modification lies in the fact that in our case the oscillations of the Hamiltonian system and the random fluctuations in (1.4) are strongly coupled, and are both of order \( O(1/\varepsilon^2) \). The oscillatory behavior of the Hamiltonian system and random fluctuations fully interact with each other and produce an effective diffusion coefficient involving the correlation of \( \partial_y \check{H}(\check{X}^{x,y}_u) \) with \( \partial_y H(\check{x}, \check{y}) \) for pairs of points on the same orbit. If the oscillatory behavior of the Hamiltonian system is slower than the random fluctuations, then their coupling is weak and we would obtain the same generators as in [FW98]. This will arise if, for example, we consider white noise in (1.4) as the driving random force. This dichotomy of limiting behaviors between strongly and weakly coupled setups is well-known in other approximation-diffusion limits of random differential equations with periodic components, (for instance, compare and contrast Theorems 6.4 and 6.5 in [FGPS107]). In the present work, we need the oscillations of the Hamiltonian system and the random fluctuations to be of the same order; otherwise, the shift \( \check{X}^{x,y}_u \), and hence \( \Sigma \), is not well-defined. We hope to present a proof of this phenomenon in future work.
1.3. Plan of this paper. In Sect. 2 we prove Proposition 1.1, modulo two computational lemmas. The proofs of these lemmas are relegated to Appendix A. The proof of Proposition 1.2 is presented in appendix B. In Sect. 3 we construct the noise process $v$ that will be used in Theorem 1.3, and establish basic properties of the noise. In Sect. 4 we state a precise, more general, version of Theorem 1.3 (Theorems 4.2 and 4.3), and prove Theorem 1.3. We prove regularity of the coefficients of the limiting equation in Theorem 4.2 (Eq. (4.13)) in Appendix C and study the limiting equation itself in Sect. 5. Finally, we prove Theorem 4.2 in Sects. 6, 7 and 8.

2. Proof of Convergence for a Quadratic Hamiltonian

This section is devoted to proving Proposition 1.1. The main idea behind the proof is that when $H$ is quadratic, the deterministic dynamics rotates with constant angular speed in all trajectories. Performing a spatial rotation will now reduce the problem to studying convergence of Gaussian processes, which can be resolved by computing covariances.

Proof of Proposition 1.1. When $H$ is given by (1.6), we use Duhamel’s formula to write the solution of (1.4) as

$$X_\varepsilon^t = \left( \begin{array}{cc} \cos(t/\varepsilon^2) & \sin(t/\varepsilon^2) \\ -\sin(t/\varepsilon^2) & \cos(t/\varepsilon^2) \end{array} \right) X_0 + \frac{1}{\varepsilon} \int_0^t v\left(\frac{\tau}{\varepsilon^2}\right) \left( \begin{array}{c} \sin(\frac{t-\tau}{\varepsilon^2}) \\ \cos(\frac{t-\tau}{\varepsilon^2}) \end{array} \right) d\tau.$$ 

Let $M(t)$ be the rotation matrix

$$M(t) \overset{\text{def}}{=} \left( \begin{array}{cc} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{array} \right),$$

and define the rotated process $Y_\varepsilon^t$ by $Y_\varepsilon^t = M(-t/\varepsilon^2) X_\varepsilon^t$. Clearly $H(X_\varepsilon) = H(Y_\varepsilon)$, and so convergence of the processes $H(X_\varepsilon)$ reduces to convergence of the processes $H(Y_\varepsilon)$.

We claim that the processes $Y_\varepsilon^t$ itself converge to a Brownian motion as $\varepsilon \to 0$. To see this, observe

$$Y_\varepsilon^t = X_0 + \frac{1}{\varepsilon} \int_0^t v\left(\frac{\tau}{\varepsilon^2}\right) \left( -\sin\left(\frac{t-\tau}{\varepsilon^2}\right) \cos\left(\frac{t-\tau}{\varepsilon^2}\right) \right) d\tau.$$

We will now show that the second term above converges to a Brownian motion. For this, define

$$w_1^\varepsilon(t) \overset{\text{def}}{=} \frac{1}{\varepsilon} \int_0^t v\left(\frac{\tau}{\varepsilon^2}\right) \sin\left(\frac{\tau}{\varepsilon^2}\right) d\tau,$$

$$w_2^\varepsilon(t) \overset{\text{def}}{=} \frac{1}{\varepsilon} \int_0^t v\left(\frac{\tau}{\varepsilon^2}\right) \cos\left(\frac{\tau}{\varepsilon^2}\right) d\tau.$$ (2.1)

Note, the noise $v$ (when normalized similarly by $\sigma(\varepsilon)$ without the periodic terms) converges to fBm. However, the expressions above use a normalization factor of $\varepsilon$ instead of $\sigma(\varepsilon)$, and have an oscillatory factor. For this reason we claim $(w_1^\varepsilon, w_2^\varepsilon)$ converges to a Brownian motion. To prove convergence of $w_1^\varepsilon, w_2^\varepsilon$, we first state two lemmas:
Lemma 2.1. For any $T > 0$, there exists a constant $C > 0$ such that for every $s,t \in [0,T]$, $i \in \{1,2\}$, we have
\[
E(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2 \leq \begin{cases} 
C|t - s|^{1-\gamma} & \gamma \in (0,1), \\
C|t - s|^{2-\gamma} & \gamma \in [1,2).
\end{cases} \tag{2.3}
\]

Lemma 2.2. For every $s,t \geq 0$, and $i \in \{1,2\}$ we have
\[
\lim_{\varepsilon \to 0} E(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2 = D_{11}|t - s|,
\]
\[
\lim_{\varepsilon \to 0} E(w_1^\varepsilon(t) - w_1^\varepsilon(s))(w_2^\varepsilon(t) - w_2^\varepsilon(s)) = 0. \tag{2.5}
\]

The proofs of both Lemmas 2.1 and 2.2 are lengthy, but direct computations. Thus, for clarity of presentation, we postpone the proofs to Appendix A. Once Lemmas 2.1 and 2.2 are established, the proof of Proposition 1.1 follows quickly. Indeed, Lemma 2.1 implies that the family $(Y_\varepsilon)_{\varepsilon > 0}$ is tight on $C([0,T], \mathbb{R}^2)$.

To see this, it suffices to show that the processes $w_i^\varepsilon$ are tight. Without loss of generality suppose $T = 1$, and suppose $\gamma \in (0,1)$ (the case when $\gamma \in [1,2)$ is similar). Choose an integer $M$ such that $M\gamma > M(1-\gamma) > 1$. Since $w_i^\varepsilon$ is Gaussian process, we know
\[
E(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2M = C_M[E(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2]^M
\]
for some constant $C_M$. We will allow $C_M$ to change from line to line as long as it only depends on $M$, and remains independent of $\varepsilon$. The above implies
\[
E(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2M \leq C_M|t - s|^{M\gamma}. \tag{2.6}
\]
Since $M\gamma > 1$ by choice, Kolmogorov’s criterion (see for instance [EK86] Proposition 10.3) implies that the family $w_i^\varepsilon$ is tight. This implies that the processes $Y_\varepsilon$ converge in distribution along a subsequence to a continuous process $Y$.

Since each process $Y_\varepsilon$ is a Gaussian process, the limiting process $Y$ must also be a Gaussian process. Now Lemma 2.2 implies $Y$ is a Brownian motion in $\mathbb{R}^2$ with $Y_0 = X_0$ and covariance matrix $D$. Since the law of the limiting process is uniquely determined, the family $Y_\varepsilon$ itself must converge in distribution, without having to select a subsequence. Finally, since $H(X_\varepsilon^\varepsilon) = H(Y_\varepsilon^\varepsilon)$, we obtain convergence of $(H(X_\varepsilon^\varepsilon))_{\varepsilon > 0}$ as claimed. \[\square\]

3. Construction of the Noise

In this section, we will construct noise $v$ that will be used in Theorem 1.3, and establish a few properties that will be used in the course of the proof of Theorem 1.3. We require the covariance function $R$ to be of the form
\[
R(t) \overset{\text{def}}{=} \int_S r(p)e^{-\mu|p|^2|t|} dp, \tag{3.1}
\]
where $S = (-r_s, r_s)$ is a symmetric, bounded, open interval, $\mu$, $\beta > 0$ are constants, and $r: S \setminus \{0\} \to [0, \infty)$ is defined by
\[
r(p) \overset{\text{def}}{=} \frac{\lambda(p)}{|p|^{2\alpha}}.
\]
Here $\lambda: S \to \mathbb{R}$ is a smooth bounded even function such that $\lambda(0) \neq 0$ and
\[
\int_S r(p) \, dp > 0,
\]
and
\[
\alpha < \frac{1}{2}, \quad (3.2)
\]

We will now construct a stationary Gaussian process $v$ with covariance function $R$. The chosen form of $R$ allows us to construct $v$ by superimposing Ornstein–Uhlenbeck processes as follows. Let $\xi$ be 2D white noise and define a Gaussian random measure $\mathcal{B}$ by
\[
\mathcal{B}(du, dp) = \sqrt{2\mu r(p) |p|^\beta} \mathbf{1}_S(p) \xi(du, dp).
\]
Clearly, the covariance of $\mathcal{B}$ is given by
\[
E\mathcal{B}(du, dp) \mathcal{B}(du', dp') = 2\mu r(p) |p|^{2\beta} \delta(u - u') \delta(p - p') \, du \, du' \, dp \, dp'.
\]
Now define the measure-valued Gaussian random process $V$ by
\[
V(t, dp) = \int_{-\infty}^t e^{-\mu |p|^{2\beta} (t-u)} \mathcal{B}(du, dp),
\]
and finally define the noise $v$ by
\[
v(t) = \int_S V(t, dp).
\]

We now establish a few properties of the process $v$ that will be used in the proof of Theorem 1.3.

**Lemma 3.1.** The process $v$ defined above is a stationary Gaussian process with covariance function $R$.

**Proof.** Clearly $v$ is a Gaussian process. To see that $v$ is stationary with covariance function $R$ we choose any $s \leq t$ and compute
\[
Ev(s)v(t) = \int_{-\infty}^s \int_S \exp \left( -\mu \left( |p|^{2\beta} (s-u) + |p|^{2\beta} (t-u) \right) \right) \cdot 2\mu |p|^{2\beta} r(p) \, du \, dp
\]
\[
= \int_S r(p) e^{-\mu |p|^{2\beta} (t-s)} \, dp = R(t-s). \quad \square
\]

**Lemma 3.2.** If $R$ is given by (3.1), then
\[
\lim_{t \to \infty} t^\gamma R(t) = c_0
\]
where $c_0$ and $\gamma$ are defined by
\[
\gamma \overset{\text{def}}{=} \frac{1 - 2\alpha}{2\beta} \quad \text{and} \quad c_0 \overset{\text{def}}{=} \lambda(0) \int_{p \in \mathbb{R}} \frac{e^{-\mu |p|^{2\beta}}}{|p|^{2\alpha}} \, dp.
\]
Proof. By a change of variable, note that
\[ t^\gamma R(t) = t^\gamma \int_S \frac{\lambda(p)}{|p|^{2\alpha}} e^{-\mu |p|^{2\beta} t} dp = \int_{(t^{1/(2\beta)}) S} \frac{1}{|p'|^{2\alpha}} \lambda\left(\frac{p'}{t^{1/(2\beta)}}\right) e^{-\mu |p'|^{2\beta}} dp', \]
which converges to \( c_0 \) as \( t \to \infty \). \( \square \)

Remark 3.1. Note that (3.2) already implies \( \gamma > 0 \). If additionally we assume
\[ 2\alpha + 4\beta > 1 \quad (3.3) \]
then we will also have \( \gamma < 2 \). In this case the process \( v \) can be renormalized as in (1.11) to converge to a fBm. For Theorem 1.3, however, the assumption (3.3) is unnecessary.

Let \( \{G_t\} \) be the augmented filtration generated by \( V \). That is,
\[ G_t \overset{\text{def}}{=} \sigma (\mathcal{N} \cup \{\sigma(V(s, \cdot)) | 0 \leq s \leq t\}) , \quad (3.4) \]
where \( \mathcal{N} \) is the set of all null sets known at time infinity.

Our next lemma computes conditional expectations of \( V(t, dp) \) with respect to the filtration \( G \).

Lemma 3.3. We have for any \( t, h \geq 0 \)
\[ E\left[ V(t + h, dp) | G_t \right] = e^{-\mu |p|^{2\beta} h} V(t, dp) \quad (3.5) \]
and
\[ E\left[ V(t + h, dp) V(t + h, dq) | G_t \right] - E\left[ V(t + h, dp) | G_t \right] E\left[ V(t + h, dq) | G_t \right] = (1 - e^{-2\mu |p|^{2\beta} h}) r(p) \delta(p - q) dp dq . \quad (3.6) \]

Our last result concerns the boundedness of our noise process.

Lemma 3.4. Let \( T > 0, M > 0, \)
\[ D_{k,M} \overset{\text{def}}{=} [0, T] \times L^\infty([0, T], W_{k,M}) \]
with
\[ W_{k,M} \overset{\text{def}}{=} \{ \varphi \in W_1^{1,k}(S) : \|\varphi\|_{W^{1,k}} \leq M \} \]
and where \( W_1^{1,k}(S) \) stands for the Sobolev space with \( k \in (1, \infty) \). We have
\[ E\left[ \sup_{(t, \varphi) \in D_{k,M}} \left| V\left(\frac{t}{\varepsilon^2}, \varphi(t, \cdot)\right)\right| \right] \leq C + \frac{C(\varepsilon)}{\varepsilon} , \quad (3.7) \]
and for any \( n \in \mathbb{N}^* \)
\[ \sup_{\varepsilon} \sup_{t \in [0,T]} E\left[ \sup_{\varphi \in W_{k,M}} \left| V\left(\frac{t}{\varepsilon^2}, \varphi\right)\right|^n \right] \leq C_n , \quad (3.8) \]
where \( C, C_n \) and \( C(\varepsilon) \) are three positive constants where the latter satisfies
\[ \lim_{\varepsilon \to 0} C(\varepsilon) = 0. \]
Here the notation $V\left(\frac{t}{\varepsilon^2}, \varphi\right)$ denotes integral of a function $\varphi(p)$ with respect to the measure $V\left(\frac{t}{\varepsilon^2}, dp\right)$. That is,

$$
V\left(\frac{t}{\varepsilon^2}, \varphi\right) \overset{\text{def}}{=} \int_S \varphi(p) \; V\left(\frac{t}{\varepsilon^2}, dp\right),
$$

and we will drop the $S$ in the above notation for simplicity unless specified otherwise. The proofs of these two Lemmas are straightforward, but computationally involved. For brevity, we do not present them here, but instead refer the reader to Appendices C and D in [GILN21], where they are presented in detail.

4. The Main Theorem

One canonical way to analyze integrable Hamiltonian systems is to use a set of action-angle coordinates. These coordinates separate the slow and fast motion, and preserve the Hamiltonian structure.

4.1. Action-angle coordinates. The Liouville-Arnold theorem [Arn89] asserts that there exists a symplectic canonical transformation $\varphi: X = (x, y) \mapsto (I, \theta) \in \mathbb{R} \times \mathbb{T}$, where the action variable $I$ and the angle variable $\theta$ satisfy

$$
K(I) = H(x, y) \quad \text{and} \quad \{I, \theta\} = 1, \quad (4.1)
$$

where $K$ is a given one variable smooth enough increasing function such that $K(0) = 0$, $\{\cdot, \cdot\}$ stands for the standard Poisson bracket and is defined by

$$
\{g, h\} = \partial_x g \partial_y h - \partial_y g \partial_x h,
$$

for our Hamiltonian system with one degree of freedom. The relation on the right-hand side of (4.1) can be also stated as

$$
\nabla I \cdot \nabla \theta = 1.
$$

Note that in the action-angle coordinate the Hamiltonian is a function of the action coordinate alone.

The existence of such transformation $\varphi$ is guaranteed by the Liouville-Arnold theorem, and can be constructed through a generating function [Arn89, Sect. 50 pp. 281], and implicitly defined by the relations

$$
y = \partial_x S(I, x), \quad \theta = \partial_I S(I, x), \quad \text{and} \quad H(x, \partial_x S(I, x)) = K(I). \quad (4.2)
$$

Such a construction is not unique. For convenience in the forthcoming analysis we will choose $S(I, 0) = 0$ so that

$$
S(I, x) = \int_0^x \partial_x S(I, x') \; dx',
$$

$$
\partial_I S(I, x) = \int_0^x \partial_{Ix'} S(I, x') \; dx',
$$

and hence

$$
\partial_I S(I, 0) = 0.
$$
The meaning of the second relation is that the zero angle corresponds to either the positive part of the $y$-axis or the negative part. Consequently, we have for any $I > 0$

$$\varphi_1^{-1}(I, \theta = 0) = 0 \quad \text{and} \quad \varphi_2^{-1}(I, \theta = 0) \neq 0,$$

(4.3)

and in particular,

$$\partial_I \varphi_1^{-1}(I, \theta = 0) = 0. \quad (4.4)$$

The identity (4.3) follows from the fact that $K$ vanishes only for $I = 0$. For any $I > 0$, the $x$ and $y$ components cannot simultaneously vanish.

Now let us write $\varphi = (\varphi_1, \varphi_2)$ and define

$$I_t \overset{\text{def}}{=} I(X_t) = \varphi_1(X_t), \quad \text{and} \quad \theta_t \overset{\text{def}}{=} \theta(X_t) = \varphi_2(X_t).$$

In the absence of random fluctuations (i.e. when $v = 0$), the Hamiltonian system $\dot{X}_t = \nabla_{\perp} H(X_t)$ becomes

$$\dot{I}_t = 0, \quad \text{and} \quad \dot{\theta}_t = \omega(I_t), \quad \text{with} \quad \omega(I) = K'(I).$$

In the presence of random fluctuations, the Hamiltonian system (1.3) becomes

$$\dot{I}_t = \varepsilon v(t)a(I_t, \theta_t), \quad \text{and} \quad \dot{\theta}_t = \omega(I_t) + \varepsilon v(t)b(I_t, \theta_t), \quad (4.5)$$

where

$$a = e_2 \cdot \nabla \varphi_1 \circ \varphi^{-1}, \quad \text{and} \quad b = e_2 \cdot \nabla \varphi_2 \circ \varphi^{-1}. \quad (4.6)$$

**Example 4.1.** For illustration purposes, we now explicitly compute the action-angle variables when the Hamiltonian is quadratic. From (4.2), we have

$$x = \sqrt{\frac{I}{\pi}} \cos \left(2\pi \theta - \frac{\pi}{2}\right) \quad \text{and} \quad y = \sqrt{\frac{I}{\pi}} \sin \left(2\pi \theta - \frac{\pi}{2}\right),$$

where $\theta$ gives the angle of $X = (x, y)^T$ from the negative part of the vertical axis. One can then see that the action variable is a multiple of the Hamiltonian,

$$K(I) = \frac{I}{2\pi} \quad \text{and} \quad \omega(I) = \frac{1}{2\pi},$$

while the angle variable corresponds to the angle on a trajectory, which is a circle in the case of $H(X) = |X|^2/2$, with period 1. In this case, the functions $a$ and $b$ are as follows

$$a(I, \theta) = 2\sqrt{\pi I} \sin \left(2\pi \theta - \frac{\pi}{2}\right) \quad \text{and} \quad b(I, \theta) = \frac{1}{2\sqrt{\pi I}} \cos \left(2\pi \theta - \frac{\pi}{2}\right). \quad (4.6)$$
4.2. The main theorem. To study the long-time behavior of (4.5), we let

\[ I^\varepsilon_t = I_{t/\varepsilon^2} \quad \text{and} \quad \theta^\varepsilon_t = \theta_{t/\varepsilon^2}. \]

Now (4.5) becomes

\[ \dot{I}^\varepsilon_t = \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) a(I^\varepsilon_t, \theta^\varepsilon_t), \quad \dot{\theta}^\varepsilon_t = \frac{\omega(I^\varepsilon_t)}{\varepsilon^2} + \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) b(I^\varepsilon_t, \theta^\varepsilon_t), \]

with \( I^\varepsilon_0 = I_0 \) and \( \theta^\varepsilon_0 = \theta_0 \). Using the approach in [Gar97] to separate fast and slow motions, we study this system by splitting the angle variable \( \theta^\varepsilon_t \) into two parts

\[ \theta^\varepsilon_t = \psi^\varepsilon_t + \tau^\varepsilon_t. \]

The evolution of \( \tau^\varepsilon_t \) (fast motion) is obtained by averaging (4.7) over the angular coordinate, and \( \psi^\varepsilon_t \) (slow motion) is the remainder. That is we require

\[ \dot{\tau}^\varepsilon_t = \frac{\omega(I^\varepsilon_t)}{\varepsilon^2} + \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) \langle b(I^\varepsilon_t, \cdot) \rangle \]

with initial condition \( \tau^\varepsilon_0 = 0 \), and where we have the notation

\[ \langle g \rangle \overset{\text{def}}{=} \int_{\theta=0}^{1} g(\theta) \, d\theta. \]

Considering only the slow motion variables \( (I^\varepsilon_t, \psi^\varepsilon_t) \), the system (4.7) becomes

\[ \begin{align*}
\dot{I}^\varepsilon_t &= \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) A(I^\varepsilon_t, \psi^\varepsilon_t, \tau^\varepsilon_t) \\
\dot{\psi}^\varepsilon_t &= \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) B(I^\varepsilon_t, \psi^\varepsilon_t, \tau^\varepsilon_t)
\end{align*} \]

with

\[ A(I, \psi, \tau) \overset{\text{def}}{=} a(I, \psi + \tau) \quad \text{and} \quad B(I, \psi, \tau) \overset{\text{def}}{=} b(I, \psi + \tau) - \langle b(I, \cdot) \rangle. \]

The above equations are coupled with the initial conditions \( I^\varepsilon_0 = I_0 \), and \( \psi^\varepsilon_0 = \theta_0 \). Note that \( A \) and \( B \) are both 1-periodic and mean-zero with respect to \( \tau \). This latter property is mandatory to deal with the long-range correlation case as illustrated in Proposition 1.1. In fact, if the system possesses a component with zero frequency, a noise with long-range correlations will charge this component and cause the system to blow up as \( \varepsilon \to 0 \).

