Euro. Jnl of Applied Mathematics (2020), *vol.* 31, *pp.* 709–736
doi:10.1017/S0956792519000251

Mass transport in Fokker–Planck equations with tilted periodic potential

MICHAEL HERRMANN[1] and BARBARA NIETHAMMER[2]

[1]*Institute of Computational Mathematics, Technische Universität Braunschweig*
Universitätsplatz 2, 38106 Braunschweig, Germany
email: michael.herrmann@tu-braunschweig.de
[2]*Institute of Applied Mathematics, University of Bonn*
Endenicher Allee 60, 53115 Bonn, Germany
email: niethammer@iam.uni-bonn.de

(*Received 22 January 2018*; *revised 10 June 2019*; *accepted 23 August 2019*; *first published online 30 September 2019*)

We consider Fokker–Planck equations with tilted periodic potential in the subcritical regime and characterise the spatio-temporal dynamics of the partial masses in the limit of vanishing diffusion. Our convergence proof relies on suitably defined substitute masses and bounds the approximation error using the energy-dissipation relation of the underlying Wasserstein gradient structure. In the appendix, we also discuss the case of an asymmetric double-well potential and derive the corresponding limit dynamics in an elementary way.

Key words: Fokker–Planck equations with tilted period potential, model reduction for multi-scale dynamical systems, asymptotic analysis of singular limits

2010 Mathematics Subject Classification: 35Q84 (Primary); 35B25, 35B40 (Secondary)

1 Introduction

We study the Fokker–Planck equation

$$\tau\big(\partial_t\varrho(t,\,x,\,p)-\Delta_x\varrho(t,\,x,\,p)\big)=\nu^2\partial_p^2\varrho(t,\,x,\,p)+\partial_p\big(\big(H'(p)-\sigma\big)\varrho(t,\,x,\,p)\big),\tag{1.1}$$

with small parameters τ and ν. Here, t and $x\in\mathbb{R}^n$ denote the time and space variable, respectively, $p\in\mathbb{R}$ stands for an internal but scalar state variable and the unknown ϱ is supposed to be non-negative and normalised by

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}}\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x=1\,.\tag{1.2}$$

In the spatially homogeneous situation – that is, without any x-dependence – the resulting partial differential equation (PDE) for ϱ as function of t and p is also called Kramers–Smoluchowski equation and can be viewed as the deterministic equation for the probability distribution of a stochastic particle systems. The latter is governed by the overdamped Langevin or Smoluchowski equation

$$\tau\,\mathrm{d}p=\big(\sigma-H'(p)\big)\,\mathrm{d}t+\sqrt{2\nu^2}\,\mathrm{d}W\,,\tag{1.3}$$

where W represents a standard Wiener Process related to Brownian motion in p-space.

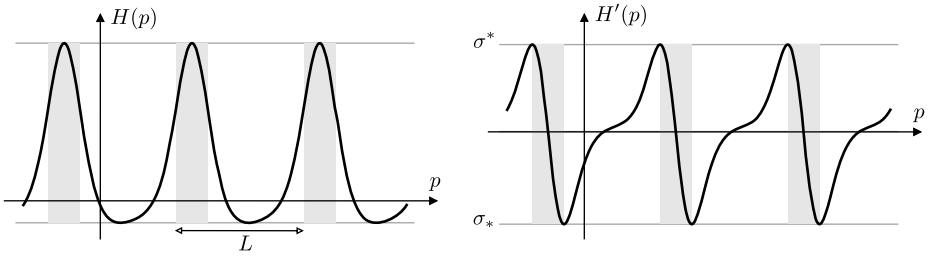

FIGURE 1. *Left panel*. Example of an L-periodic potential H as in Assumption 1.1. The grey boxes indicate the spinodal regions in which H is concave. *Right panel*. The local extrema of H' are denoted by σ_* and σ^*. In this paper, we always assume that the tilting parameter σ is restricted by (1.5) so that the effective potential H_{eff} admits equidistant wells as illustrated in the left panel of Figure 2.