It is straightforward to see this mean-zero property in \( \tau \) for \( B \), but for \( A \) we use that

\[ \text{Jac} \varphi^{-1}(I, \theta) = [\text{Jac} \varphi(\varphi^{-1}(I, \theta))]^{-1} \]

and that \( \text{det} \text{Jac} \varphi(X) = 1 \) since \( \varphi \) is a symplectic transformation, to obtain the following relations

\[ \begin{align*}
\partial_I \varphi^{-1}_1(I, \theta) &= \partial_y \varphi_2(X), \quad \partial_I \varphi^{-1}_2(I, \theta) = -\partial_x \varphi_2(X), \\
\partial_{\theta} \varphi^{-1}_1(I, \theta) &= -\partial_y \varphi_1(X), \quad \partial_{\theta} \varphi^{-1}_2(I, \theta) = \partial_x \varphi_1(X).
\end{align*} \]

Therefore, we have

\[ a(I, \theta) = -\partial_{\theta} \varphi^{-1}_1(I, \theta) \]
which is clearly mean-zero with respect to $\theta$ since we have the derivative of a periodic function. Note also that from these relations the mean-zero property in $\theta$ for $b$ is not clear. This is the reason why we introduce the compensation $\langle b \rangle$ in the definition of $B$.

Our main result obtains the limiting behavior of (4.9) as $\varepsilon \to 0$ under the following assumptions.

- The function $K$ is smooth, and
  \[
  \inf_{I \geq 0} K'(I) = \inf_{I \geq 0} \omega(I) > \omega_0 > 0
  \]
  (4.11)

  for some strictly positive number $\omega_0$.

- There exist $r > 0$, and positive constants $c_{1,r}, c_{2,r} > 0$ such that for any $I \in (0, r)$
  \[
  c_{1,r} I \leq K(I) \leq c_{2,r} I \quad \text{and} \quad |\omega'(I)| \leq \frac{c_{2,r}}{I}
  \]
  (4.12)

Note that these conditions imply that $K(0) = 0$ and that $K$ is an increasing function in $I$. These assumptions are not too restrictive since any Hamiltonian satisfies (4.11) and (4.12) near non-degenerate critical points.

The following proposition is a consequence of the above assumptions, which concerns bounds on the function $a$. The proof is given in Appendix C.

**Proposition 4.1.** Under (4.11) and (4.12) there exist $r > 0$ and a constant $C_r > 0$ such that for any $I \in (0, r)$ we have

\[
\sup_{\theta \in T} \left( |a(I, \theta)| + |\partial_\theta a(I, \theta)| \right) \leq C_r \sqrt{I},
\]

and

\[
\sup_{\theta \in T} |\partial_\theta a(I, \theta)| \leq \frac{C_r}{\sqrt{I}}.
\]

The proposition mainly describes the behavior of $a, \partial_\theta a$ and $\partial_\theta a$ around $I = 0$. These bounds will allow us to prove the main result of the paper that characterizes the limiting process of the sequence of processes $(I^\varepsilon, \psi^\varepsilon)_{\varepsilon > 0}$.

**Theorem 4.2.** Assume (4.11) and (4.12) hold, the family $(I^\varepsilon, \psi^\varepsilon)_{\varepsilon > 0}$ (defined in (4.8)–(4.9)) converges in distribution in $C([0, \infty), \mathbb{R}^2)$ to a process $(I_t, \psi_t)_{t \geq 0}$, where $(I_t)_{t \geq 0}$ is the unique weak solution of the SDE

\[
dI_t = \int_0^1 a(I_t, \tau) dW_t(I_t, \tau) d\tau
+ \left[ \int_{u=0}^\infty R(u) \int_0^1 \partial_\tau \left( a(I, \tau + \omega(I)u) \right) |_{I=t} d\tau a(I_t, \tau)
  + a(I_t, \tau + \omega(I)u) \partial_\tau a(I_t, \tau) d\tau du \right] dt,
\]

with initial condition $I_{t=0} = I_0$. Here, $W$ is a real valued Brownian field with covariance function

\[
E[W_t(y, \phi_1)W_s(y, \phi_2)] = t \wedge s \int_{u=0}^\infty R(u)
\times \int_0^1 \phi_1(\tau + \omega(y)u) \phi_2(\tau) + \phi_1(\tau) \phi_2(\tau + \omega(y)u) d\tau du,
\]

(4.14)
for any \( \phi_1, \phi_2 \in L_0^2(\mathbb{T}) \) where
\[
L_0^2(\mathbb{T}) \overset{\text{def}}{=} \{ \phi \in L^2(\mathbb{T}) : \langle \phi \rangle = 0 \}.
\]
Explicitly, we can write
\[
W_t(y, \tau) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}^*} e^{2i\pi n \tau} R^{1/2}_n(y)(B^1_{t,n} - i B^2_{t,n})
\]
where \( B^1_{t,n} \) and \( B^2_{t,n} \) are two independent real valued cylindrical Brownian motions satisfying \( B^1_{t,-n} = B^1_{t,n} \), \( B^2_{t,-n} = -B^2_{t,n} \), and \( R_n = 2 \int_0^\infty R(u) \cos(2\pi n \omega(I)u) \, du \).
Also, we have
\[
d \psi_t = \int_0^1 b(I_t, \tau) \, dW_t(I_t, \tau) \, d\tau
+ \left[ \int_0^\infty R(u) \int_0^1 \frac{\partial I}{\partial \tau} (b(I, \tau + \omega(I)u))|_{I=I_t} a(I_t, \tau)
+ b(I_t, \tau + \omega(I_t)u) \frac{\partial I}{\partial \tau} a(I_t, \tau) \, d\tau \, du \right] d\tau,
\]
with \( \psi_{t=0} = \theta_0 \).

Note that the action variable \( I \) does not depend on the slow angular motion \( \psi \). However, the distribution of the slow angular motion \( \psi \) is completely determined by the motion of the action variable \( I \) and the Brownian field.

We also note that the functions \( \frac{\partial I}{\partial \tau} a \) and \( \frac{\partial I}{\partial \tau} b \) appearing in (4.13) and (4.15) may be singular at \( I = 0 \). Nevertheless, since the minimum of \( H \) is non-degenerate, the function \( a \frac{\partial I}{\partial \tau} a \) has no singularity at \( I = 0 \) and there is no singularity on the right-hand side of Eq. (4.13). The term \( a \frac{\partial I}{\partial \tau} b \) appearing in (4.15) may still have an \( O(I^{-1/2}) \) singularity at \( I = 0 \). However, since the point \( I = 0 \) is inaccessible (Proposition 5.1), the right-hand side of (4.15) is well defined.

We now use Theorem 4.2 to prove Theorem 1.3.

Proof of Theorem 1.3. Using action angle coordinates, we convert the system (1.4) to (4.9). Since the Hamiltonian \( H \) has exactly one non-degenerate critical point, the assumptions (4.11)–(4.12) are satisfied. Now by Theorem 4.2 we see that the family of processes \( (I^\varepsilon, \psi^\varepsilon)_{\varepsilon>0} \) converges in distribution to the pair \( (I, \psi) \) which solves (4.13), (4.15). Since \( H(X^\varepsilon) = K(I^\varepsilon) \), we now obtain Theorem 1.3 by considering \( X = K(I) \).

The infinitesimal generator for \( I \) can be written in the form
\[
\frac{1}{2} (\frac{\partial I}{\partial \tau} \tilde{\Sigma}(I) \frac{\partial I}{\partial \tau})
\]
where
\[
\tilde{\Sigma}(I) = 2 \int_0^\infty R(u) \int_0^1 a(I, \tau + \omega(I)u) a(I, \tau) \, du \, d\tau.
\]
Now, using (C.1), (C.5) and that \( \varphi^{-1}(I, \tau + \omega(I)u) = \tilde{X}^{x,y}_u \) for \( \varphi^{-1}(I, \tau) = (x, y) \), together with \( \omega(I) = 1/A(X) \), we have
\[
\tilde{\Sigma}(I) = \frac{2}{\omega(I)} \int_0^\infty du \, R(u) \int_{\mathcal{H} = x} \partial_y H(\tilde{X}_u^{x,y}) \partial_y H(x, y) \|\nabla H(x, y)\|^{-1} \, dl(x, y).
\]
Finally, using \( \partial I = \omega(I) \partial X \) provides the desired result. \( \square \)
For the quadratic Hamiltonian we can explicitly derive the stochastic differential equations for \((I_t)_{t \geq 0}\) and \((\psi_t)_{t \geq 0}\). Using Eq. (4.6) we obtain
\[
dI_t = 2\sqrt{mI_t} dB^1_t + 2m dt \quad \text{and} \quad d\psi_t = \frac{m}{2\pi \sqrt{I_t}} dB^2_t \tag{4.16}
\]
with
\[
m = \pi \int_0^{\infty} R(u) \cos(u) \, du,
\]
and where \(B^1_t\) and \(B^2_t\) are two independent standard Brownian motions. One can easily remark that \((I_t/m)_{t \geq 0}\) is a 2-dimensional squared Bessel process, so that \((I_t/2\pi)_{t \geq 0}\) has the same law as \(|W_t|^2/2_{t \geq 0}\), for \(W\) being a 2-dimensional Brownian motion with covariance matrix given by (1.9). Then, we recover the result of Proposition 1.1.

The proof of Theorem 4.2 will be given in Sect. 6.

4.3. Generalization for Hamiltonian systems with one degree of freedom and 2-dimensional noises. Theorem 4.2 can be readily extended to general Hamiltonian systems with one degree of freedom and 2-dimensional noises. In fact, our proof does not depend on the particular shapes of Hamiltonians for a one-degree of freedom oscillator problems. Also, considering 2-dimensional noises brings no additional difficulties except more tedious notations in the proof of Theorem 4.2. Here we state this extension of Theorem 4.2 for the sake of completeness, but the detailed proof will be omitted.

We consider the following one-degree of freedom Hamiltonian system
\[
X^\varepsilon_t = \frac{1}{\varepsilon} \nabla^\perp H(X^\varepsilon_t) + \frac{1}{\varepsilon} v(t^\varepsilon), \quad X^\varepsilon_0 = X_0,
\]
where the Hamiltonian \(H\) is as in Theorem 4.2, and \(v\) is a 2-dimensional noise with covariance function
\[
R_{jl}(t-s) \overset{\text{def}}{=} E[v_j(t)v_l(s)] = \int_S \lambda'_{jl}(p)e^{-\mu|p|^2|t-s|} dp \quad j, l \in \{1, 2\},
\]
where \((\lambda'_{jl})_{j,l \in \{1, 2\}}\) is a \(2 \times 2\) symmetric matrix valued function with assumptions similar to the ones considered in Sect. 3.

Introducing the same action-angle coordinates as in Sect. 4.1, they now satisfy
\[
\dot{I}_t^\varepsilon = \frac{1}{\varepsilon} v\left(\frac{t}{\varepsilon^2}\right) \cdot A(I_t^\varepsilon, \psi_t^\varepsilon, \tau_t^\varepsilon) \quad \text{and} \quad \dot{\psi}_t^\varepsilon = \frac{1}{\varepsilon} v\left(\frac{t}{\varepsilon^2}\right) \cdot B(I_t^\varepsilon, \psi_t^\varepsilon, \tau_t^\varepsilon),
\]
with
\[
\dot{\tau}_t^\varepsilon = \frac{\omega(I_t^\varepsilon)}{\varepsilon^2} + \frac{1}{\varepsilon} v\left(\frac{t}{\varepsilon^2}\right) \cdot \langle b(I_t^\varepsilon, \cdot) \rangle,
\]
and where
\[
A(I, \psi, \tau) \overset{\text{def}}{=} \begin{pmatrix} a_1(I, \psi + \tau) \\ a_2(I, \psi + \tau) \end{pmatrix} \quad \text{and} \quad B(I, \psi, \tau) \overset{\text{def}}{=} \begin{pmatrix} b_1(I, \psi + \tau) - \langle b_1(I, \cdot) \rangle \\ b_2(I, \psi + \tau) - \langle b_2(I, \cdot) \rangle \end{pmatrix},
\]
with
\[
a_j(I, \theta) = e_j \cdot \nabla \varphi_1(\varphi^{-1}(I, \theta)) \quad \text{and} \quad b_j(I, \theta) = e_j \cdot \nabla \varphi_2(\varphi^{-1}(I, \theta)),
\]
for \(j \in \{1, 2\}.

The extension of Theorem 4.2 is as follows.
Theorem 4.3. The family \((I^\varepsilon, \psi^\varepsilon)_{\varepsilon > 0}\) converges in distribution in \(C([0, \infty), \mathbb{R}^2)\) to a process \((I_t, \psi_t)_{t \geq 0}\), where \((I_t)_{t \geq 0}\) is the unique weak solution of the SDE

\[
dI_t = \int_{\tau=0}^{1} a(I_t, \tau) \cdot dW_t(I_t, \tau) d\tau + \int_{u=0}^{\infty} \int_{\tau=0}^{1} \partial_I \left( a^T(I, \tau + \omega(I)u) \right)_{|I=I_t} R(u) a(I_t, \tau) d\tau \bigg|_{I=I_t} du + a^T(I_t, \tau + \omega(I)u) R(u) \partial_I a(I_t, \tau) d\tau du dt,
\]

with initial condition \(I_{t=0} = I_0\). Here, \(W\) is a 2-dimensional real valued Brownian field with covariance function

\[
E[W^j_t(y, \phi_1) W^l_s(y, \phi_2)] = (t \land s) \int_{u=0}^{\infty} R_{jl}(u) \phi_1(\tau + \omega(y)u) \phi_2(\tau) + \phi_1(\tau) \phi_2(\tau + \omega(y)u) d\tau du,
\]

for any \(j, l \in \{1, 2\}\) and \(\phi_1, \phi_2 \in L^2_0(\mathbb{T})\) where

\[
L^2_0(\mathbb{T}) \overset{\text{def}}{=} \{ \phi \in L^2(\mathbb{T}) : \langle \phi \rangle = 0 \}.
\]

Also, we have

\[
d\psi_t = \int_{\tau=0}^{1} b(I_t, \tau) \cdot dW_t(I_t, \tau) d\tau + \int_{u=0}^{\infty} \int_{\tau=0}^{1} \partial_I \left( b^T(I, \tau + \omega(I)u) \right)_{|I=I_t} R(u) a(I_t, \tau) d\tau \bigg|_{I=I_t} du + b^T(I_t, \tau + \omega(I)u) R(u) \partial_I a(I_t, \tau) d\tau du dt,
\]

with \(\psi_{t=0} = \theta_0\).

5. Properties of the Limiting Equation for the Action Variable

In this section, we study properties of the Eq. (4.13). The main issue concerns the behavior of \(a\) at \(I = 0\), but we will see how to define a unique global solution for this equation that does not reach 0 with probability one. In the case of a quadratic Hamiltonian, we have seen in (4.16) that \((I_t)_{t \geq 0}\) is related to a 2-dimensional squared Bessel process (see (4.16)) which is known to reach 0 with probability zero [KS91, Proposition 3.22 p. 161].

5.1. Existence of a solution. First, we remark that the function \(a\) is not Lipschitz in \(I\) as it typically has a square-root singularity near 0. So it is not immediately apparent that Eq. (4.13) has solutions. To construct solutions to Eq. (4.13), define

\[
a_n(I) \overset{\text{def}}{=} \int_{\tau=0}^{1} a(I, \tau) e^{-2i\pi n \tau} d\tau
\]

\[
R_n(I) \overset{\text{def}}{=} \int_{u=0}^{\infty} R(u) \cos(2\pi n \omega(I)u) du.
\]
and let

\[ a(I) \overset{\text{def}}{=} \sum_{n \in \mathbb{Z}^*} |a_n(I)|^2 R_n(I) \]

\[ = 2 \int_{u=0}^{\infty} R(u) \int_{\tau=0}^{1} a(I, \tau) a(I, \tau + \omega(I) u) \, d\tau \, du, \quad (5.3) \]

\[ b(I) \overset{\text{def}}{=} \int_{u=0}^{\infty} R(u) \int_{\tau=0}^{1} \left[ \frac{\partial I}{\partial \tau} (a(I, \tau + \omega(I) u)) a(I, \tau) + \partial_I a(I, \tau) a(I, \tau + \omega(I) u) \right] \, d\tau \, du. \quad (5.4) \]

Note that equality in (5.3) is justified by using the Fourier series expansion for \( a(I, \tau) \), and using the fact that \( a(I, \tau) \) is mean-zero in \( \tau \). That is,

\[ a(I, \tau) = \sum_{n \in \mathbb{Z}^*} e^{2i\pi n \tau} a_n(I). \]

**Definition 5.1.** A weak solution in the interval \((0, \infty)\) to Eq. (4.13) is a filtered probability space \((\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})\) satisfying the usual conditions and a continuous pair of processes \((I_t, W_t)\) such that the followings hold.

1. The process \( I \) takes values in \([0, \infty)\) with \( I_0 \in (0, \infty) \).
2. The process \( W \) is a standard \( \mathcal{F}_t \)-adapted Brownian motion on the Hilbert space \( L_0^2(\mathbb{T}) \) with covariance function given by (4.14).
3. For any \( t, M > 0 \) we have

\[ \int_{s=0}^{t \wedge \xi_M} (a(I_s) + |b(I_s)|) \, ds < \infty, \]

\[ I_{t \wedge \xi_M} = I_0 + \int_{s=0}^{t \wedge \xi_M} a(I_s, \tau) \, dW_s(I_s, \tau) + \int_{s=0}^{t \wedge \xi_M} b(I_s) \, ds \quad \forall t \geq 0, \quad (5.6) \]

almost surely. Here \( \xi_M \overset{\text{def}}{=} \xi_{1,M} \wedge \xi_{2,M} \) where \( \xi_{1,M} \) and \( \xi_{2,M} \) are defined by

\[ \xi_{1,M} \overset{\text{def}}{=} \inf \{ t \geq 0 \mid |I_t| \geq M \} \]

\[ \xi_{2,M} \overset{\text{def}}{=} \inf \{ t \geq 0 \mid |I_t| \leq 1/M \}. \]

For notational convenience we denote

\[ \xi_{\infty} \overset{\text{def}}{=} \lim_{M \to \infty} \xi_M = \inf \{ t \geq 0 \mid I_t \notin (0, \infty) \} \]

(5.7)

to be the first exit time of \( I \) from \((0, \infty)\).

We will now work with the canonical probability space \( \mathcal{C}([0, \infty), \mathbb{R}) \) and canonical filtration

\[ \mathcal{M}_t = \sigma(h_s, 0 \leq s \leq t). \]

To construct a solution to (4.13), we note first that when restricted to the interval \([1/M, M]\), the function \( a \) is Lipschitz. Thus we know (4.13) has a strong solution when truncated to this interval (see Proposition 6.1, below). If we denote this truncated solution by
we obtain an increasing sequence of stopping times $\xi_M$, an increasing family of $\sigma$-algebras $\mathcal{M}_{\xi_M}$, and a sequence of probability measures $P_1^M$ such that

$$P_1^{M+1} = P_1^M \quad \text{on} \quad \mathcal{M}_{\xi_M} \quad \forall M > 0.$$ 

As a result, we can define $P_0$ on $\bigcup_M \mathcal{M}_{\xi_M}$ by

$$P_0^0(O) = P_1^M(O) \quad \text{for} \quad O \in \mathcal{M}_{\xi_M},$$

and extend $P_0^0$ to a probability measure on the $\sigma$-algebra $\mathcal{M}_{\xi_\infty}$ such that for any $M \in \mathbb{N}$ we have

$$P_0^0 = P_1^M \quad \text{on} \quad \mathcal{M}_{\xi_M}. \quad (5.8)$$

Note that

$$\mathcal{M}_{\xi_\infty} = \sigma \left( \bigcup_M \mathcal{M}_{\xi_M} \right), \quad (5.9)$$

which can be easily verified from the fact that

$$\mathcal{M}_{\xi_\infty} = \sigma (h_I \wedge \xi_\infty, t \geq 0),$$

(see for instance [SV06, Lemma 1.3.3 p. 33]). As a result $P_0^0$ provides a solution to (4.13), that we can denote by $I = (I_t)_{t \geq 0}$, in the sense of Definition 5.1. Note that the process $(I_t)_{t \geq 0}$ is nonnegative with probability one thanks to the Portmanteau theorem [Bil99, Theorem 2.1 p. 16].

In the next two sections, we will first show (using the bounds in Proposition 4.1) that $P_0^0(\xi_\infty = \infty) = 1$. That is, the process $I$ does not reach $0$ or blow up in finite time. This will show that $P_0^0$ is the unique extension on $\mathcal{M} \overset{\text{def}}{=} \sigma (\bigcup_{t \geq 0} \mathcal{M}_t)$ satisfying (5.8). Additionally, we will show that $I = (I_t)_{t \geq 0}$ is the unique global solution to (4.13).

### 5.2. Non-explosion and inaccessibility to $0$.

Our aim in this section is to show that the process $I$ can not explode in finite time nor reach $0$.

**Proposition 5.1.** Consider a weak solution to (4.13) (as in definition 5.1), and let $\xi_\infty$ be as in (5.7). Then

$$P_0^0(\xi_\infty = \infty) = 1.$$ 

This result together with (5.8) implies that for any $t \geq 0$

$$\lim_{M \to \infty} P_1^M(\xi_M \leq t) = \lim_{M \to \infty} P_0^0(\xi_M \leq t) = 0.$$ 

Therefore, [SV06, Theorem 1.3.5 p. 34] guarantees that $P_0^0$ is the unique extension satisfying (5.8).

Before proving Proposition 5.1 we need to introduce the scale function

$$p(x) \overset{\text{def}}{=} \int_1^x \exp \left( -2 \int_1^x \frac{b(v)}{a(v)} \, dv \right) d\xi, \quad x \in (0, \infty)$$
where \( a \) and \( b \) are defined by (5.3) and (5.4) respectively. Observe
\[
p''(x) = -\frac{b(x)}{a(x)} p'(x).
\]

Standard results express the exit probability of \( I \) in terms of the scale function, and we prove them in our context for completeness.

**Lemma 5.1.** For any \( 0 < x_- < x_+ < \infty \) we have
\[
P(I_{\xi_{x_-,x_+}} = x_-) = \frac{p(x_+) - p(I_0)}{p(x_+) - p(x_-)} \quad \text{and} \quad P(I_{\xi_{x_-,x_+}} = x_+) = \frac{p(I_0) - p(x_-)}{p(x_+) - p(x_-)},
\]

where \( \xi_{x_-,x_+} = \xi_{1,x_+} \wedge \xi_{2,x_-} \).

**Proof of Lemma 5.1.** The proof consists of applying the Itô formula for \( p(I_t) \) with \( t \leq \xi_{x_-,x_+} \leq \xi_{\infty} \). To do this we rewrite the stochastic integral of (4.13) in the Fourier domain, that is
\[
\int_{s=0}^{t} \int_{\tau=0}^{1} a(I_s, \tau) dW_s(I_s, \tau) d\tau = \int_{s=0}^{t} \sum_{n \in \mathbb{Z}^+} a_n(I_s) R_n^{1/2}(I_s) \frac{1}{\sqrt{2}} (dB_{s,n}^1 - i dB_{s,n}^2),
\]

providing an explicit semi-martingale representation for \((I_t)_{t \geq 0}\). Therefore, applying the Itô formula to \( p(I_t) \) we have
\[
p(I_t \wedge \xi_{x_-,x_+}) = p(I_0) + \int_0^{t \wedge \xi_{x_-,x_+}} p'(I_s) dI_s + \frac{1}{2} \int_0^{t \wedge \xi_{x_-,x_+}} p''(I_s) d(I_s)
\]
\[
= p(I_0) + \int_0^{t \wedge \xi_{x_-,x_+}} p'(I_s) \sum_{n \in \mathbb{Z}^+} a_n(I_s) R_n^{1/2}(I_s) \frac{1}{\sqrt{2}} (dB_{s,n}^1 - i dB_{s,n}^2) + \int_0^{t \wedge \xi_{x_-,x_+}} p'(I_s) b(I_s) \, ds
\]
\[
+ \frac{1}{2} \int_0^{t \wedge \xi_{x_-,x_+}} -2 \frac{b(I_s)}{a(I_s)} p'(I_s) \sum_{n \in \mathbb{Z}^+} |a_n(I_s)|^2 R_n(I_s) \, ds
\]
\[
= p(I_0) + \int_0^{t \wedge \xi_{x_-,x_+}} p'(I_s) \sum_{n \in \mathbb{Z}^+} a_n(I_s) R_n^{1/2}(I_s) \frac{1}{\sqrt{2}} (dB_{s,n}^1 - i dB_{s,n}^2),
\]

thanks to (5.10) and (5.3). Note that the stochastic integral on the right-hand side of the last line is a martingale starting from 0, so by taking the expectation, and passing to the limit \( t \to \infty \) we obtain
\[
p(I_0) = E[p(I_{\xi_{x_-,x_+}})] = p(x_-)P(I_{\xi_{x_-,x_+}} = x_-) + p(x_+)P(I_{\xi_{x_-,x_+}} = x_+).
\]

Solving for \( P(I_{\xi_{x_-,x_+}} = x_-) \) and \( P(I_{\xi_{x_-,x_+}} = x_+ \) we obtain the desired result. \( \square \)

With this lemma, we can now turn to the proof of Proposition 5.1.
Proof of Proposition 5.1. Let us start with the following remark. The term $b$ can be written as

$$b(I) = \frac{1}{2} \partial_I a(I),$$

so that

$$p(0_+) = \lim_{x \to 0_+} \int_1^x \exp \left( -2 \int_1^x \frac{b(\nu)}{a(\nu)} d\nu \right) d\xi$$

$$= \lim_{x \to 0_+} a(1) \int_1^x \exp (-\ln(a(\xi))) d\xi$$

$$= \lim_{x \to 0_+} a(1) \int_1^x \frac{d\xi}{a(\xi)}.$$

Now, using (5.3) along with Proposition 4.1, we have

$$0 \leq a(\xi) \leq C \xi$$

since for any $n \in \mathbb{Z}^*$

$$0 \leq R_n(\xi) = 2 \int_0^\infty R(u) \cos(2\pi n \omega(\xi) u) du$$

$$\leq 2 \int_0^\infty \left( \int_S r(p) e^{-\mu|p|^{2\beta} u} dp \right) \cos(2\pi n \omega(\xi) u) du$$

$$= 2 \int_S dp r(p) \left( \int_0^\infty \cos(2\pi n \omega(\xi) u) e^{-\mu|p|^{2\beta} u} du \right)$$

$$\leq \int_S dp r(p) \frac{2\mu|p|^{2\beta}}{\mu^2|p|^{4\beta} + 4\pi^2 n^2 \omega^2(\xi)} \leq \frac{C_{\omega_0}}{n^2} \int_S |p|^{2\beta} r(p) dp. \quad (5.11)$$

As a result, we have

$$p(0_+) \leq \lim_{x \to 0_+} a(1) \int_1^x \frac{d\xi}{C \xi} = -\infty,$$

which implies $p(0_+) = -\infty$. To conclude the proof we consider two cases, that is $p(\infty) = \infty$ and $p(\infty) < \infty$. Following the same lines as [KS91, Proposition 5.22 p. 345], if $p(\infty) = \infty$ we directly have the desired result, and if $p(\infty) < \infty$ we have

$$\lim_{x_+ \to \infty} \lim_{x_+ \to \infty} P(I_{\xi_1, x_+} = x_+) = \lim_{x_+ \to \infty} \lim_{x_+ \to \infty} \frac{p(l_0) - p(x_-)}{p(x_+) - p(x_-)} = 1.$$

This implies $P(\xi_{\infty} = \xi_{1, \infty}) = 1$ meaning that if $I_t$ exits the interval $(0, \infty)$, it almost surely does so at $\infty$. Here

$$\xi_{1, \infty} = \lim_{M \to \infty} \xi_{1, M}.$$

To prove that $P(\xi_{1, \infty} = \infty) = 1$, and then conclude the proof of the Proposition 5.1, we need the following result. \qed
Lemma 5.2. For any $T > 0$ we have

$$\lim_{M \to \infty} P \left( \sup_{t \in [0, T]} I_{t \wedge \zeta_M} \geq M \right) = 0.$$ 

Indeed, writing

$$P(\zeta_1, \infty < \infty) = \lim_{T \to \infty} P(\zeta_1, \infty \leq T),$$

$$= \lim_{T \to \infty} \lim_{M_0 \to \infty} \lim_{M \to \infty} P \left( \zeta_1, \infty \leq T, \inf_{t \in [0, \zeta_M]} I_t > 1/M_0 \right),$$

To justify the second equality, we note that if $\zeta_1, \infty$ is finite, then $\inf_{t \in [0, \zeta_M]} I_t$ must be positive since $\zeta_1, \infty = \zeta_\infty$ almost surely. We have for any $M > M_0$

$$P \left( \zeta_1, M \leq T, \inf_{t \in [0, \zeta_M]} I_t > 1/M_0 \right) \leq P \left( \sup_{t \in [0, T]} I_{t \wedge \zeta_M} \geq M \right).$$

Therefore, Lemma 5.2 will imply $P(\zeta_1, \infty < \infty) = 0$.

**Proof of Lemma 5.2.** First we remark that using Chebychev’s inequality

$$P \left( \sup_{t \in [0, T]} I_{t \wedge \zeta_M} \geq M \right) = P \left( \sup_{t \in [0, T]} I_{t \wedge \zeta_M}^2 \geq M^2 \right) \leq \frac{1}{M^2} E \sup_{t \in [0, T]} I_{t \wedge \zeta_M}^2. \quad (5.12)$$

Let $T' \leq T$, we also have

$$E \sup_{t \in [0, T']} I_{t \wedge \zeta_M}^2 \leq 3I_0 + 3E \sup_{t \in [0, T']} N_t^2 + 3E \sup_{t \in [0, T']} \left( \int_0^{t \wedge \zeta_M} b(I_s) ds \right)^2$$

$$= 3I_0 + J_{T'} + K_{T'},$$

and where, using (5.6)

$$N_{t \wedge \zeta_M} \overset{\text{def}}{=} I_{t \wedge \zeta_M} - I_0 - \int_0^{t \wedge \zeta_M} b(I_s) ds = \int_0^{t \wedge \zeta_M} a(I_s, \tau) dW_s(I_s, \tau) d\tau$$

is a martingale with quadratic variation

$$\langle N \rangle_t = \int_0^{t \wedge \zeta_M} a(I_s) ds.$$ 

For the term $J_{T'}$ we apply Burkholder-Davis-Gundy inequality [KS91, Theorem 3.28 p. 166] together with (5.11) from which we obtain

$$J_{T'} \leq C E \int_0^{T' \wedge \zeta_M} a(I_s) ds \leq C \int_0^{T'} E[a(I_{s \wedge \zeta_M})] ds.$$
\[
C_{\omega_0} \int_S |p|^{2\beta} |r(p)| \, dp \int_{s=0}^{T'} \int_{\tau=0}^{1} E[|a(I_{s \wedge \xi_M}, \tau)|^2] \, d\tau \, ds.
\]

Moreover, using Proposition 4.1, we have

\[
\int_0^1 E[|a(I_{s \wedge \xi_M}, \tau)|^2] \, d\tau = \int_0^1 \left( E[|a(I_{s \wedge \xi_M}, \tau)|^2 1_{(I_{s \wedge \xi_M} \leq r)}] + E[|a(I_{s \wedge \xi_M}, \tau)|^2 1_{(I_{s \wedge \xi_M} > r)}] \right) \, d\tau
\leq \int_0^1 \left( r + C_T E I_{s \wedge \xi_M}^2 \right) \, d\tau
\leq r C + C_T E \sup_{t \in [0,s]} I_{t \wedge \xi_M}^2,
\]

and hence

\[
J_{T'} \leq C_1 T + C_2 \int_0^{T'} E \sup_{t \in [0,s]} I_{t \wedge \xi_M}^2 \, ds.
\]

Now, for \( K_{T'} \), we use Cauchy-Schwarz inequality

\[
E \sup_{t \in [0,T']} \left( \int_0^{t \wedge \xi_M} b(I_s) \, ds \right)^2 \leq T E \sup_{t \in [0,T']} \int_0^{t \wedge \xi_M} b^2(I_s) \, ds
\leq TE \int_0^{T'} b^2(I_{s \wedge \xi_M}) \, ds.
\]

From Proposition 4.1, for \( I \leq r \), we again use the Cauchy-Schwarz inequality

\[
|b(I)| \leq C |\omega(J)| \sqrt{\int_0^1 |a(I, \tau)|^2 \, d\tau \int_0^1 |\partial_\tau a(I, \tau)|^2 \, d\tau} + C_{\omega_0} \sqrt{\int_0^1 |a(I, \tau)|^2 \, d\tau \int_0^1 |\partial_I a(I, \tau)|^2 \, d\tau}
\leq C.
\]

As a result,

\[
E \int_0^{T'} b^2(I_{s \wedge \xi_M}) \, ds = E \int_0^{T'} b^2(I_{s \wedge \xi_M}) 1_{(I_{s \wedge \xi_M} \leq r)} \, ds
+ E \int_0^{T'} b^2(I_{s \wedge \xi_M}) 1_{(I_{s \wedge \xi_M} > r)} \, ds
\leq C_1 T + C_2 r E \sup_{t \in [0,s]} I_{t \wedge \xi_M}^2 \, ds,
\]

and

\[
K_{T'} \leq C_1 T^2 + C_2 T \int_0^{T'} E \sup_{t \in [0,s]} I_{t \wedge \xi_M}^2 \, ds.
\]
Finally, we have
\[
E \sup_{t \in [0, T']} I_t^2 \leq C_1 (1 + T + T^2) + C_2 (1 + T) \int_0^{T'} E \left( \sup_{t \in [0, s]} I_t^2 \right) ds
\]
and then by Gronwall’s inequality with \( T' = T \)
\[
E \sup_{t \in [0, T]} I_t^2 \leq C_1 (1 + T + T^2) e^{C_2 (1 + T) T}.
\]
This concludes the proof of the Lemma by letting \( M \to \infty \) in (5.12). Consequently, the proof of Proposition 5.1 is also complete.

5.3. Uniqueness of solutions. The weak uniqueness property for Eq. (4.13) comes from the fact that this equation has strong uniqueness on \( \mathcal{M}_{\xi_2, M} \) for any \( M \), and then weak uniqueness on \( \mathcal{M}_{\xi_2, M} \) for any \( M \) [KS91, Proposition 3.20 p. 309], since we avoid the lack of regularity at 0 for the action variable \( I \). However, according to Proposition 5.1, we have \( \lim_{M \to \infty} P^0_0(\xi_2, M \leq T) = 0 \) for any \( T > 0 \) which guarantees the weak uniqueness thanks to [SV06, Theorem 1.3.5 p. 34].

6. Proof of the Main Theorem (Theorem 4.2)

6.1. Truncated process. We first describe the strategy used to prove Theorem 4.2. In addition to the difficulties concerning the behavior of \( a \) around \( I = 0 \) that we mentioned in Theorem 4.2, we have no \textit{apriori} estimates for the process \((I^\varepsilon_t, \psi^\varepsilon_t)_{t \geq 0}\) that are uniform in \( \varepsilon \), in probability. This is a problem when trying to apply [Kus84, Theorem 4 p. 48] directly to prove the tightness property, and when trying to identify the law of the subsequential limits. To bypass this we follow the strategy developed in [SV06, Chap. 11] by introducing a \textit{truncated process} that does not suffer from the above problems. To relate the truncated process with the original one we introduce a family of stopping times. As we will see, these times go to \( \infty \) when we remove the truncation and then finally obtain the weak convergence of the original processes.