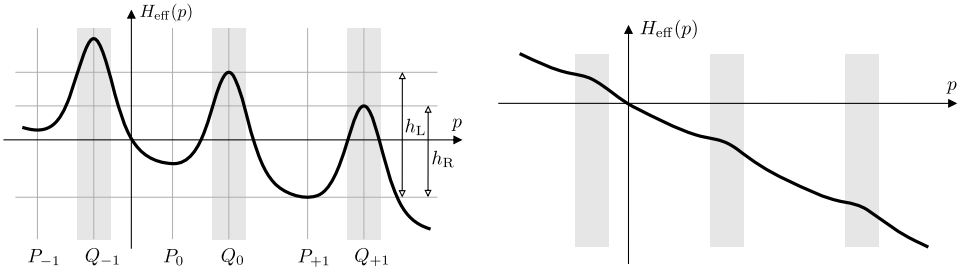

FIGURE 2. *Left panel*. In the subcritical regime $\sigma_* < \sigma < \sigma^*$ studied in this paper, the effective potential H_{eff} from (1.4) admits multiple wells (here depicted for $0 < \sigma < \sigma^*$) with local minima and maxima located at the positions P_j and Q_j, respectively, where $j \in \mathbb{Z}$ and $P_j < Q_j < P_{j+1}$. The grey boxes indicate the spinodal regions. *Right panel*. Our results do not cover the supercritical regime $\sigma > \sigma^*$ since the effective potential has no wells anymore. The mass transfer is therefore very fast; see Appendix A.

In what follows we always suppose that the potential H is a smooth and periodic function in p; see Figure 1 for an illustration. The particles, however, move in the effective potential

$$H_{\mathrm{eff}}(p) = H(p) - \sigma p. \tag{1.4}$$

due to the presence of the tilting parameter $\sigma \in \mathbb{R}$, which is assumed to be independent of ν. As depicted in Figure 2, the properties of H_{eff} strongly depend on the choice of σ, where the critical values σ_* and σ^* denote the global minimum and maximum of H', respectively. In the supercritical regime, we have either $\sigma < \sigma_*$ or $\sigma > \sigma^*$, so H_{eff} is either strictly increasing or decreasing. In the subcritical regime $\sigma_* < \sigma < \sigma^*$, however, the effective potential possesses several wells which represent metastable traps for the stochastic particle dynamics (1.3).

Overdamped Langevin equations with tilted periodic potential and the corresponding Fokker–Planck equations appear in many branches of mathematics and the sciences. There exists a quite exhaustive literature; see, for instance, [27, 20, 26, 5, 9, 19] and references therein, and applications range from plasma and solid state physics, materials science and electrical engineering to molecular motors, surface diffusion, Brownian ratches and Josephson junctions.

Another interpretation of the Fokker–Planck operator is reminiscent of Kramers' seminal paper [18] on kinetic models for chemical reactions. More precisely, we can regard each local minimum of H_{eff} as a local equilibrium of a reacting systems, and the random walk in p-direction

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

on the particle level — or equivalently, the term $\nu^2\partial_p^2$ – describes reactions; this means transitions between adjacent wells of the effective potential. Notice, however, that the PDE (1.1) also involves Δ_x. This term describes an additional diffusion with respect to the physical space variable x, where the latter has been scaled so that the corresponding diffusion constant equals the small relaxation time in front of the time derivative ∂_t. Although such combined Fokker–Planck equations for reaction-diffusion systems are less common in the literature, they appear naturally in the description of spatially extended dynamical systems and the extra term does not complicate the mathematical analysis very much. Moreover, we regard (1.1) merely as a toy model to study some aspects of multi-scale analysis and asymptotic model reduction for driven particle systems.

In the present paper, we concentrate on the subcritical regime and study the singular limit $\nu \to 0$ on the level of the Fokker–Planck equation. In particular, we derive a dynamical limit model which is still infinite-dimensional but simpler and more regular than (1.1) as it does not involve any small parameter.

Before we describe our findings and methods in more detail, we emphasise that both the supercritical and the subcritical regimes of (1.1) have been studied intensively in the physics community. There, the main focus is on the long-time behaviour of the effective velocity and the effective diffusion tensor. These quantities are completely determined by the first and the second p-moment of ϱ and their averaged growth in time can be computed in many situations; see our discussion at the end of Section 2.4, [20, 26, 29] for an overview (including more general models), and [28, 12, 19, 7] for related rigorous results. Our contribution consists in the derivation of a refined model for the limit dynamics that accounts for the mass inside of each well and in the presentation of a particular proof strategy.

1.1 Effective mass transport in the subcritical regime

Throughout this paper, we suppose that the potential H has the following properties.

Assumption 1.1 (periodic part of the energy landscape) *The potential H is L-periodic and sufficiently smooth such that*

$$\sigma_* := \min_{p\in\mathbb{R}} H'(p)\,, \qquad \sigma^* := \max_{p\in\mathbb{R}} H'(p)\,, \qquad \zeta := \sup_{p\in\mathbb{R}} \left|H'''(p)\right|.$$

are well-defined real numbers. Moreover, H' is unimodal and non-degenerate in the sense that each critical point is a global extremum – that is, $H''(p) = 0$ implies $H'(p) \in \{\sigma_*, \sigma^*\}$.

A prototypical example of Assumption 1.1 is

$$H(x) = G\big(\sin(x)\big)\,,$$

where $G : \mathbb{R} \to \mathbb{R}$ is a smooth and strictly increasing function, and a more asymmetric example is depicted in Figure 1. As mentioned above, we restrict our considerations to the subcritical regime. This means we fix σ independent of ν with

$$\sigma_* < \sigma < \sigma^*. \tag{1.5}$$

so that the effective potential from (1.4) is tilted to the left for $\sigma_* < \sigma < 0$ but tilted to the right for $0 < \sigma < \sigma^*$. The constraint (1.5) guarantees that H_{eff} admits an infinite number of local minima

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

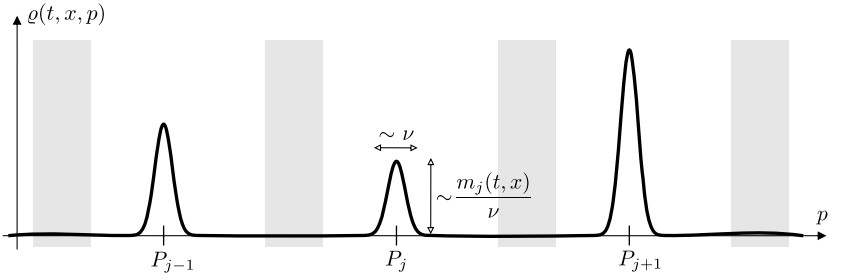

FIGURE 3. Cartoon of the Fokker–Planck solution for small $0 < \nu \ll 1$. The function $p \mapsto \varrho(t, x, p)$ is basically the superposition of infinitely many narrow peaks, where the jth peak is localised at $p = P_j$ and carries mass $m_j(t, x)$. These peaks do not move but exchange mass according to the limit dynamics (1.11) or (1.12).

and maxima, whose positions are denoted by P_j and Q_j, respectively. These positions depend on σ but the periodicity of H guarantees that $P_j = P_0 + jL$ and $Q_j = Q_0 + jL$ for all $j \in \mathbb{Z}$; see Figure 2 for an illustration.

For any $j \in \mathbb{Z}$, we define the partial mass

$$m_j(t, x) := \int_{Q_{j-1}}^{Q_j} \varrho(t, x, p)\, \mathrm{d}p\,, \tag{1.6}$$

which quantifies at any (t, x) the amount of mass that is contained in the well around the local minimum P_j. The PDE (1.1) implies that the pointwise total mass

$$m(t, x) := \sum_{j \in \mathbb{Z}} m_j(t, x), \tag{1.7}$$

diffuses in x-space according to

$$\partial_t m(t, x) - \Delta_x m(t, x) = 0\,,$$

but it remains to understand the spatio-temporal dynamics of m_j. This problem is well-understood on the heuristic level and the key arguments for small ν can be summarised as follows. Due to the deterministic part in the Brownian motion, it is very likely to find particles near one of the local minima. In other words, $\varrho(t, x, \cdot)$ consists of infinitely many localised peaks and we can approximate

$$\varrho(t, x, p) \approx \sum_{j \in \mathbb{Z}} m_j(t, x) \delta_{P_j}(p). \tag{1.8}$$

at least in weak* sense with Dirac distributions on the right-hand side; see Figure 3 for a schematic representation. The small diffusion in p-direction, however, guarantees that each peak has width of order $O(\nu)$ and that particles can cross the energy barriers at the local maxima of H_{eff} due to random fluctuations. For fixed x, this gives rise to a hopping process between the different wells whose characteristic timescales can be computed asymptotically by Kramers celebrated