Let \( M > 0 \) and consider the process \((I^\varepsilon_t, \psi^\varepsilon_t)_{t \geq 0}\) which is the unique solution to
\[
\begin{align*}
\dot{I}^\varepsilon_t &= \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) \phi_M \left(I^\varepsilon_t, \psi^\varepsilon_t\right) a \left(I^\varepsilon_t, \psi^\varepsilon_t + \tau^\varepsilon_t\right) \\
\dot{\psi}^\varepsilon_t &= \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) \phi_M \left(I^\varepsilon_t, \psi^\varepsilon_t\right) b \left(I^\varepsilon_t, \psi^\varepsilon_t + \tau^\varepsilon_t\right) - \left< b \left(I^\varepsilon_t, \cdot \right) \right>
\end{align*}
\]
(6.1)
coupled with the initial conditions \( I^\varepsilon_0 = I_0 \), and \( \psi^\varepsilon_0 = \theta_0 \), with
\[
\tau^\varepsilon_t = \omega \left( I^\varepsilon_t, \varepsilon \right) - \frac{1}{\varepsilon} v \left( \frac{t}{\varepsilon^2} \right) \phi_M \left(I^\varepsilon_t, \psi^\varepsilon_t\right) \left< b \left(I^\varepsilon_t, \cdot \right) \right>.
\]
Here \( \phi_M \) is a smooth function on \( \mathbb{R}^2 \) such that
\[
0 \leq \phi_M \leq 1, \quad \phi_M(I, \psi) = 1 \quad \text{if} \quad 1/M \leq I \leq M \quad \text{and} \quad |\psi| \leq M,
\]
\[ \phi_M(I, \psi) = 0 \text{ if } I > 2M \text{ or } I < 1/(2M) \text{ or } |\psi| > 2M. \]

Thanks to the truncation provided by the cutoff function \( \phi_M \), for both action and angle variables, the process \((I^\varepsilon, M^\varepsilon, \psi^\varepsilon, M^\varepsilon)_{t \geq 0} \) does not suffer from the regularity problem for the action variable around 0 as mentioned above. Also, its convergence in distribution in \( C([0, \infty), \mathbb{R}^2) \) can be proved using the perturbed test function method and martingale properties \([Kus84, PSV76]\) thanks to the boundedness of the process. All these points are developed more precisely in Sect. 7.

To simplify the notation, we now omit the superscript \( M \) and denote
\[ Y_t^\varepsilon \overset{\text{def}}{=} \left( I_t^\varepsilon, \psi_t^\varepsilon, M_t^\varepsilon \right). \]

Now the system (6.1) can be rewritten as
\[ \dot{Y}_t^\varepsilon = \frac{1}{\varepsilon} F(Y_t^\varepsilon, \tau_t^\varepsilon), \quad (6.2) \]
with
\[ \tau_t^\varepsilon = \frac{\omega(Y_t^\varepsilon)}{\varepsilon^2} + \frac{1}{\varepsilon} G(Y_t^\varepsilon). \quad (6.3) \]

Here
\[ F(Y, \tau) \overset{\text{def}}{=} \left( \phi_M(I, \psi)a(I, \psi + \tau) \right) + \phi_M(I, \psi)(b(I, \psi + \tau) - \langle b(I, \cdot) \rangle), \quad (6.4) \]
and \( a, b \) are defined in Eq. (4.5). Note that \( F \) is a smooth bounded function (with bounded derivatives) 1-periodic and mean-zero with respect to \( \tau \) (the truncation only affects \( \psi \), not \( \tau \)), and \( \omega(Y_t^\varepsilon) = \omega(I_t^\varepsilon, M_t^\varepsilon) \).

To relate the truncated and original processes we introduce first some notations. In the remaining of the paper, the space of all possible outcomes is \( C([0, \infty), \mathbb{R}^2) \), and we denote the corresponding canonical filtration by \( \mathcal{M}_t = \sigma(y_s, 0 \leq s \leq t) \). The laws of the truncated and original processes on \( C([0, \infty), \mathbb{R}^2) \) will be denoted respectively by \( P^\varepsilon, M \) and \( P^\varepsilon \). Now, we consider the stopping times
\[ \eta_{1, M}(y) \overset{\text{def}}{=} \inf \{ t \geq 0 : \| y_t \| \geq M \} \quad \text{and} \quad \eta_{2, M}(y) \overset{\text{def}}{=} \inf \{ t \geq 0 : |y_t^1| \leq 1/M \}, \]
for any \( y = (y^1, y^2) \in C([0, \infty), \mathbb{R}^2) \) and
\[ \eta_M \overset{\text{def}}{=} \eta_{1, M} \wedge \eta_{2, M}. \quad (6.5) \]

It is not hard to see that \( \eta_{1, M} \) and \( \eta_{2, M} \) are lower semi-continuous so that \( \eta_M \) is also lower semi-continuous, that is their lower level sets are closed subsets of \( C([0, \infty), \mathbb{R}^2) \). The latter property will be used to obtain the convergence of the original process from the truncated one.
From the definition of the above stopping times and the cutoff function $\phi_M$ it is clear that
\[ p^{e,M} = p^e \quad \text{on} \quad \mathcal{M}_{\eta M}. \] (6.6)
This is the relation that links the truncated and original processes and that will be also used below in Sect. 6.2 to show the convergence of the original process.

For now, we have the following convergence results for the truncated process. We return the index $M$ for a moment to emphasize the truncation.

**Proposition 6.1.** The family $(Y^e)_{e > 0} \overset{\text{def}}{=} (I^{e,M}, \psi^{e,M})_{e > 0}$ converges in distribution in $\mathcal{C}([0, \infty), \mathbb{R}^2)$ to the unique solution $(Y^M_t)_{t \geq 0} = (I^M_t, \psi^M_t)_{t \geq 0}$ to the martingale problem with generator defined by
\[ Lh(y) \overset{\text{def}}{=} \int_{\tau=0}^1 \int_{u=0}^\infty R(u) F(0, y, \tau) \cdot \nabla_y (F(u, y, \tau) \cdot \nabla_y h(y)) \, du \, d\tau, \] (6.7)
where $\tilde{F}(u, y, \tau) \overset{\text{def}}{=} F(y, \tau + \omega(y)u)$, with $F$ defined in (6.4) and starting point $(I_0, \psi_0)$.

From this characterization, we can deduce the SDEs satisfied by the limiting process $Y^M_t = (I^M_t, \psi^M_t)$ up to the stopping time $\eta_M$.

**Proposition 6.2.** Denoting $\tilde{I}^M_t \overset{\text{def}}{=} I^M_{t \wedge \eta M}$, we have
\[ d\tilde{I}^M_t = \int_{\tau=0}^1 a(\tilde{I}^M_t, \tau) \, dW_{t \wedge \eta M}(\tilde{I}^M_t, \tau) \, d\tau + \int_{u=0}^\infty R(u) \left( \int_{\tau=0}^1 \partial_I a(I, \tau + \omega(I)u) \big|_{I = \tilde{I}^M_t} a(\tilde{I}^M_t, \tau) \right) \, du \, d\tau \] (6.8)
with $\tilde{I}^M_{t=0} = I_0$, and where $W$ is defined in (4.14). Also, we have
\[ d\psi^M_{t \wedge \eta M} = \int_{\tau=0}^1 b(\tilde{I}^M_t, \tau) \, dW_{t \wedge \eta M}(\tilde{I}^M_t, \tau) \, d\tau \] (6.9)
with $\psi^M_{t=0} = \theta_0$.

As we can see, this result is similar to the one of Theorem 4.2, with coefficients $a$ and $b$ themselves, but involving the cutoff $M$ through the stopping time $\eta_M$. The proofs of Proposition 6.1 and 6.2 are presented in Sects. 7 and 8, respectively. In the next section, we will show how these two propositions can be used to prove Theorem 4.2.
6.2. Convergence of \((I^\epsilon, \psi^\epsilon)_{\epsilon > 0}\). In order to prove Theorem 4.2, we need to remove the cutoff \(M\) to show the convergence in distribution of \((I^\epsilon, \psi^\epsilon)_{\epsilon > 0}\) to \((I, \psi_I)_{t \geq 0}\), where \(I, \psi\) are defined by (4.13)-(4.15). This is similar to Lemma 11.1.1 (p. 262) in [SV06], and we adapt the proof to our situation.

In what follows, we denote by \(P^M\) and \(P^0\) the distribution of \(Y^M = (I^M_t, \psi^M_t)_{t \geq 0}\) (defined in Proposition 6.1) and \(Y = (I_t, \psi_I)_{t \geq 0}\) (defined in Proposition 4.2), respectively. Note that according to (6.9) and (4.15) the distributions \(P^M\) and \(P^0\) are completely determined by their first marginal \(P^M_1\) and \(P^0_1\). Also, in view of (6.8) and (6.9), it is straightforward to see that (with \(\eta_M\) defined in (6.5))

\[
P^0 = P^M \quad \text{on } \mathcal{M}_{\eta_M}. \quad (6.10)
\]

To prove this convergence, let \(\Phi\) be a continuous bounded function on \(C([0, \infty), \mathbb{R}^2)\), and write

\[
E^{P^\epsilon} [\Phi(y_t, t \in [0, T])] - E^P [\Phi(y_t, t \in [0, T])]
\]

\[
= E^{P^\epsilon} [\Phi(y_t, t \in [0, T])] - E^{P^{\epsilon,M}} [\Phi(y_t, t \in [0, T])]
\]

\[
+ E^{P^{\epsilon,M}} [\Phi(y_t, t \in [0, T])] - E^{P^M} [\Phi(y_t, t \in [0, T])]
\]

\[
+ E^{P^M} [\Phi(y_t, t \in [0, T])] - E^{P^0} [\Phi(y_t, t \in [0, T])].
\]

Considering the following decomposition

\[
E^{P^\epsilon} [\Phi(y_t, t \in [0, T])] = E^{P^\epsilon} [\Phi(y_t, t \in [0, T])1_{[\eta_M \leq T]}]
\]

\[
+ E^{P^\epsilon} [\Phi(y_t, t \in [0, T])1_{[\eta_M > T]}]
\]

with

\[
E^{P^\epsilon} [\Phi(y_t, t \in [0, T])1_{[\eta_M \leq T]}] = E^{P^\epsilon} [\Phi(y_t \wedge \eta_M, t \in [0, T])1_{[\eta_M \leq T]}]
\]

\[
= E^{P^{\epsilon,M}} [\Phi(y_t \wedge \eta_M, t \in [0, T])1_{[\eta_M \geq T]}]
\]

\[
= E^{P^{\epsilon,M}} [\Phi(y_t, t \in [0, T])1_{[\eta_M > T]}]
\]

\[
= E^{P^{\epsilon,M}} [\Phi(y_t, t \in [0, T])] - E^{P^{\epsilon,M}} [\Phi(y_t, t \in [0, T])1_{[\eta_M \leq T]}].
\]

Here we used (6.6) in the second line. Using (6.6) again we note

\[
E^{P^\epsilon} [\Phi(y_t, t \in [0, T])] - E^{P^{\epsilon,M}} [\Phi(y_t, t \in [0, T])]
\]

\[
\leq \|\Phi\|_{\infty} \{P^\epsilon(\eta_M \leq T) + P^{\epsilon,M}(\eta_M \leq T)\}
\]

\[
\leq 2\|\Phi\|_{\infty} P^{\epsilon,M}(\eta_M \leq T).
\]

The same results hold for \(P^0\) and \(P^M\) instead of \(P^\epsilon\) and \(P^{\epsilon,M}\) respectively, but using (6.10) instead of (6.6). As a result, we obtain

\[
\limsup_{\epsilon \to 0} \left| E^{P^\epsilon} [\Phi(y_t, t \in [0, T])] - E^{P^0} [\Phi(y_t, t \in [0, T])] \right|
\]

\[
\leq 2\|\Phi\|_{\infty} \limsup_{\epsilon \to 0} P^{\epsilon,M}(\eta_M \leq T)
\]

\[
+ 2\|\Phi\|_{\infty} P^0(\eta_M \leq T).
\]
Notice that the stopping time $\eta_M$ is lower semi-continuous on $C((0, \infty), \mathbb{R}^2)$, so that $(\eta_M \leq M) \in \mathcal{M}_{\eta_M}$ is a closed subset of $C((0, \infty), \mathbb{R}^2)$. According to the Portmanteau theorem [Bil99, Theorem 2.1 p. 16] we have
\[
\limsup_{\varepsilon \to 0} |E^{\varepsilon}[\Phi(y_t, t \in [0, T])] - E^{0}[\Phi(y_t, t \in [0, T])]| \leq 4\|\Phi\|_{L^\infty} P^0(\eta_M \leq T),
\]
for any $M$. Moreover, we have
\[
P^0(\eta_M \leq T) \leq P^0(\zeta_M \leq T) + P^0(\zeta'_M \leq T).
\]
where for any any $y = (y^1, y^2) \in C((0, \infty), \mathbb{R}^2)$ we define
\[
\zeta'_M(y) \overset{\text{def}}{=} \inf\{t \geq 0 : |y^2_t| \geq M\}.
\]
By Proposition 5.1 we know
\[
P^0(\zeta_M \leq T) = P^0(\zeta_M \leq T) \longrightarrow 0.
\]
For the other term, let $0 < \eta < 1$ and write
\[
P^0(\zeta'_M \leq T) \leq P\left(\sup_{t \in [0, T]} |\psi_t| \geq M, \inf_{t \in [0, T]} I_t > \eta, \sup_{t \in [0, T]} I_t < \frac{1}{\eta}\right)
\]
\[+ P\left(\sup_{t \in [0, T]} I_t \geq \frac{1}{\eta}\right) + P\left(\inf_{t \in [0, T]} I_t \leq \eta\right),
\]
then using Markov inequality and (4.15), we obtain the following for some $C_\eta > 0$
\[
P\left(\sup_{t \in [0, T]} |\psi_t| \geq M, \inf_{t \in [0, T]} I_t > \eta, \sup_{t \in [0, T]} I_t < \frac{1}{\eta}\right)
\]
\[\leq \frac{1}{M^2} E\left[\sup_{t \in [0, T]} |\psi_t|^2 1_{\inf_{t \in [0, T]} I_t > \eta, \sup_{t \in [0, T]} I_t < \frac{1}{\eta}}\right]
\]
\[\leq \frac{1}{M^2} E\left[\sup_{t \in [0, T]} \int_0^1 \int_0^t b(I_s, \tau) dW_s(I_s, \tau) d\tau\right]^2 1_{\inf_{t \in [0, T]} I_t > \eta, \sup_{t \in [0, T]} I_t < \frac{1}{\eta}}
\]
\[+ \frac{C_\eta}{M^2}.
\]
Indeed, since we are on the event $\{\inf_{t \in [0, T]} I_t > \eta, \sup_{t \in [0, T]} I_t < \frac{1}{\eta}\}$ we avoid the singularity of the function $b$ at $I = 0$ and keep $I$ bounded. Therefore, we can directly bound the drift term in (4.15) by bounding continuous functions $a, b, \partial_{\eta} a, \partial_{\eta} b$ by some constant $C_\eta$. We will subsequently let $C_\eta$ be a constant that only depends on $\eta$ whose value may change from line to line. The details are as follows.

Write the Fourier expansions of $a$ and $b$ as
\[
a(I, \tau) = \sum_{n \in \mathbb{Z}} e^{2i\pi n \tau} a_n(I), \quad b(I, \tau) = \sum_{n \in \mathbb{Z}} e^{2i\pi n \tau} b_n(I).
\]
Considering only the first integrand in the drift term in (4.15) and substituting the above expansions, we obtain
\[
\int_{\tau=0}^{\infty} R(u) \int_{\tau=0}^{1} \partial_I \left(b(I, \tau + \omega(I)u)\right) |I| I_t a(I_t, \tau) d\tau du
\]
\begin{align*}
&= \int_{u=0}^{\infty} R(u) \int_{\tau=0}^{1} \left[ \sum_{n \in \mathbb{Z}} \left( (\partial_fb)_n(I_t) + 2i\pi n\omega'(I_t) u b_n(I_t) \right) e^{2i\pi n(\tau + \omega(I_t)u)} \right] \left( \sum_{n' \in \mathbb{Z}} a_{n'}(I_t) e^{2i\pi n'\tau} \right) d\tau \, du \\
&= \int_{u=0}^{\infty} R(u) \sum_{n \in \mathbb{Z}^*} \left( (\partial_fb)_n(I_t) + 2i\pi n\omega'(I_t) u b_n(I_t) \right) a_{-n}(I_t) e^{2i\pi n\omega(I_t)u} \, du \quad (6.11) \\
&= R_1 + R_2, \quad (6.12)
\end{align*}

where

\begin{align*}
R_1 &\stackrel{\text{def}}{=} \int_{u=0}^{\infty} R(u) \sum_{n \in \mathbb{Z}^*} (\partial_fb)_n(I_t) a_{-n}(I_t) \cos(2\pi n\omega(I_t)u) \, du \\
R_2 &\stackrel{\text{def}}{=} -\int_{u=0}^{\infty} R(u) \sum_{n \in \mathbb{Z}^*} (\partial_fb)_n(I_t) a_{-n}(I_t) \cdot 2\pi n\omega'(I_t) u \sin(2\pi n\omega(I_t)u) \, du.
\end{align*}

Here (6.11) followed by expanding the double sum in $n$ and $n'$, integrating in $\tau$, and observing that the only terms that do not vanish are those with $n + n' = 0$. The last equality (6.12) followed since the expression is real-valued and so must equal its real part.

Using (5.11) to bound the integral in $u$ and bounding $(\partial fb)_n(I_t), a_{-n}(I_t)$ by some constant $C_\eta$, we have that

$$|R_1| \leq C_\eta \sum_{n \in \mathbb{Z}^*} \frac{1}{n^2} \leq C_\eta.$$ 

Similarly, we can estimate that

$$\begin{align*}
&\int_{u=0}^{\infty} R(u) \cdot 2\pi nu \sin(2\pi n\omega(I_t)u) \, du \\
&= \int_{u=0}^{\infty} \left( \int_S r(p)e^{-\mu|p|^{2\beta}u} \, dp \right) \cdot 2\pi nu \sin(2\pi n\omega(I_t)u) \, du \\
&= \int_S dp \, r(p) \left( \int_{u=0}^{\infty} 2\pi nu \sin(2\pi n\omega(I_t)u)e^{-\mu|p|^{2\beta}u} \, du \right) \\
&= 2 \int_S dp \, r(p) \frac{(2\pi n)^2 \omega(I_t)\mu |p|^{2\beta}}{\left( \mu^2 |p|^{4\beta} + 4\pi^2 n^2 \omega^2(I_t) \right)^2} \leq \frac{C_{\omega_0}}{n^2} \int_S |p|^{2\beta} r(p) \, dp.
\end{align*}$$

Therefore, we also have

$$|R_2| \leq C_\eta \sum_{n \in \mathbb{Z}^*} \frac{1}{n^2} \leq C_\eta.$$ 

Combining these estimates we have

$$\left| \int_{u=0}^{\infty} R(u) \int_{\tau=0}^{1} \partial_f(b(I, \tau + \omega(I)u)) \big|_{I=I_t} a(I_t, \tau) \right| \leq C_\eta.$$
\[ + b(I_t, \tau + \omega(I_t) u) \partial f_a(I_t, \tau) d\tau du \leq C_\eta \sum_{n \in \mathbb{Z}} \frac{1}{n^2} \leq C_\eta. \]

For the stochastic integral term, we introduce a smooth function \( b^\eta \) such that for any \( \tau \in (0, 1) \) we have
\[
b^\eta(I, \tau) = b(I, \tau) \quad \text{if} \quad I > \eta \quad \text{and} \quad b^\eta(I, \tau) = 0 \quad \text{if} \quad I < \frac{\eta}{2}.
\]

Then
\[
E \left[ \sup_{t \in [0, T]} \left| \int_{\tau=0}^{t} \int_{s=0}^{t} b(I_s, \tau) dW_s(I_s, \tau) d\tau \right|^2 \mathbf{1}_{\{ \inf_{t \in [0, T]} I_t \geq \eta, \ \sup_{t \in [0, T]} I_t < \frac{1}{\eta} \}} \right]
\]
\[
\leq C_\eta,
\]
where the last inequality follows from the Burkholder-Davis-Gundy inequality to bound the supremum of a continuous local martingale. Finally, for any \( 0 < \eta < 1 \) we have
\[
\limsup_{M \to \infty} P(\zeta_M \leq T) \leq P\left( \sup_{t \in [0, T]} I_t \geq \frac{1}{\eta} \right) + P\left( \inf_{t \in [0, T]} I_t \leq \eta \right),
\]
and
\[
\lim_{\eta \to 0} P\left( \inf_{t \in [0, T]} I_t \leq \eta \right) = P\left( \inf_{t \in [0, T]} I_t = 0 \right) = 0,
\]
\[
\lim_{\eta \to 0} P\left( \sup_{t \in [0, T]} I_t \geq \frac{1}{\eta} \right) = 0 \quad \text{(by Lemma 5.2)}.
\]
As a result, we obtain
\[
\lim_{\varepsilon \to 0} E^{P_\varepsilon}\left[ \Phi(y_t, \ t \in [0, T]) \right] = E^{P_0}\left[ \Phi(y_t, \ t \in [0, T]) \right],
\]
which concludes the proof of Theorem 4.2.

It remains to prove Proposition 6.1, and we do this in the next section.

7. Proof of Proposition 6.1

We will prove convergence of the family \((Y^\varepsilon)_{\varepsilon \geq 0}\) by first showing tightness, and then showing all subsequential limits are solutions of a well-posed martingale problem with generator \((6.7)\). This follows the classical approach to prove weak convergence of a sequence of diffusions using martingale problem that is well developed in [SV06]. However, because the processes involved are not diffusions (and often not even Markovian), we use pseudo-generators and the perturbed test function method (see for instance [Kus84]).
7.1. Pseudo-generator and perturbed test function method. In this section, we follow [Kus84, Chap. 3]. Even though the pair of processes \((Y^\varepsilon_t, \tau^\varepsilon_t)_{t \geq 0}\) is Markov for each \(\varepsilon > 0\), the process \((Y^\varepsilon_t)_{t \geq 0}\) by itself is not Markov. Since analyzing the pair \((Y^\varepsilon_t, \tau^\varepsilon_t)_{t \geq 0}\) is difficult, we study the process \((Y^\varepsilon_t)_{t \geq 0}\) on its own by using the pseudo-generator. This allows us to apply martingale techniques to the non-Markovian process \((Y^\varepsilon_t)_{t \geq 0}\).

7.1.1. Pseudo-generator Let

\[ G^\varepsilon_t \overset{\text{def}}{=} \mathcal{G}^\varepsilon_t, \]

where \(G^\varepsilon_t\) is defined by (3.4), and \(\mathcal{S}^\varepsilon\) be the set of all measurable functions \(h(t)\), adapted to the filtration \((\mathcal{G}^\varepsilon_t)\), for which \(\sup_{t \leq T} E[|h(t)|] < +\infty\), and where \(T > 0\) is fixed. The p-lim and the pseudo-generator are defined as follows. Let \(h\) and \(h^\delta\) in \(\mathcal{S}^\varepsilon\) for all \(\delta > 0\). We say that \(h = \lim_{\delta \to 0} h^\delta\) if

\[ \sup_{t, \delta} E|h^\delta(t)| < +\infty \quad \text{and} \quad \lim_{\delta \to 0} E|h^\delta(t) - h(t)| = 0 \quad \forall t \geq 0. \]

We say that \(h \in \mathcal{D}(\mathcal{A}^\varepsilon)\) the domain of \(\mathcal{A}^\varepsilon\) and \(\mathcal{A}^\varepsilon h = g\) if \(h\) and \(g\) are in \(\mathcal{S}^\varepsilon\) and

\[ \lim_{\delta \to 0} \left( \frac{E^\varepsilon_t h(t + \delta) - h(t)}{\delta} - g(t) \right) = 0, \]

where \(E^\varepsilon_t\) is the conditional expectation given \(\mathcal{G}^\varepsilon_t\). The difference with respect to the usual infinitesimal generator is that we have average limit (via p-lim) rather than pointwise limit. This greatly helps us average out the noise in the limit \(\varepsilon \to 0\). A useful result [Kus84, Theorem 1, p. 39] about the pseudo-generator \(\mathcal{A}^\varepsilon\) that will be central in our proof is the following.

**Proposition 7.1** (Theorem 1, p. 30 in [Kus84]). Let \(h \in \mathcal{D}(\mathcal{A}^\varepsilon)\). Then

\[ M^\varepsilon_t(t) = h(t) - h(0) - \int_0^t \mathcal{A}^\varepsilon h(u) \, du \]

is a \((\mathcal{G}^\varepsilon_t)\)-martingale.

In other words, we introduce an operator \(\mathcal{A}^\varepsilon\) on functions of \(Y^\varepsilon\) having similar properties to those of an infinitesimal operator of a Markov process; more specifically, it satisfies a martingale problem.

7.1.2. Perturbed test function method The main idea behind the perturbed test function method is the following. To characterize the limiting process through a martingale problem with generator \(\mathcal{A}\), we can then try to compare \(\mathcal{A}^\varepsilon h\) with \(\mathcal{A}h\) for some test function \(h\). Unfortunately, \(\mathcal{A}^\varepsilon h\) may have singular terms in \(\varepsilon\), and we cannot proceed that way directly. Instead, try and extract the effective statistical behavior of the system under consideration from \(\mathcal{A}^\varepsilon\) by introducing \(h^\varepsilon\) an appropriate perturbation of \(h\). In other words, if we can show \(\lim_{\varepsilon \to 0} E[|h^\varepsilon(t) - h(t)|] = 0\) and \(\lim_{\varepsilon \to 0} E[|\mathcal{A}^\varepsilon h^\varepsilon(t) - \mathcal{A}h(Y^\varepsilon_t)|] = 0\) for each \(t\), then we can use Proposition 7.1 to conclude that the limiting process \((Y^\varepsilon_t)_{t \geq 0}\) satisfies the martingale problem with generator \(\mathcal{A}\).

To depict how the perturbation \(h^\varepsilon\) is built, let us briefly remind its construction in the simpler situation

\[ \dot{Y}^\varepsilon_t = \frac{1}{\varepsilon} F(Y^\varepsilon_t, v), \left(\frac{t}{\varepsilon^2}\right), \frac{t}{\varepsilon^2} \quad \text{and} \quad F(y, v, \tau) = v F(y, \tau), \]
An Oscillator Driven by Algebraically...

where \( \nu \) is a centered ergodic Markov process with invariant measure \( \mu \), and generator \( \mathcal{L}_\nu \) satisfying the Fredholm alternative. In this situation, to apply the perturbed test function method, we consider the extended Markov process \( Z^\varepsilon_t = (Y^\varepsilon_t, \nu(t/\varepsilon^2), \tau_{t/\varepsilon^2}) \) whose infinitesimal generator is given by

\[
\mathcal{L}^\varepsilon = \frac{1}{\varepsilon^2} Q + \frac{1}{\varepsilon} \mathbf{F}(y, v, \tau) \cdot \nabla_y \quad \text{with} \quad Q = (\mathcal{L}_\nu + \frac{\partial}{\partial \tau}).
\]

Looking for a perturbation \( h^\varepsilon \) of \( h \) of the form

\[
h^\varepsilon(y, v, \tau) = h(y) + \varepsilon h_1(y, v, \tau) + \varepsilon^2 h_2(y, v, \tau),
\]

we have

\[
\mathcal{L}^\varepsilon h^\varepsilon(y, v, \tau) = \frac{1}{\varepsilon}(Q h_1(y, v, \tau) + \mathbf{F}(y, v, \tau) \cdot \nabla_y h(y)) + Q h_2(y, v, \tau) + \mathbf{F}(y, v, \tau) \cdot \nabla_y h_1(y, v, \tau) + \mathcal{O}(\varepsilon).
\] (7.1)

To cancel the singular term we have to solve a Poisson problem of the form

\[
Q p = -q(y, v, \tau).
\]

Assuming the centering condition

\[
\int_0^1 E_\mu[q(y, \nu(0), \tau)] d\tau = 0,
\]

where the expectation is taken with respect to the invariant measure \( \mu \) of the process \( \nu \), the Poisson problem admits

\[
p = \int_0^\infty e^{u Q} q du
\]
as a solution, so that

\[
p(Z^\varepsilon_t) = \int_0^\infty E_\mu[q(Y^\varepsilon_t, \nu(u + \frac{t}{\varepsilon^2}), u + \frac{t}{\varepsilon^2}) \mid v(s/\varepsilon^2), s \leq t] du,
\]
from the Markov property. Remembering that \( \nu \) is centered and \( F \) is mean zero with respect to \( \tau \), we then have

\[
h_1(Z^\varepsilon_t) = \int_0^\infty E_\mu[\mathbf{F}(Y^\varepsilon_t, \nu(u + \frac{t}{\varepsilon^2}), u + \frac{t}{\varepsilon^2}) \cdot \nabla_y h(Y^\varepsilon_t) \mid v(s/\varepsilon^2), s \leq t] du. \] (7.2)

To treat the term of order 1 in (7.1), which still has fast oscillations through \( h_1 \), we solve again the Poisson problem but this time with

\[
q(y, v, \tau) = \mathbf{F}(y, v, \tau) \cdot \nabla_y h_1(y, v, \tau)
\]

\[
- \int_0^1 E_\mu[\mathbf{F}(y, \nu(0), \tau') \cdot \nabla_y h_1(y, \nu(0), \tau')] d\tau', \] (7.3)
which satisfies the above centering condition. As a result, we can chose
\[ h_2(Z_t^\varepsilon) = \int_0^\infty E\left[ (\mathbf{F} \cdot \nabla_y h_1) \left( Y_t^\varepsilon, v \left( u + \frac{t}{\varepsilon^2} \right), u + \frac{t}{\varepsilon^2} \right) \right] v(s/\varepsilon^2), s \leq t \] 
\[ - \int_0^1 E_\mu \left[ \mathbf{F}(Y_t^\varepsilon, v(0), \tau') \cdot \nabla_y h_1(Y_t^\varepsilon, v(0), \tau') \right] d\tau'. \]

Finally, due to the centering in (7.3), the expansion (7.1) becomes
\[ L^\varepsilon h^\varepsilon(Z_t^\varepsilon) = \int_0^1 \int_0^\infty E_\mu \left[ \mathbf{F}(Y_t^\varepsilon, v(0), \tau) \cdot \nabla_y \left( \mathbf{F}(Y_t^\varepsilon, v(u), \tau + u) \cdot \nabla_y h(Y_t^\varepsilon) \right) \right] du d\tau + O(\varepsilon), \]

and we obtain the same expression as in (6.7).

However, considering non-Markovian fluctuations, as well as long-range correlations, and a more complex periodic structure, make the construction of \( h^\varepsilon \) more delicate. The main concern is to handle the non-integrability at infinity of the correlation function of \( V \). To overcome this difficulty we must take advantage of the periodic structure as in the proof of Proposition 1.1. To illustrate the difficulty, let us take a look at the first corrector. Let \( h \) be a bounded smooth function on \( \mathbb{R}^2 \) with bounded derivatives, and set \( h_0^\varepsilon(t) = h(Y_t^\varepsilon) \). The pseudo-generator at \( h_0^\varepsilon \) is then given by
\[ \mathcal{A}^\varepsilon h_0^\varepsilon(t) = \frac{1}{\varepsilon} \int V \left( \frac{t}{\varepsilon^2}, dp \right) F(Y_t^\varepsilon, \tau_t^\varepsilon) \cdot \nabla h(Y_t^\varepsilon), \]

which is simply differentiating \( h_0^\varepsilon(t) \) with respect to \( t \) (the derivative \( \dot{Y}_t^\varepsilon \) is stated in Eq. (6.2)). The goal now is to modify the test function \( h_0^\varepsilon \) using a small perturbation \( h_1^\varepsilon \) so that the pseudo-generator \( \mathcal{A}^\varepsilon (h_0^\varepsilon + h_1^\varepsilon) \) does not blow up in \( \varepsilon \) anymore. This is equivalent to canceling the singular term in (7.1), and following (7.2) with a slight adaptation to our periodic structure, we chose
\[ h_1^\varepsilon(t) = \varepsilon \int_0^\infty du \int E_t^\varepsilon \left[ V \left( u + \frac{t}{\varepsilon^2}, dp \right) F \left( Y_t^\varepsilon, \tau_t^\varepsilon + u \omega(Y_t^\varepsilon) \right) \right] \cdot \nabla h(Y_t^\varepsilon) \]
\[ = \varepsilon \int_0^\infty du \int \sum_{n \in \mathbb{Z}^*} E_t^\varepsilon \left[ V \left( u + \frac{t}{\varepsilon^2}, dp \right) \right] e^{2i\pi n (\tau_t^\varepsilon + u \omega(Y_t^\varepsilon))} F_n(Y_t^\varepsilon) \cdot \nabla h(Y_t^\varepsilon), \]

where
\[ F_n(y) \overset{\text{def}}{=} \int_0^1 F(y, \tau) e^{-2i\pi n \tau} d\tau. \]

According to Lemma 3.3, we have
\[ E_t^\varepsilon V \left( u + \frac{t}{\varepsilon^2}, dp \right) = e^{-\mu |p|^{2\beta}} V \left( \frac{t}{\varepsilon^2}, dp \right), \]
so that
\[ h_1^\varepsilon(t) = \varepsilon \sum_{n \in \mathbb{Z}^*} e^{2i\pi n \tau_t^\varepsilon} \int V \left( \frac{t}{\varepsilon^2}, dp \right) F_n(Y_t^\varepsilon) \cdot \nabla h(Y_t^\varepsilon) \int_0^\infty e^{-\mu |p|^{2\beta}} e^{2i\pi n \omega(Y_t^\varepsilon)u} du \]
\[ = \varepsilon \sum_{n \in \mathbb{Z}^*} e^{2i\pi n \tau_t^\varepsilon} \int \frac{V(t/\varepsilon^2, dp)}{\mu |p|^{2\beta} - 2i\pi n \omega(Y_t^\varepsilon)} F_n(Y_t^\varepsilon) \cdot \nabla h(Y_t^\varepsilon). \]
This corrector is well defined since we sum over \( n \neq 0 \), because \( F \) has mean zero respect to the \( \tau \), and providing a nonnull imaginary part in the ratio of the later equality. This imaginary part is provided by the periodicity of \( F \), and without it \((n = 0)\) we would have a term of the form

\[
\int_S \frac{\lambda(p)}{|p|^{2\alpha+2\beta}} dp = \infty
\]

when looking at the first moment of \( h_1^\varepsilon(t) \) because of the long-range correlations. The same problem holds true for

\[
\mathcal{A}_\varepsilon (h_0^\varepsilon + h_1^\varepsilon) = \sum_{n,m \in \mathbb{Z}^*} e^{2i \pi (n+m) \tau_1^\varepsilon} \int \int \frac{V(t/\varepsilon^2, dp) V(t/\varepsilon^2, dp')}{\mu |p|^{2\beta} - 2i \pi n \omega(Y_1^\varepsilon)} \cdot \left( F_n(Y_1^\varepsilon)^T \nabla h(Y_1^\varepsilon) F_m(Y_1^\varepsilon) + \nabla h(Y_1^\varepsilon) \text{Jac} F_n(Y_1^\varepsilon) F_m(Y_1^\varepsilon) \right) \\
- \frac{2i \pi n}{\mu |p|^{2\beta} - 2i \pi n \omega(Y_1^\varepsilon)} F_n(Y_1^\varepsilon) \cdot \nabla h(Y_1^\varepsilon) F_m(Y_1^\varepsilon) \\
+ \sum_{n \in \mathbb{Z}^*} 2i \pi n e^{2i \pi n \tau_1^\varepsilon} \int \int \frac{V(t/\varepsilon^2, dp) V(t/\varepsilon^2, dp')}{\mu |p|^{2\beta} - 2i \pi n \omega(Y_1^\varepsilon)} \cdot F_n(Y_1^\varepsilon) \cdot \nabla h(Y_1^\varepsilon) G(Y_1^\varepsilon),
\]

but this time with a term of the form

\[
\int \int \frac{\lambda(p)}{|p|^{2\alpha+2\beta}} dp = \infty
\]

if we do not consider the periodic structure in our analysis. This explains why this mean-zero condition in \( \tau \) is crucial to the analysis. An additional problem is the dependence of \( \omega \) on the state \( Y_1^\varepsilon \) that does not allow us to claim that \( h_1^\varepsilon \) (or \( \mathcal{A}_\varepsilon (h_0^\varepsilon + h_1^\varepsilon) \)) is almost like a sum (or a sum of products) of Gaussian random fields by the integration in \( p \), and then claim boundedness properties using standard results of [AT07]. This latter difficulty is handled by the technical Lemma 3.4. Similar technical problems will rise when looking at the second-order correctors and their corresponding pseudo-generator, and this is the reason why the long-range correlation property requires careful analysis in the perturbed test function method.

7.2. Tightness. The tightness of \((Y_\varepsilon)_{\varepsilon > 0}\), seen as a family of continuous-time processes, implies weak subsequential convergence of stochastic processes. According to [Bil99, Theorem 13.4] it suffices to prove the tightness of \((Y_\varepsilon)_{\varepsilon > 0}\) in \( \mathcal{D}([0, T], \mathbb{R}^2) \) (which is the set of càdlàg functions with values in \( \mathbb{R}^2 \) equipped with the Skorokhod topology). Before going further, we remind the reader that \((Y_\varepsilon)_{\varepsilon > 0}\) is defined via a truncated process in Eq. (6.2) and there is an omitted index \( M \) in the functions \( F \) and \( G \) to indicate the truncation. This allows us to bound several expressions in the proofs in terms of \( M \).

**Proposition 7.2.** The family \((Y_\varepsilon)_{\varepsilon > 0}\) is tight in \( \mathcal{D}([0, T], \mathbb{R}^2) \).

The proof of Proposition 7.2 consists of applying [Kus84, Theorem 4, p. 48] thanks to the two following lemmas whose complete proofs are provided in [GILN21, Section 7.2]:
Lemma 7.1. For any $T > 0$, and $\eta > 0$

$$\lim_{\varepsilon \to 0} P\left( \sup_{t \in [0, T]} |h_1^\varepsilon(t)| > \eta \right) = 0 \quad \text{and} \quad \lim_{\varepsilon \to 0} \sup_{t \in [0, T]} E|h_1^\varepsilon(t)| = 0.$$ 

Lemma 7.2. For any $T > 0$, $\{A^\varepsilon(h_0^\varepsilon + h_1^\varepsilon)(t), \varepsilon \in (0, 1), 0 \leq t \leq T \}$ is uniformly integrable.

7.3. Identification of the limit. All the limit points of $(Y^\varepsilon)_{\varepsilon > 0}$ are identified via a well-posed martingale problem as stated in the following proposition. By abuse of notations, we still denote by $(Y^\varepsilon)_{\varepsilon > 0}$ a converging subsequence and by $(Y_t)_{t \geq 0}$ a limit point.

Proposition 7.3. All the limit points $(Y_t)_{t \geq 0}$ of $(Y^\varepsilon)_{\varepsilon > 0}$ are solutions of a well-posed martingale problem with generator

$$\mathcal{L}h(y) = \int_0^1 d\tau \int_0^\infty du \, R(u) \tilde{F}(0, y, \tau) \cdot \nabla_y (\tilde{F}(u, y, \tau) \cdot \nabla_y h(y)) \quad (7.8)$$

with

$$\tilde{F}(u, y, \tau) = F(y, \tau + \omega(y)u).$$

To prove this proposition we use the notion of pseudo-generator introduced in Sect. 7.1.1 and the perturbed test function technique that we have already used for the tightness by introducing additional correctors. Thanks to the pseudo-generator we can characterize the subsequential limits of $(Y^\varepsilon)_{\varepsilon > 0}$ as solutions of a well-posed martingale problem.