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

formula from [18]. More precisely, in the limit $\nu \to 0$ the expected time for a jump to the next well on the left and on the right is given by

$$\tau c_{\mathrm{K}}^{-1} \exp\left(\frac{h_{\mathrm{L}}}{\nu^2}\right) \qquad \text{and} \qquad \tau c_{\mathrm{K}}^{-1} \exp\left(\frac{h_{\mathrm{R}}}{\nu^2}\right),$$

respectively, and the periodicity of H implies that the energy barriers

$$h_{\mathrm{L}} := H_{\mathrm{eff}}\big(Q_{j-1}\big) - H_{\mathrm{eff}}\big(P_j\big)\,, \qquad h_{\mathrm{R}} := H_{\mathrm{eff}}\big(Q_j\big) - H_{\mathrm{eff}}\big(P_j\big),$$

are actually independent of j; see again Figure 2. Moreover, the Kramers constant

$$c_{\mathrm{K}} := \frac{\sqrt{\big|H''\big(P_j\big)H''\big(Q_j\big)\big|}}{2\pi}. \tag{1.9}$$

is also independent of j and is the same for jumps to the left and to the right. This motivates the following choice of the timescale.

Assumption 1.2 (choice of τ) *For fixed σ as in (1.5), we set*

$$\tau := c_{\mathrm{K}} \exp\left(-\frac{\min\{h_{\mathrm{L}},\, h_{\mathrm{R}}\}}{\nu^2}\right). \tag{1.10}$$

where $\nu > 0$ is the small but free parameter.

Due to the informal discussion about the characteristic Kramers timescales for the aforementioned hopping process, we can formulate the expected limit dynamics depending on whether the value of the tilting parameter σ favours transport to the left or transport to the right.

Result (effective mass transport in the subcritical regime) *In the limit $\nu \to 0$, the partial masses evolve according to*

$$\partial_t m_j(t,\, x) - \Delta_x m_j(t,\, x) = \begin{cases} m_{j+1}(t,\, x) - m_j(t,\, x) & \text{for } \sigma_* < \sigma < 0, \\ m_{j-1}(t,\, x) - m_j(t,\, x) & \text{for } 0 < \sigma < \sigma^*, \end{cases} \tag{1.11}$$

and

$$\partial_t m_j(t,\, x) - \Delta_x m_j(t,\, x) = m_{j-1}(t,\, x) + \kappa m_{j+1}(t,\, x) - (1+\kappa) m_j(t,\, x) \quad \text{for} \quad \sigma = 0\,. \tag{1.12}$$

where the constant κ depends only on the properties of H and can be computed explicitly.

Our goal in this paper is to justify the limit model for the partial masses rigorously in a purely analytical framework with no appeal to probabilistic techniques. It should also be possible to justify the lattice equations (1.11) and (1.12) using standard methods from stochastic analysis (such as Large Deviation Principles) but we are not aware of any reference.

We further mention that the fundamental solution to the linear limit model can be computed explicitly. For instance, assuming $0 < \sigma < \sigma_*$ and that the entire initial mass is concentrated at $j = j_0$ and $x = x_0$, we readily verify that the corresponding solution to (1.12) is given by

$$m_j(t,\, x) = K_{\mathrm{heat}}(t,\, x - x_0) \cdot K_{\mathrm{pois}}(t,\, j - j_0)\,, \tag{1.13}$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

where

$$K_{\mathrm{heat}}(t,\,x) = (4\pi t)^{-n/2}\exp\left(-\frac{x^2}{4t}\right) \quad \text{and} \quad K_{\mathrm{pois}}(t,\,j) = \begin{cases} 0 & \text{for } j < 0 \\ \dfrac{t^j \exp(-t)}{j!} & \text{for } j \geq 0 \end{cases}.$$

represent the heat kernel and the Poisson point process, respectively.

1.2 Wasserstein gradient structure and proof strategy

The PDE (1.1) can be regarded as a Wasserstein gradient flow on the space of probability measures since it can be written as

$$\tau\,\partial_t \varrho = \left(\tau^{1/2}\nu^{-1}\partial_x + \partial_p\right)\Big(\varrho\,\left(\tau^{1/2}\nu^{-1}\partial_x + \partial_p\right)\partial_\varrho \mathcal{E}\Big)\,,$$

where $\mathcal{E}$ abbreviates the free energy of the system and ∂_ϱ denotes the functional derivative. In particular, with

$$\mathcal{E}(t) := \int_{\mathbb{R}^n}\int_{\mathbb{R}} \nu^2 \varrho(t,\,x,\,p)\ln \varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x + \int_{\mathbb{R}^n}\int_{\mathbb{R}} \big(H(p) - \sigma p\big)\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x, \tag{1.14}$$

we readily verify by direct computations the energy balance

$$\tau\dot{\mathcal{E}}(t) = -\tau\,\nu^2\,\mathcal{C}(t) - \nu^4\,\mathcal{D}(t)\,, \tag{1.15}$$

where

$$\mathcal{C}(t) := \int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{\big(\nabla_x \varrho(t,\,x,\,p)\big)^2}{\varrho(t,\,x,\,p)}\,\mathrm{d}p\,\mathrm{d}x,$$

and

$$\mathcal{D}(t) := \int_{\mathbb{R}^n}\int_{\mathbb{R}} \frac{\big(\partial_p \varrho(t,\,x,\,p) + \nu^{-2}\big(H'(p) - \sigma\big)\varrho(t,\,x,\,p)\big)^2}{\varrho(t,\,x,\,p)}\,\mathrm{d}p\,\mathrm{d}x. \tag{1.16}$$

yield the total dissipations due to the Brownian motion of particles in the x- and the p-direction, respectively.

The variational interpretation of Fokker–Planck equations like (1.1) has been first described in [16] and attracted a lot of attention during the last decades, especially for Fokker–Planck equations that admit a unique equilibrium corresponding to a global minimiser of the energy. This is, however, not true for tilted periodic potentials because the system can constantly lower its total energy by transporting mass towards $p = -\infty$ (for $\sigma < 0$) or $p = +\infty$ (for $\sigma > 0$), and thus there exists neither a lower bound for the energy nor a steady state for the gradient flow. The energy-dissipation relation (1.15) is nevertheless very useful as it provides a temporal L^1-bounds for the total dissipation on each finite time interval.

The gradient flow perspective has also been used to study the diffusive mass transfers in Fokker–Planck equations with double-well potential, for which the effective dynamics in the

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

limit $\nu \to 0$ is a scalar ordinary differential equation (ODE) that governs the mass flux though the single barrier which separates the two wells. Since our work on tilted periodic potentials has much in common with this problem, we discuss the recent literature in Appendix B and sketch how our method can be applied to the case of a double-well potential. One advantage of our approach is that it covers also asymmetric energy landscapes while most of the recent gradient flow results are restricted to even functions H. We also mention that potentials with finitely many wells having the same energy are studied in [22]. This situation shares some similarities with the untilted case $\sigma = 0$ in our paper but the analytic techniques are rather different as they rely on a careful spectral analysis of the Fokker–Planck operator; see also [5] for related work.

Our approach to the asymptotic justification of the limit dynamics consists of three main steps, which can informally be described as follows:

(1) *Effective dynamics of substitute masses:* We first identify two different approximations of the partial masses such that the time derivative of the first substitute mass can be expressed in terms of the second one. In this way, we obtain dynamical relations which resemble the lattice equations (1.11) and (1.12) up to certain error terms. The details are presented in Section 2.2 and rely on the balance equations of carefully chosen moment integrals of ϱ as well as the asymptotic auxiliary results and the local equilibrium densities from Section 2.1.

(2) *Dissipation bounds approximation error:* Another key argument is that the difference between the partial masses and their substitutes can be controlled by the Wasserstein dissipation. More precisely, we show in Section 2.3 for given t that almost all mass is in fact contained in the vicinity of the local minima $p = P_j$ provided that $\mathcal{D}(t)$ from (1.16) is sufficiently small. Similar mass-dissipation estimates have been used in [15].

(3) *Energy balance bounds dissipation:* We finally prove in Section 2.4 that (1.15) implies that $\mathcal{D}$ is small in an L^1-sense and hence, loosely speaking, also at most of the times t. This results hinges on lower bounds for $\mathcal{E}(t)$ and hence on upper bounds for the modulus of

$$\mathcal{P}(t) := \int_{\mathbb{R}^n} \int_{\mathbb{R}} p\varrho(t, x, p)\, \mathrm{d}p\, \mathrm{d}x\,. \tag{1.17}$$

but the latter can de deduced from the moment integrals for the substitute masses.

All partial results are combined in the proof of Theorem 2.9 and imply a rather elementary justification of the lattice model for the partial masses. Moreover, the authors believe that most of the key arguments can also be applied to other types of Fokker–Planck equations; see the appendices for first examples. Another, more challenging equation is the non-local variant of (1.1), in which σ is not given *a priori* but enters as the time-dependent Lagrangian multiplier of a dynamical constraint; see [14, 15] for a related problem.