The complete details for the proof of Proposition 7.3 are provided in [GILN21, Section 7.3], but its outline is as follows. Recalling (7.5) we saw that $A^\varepsilon(h_0^\varepsilon)$ has a singular $1/\varepsilon$ term, and modified the test function $h_0^\varepsilon$ with a small perturbation $h_1^\varepsilon$ to remove this singular term. As a result, $A^\varepsilon(h_0^\varepsilon + h_1^\varepsilon)$ (given by (7.7)) is not singular anymore. However, it is not yet the generator of a martingale problem as $A^\varepsilon(h_0^\varepsilon + h_1^\varepsilon)$ still has oscillatory terms in the form of $V(t/\varepsilon^2, dp)V(t/\varepsilon^2, dp')$ and $e^{2i\pi(n+m)\tau^\varepsilon_i}$ for $n + m \neq 0$. However, these terms essentially vanish as $\varepsilon \to 0$ by introducing additional correctors to cancel these oscillations. We first choose $h_2^\varepsilon$ following the idea of (7.4) to average the random fluctuations

$$h_2^\varepsilon(t) = \varepsilon^2 \int_0^\infty \int_{Z^*} du \sum_{n, m \in Z^* \atop n + m \neq 0} e^{2i\pi(n+m)(\tau^\varepsilon_i + u \omega(Y^\varepsilon_t))} W_{n,m}(p, Y^\varepsilon_t)$$

\begin{align*}
\cdot \iint \left( E^{\varepsilon_i} \left[ V\left(u + \frac{t}{\varepsilon^2}, dp\right)V\left(u + \frac{t}{\varepsilon^2}, dp'\right) - E[V(0, dp)V(0, dp')] \right] \right)
\end{align*}

$$= \varepsilon^2 \sum_{n, m \in Z^* \atop n + m \neq 0} e^{2i\pi(n+m)\tau^\varepsilon_i}$$

\begin{align*}
\times \iint \frac{V(t/\varepsilon^2, dp)V(t/\varepsilon^2, dp') - r(p)\delta(p - p') dp dp'}{\mu(|p|^{2\beta} + |p'|^{2\beta}) - 2i\pi(n + m)\omega(Y^\varepsilon_t)} W_{n,m}(p, Y^\varepsilon_t),
\end{align*}
according to formula (3.6), and $h^\varepsilon_3$ to average out the complex exponentials

$$h^\varepsilon_3(t) = \varepsilon^2 \sum_{n,m \in \mathbb{Z}^*} \frac{e^{2i\pi(n+m)t}}{\varepsilon^3 - 2i\pi(n+m)\omega(Y^\varepsilon_t)} \int dp \, r(p) W_{n,m}(p, Y^\varepsilon_t).$$

For these two correctors we have

$$W_{n,m}(p, Y^\varepsilon_t) \overset{\text{def}}{=} \frac{1}{\langle |p|^2 \rangle^{2\beta} - 2i\pi n\omega(Y^\varepsilon_t)} \left( F_n(Y^\varepsilon_t)^T \nabla^2 h(Y^\varepsilon_t) F_m(Y^\varepsilon_t) 
+ \nabla h(Y^\varepsilon_t)^T \text{Jac} F_n(Y^\varepsilon_t) F_m(Y^\varepsilon_t) \right)$$

$$- \frac{2i\pi n}{\langle |p|^2 \rangle^{2\beta} - 2i\pi n\omega(Y^\varepsilon_t)}^2 F_n(Y^\varepsilon_t) \cdot \nabla h(Y^\varepsilon_t) F_m(Y^\varepsilon_t) \cdot \nabla \omega(Y^\varepsilon_t).$$

After we have constructed the perturbed test function

$$h^\varepsilon(t) = h^\varepsilon_0(t) + h^\varepsilon_1(t) + h^\varepsilon_2(t) + h^\varepsilon_3(t),$$

we show for the pseudo-generator that

$$A^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1 + h^\varepsilon_2 + h^\varepsilon_3)(t) = A^\varepsilon_2(t) + o(1)$$

with

$$A^\varepsilon_2(t) = \sum_{n \in \mathbb{Z}^*} \int \int V\left(\frac{t}{\varepsilon^2}, dp\right) V\left(\frac{t}{\varepsilon^2}, dp'\right) W_{n,-n}(p, Y^\varepsilon_t).$$

The expression $A^\varepsilon_2$ no longer has complex oscillatory components as it only picks up the indices with $n + m = 0$ in $A^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1)$, and for which we have (Lemma 7.6 in [GILN21])

$$\lim_{\varepsilon \to 0} E \left| \int_0^t (A^\varepsilon_2(u) - \mathcal{L} h(Y^\varepsilon_u)) \, du \right| = 0,$$

for any $t \geq 0$, with

$$\mathcal{L} h(y) = \sum_{n \in \mathbb{Z}^*} \int dp \, r(p) W_n(p, y),$$

and where $\mathcal{L}$ is defined in (7.8). To prove this latter convergence, corresponding to averaging the remaining random fluctuations, one can remark that there is no remaining periodic structure to handle the long-range correlation property by adding another corrector in the perturbed test function method. However, $A^\varepsilon_2$ is not a singular term so that we can use a more standard averaging procedure, which is provided in [GILN21, Sect. 7.7].

Using Theorem 7.1 we see that for all $N \geq 1$, any bounded continuous function $\Psi$, and every sequence $0 < s_1 < \cdots < s_N \leq s < t$ we have

$$E\left[ \Psi(Y^\varepsilon_{s_1}, \ldots, Y^\varepsilon_{s_N}) \left( M^\varepsilon_h(t) - M^\varepsilon_h(s) \right) \right] = 0,$$

so that

$$\lim_{\varepsilon \to 0} E\left[ \Psi(Y^\varepsilon_{s_1}, \ldots, Y^\varepsilon_{s_N}) \left( h(Y^\varepsilon_t) - h(Y^\varepsilon_s) - \int_s^t \mathcal{L} h(u) \, du \right) \right] = 0.$$
Combining with \( Y^\varepsilon_{t=0} = Y_0 \) for every \( \varepsilon > 0 \), the Eq. (7.10) implies that any subsequential limit \( (Y_t)_{t \ge 0} \) of \( (Y^\varepsilon)_{\varepsilon > 0} \) is a solution of the martingale problem for the generator \( L \) with initial condition \( Y_t = Y_0 \) for every \( t \ge 0 \). We claim in the next section that the solution to this martingale problem is unique. Given this, along with the tightness of Proposition 7.2, Prokhorov’s theorem implies the weak convergence of the sequence of processes \( (Y^\varepsilon)_{\varepsilon > 0} \) to the process \( (Y_t)_{t \ge 0} \) as desired. This completes the proof of Proposition 6.1.

7.4. Well-posedness of the martingale problem. We devote this section to showing that the martingale problem with generator \( L \) (see Eq. (7.8)) is well-posed. First, we rewrite \( L h(y) \), with \( y = (y_1, y_2) \), as

\[
L h(y) = \frac{1}{2} \sum_{j,l=1}^2 a_{jl}(y) \partial_{y_j y_l}^2 h(y) + \sum_{j=1}^2 \partial_{y_j} h(y) b_j(y),
\]

where

\[
a_{jl}(y) \overset{\text{def}}{=} \sum_{n \in \mathbb{Z}^*} R_n(y) F_{j,n}(y) F_{l,-n}(y) = \sum_{n \in \mathbb{Z}^*} R_n(y) F_{j,n}(y) F_{l,n}(y)
\]

(7.11)

with (equivalently as in formula (5.2))

\[
R_n(y) \overset{\text{def}}{=} \int dp \, r(p) \left( \frac{1}{\mu |p|^{2\beta}} - 2i\pi n \omega(y) \right) + \frac{1}{\mu |p|^{2\beta} + 2i\pi n \omega(y)}
\]

\[
= \int dp \, r(p) \frac{2\mu |p|^{2\beta}}{\mu |p|^{4\beta} + 4\pi^2 n^2 \omega^2(y)}
\]

\[
= 2 \int_0^\infty du \, E[v(u) v(0)] \cos(2\pi n \omega(y) u),
\]

(7.12)

and

\[
b_j(y) \overset{\text{def}}{=} \sum_{n \in \mathbb{Z}^*} \int dp \, \frac{r(p)}{\mu |p|^{2\beta} - 2i\pi n \omega(y)}
\]

\[
\cdot \left( \sum_{l=1}^2 \left( \partial_{y_l} F_{j,n}(y) F_{l,-n}(y) - \frac{2i\pi n}{\mu |p|^{2\beta} - 2i\pi n \omega(y)} F_{j,n}(y) F_{l,-n}(y) \partial_{y_l} \omega(y) \right) \right).
\]

(7.13)

The uniqueness of the martingale problem is guaranteed by [KS91, Corollary 4.9 p. 317, Proposition 3.20 p. 309, and Theorem 2.5 p. 287] together with [Fri06, Theorem 1.2 p. 129], where the latter provides the locally Lipschitz behavior of a nonnegative square root. It is not hard to see that \( b \) is locally Lipschitz with respect to \( y \), as well as \( a(y) \) is a symmetric nonnegative matrix, which is \( \mathbb{C}^2(\mathbb{R}^2) \) with respect to \( y \). The smoothness property is direct from (7.11) and (7.12). Regarding the nonnegativity, we have for any \( x \in \mathbb{R}^2 \)

\[
x^T a(y) x = \langle \sigma^*(y)(x), \sigma^*(y)(x) \rangle_{\mathcal{H}} \ge 0
\]

(7.14)

with the Hilbert space

\[
\mathcal{H} = \left\{ u : \mathbb{Z}^* \to \mathbb{C} : \overline{u_{-n}} = u_n \quad \text{and} \quad \sum_{n \in \mathbb{Z}^*} |u_n|^2 < \infty \right\},
\]
the adjoint \( \sigma_M^* \) is an operator from \( \mathbb{R}^2 \) to \( \mathcal{H} \) defined by
\[
\sigma_M^*(y)(n, x) \overset{\text{def}}{=} R_n^{1/2}(y) F_{1,n}(y)x_1 + R_n^{1/2}(y) F_{2,n}(y)x_2,
\]
and \( \sigma \), from \( \mathcal{H} \) to \( \mathbb{R}^2 \), is defined by
\[
\sigma(y)(u, j) \overset{\text{def}}{=} \sum_{n \in \mathbb{Z}^*} R_n^{1/2}(y) F_{j,n}(y)u_{-n} \quad j = 1, 2. \tag{7.15}
\]
Hence, we have an explicit self-adjoint Hilbert-Schmidt square root for \( a \) that we use in Sect. 8 to derive a SDE for the process \((Y_t)_{t \geq 0}\). It is also clear that \( \sigma \) is locally Lipschitz with respect to \( y \), so that we can conclude the uniqueness of the martingale problem.

Let us finish by rewriting the martingale problem in a more convenient way and let us start with \( b \). Using the fact that
\[
-2i\pi n \partial_{y_l} \omega(y) = \partial_{y_l} \int_0^\infty u \ e^{-(\mu|p|^{2\beta}-2i\pi n \omega(y))u} \ du
\]
we sum Fourier series in (7.13) to obtain
\[
b(y) = \int_0^1 d\tau \int_0^\infty du \ R(u) \text{Jac}_\chi(\tilde{F}(u, y, \tau)) \tilde{F}(0, y, \tau),
\]
where
\[
\tilde{F}(u, y, \tau) = F(y, \tau + \omega(y)u).
\]
For \( a \) we have the following formulation
\[
a_{jl}(y) = \int_0^1 d\tau \int_0^\infty du \ R(u) \left( \tilde{F}_j(u, y, \tau) \tilde{F}_l(0, y, \tau) + \tilde{F}_j(0, y, \tau) \tilde{F}_l(u, y, \tau) \right)
\]
\[
= 2 \int_0^1 d\tau \int_0^\infty du \ R(u) \tilde{F}_j(u, y, \tau) \tilde{F}_l(0, y, \tau), \quad j, l \in \{1, 2\}.
\]
Finally, to conclude, we just have to remark that according to (6.4), as well as the definition of \( a \) involving the integration of the periodic variable over a whole period, we obtain the formulation (7.8).

8. Proof of Proposition 6.2

The proof of this proposition follows the treatment in [KS91, Proposition 4.6 p. 315] and consists of deriving the SDE corresponding to the martingale problem with generator (6.7). To prove this correspondence, we consider first the function \( h_j(y) = y_j \) for any \( j \in \{1, 2\} \), so that
\[
M_j^0(t) = Y_{j,t} - Y_{j,0} - \int_0^t b_j(Y_s) \, ds \tag{8.1}
\]
is a martingale, where \( b \) is defined as \( b^M \) in (7.13). Proceeding the same way choosing \( h_{jl}(y) = y_j y_l \) one can show that for any \( j, l \in \{1, 2\} \)

\[
M^0_j(t) M^0_l(t) - \int_0^t a_{jl}(Y_s) \, ds
\]

is a martingale. (Recall, \( a \) is defined by (7.11).) Therefore \( M^0 \) is a martingale with quadratic variation given by

\[
t \mapsto \int_0^t a_{jl}(Y_s) \, ds.
\]

Using (7.14), we note \( a = \sigma \sigma^* \), where \( \sigma \) is defined by (7.15). We can now apply a standard martingale representation theorem involving Hilbert spaces (see [DPZ96, Theorem 8.2] for instance). This guarantees the existence of a complex valued cylindrical Brownian motion \( B \) (\( B = (B^1 - i B^2)/\sqrt{2} \), with \( B^1 \) and \( B^2 \) being two independent real valued cylindrical Brownian motions) on the Hilbert space \( T \), possibly defined on an extension of the probability space under consideration, and such that with probability one

\[
M^0(t) = \int_0^t \sigma(Y_s) \, dB_s
\]

for all \( t \geq 0 \). As a result, from (8.1) we obtain the following SDE for \( (Y_t)_{t \geq 0} \),

\[
dY_t = \sigma(Y_t) \, dB_t + b(Y_t) \, dt.
\]

Note that, because of symmetries, and that for the cylindrical Brownian motion on \( T \) we necessarily have \( B^1_{-n} = B^1_n \) and \( B^2_{-n} = -B^2_n \), this defines a real valued equation.

Regarding \( \sigma \), one can write

\[
\sigma_j(Y_t) dB_t = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}^*} \tilde{F}_{j,n}(Y_t) R^{1/2}_{n}(Y_t)(dB_{t,n}^1 - i dB_{t,n}^2)
\]

(8.2)

where

\[
dW_t(y, \tau) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}^*} e^{2i\pi n \tau} R^{1/2}_{n}(y)(dB_{t,n}^1 - i dB_{t,n}^2)
\]

is a real valued Brownian field with covariance function

\[
EW_t(y, \tau) W_s(y, \tau') = t \wedge s \sum_{n \in \mathbb{Z}^*} e^{2i\pi n (\tau - \tau')} R_n(y)
\]

\[
= t \wedge s \sum_{n \in \mathbb{Z}^*} e^{2i\pi n (\tau - \tau')} \int dp \, r(p)
\]

\[
\cdot \int_0^\infty du \left[ e^{-(|p|^{2\beta} + 2i\pi n \omega(y)) u} + e^{-(|p'|^{2\beta} - 2i\pi n \omega(y)) u} \right]
\]

\[
= t \wedge s \int_0^\infty du \, R(u)
\]

\[
\cdot \sum_{n \in \mathbb{Z}^*} \left[ e^{2i\pi n (\tau - \tau' + \omega(y) u)} + e^{2i\pi n (\tau - \tau' - \omega(y) u)} \right].
\]
so that for any test functions \( \phi_1, \phi_2 \) in the Hilbert space

\[
L^2_0(\mathbb{T}) = \left\{ \phi \in L^2(\mathbb{T}) : \int_0^1 \phi(\tau) d\tau = 0 \right\},
\]

the covariance function reads

\[
E W_t(y, \phi_1) W_s(y, \phi_2) = t \wedge s \int_0^\infty du \, R(u) \cdot \int_0^1 d\tau \left[ \phi_1(\tau + \omega(y)u)\phi_2(\tau) + \phi_1(\tau)\phi_2(\tau + \omega(y)u) \right].
\]

To determine the SDE satisfied by \( (I^M_t)_{t \geq 0} \) we just have to consider the first coordinate of \( (Y_t)_{t \geq 0} \). Then, we have

\[
dI^M_t = \int_0^1 d\tau \, a^M(I^M_t, \tau) \, dW_t(I^M_t, \tau) + \int_0^1 d\tau \int_0^\infty du \, R(u) \tilde{A}(u, I^M_t, \tau) \, dt
\]

with

\[
\tilde{A}(u, I, \tau) = \partial_I (a^M(I, \tau + \omega(I)u)) a^M(I, \tau) + \partial_\tau a^M(I, \tau + \omega(I)u) \left( b^M(I, \tau) - \left\{ b^M(I) \right\} \right).
\]

Using the fact that \( a^M \) is mean-zero in \( \tau \), and integrating by parts, we have

\[
\int_0^1 d\tau \, \partial_\tau a^M(I, \tau + \omega(I)u) \left( b^M(I, \tau) - \left\{ b^M(I) \right\} \right)
= \int_0^1 d\tau \, \partial_\tau a^M(I, \tau + \omega(I)u) b^M(I, \tau)
= - \int_0^1 d\tau \, a^M(I, \tau + \omega(I)u) \partial_\tau a^M(I, \tau).
\]

Now, we restrict our study for \( t \leq \eta_M \) so that, according to (4.10), we have for \( I \in (1/M, M) \)

\[
\partial_\tau b^M(I, \tau) = \partial_\tau b(I, \tau) = \partial_\tau \partial_I \varphi^{-1}(I, \tau) = \partial_I \partial_\tau \varphi^{-1}(I, \tau) = - \partial_I a(I, \tau),
\]

so that

\[
\int_0^1 d\tau \, \tilde{A}(u, I, \tau) = \int_0^1 d\tau \, \partial_I \left( a(I, \tau + \omega(I)u) \right) a(I, \tau) + a(I, \tau + \omega(I)u) \partial_I a(I, \tau).
\]

We then obtain the desired result for the stopped process \( \tilde{I}^M_t = I^M_{t \wedge \eta_M} \). Now for the process \( \psi^M_{t \wedge \eta_M} \), we look at the second coordinate of \( (Y_t)_{t \geq 0} \), and we have

\[
d\psi^M_t = \int_0^1 d\tau \, b^M(I^M_t, \tau) \, dW_t(I^M_t, \tau) + \int_0^1 d\tau \int_0^\infty du \, R(u) \tilde{B}(u, I^M_t, \tau) \, dt
\]
with
\[ \tilde{B}(u, I, \tau) = \partial_I (b^M(I, \tau + \omega(I)u) - \langle b^M(I) \rangle) + \partial_\tau (b^M(I, \tau + \omega(I)u) - \langle b^M(I) \rangle (b^M(I, \tau) - \langle b^M(I) \rangle). \]

Using again that \( a^M \) and \( \partial_\tau b^M \) are mean-zero in \( \tau \), together with an integration by parts, we obtain
\[ \int_0^1 d\tau \tilde{B}(u, I, \tau) = \int_0^1 d\tau \partial_I (b^M(I, \tau + \omega(I)u))a^M(I, \tau) \]
\[ - \int_0^1 d\tau b^M(I, \tau + \omega(I)u)\partial_\tau b^M(I, \tau). \]

Since \( I \in (1/M, M) \), we know
\[ \partial_\tau b^M(I, \tau) = -\partial_I a(I, \tau), \]
the proof of Proposition 6.2 is complete.

Acknowledgments. The authors thank Lenya Ryzhik for suggesting this problem to us, and for many helpful discussions. This work has been partially supported by the National Science Foundation under grants NSF DMS-1813943, DMS-1515187 to AN and HL, NSF DMS-1814147 to GI, and the Center for Nonlinear Analysis.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

A Proofs of Lemmas 2.1 and 2.2

In this appendix we prove Lemmas 2.1 and 2.2. While the proofs are a bit lengthy, they are a direct computation.

Proof of Lemma 2.1. From (2.1) we note
\[ w_1^\epsilon(t) = \epsilon \int_0^{t/\epsilon^2} \sin(\tau)v(\tau) d\tau, \quad (A.1) \]
and hence
\[ E(w_1^\epsilon(t) - w_1^\epsilon(s))^2 = \epsilon^2 \int_{s/\epsilon^2}^{t/\epsilon^2} \int_{s/\epsilon^2}^{t/\epsilon^2} \sin(\theta) \sin(\tau)R(\theta - \tau) d\theta d\tau \]
\[ = \epsilon^2 \int_{s/\epsilon^2}^{t/\epsilon^2} \int_{s/\epsilon^2}^{t/\epsilon^2} [\cos(\theta - \tau) - \cos(\theta + \tau)] R(\theta - \tau) d\theta d\tau \]
\[ = \frac{1}{2} \left( I_1^\epsilon(s, t) - I_2^\epsilon(s, t) \right), \quad \text{(A.2)} \]
where

\[ I_{1}^{\varepsilon}(s, t) \overset{\text{def}}{=} \varepsilon^{2} \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \cos(\theta - \tau) R(\theta - \tau) d\theta d\tau, \]

and \[ I_{2}^{\varepsilon}(s, t) \overset{\text{def}}{=} \varepsilon^{2} \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \cos(\theta + \tau) R(\theta - \tau) d\theta d\tau. \]

We first analyze \( I_{1}^{\varepsilon}(s, t) \). Making a change of variables \( z = \theta - \tau \) and integrating by parts with respect to variable \( \tau \), we have

\[
I_{1}^{\varepsilon}(s, t) = \varepsilon^{2} \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \cos(\frac{s}{\varepsilon^{2}} - \tau) \cos(\frac{t}{\varepsilon^{2}} - \tau) R(\frac{s}{\varepsilon^{2}} - \tau) d\tau.
\]

The first term on the right reduces to

\[
\varepsilon^{2} \left[ \tau \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \cos(\frac{s}{\varepsilon^{2}} - \tau) R(\frac{s}{\varepsilon^{2}} - \tau) d\tau \right]_{\tau=s/\varepsilon^{2}} = (t - s) \int_{0}^{(t-s)/\varepsilon^{2}} \cos(\frac{s}{\varepsilon^{2}} - \tau) R(\frac{s}{\varepsilon^{2}} - \tau) d\tau.
\]

For the second term on the right, we note

\[
\varepsilon^{2} \left[ \tau \int_{s/\varepsilon^{2}}^{t/\varepsilon^{2}} \cos(\frac{t}{\varepsilon^{2}} - \tau) R(\frac{t}{\varepsilon^{2}} - \tau) d\tau \right]_{\tau=s/\varepsilon^{2}} = t \int_{0}^{(t-s)/\varepsilon^{2}} \cos(\frac{t}{\varepsilon^{2}} - \tau) R(\frac{t}{\varepsilon^{2}} - \tau) d\tau.
\]

Hence,

\[
I_{1}^{\varepsilon}(s, t) = 2(t - s) \int_{0}^{(t-s)/\varepsilon^{2}} \cos(\frac{s}{\varepsilon^{2}} - \tau) R(\frac{s}{\varepsilon^{2}} - \tau) d\tau - 2\varepsilon^{2} \int_{0}^{(t-s)/\varepsilon^{2}} \tau \cos(\frac{s}{\varepsilon^{2}} - \tau) R(\frac{s}{\varepsilon^{2}} - \tau) d\tau = I_{11}^{\varepsilon} - 2I_{12}^{\varepsilon}.
\]
Clearly
\[ |I_{11}^\epsilon| \leq 2(t - s) \sup_{x \in \mathbb{R}} \int_0^x \cos(z) R(z) \, dz. \tag{A.4} \]

Note that the supremum on the right-hand side is finite for all \( \gamma \in (0, 2) \), as by Remark 1.1 we know \( \lim_{x \to \infty} \int_0^x \cos(z) R(z) \, dz \) exists and is finite.

To estimate \( I_{12} \), we divide the analysis into cases.

**Case I:** \( \gamma \in (0, 1) \). Integrating by parts we note
\[ I_{12}^\epsilon = \epsilon^2 \left[ \cos(z) R(z) \right]_{z=0}^{z=(t-s)/\epsilon^2} - \epsilon^2 \int_0^{(t-s)/\epsilon^2} \cos(z) (R'(z)z + R(z)) \, dz \]
\[ = \epsilon^2 \gamma (t-s)^{1-\gamma} L\left( \frac{1}{\epsilon^2} \right) - \epsilon^2 \int_0^{(t-s)/\epsilon^2} \cos(z) (R'(z)z + R(z)) \, dz. \tag{A.5} \]

For the first term on the right we note that
\[ \epsilon^2 \gamma (t-s)^{1-\gamma} L\left( \frac{1}{\epsilon^2} \right) \leq 2 \epsilon^2 \gamma L\left( \frac{1}{\epsilon^2} \right) (t-s)^{1-\gamma}, \]
for all sufficiently small \( \epsilon \).

For the second term, we note that (1.8) implies
\[ |R'(z)z| = |L'(z)z^{1-\gamma} - \gamma L(z)z^{-\gamma}| \leq 2|R(z)|, \]
for all sufficiently large \( z \).

Thus
\[ \epsilon^2 \int_0^{(t-s)/\epsilon^2} \cos(z) (R'(z)z + R(z)) \, dz \leq 3 \epsilon^2 \int_0^{(t-s)/\epsilon^2} |R(z)| \, dz \]
\[ \leq 4 \epsilon^2 \gamma L\left( \frac{1}{\epsilon^2} \right) (t-s)^{1-\gamma}, \tag{A.6} \]
where the last inequality followed from Karamata’s theorem [BIKS18, Th 6.2.1], and the fact that \( L \) is slowly varying. Combining these, we see
\[ |I_{12}^\epsilon| \leq 10 \epsilon^2 \gamma L\left( \frac{1}{\epsilon^2} \right) (t-s)^{1-\gamma}, \tag{A.7} \]
for all sufficiently small \( \epsilon \).

**Case II:** \( \gamma = 1 \). We follow the proof in the case \( \gamma < 1 \), with a few minor changes.

For the first term on the right of (A.5), we note that \( L(z) \leq 2R(0)z \) for all sufficiently small \( z \). Since \( L \) is slowly varying at infinity we must certainly have \( L(z) \leq z \) for all sufficiently large \( z \). Hence we can find a finite constant \( C \) such that \( L(z) \leq Cz \) for all \( z \geq 0 \). As a result, we have
\[ \epsilon^2 L\left( \frac{t-s}{\epsilon^2} \right) \leq C(t-s). \]
For the remaining terms, note that the function \( g(x) \overset{\text{def}}{=} \int_{0}^{x} R(z) \, dz \) is slowly varying (see for instance [BGT89] Proposition 1.5.9a, p. 26). Thus using (A.6) we see that

\[
\varepsilon^2 \int_{0}^{(t-s)/\varepsilon^2} \sin(z)(R'(z)z + R(z)) \, dz \leq 3\varepsilon^2 \int_{0}^{(t-s)/\varepsilon^2} |R(z)| \, dz \quad (A.8)
\]

\[
= 3\varepsilon^2 g \left( \frac{t - s}{\varepsilon^2} \right) . \quad (A.9)
\]

Since \( g(0) = 0 \), \( g'(0) = R(0) < \infty \), and \( g(z) \leq Cz \) for all \( z \geq 0 \) and a finite constant \( C \). (As before, we allow the constant \( C \) to change from line to line, provided it does not depend on \( \varepsilon, t \) and \( s \).) Consequently,

\[
|I_{21}^\varepsilon| \leq C(t - s),
\]

as desired.

**Case III:** \( \gamma \in (1, 2) \). Directly integrating \( I_{12}^\varepsilon \) and using Karamata’s theorem we see

\[
|I_{12}^\varepsilon| \leq 4\varepsilon^2(\gamma - 1) L \left( \frac{1}{\varepsilon^2} \right) (t - s)^{2-\gamma}.
\]

We now turn our attention to the term \( I_{2}^\varepsilon \). Substituting \( z = \theta - \tau \), we note

\[
I_{2}^\varepsilon = \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} \cos(z + 2\tau) R(z) \, dz \, d\tau
\]

\[
= \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} [\cos(z) \cos(2\tau) - \sin(z) \sin(2\tau)] R(z) \, dz \, d\tau \quad (A.10)
\]

\[
= I_{21}^\varepsilon - I_{22}^\varepsilon , \quad (A.11)
\]

where

\[
I_{21}^\varepsilon \overset{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} \cos(z) \cos(2\tau) R(z) \, dz \, d\tau ,
\]

\[
I_{22}^\varepsilon \overset{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} \sin(z) \sin(2\tau) R(z) \, dz \, d\tau . \quad (A.12)
\]

We again divide the analysis into cases.

**Case I:** \( \gamma \in (0, 1) \). Integrating by parts with respect to \( \tau \), we observe

\[
I_{21}^\varepsilon = \frac{\varepsilon^2}{2} \left[ \sin(2\tau) \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} \cos(z) R(z) \, dz \right]_{\tau=s/\varepsilon^2}^{\tau=t/\varepsilon^2} \quad (A.13)
\]

\[
- \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \left( \cos \left( \frac{s}{\varepsilon^2} - \tau \right) \frac{s}{\varepsilon^2} - \tau \right) R \left( \frac{s}{\varepsilon^2} - \tau \right) \sin(2\tau) \, d\tau . \quad (A.14)
\]
and hence by Karamata’s theorem,
\[ |I_{21}| \leq C \varepsilon^2 \int_0^{(t-s)/\varepsilon^2} |R(z)| \, dz \leq C \varepsilon^{2\gamma} L \left( \frac{1}{\varepsilon^2} \right) (t-s)^{1-\gamma}, \]  
(A.15)

for some finite constant C.

**Case II:** \( \gamma \in (0, 1) \). As in the previous case, we know from (A.15) that
\[ |I_{21}| \leq C \varepsilon^2 \int_0^{(t-s)/\varepsilon^2} |R(z)| \, dz. \]

Using the same argument as that used to bound (A.9) we obtain
\[ |I_{21}| \leq C (t-s) \]
as desired.

**Case III:** \( \gamma \in (1, 2) \). As before, set \( g(x) \overset{\text{def}}{=} \int_0^x |R(z)| \, dz \), and note
\[ |I_{21}| \leq \varepsilon^2 \int_0^{t/\varepsilon^2} \left( g \left( \frac{t}{\varepsilon^2} - \tau \right) - g \left( \frac{s}{\varepsilon^2} - \tau \right) \right) \, d\tau = 2\varepsilon^2 \int_0^{(t-s)/\varepsilon^2} g(z) \, dz. \]
Let \( G(x) = \int_0^x g(z) \, dz \) and note that \( G \) is regularly varying with index \( 2 - \gamma \). Thus
\[ |I_{21}| \leq 2\varepsilon^2 G \left( \frac{t-s}{\varepsilon^2} \right) = 3\varepsilon^{2\gamma-2} \tilde{L} \left( \frac{1}{\varepsilon^2} \right) (t-s)^{2-\gamma} \]
The estimates for \( I_{22} \) are identical to those for \( I_{21} \). Combining the above estimates we obtain (2.3) for \( i = 1 \). The proof when \( i = 2 \) is identical, and this finishes the proof of Lemma 2.1.

**Proof of Lemma 2.2.** We first prove Eq. (2.4) for \( i = 1 \). To do this, we follow the same computation as in the proof of Lemma 2.1 up to (A.3). Now note
\[ \lim_{\varepsilon \to 0} I_{11}^\varepsilon = 2 \int_0^\infty \cos(z) R(z) \, dz, \]
where the above integral converges absolutely for \( \gamma \in (1, 2) \), and conditionally for \( \gamma \in (0, 1] \) (see Remark 1.1). When \( \gamma \neq 1 \), the proof of Lemma 2.1 already shows that \( I_{12}^\varepsilon, I_{21}^\varepsilon \) and \( I_{22}^\varepsilon \) all vanish as \( \varepsilon \to 0 \). When \( \gamma = 1 \), the proof of Lemma 2.1 shows
\[ |I_{12}^\varepsilon| + |I_{21}^\varepsilon| \leq C \varepsilon^2 \int_0^{(t-s)/\varepsilon^2} |R(z)| \, dz = C \varepsilon^{2\gamma} \tilde{L} \left( \frac{1}{\varepsilon^2} \right), \]
where \( g(x) = \int_0^x |R(z)| \, dz \). Since \( g \) is slowly varying this vanishes as \( \varepsilon \to 0 \). This proves (2.4) as claimed.

To prove (2.5), we note
\[ E(w_1^\varepsilon(t) - w_1^\varepsilon(s))(w_2^\varepsilon(t) - w_2^\varepsilon(s)) \]
\[ = \varepsilon^2 \int_{t/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} R(\theta - \tau) \sin(\theta) \cos(\tau) \, d\theta \, d\tau \]
\[
  \tilde{I}_1^\varepsilon + \tilde{I}_2^\varepsilon,
\]

where

\[
  \tilde{I}_1^\varepsilon \overset{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\tau - \theta') R(\tau' - \theta) \, d\theta' \, d\tau',
\]

\[
  \tilde{I}_2^\varepsilon \overset{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\tau' + \theta') R(\tau' - \theta) \, d\theta' \, d\tau'.
\]

We claim \( \tilde{I}_1^\varepsilon = 0 \). To see this, make the changes of variables

\[
  \theta' \overset{\text{def}}{=} \frac{(t + s)}{\varepsilon^2} - \theta, \quad \tau' \overset{\text{def}}{=} \frac{(t + s)}{\varepsilon^2} - \tau,
\]

and rewrite \( \tilde{I}_1^\varepsilon \) as

\[
  \tilde{I}_1^\varepsilon = \varepsilon^2 \int_{t/\varepsilon^2}^{s/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\tau' - \theta') R(\tau' - \theta') \, d\theta' \, d\tau'
\]

\[
  = -\varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\theta' - \tau') R(\theta' - \tau') \, d\theta' \, d\tau'
\]

\[
  = -\tilde{I}_1^\varepsilon.
\]

This implies \( \tilde{I}_1^\varepsilon = 0 \) as claimed.

We now turn our attention to \( \tilde{I}_2^\varepsilon \). Making the change of variable \( z = \theta - \tau \) we note

\[
  \tilde{I}_2^\varepsilon = \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(2\tau) R(z) \, dz \, d\tau
\]

\[
  = \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} (\sin(z) \cos(2\tau) + \cos(z) \sin(2\tau)) R(z) \, dz \, d\tau.
\]

This is similar to the expression for \( I_2^\varepsilon \) (Eq. (A.10)) and following the proof of Lemma 2.1 we see \( \tilde{I}_2^\varepsilon \to 0 \) as \( \varepsilon \to 0 \). Since \( \tilde{I}_1^\varepsilon \) and \( \tilde{I}_2^\varepsilon \) both vanish as \( \varepsilon \to 0 \) we obtain (2.5) as claimed.

\[\Box\]

B Proof of Proposition 1.2

In this section we show convergence of the rescaled processes \( u^\varepsilon \) (Eq. (1.11)) to fBm. The result is similar to well known results [Taq75, Taq77, Mar05], and the proof is presented here for completeness.
**Proof of Proposition 1.2.** when \( \gamma \in (0, 1) \) Using (1.5) and (1.11) we note

\[
E(u^\varepsilon(t) - u^\varepsilon(s))^2 = \frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^t \int_{r_2=s}^{r_1} R \left( \frac{r_1 - r_2}{\varepsilon^2} \right) \, dr_2 \, dr_1.
\]

By (1.7) and the uniform convergence theorem [BGT89, Th. 1.2.1], we know

\[
\frac{1}{\sigma(\varepsilon)^2} \left| R \left( \frac{r_1 - r_2}{\varepsilon^2} \right) \right| = \frac{1}{L(\varepsilon^{-2}) |r_1 - r_2|^{\gamma}} L \left( \frac{r_1 - r_2}{\varepsilon^2} \right) \leq \frac{2}{(r_1 - r_2)^\gamma}, \tag{B.1}
\]

for all \( r_1, r_2 \in [s, t] \) and all sufficiently small \( \varepsilon \). Hence

\[
E(u^\varepsilon(t) - u^\varepsilon(s))^2 \leq 4 \int_{r_1=s}^t \int_{r_2=s}^{r_1} \frac{dr_2 \, dr_1}{(r_1 - r_2)^\gamma} = \frac{4(t - s)^{2-\gamma}}{(1 - \gamma)(2 - \gamma)}. \tag{B.2}
\]

Moreover (1.7), (B.1) and the dominated convergence theorem imply

\[
\lim_{\varepsilon \to 0} E(u^\varepsilon(t) - u^\varepsilon(s))^2 = 2 \int_{r_1=s}^t \int_{r_2=s}^{r_1} \frac{dr_2 \, dr_1}{(r_1 - r_2)^\gamma} = \frac{(t - s)^{2H}}{(2H - 1)^H}.
\]

Since \( u^\varepsilon \) is Gaussian this implies that the finite-dimensional distributions of \( u^\varepsilon \) converge to that of \( \sigma_H B^H \). For convergence in law, note (B.2) implies

\[
E[(u^\varepsilon(s) - u^\varepsilon(r))^2(u^\varepsilon(t) - u^\varepsilon(s))^2] \leq E[(u^\varepsilon(s) - u^\varepsilon(r))^4]^{1/2} E[(u^\varepsilon(t) - u^\varepsilon(s))^4]^{1/2} \leq 3E[(u^\varepsilon(s) - u^\varepsilon(r))^2]E[(u^\varepsilon(t) - u^\varepsilon(s))^2] \leq C|t - r|^{4H}.
\]

Since \( 4H > 1 \), Theorems 13.4 and 13.5 in [Bil99] imply that \( u^\varepsilon \) converges to \( \sigma_H B^H \) in law. \( \square \)

**Proof of Proposition 1.2.** when \( \gamma \in (1, 2) \) Using (1.5), (1.11) and a change of variable we note

\[
E(u^\varepsilon(t) - u^\varepsilon(s))^2 = \frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^t \int_{r_2=s}^{r_1} R \left( \frac{r_1 - r_2}{\varepsilon^2} \right) \, dr_2 \, dr_1
\]

\[
= \frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^t \int_{r_1-s}^{r_1} R \left( \frac{z}{\varepsilon^2} \right) \, dz \, dr_1
\]

\[
= -\frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^t \int_{z=r_1-s}^{\infty} R \left( \frac{z}{\varepsilon^2} \right) \, dz \, dr_1, \tag{B.3}
\]

where the last equality followed from (1.10). Since \( L \) is slowly varying, (1.7) implies

\[
\frac{1}{\sigma(\varepsilon)^2} R \left( \frac{z}{\varepsilon^2} \right) \leq R(0) \wedge \left( \frac{C}{z^{\gamma'}} \right), \tag{B.4}
\]

where \( \gamma' = (1 + \gamma)/2 \) and \( C \) is a finite constant that is independent of \( \varepsilon \). We will subsequently allow \( C \) to change from line to line, as long as it doesn’t depend on \( \varepsilon \).

Note (B.3) and (B.4) immediately imply

\[
E(u^\varepsilon(t) - u^\varepsilon(s))^2 \leq 2C \int_{r_1=s}^t \int_{z=r_1-s}^{\infty} \frac{dz \, dr_1}{z^{\gamma'}} = C(t - s)^{2-\gamma'} \tag{B.5}
\]
Moreover, by (1.7), (B.4) and the dominated convergence theorem, we see
\[
\lim_{\epsilon \to 0} E(u^\epsilon(t) - u^\epsilon(s))^2 = 2 \int_{t_1=s}^{t_1} \int_{z=r_1-s}^{\infty} \frac{dz \, dr}{z^{\gamma}} = \frac{(t-s)^2H}{(1-2H)H}.
\]
Now the remainder of the proof is identical to the case when $\gamma \in (0, 1)$.

\textbf{Proof of Proposition 1.2. when $\gamma = 1$} In this case we write
\[
E(u^\epsilon(t) - u^\epsilon(s))^2 = \frac{2}{\sigma(\epsilon)^2} \int_{r=s}^{t} \int_{z=0}^{r-s} R\left(\frac{z}{\epsilon^2}\right) \, dz \, dr
\]
The first term on the right vanishes as $\epsilon \to 0$. Using the uniform convergence theorem [BGT89, Th 1.2.1] on the second term we see
\[
\lim_{\epsilon \to 0} E(u^\epsilon(t) - u^\epsilon(s))^2 = \lim_{\epsilon \to 0} \frac{2}{L(e^{-\gamma} \ln \epsilon)} \int_{r=s}^{t} \int_{z=0}^{r-s} R\left(\frac{z}{\epsilon^2}\right) \, dz \, dr = t-s,
\]
and the convergence is uniform when $s, t$ belong to any bounded interval. Hence by the same argument as in the previous two cases we see that $(u^\epsilon)$ converges in law to $\sigma_H B_H^H$ as claimed.

\textbf{C Proof of Proposition 4.1}

The estimates for $a$, $\partial_\theta a$, and $\partial_I a$ we prove in this section are based on the following estimates on the inverse map $\varphi^{-1}$.

\textbf{Lemma C.1.} There exist $r > 0$ small enough, and a constant $C_r > 0$ such that for any $I \in (0, r)$
\[
\sup_{\theta \in \mathbb{T}} \left(\|\varphi^{-1}(I, \theta)\| + \|\partial_\theta \varphi^{-1}(I, \theta)\|\right) \leq C_r \sqrt{t},
\]
and
\[
\sup_{\theta \in \mathbb{T}} \|\partial_I \varphi^{-1}(I, \theta)\| \leq \frac{C_r}{\sqrt{t}}.
\]

The proof of this lemma is postponed to the end of this appendix.
In what follows we denote by
\[
B_r = \varphi^{-1}((0, r) \times \mathbb{T})
\]
with $r > 0$ small enough so that $B_r$ describes a small neighborhood of $(0, 0) \in \mathbb{R}^2$. 

Let us remind the reader that
\[ a(I, \theta) = e_2 \cdot \nabla \varphi_1(\varphi^{-1}(I, \theta)) = \partial_y \varphi_1(\varphi^{-1}(I, \theta)), \]
so that differentiating w.r.t. \( y \) the relation \( K(I(x, y)) = H(x, y) \) we have
\[ \partial_y I(x, y) = \partial_y \varphi_1(x, y) = \partial_y H(x, y)/\omega(I). \]
Now, using that \( \partial_y H(0, 0) = 0 \), we have for any \( I \in (0, r) \)
\[ |a(I, \theta)| \leq \omega_0^{-1} |\partial_y H(\varphi^{-1}(I, \theta)) - \partial_y H(0, 0)| \]
\[ \leq \omega_0^{-1} \sup_{B_r} \|\nabla \partial_y H\| \cdot \|\varphi^{-1}(I, \theta)\| \leq C_r \sqrt{I}. \]

**C.2 Bound for \( \partial_\theta a \).** Using that
\[ a(I, \theta) = \partial_y H(\varphi^{-1}(I, \theta))/\omega(I), \quad (C.1) \]
and now differentiating w.r.t. \( \theta \), we obtain
\[ \partial_\theta a(I, \theta) = \partial_\theta \varphi^{-1}(I, \theta) \cdot \nabla \partial_y H(\varphi^{-1}(I, \theta))/\omega(I). \]
Then, by the Cauchy-Schwarz inequality
\[ |\partial_\theta a(I, \theta)| \leq \omega_0^{-1} \|\partial_\theta \varphi^{-1}(I, \theta)\| \sup_{B_r} \|\nabla \partial_y H\| \leq C_r \sqrt{I}. \]

**C.3 Bound for \( \partial_I a \).** Differentiating (C.1) w.r.t. \( I \), we obtain
\[
\partial_I a(I, \theta) = -\frac{\omega'(I)}{\omega^2(I)} \partial_y H(\varphi^{-1}(I, \theta)) + \frac{1}{\omega(I)} \partial_I \varphi^{-1}(I, \theta) \cdot \nabla \partial_y H(\varphi^{-1}(I, \theta)) \\
\overset{\text{def}}{=} T_1 + T_2.
\]
For \( T_1 \), using (4.12) and proceeding as for the function \( a \) to deal with the term \( \partial_y H \) we have
\[ |T_1| \leq \frac{C_r}{\sqrt{I}}. \]
For \( T_2 \), applying the Cauchy-Schwarz inequality together with Lemma C.1 yields
\[ |T_2| \leq \frac{C'}{\sqrt{I}}, \]
which concludes the proof of Proposition 4.1.
C.4 Proof of lemma C.1. Before going into the proof of the lemma, let us remind the reader that $H$ admits a unique minimum at $(0, 0)$ so that $\nabla^2 H(0, 0)$ is a positive definite matrix

$$\nabla^2 H(0, 0) = \begin{pmatrix} \lambda_1 & \lambda_{12} \\ \lambda_{12} & \lambda_2 \end{pmatrix}$$

with $\lambda_1, \lambda_2 > 0$. In this case, the inner product corresponding to this matrix defines a norm on $\mathbb{R}^2$ which is equivalent to the Euclidean norm $| \cdot |$. In other words, there exist two constants $\lambda, \tilde{\lambda} > 0$ such that for any $X = (x, y)^T \in \mathbb{R}^2$

$$\lambda |X|^2 \leq X^T \nabla^2 H(0, 0) X \leq \tilde{\lambda} |X|^2.$$  \hfill (C.2)

C.4.1 Bound for $\varphi^{-1}$ Since $(0, 0)$ is the unique minimum to $H$, a Taylor expansion at the second order gives us

$$K(I) = H(X) = \frac{1}{2} X^T \nabla^2 H(0, 0) X + o(|X|^2),$$  \hfill (C.3)

with $X = (x, y)^T$. For any $X \in B_r$ we can have using (C.2) that

$$K(I) \geq \tilde{\lambda} |X|^2 \geq \tilde{\lambda} y^2$$  \hfill (C.4)

with $\tilde{\lambda} = \lambda/2 - \mu > 0$ for some small $\mu$, yielding directly

$$|\varphi_2^{-1}(I, \theta)| = |y| \leq \sqrt{K(I)/\tilde{\lambda}} = C \sqrt{K(I) - K(0)} \leq C \sup_{J \in (0, r)} \sqrt{\omega(J) \sqrt{I}} = C' \sqrt{I}.$$

In the same way, we also have

$$|\varphi_1^{-1}(I, \theta)| \leq C' \sqrt{I}.$$

C.4.2 Bound for $\partial_\theta \varphi^{-1}$ Differentiating w.r.t. $x$ and then $y$ the relation $K(I(X)) = H(X)$ we have

$$\partial_x \varphi_1(X) = \partial_x H(X)/\omega(I) \quad \text{and} \quad \partial_y \varphi_1(X) = \partial_y H(X)/\omega(I),$$

so that using (4.10) we obtain

$$\partial_\theta \varphi_2^{-1}(I, \theta) = \frac{\partial_x H(\varphi^{-1}(I, \theta))}{\omega(I)} \quad \text{and} \quad \partial_\theta \varphi_1^{-1}(I, \theta) = -\frac{\partial_y H(\varphi^{-1}(I, \theta))}{\omega(I)}.$$  \hfill (C.5)

Therefore, we have

$$\|\partial_\theta \varphi^{-1}(I, \theta)\| \leq \|\nabla H(\varphi^{-1}(I, \theta))\|/\omega_0,$$

with

$$\|\nabla H(\varphi^{-1}(I, \theta))\| = \|\nabla H(\varphi^{-1}(I, \theta)) - \nabla H(0, 0)\| \leq \sup_{B_r} \|\nabla^2 H\| \cdot \|\varphi^{-1}(I, \theta)\| \leq C_r \sqrt{I}. \hfill (C.6)$$
C.4.3 Bound for $\partial_I \varphi^{-1}$ Differentiating (C.5) w.r.t. $I$ we obtain

$$
\partial^2_I \varphi^{-1} I, \theta = -\frac{\omega'(I)}{\omega^2(I)} \partial_x H(\varphi^{-1}(I, \theta)) + \frac{1}{\omega(I)} \partial_I \varphi^{-1}(I, \theta) \cdot \nabla \partial_x H(\varphi^{-1}(I, \theta))
$$

and

$$
\partial^2_I \varphi^{-1} I, \theta = \frac{\omega'(I)}{\omega^2(I)} \partial_y H(\varphi^{-1}(I, \theta)) - \frac{1}{\omega(I)} \partial_I \varphi^{-1}(I, \theta) \cdot \nabla \partial_y H(\varphi^{-1}(I, \theta)),
$$

so that using (4.4) and (C.6)

$$
|\partial_I \varphi^{-1} I, \theta| \leq \frac{C_1}{\sqrt{I}} + \frac{1}{\omega_0} \sup_{B_r} \|\nabla^2 H\| \int_0^\theta \|\partial_I \varphi^{-1} I, \theta'\| d\theta',
$$

and in the same way for $\varphi_2^{-1}$

$$
|\partial_I \varphi_2^{-1} I, \theta| \leq \frac{C_2}{\sqrt{I}} + |\partial_I \varphi_2^{-1} I, 0| + \frac{1}{\omega_0} \sup_{B_r} \|\nabla^2 H\| \int_0^\theta \|\partial_I \varphi^{-1} I, \theta'\| d\theta'.
$$

Letting

$$
U(I, \theta) \overset{\text{def}}{=} |\partial_I \varphi^{-1} I, \theta| + |\partial_I \varphi_2^{-1} I, \theta|,
$$

and gathering the last two inequalities, we have

$$
U(I, \theta) \leq \frac{C}{\sqrt{I}} + |\partial_I \varphi_2^{-1} I, 0| + C' \int_0^\theta U(I, \theta') d\theta'.
$$

Then, applying the Gronwall lemma, we obtain

$$
\sup_{\theta \in (0, 1)} U(I, \theta) \leq \left( \frac{C}{\sqrt{I}} + |\partial_I \varphi_2^{-1} I, 0| \right) e^{C'}.
$$

Moreover, differentiating w.r.t. $I$ the relation

$$
H(\varphi^{-1}(I, \theta)) = K(I)
$$

we have

$$
\partial_I \varphi^{-1} I, \theta \partial_x H(\varphi^{-1}(I, \theta)) + \partial_I \varphi_2^{-1} I, \theta \partial_y H(\varphi^{-1}(I, \theta)) = \omega(I)
$$

and setting $\theta = 0$ we have from (4.3) and (4.4)

$$
\partial_I \varphi_2^{-1} I, 0 \partial_y H(0, \varphi_2^{-1} I, 0) = \omega(I).
$$

Also, using that

$$
\partial_y H(0, y) = \partial_y H(0, 0) + \partial^2_{yy} H(0, 0) y + o(y) = \lambda_2 y + o(y),
$$

so that for any $(0, y) \in B_r$ we have, for some small $\mu$,

$$
|\partial_y H(0, y)| \geq (\lambda_2 - \mu)|y|.
$$
As a result,
\[ |\partial_I \varphi_2^{-1}(I, 0)| \leq \frac{\omega(I)}{(\lambda_2 - \mu)|\varphi_2^{-1}(I, 0)|}, \]
and from (4.3) together with (C.3), we have
\[ K(I) = H(0, \varphi_2^{-1}(I, 0)) \leq \tilde{\lambda}_2 (\varphi_2^{-1}(I, 0))^2, \]
with \( \tilde{\lambda}_2 = \lambda_2/2 + \mu \) so that
\[ |\varphi_2^{-1}(I, 0)| \geq \sqrt{K(I)/\tilde{\lambda}_2} \geq \sqrt{c_1 I/\tilde{\lambda}_2}. \]
Finally, we obtain
\[ \sup_{\theta \in \mathbb{T}} |\partial_I \varphi_2^{-1}(I, \theta)| + |\partial_I \varphi_2^{-1}(I, \theta)| \leq \frac{C_r}{\sqrt{I}}, \]
which gives the desired result and concludes the proof of the lemma.

References

[Arn89] Arnold, V.I.: Mathematical Methods of Classical Mechanics. Springer, New York (1989)

[AT07] Adler, R.J., Taylor, J.E.: Random Fields and Geometry. Springer Monographs in Mathematics, Springer, New York (2007)

[Bak11] Bakhtin, Y.: Noisy heteroclinic networks. Probab. Theory Relat. Fields 150(1–2), 1–42 (2011). https://doi.org/10.1007/s00440-010-0264-0

[BGT89] Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge University Press, Cambridge (1989)

[BIKS18] Buldygin, V.V., Indlekofer, K.-H., Klesov, O.I., Steinebach, J.G.: Pseudo-Regularly Varying Functions and Generalized Renewal Processes. Probability Theory and Stochastic Modelling, vol. 91. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-99537-3

[Bill99] Billingsley, P.: Convergence of Probability Measures. Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc, New York (1999). https://doi.org/10.1002/9780470316962

[Brezis99] Brezis, H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer, New York (2011)

[BS67] Birman, M.Š, Solomjak, M.Z.: Piecewise-polynomial approximations of functions of the classes \( w^p_{\alpha} \). Math. USSR-Sbornik 2(3), 295–317 (1967). https://doi.org/10.1070/sm1967v002n03abeh002343

[CV21] Caillerie, N., Vovelle, J.: Diffusion-approximation for a kinetic equation with perturbed velocity redistribution process. Ann. Appl. Probab. 31(3), 1299–1335 (2021). https://doi.org/10.1214/20-aap1619

[DPZ96] Da Prato, G., Zabczyk, J.: Ergodicity for Infinite-Dimensional Systems. London Mathematical Society Lecture Note Series, vol. 229. Cambridge University Press, Cambridge (1996). https://doi.org/10.1017/CBO9780511662829

[DRV21] Debussche, A., Rosello, A., Vovelle, J.: Diffusion-approximation for a kinetic spray-like system with random forcing. Discrete Contin. Dyn. Syst.- S 14(8), 2751–2803 (2021). https://doi.org/10.3934/dcdss.2021039

[EK86] Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc, New York (1986). https://doi.org/10.1002/9780470316658

[FGPS07] Fouque, J.-P., Garnier, J., Papanicolaou, G., Solna, G.: Wave Propagation and Time Reversal in Randomly Layered Media. Stochastic Modelling and Applied Probability, vol. 56. Springer, New York (2007)

[Fri06] Friedman, A.: Stochastic Differential Equations and Applications. Dover Publications Inc, Mineola, NY (2006). Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes
[FW93] Freidlin, M.I., Wentzell, A.D.: Diffusion processes on graphs and the averaging principle. Ann. Probab. 21(4), 2215–2245 (1993). https://doi.org/10.1214/aop/1176989018

[FW94] Freidlin, M.I., Wentzell, A.D.: Random perturbations of Hamiltonian systems. Mem. Amer. Math. Soc. 109(523), viii+82 (1994). https://doi.org/10.1090/memo/0523

[FW98] Freidlin, M., Weber, M.: Random perturbations of nonlinear oscillators. Ann. Probab. 26(3), 925–967 (1998). https://doi.org/10.1002/1091-1886(199809)

[Gar97] Garnier, J.: Multi-scaled diffusion-approximation. Applications to wave propagation in random media. ESAIM Probab. Statist. 1, 183–206 (1995/97). https://doi.org/10.1051/ps:1997107

[GILN21] Gomez, C., Iyer, G., Le, H., Novikov, A.: An oscillator driven by algebraically decorrelating noise. arXiv preprint (2021) https://doi.org/10.48550/arXiv.2107.07363

[HIK+18] Hairer, M., Iyer, G., Koralov, L., Novikov, A., Pajor-Gyulai, Z.: A fractional kinetic process describing the intermediate time behaviour of cellular flows. Ann. Probab. 46(2), 897–955 (2018). https://doi.org/10.1214/17-AOP1196

[HKPG16] Hairer, M., Koralov, L., Pajor-Gyulai, Z.: From averaging to homogenization in cellular flows—an exact description of the transition. Ann. Inst. Henri Poincaré Probab. Stat. 52(4), 1592–1613 (2016). https://doi.org/10.1214/15-AIHP690

[HL20] Hairer, M., Li, X.-M.: Averaging dynamics driven by fractional Brownian motion. Ann. Probab. 48(4), 1826–1860 (2020). https://doi.org/10.1214/19-AOP1408

[HV14a] Haynes, P.H., Vanneste, J.: Dispersion in the large-deviation regime. Part 1: shear flows and periodic flows. J. Fluid Mech. 745, 321–350 (2014). https://doi.org/10.1017/jfm.2014.64

[HV14b] Haynes, P.H., Vanneste, J.: Dispersion in the large-deviation regime. Part 2. Cellular flow at large Péclet number. J. Fluid Mech. 745, 351–377 (2014). https://doi.org/10.1017/jfm.2014.65

[KNR14] Komorowski, T., Novikov, A., Ryzhik, L.: Homogenization driven by a fractional Brownian motion: the shear layer case. Multiscale Model. Simul. 12(2), 440–457 (2014). https://doi.org/10.1137/13092068X

[KS91] Karatzas, I., Shreve, S.E.: Brownian motion and stochastic calculus. Graduate Texts in Mathematics, vol. 113, 2nd edn. Springer-Verlag, New York (1991). https://doi.org/10.1007/978-1-4612-0949-2

[Kus84] Kushner, H.J.: Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory. MIT Press Series in Signal Processing, Optimization, and Control, vol. 6. MIT Press, Cambridge (1984)

[Mar05] Marty, R.: Asymptotic behavior of differential equations driven by periodic and random processes with slowly decaying correlations. ESAIM Probab. Stat. 9, 165–184 (2005). https://doi.org/10.1051/ps:2005009

[PS08] Pavliotis, G.A., Stuart, A.M.: Multiscale Methods—Averaging and Homogenization. Texts in Applied Mathematics, vol. 53. Springer, New York (2008)

[PSV76] Papanicolaou, G.C., Stroock, D., Varadhan, S.R.S.: Martingale approach to some limit theorem. Duke Univ. Math. Series III 4, 1–120 (1976)

[Roz87] Rozanov, Y.A.: Introduction to Random Processes. Springer Series in Soviet Mathematics, Springer-Verlag, Berlin (1987). https://doi.org/10.1007/978-3-642-72717-7. Translated from the Russian by Birgit Röthinger

[SV06] Stroock, D.W., Varadhan, S.R.S.: Multidimensional diffusion processes. Classics in Mathematics. Springer-Verlag, Berlin (2006). Reprint of the 1997 edition

[Taq77] Taqqu, M.S.: Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 40(3), 203–238 (1977). https://doi.org/10.1007/BF00736047

[Taq75] Taqqu, M.S.: Weak convergence to fractional Brownian motion and to the Rosenblatt process. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 31, 287–302 (1974/75). https://doi.org/10.1007/BF00532868

[TV16] Tzella, A., Vanneste, J.: Dispersion in rectangular networks: effective diffusivity and large-deviation rate function. Phys. Rev. Lett. 117, 114501 (2016). https://doi.org/10.1103/PhysRevLett.117.114501

[YJ91] Young, W.R., Jones, S.: Shear dispersion. Phys. Fluids A: Fluid Dyn. 3(5), 1087–1101 (1991)

[You88] Young, W.R.: Arrested shear dispersion and other models of anomalous diffusion. J. Fluid Mech. 193, 129–149 (1988). https://doi.org/10.1017/S0022112088002083

Communicated by M. Hairer