We finally mention that systems of weakly coupled Fokker–Planck equations with periodic coefficients arise in the theory of motor proteins. Effective dynamical models for the limit of vanishing p-diffusion have been established rigorously in [25, 23] but they concern another scaling limit and are derived by different asymptotic techniques, namely homogenisation via viscosity solutions and non-linear Hamilton-Jacobi equations. Moreover, the asymptotic drift coefficient for a similar class of models has been derived in [8] and homogenisation for stationary Fokker–Planck equations can be found in [5].

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

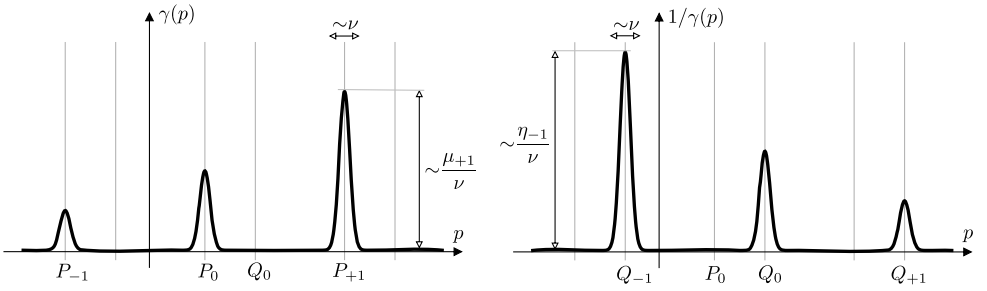

FIGURE 4. Schematic representation of the Gibbs function γ from (2.2) (*left panel*) and its reciprocal (*right panel*) for $0 < \nu \ll 1$. Each function can be approximated by a infinite superposition of equidistant peaks with width of order ν, where the mass inside each peak depends exponentially on its position; see Lemma 2.1.

2 Asymptotic analysis

To prove our main result from Section 1, we assume from now on that

$$0 \leq \sigma < \sigma^*. \tag{2.1}$$

but emphasise that the case $\sigma_* < \sigma \leq 0$ can be proven along the same lines. We also denote by C any generic constant that is independent of ν but can depend on the potential H and the choice of σ.

2.1 Preliminaries

A key quantity for our asymptotic analysis is the Gibbs function

$$\gamma(p) := \exp\left(\frac{-H(p) + \sigma p}{\nu^2}\right), \tag{2.2}$$

which is illustrated in Figure 4. Notice that γ is not integrable and this reflects the lack of non-trivial steady states. This is different to other variants of the Fokker–Planck equation – as, for instance, the case of a proper double-well potential discussed in Appendix B – in which the normalisation of γ defines the unique and globally attracting equilibrium.

Our first auxiliary result characterises the behaviour of γ and $1/\gamma$ in the intervals

$$J_j := (Q_{j-1},\, Q_j) \qquad \text{and} \qquad K_j := (P_j,\, P_{j+1}), \tag{2.3}$$

respectively, and provides a rigorous link to the exponential scaling parameter τ from the Kramers law (1.10). The derivation of the latter exploits the well-known Laplace method from the theory of asymptotic integrals; see, for instance, [3, Section 6.4, esp. equations (6.4.1) and (6.4.35)].

Lemma 2.1 (asymptotic integrals) *The scalars*

$$\mu_j := \int_{J_j} \gamma(p)\, \mathrm{d}p\,, \qquad \eta_j := \int_{K_j} \frac{1}{\gamma(p)}\, \mathrm{d}p \tag{2.4}$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

satisfy

$$\mu_j = \mu_0 \kappa^{-j}, \qquad \eta_j = \eta_0 \kappa^{+j}, \qquad \kappa := \exp\left(-\frac{\sigma L}{\nu^2}\right). \tag{2.5}$$

Moreover, we have

$$|\theta| \le C\nu^2, \qquad \theta := \frac{\tau \mu_0 \eta_0}{\nu^2} - 1. \tag{2.6}$$

for some constant C which depends on σ but not on ν.

Proof The identities (2.5) follow – thanks to the L-periodicity of H – immediately from the definition in (2.2) and (2.4). Moreover, by Laplace's method we verify

$$\mu_0 = \frac{\nu\sqrt{2\pi}}{\sqrt{|H''(P_0)|}} \exp\left(\frac{-H(P_0) + \sigma P_0}{\nu^2}\right)\left(1 \pm O\left(\nu^2\right)\right), \tag{2.7}$$

as well as

$$\eta_0 = \frac{\nu\sqrt{2\pi}}{\sqrt{|H''(Q_0)|}} \exp\left(\frac{+H(Q_0) - \sigma Q_0}{\nu^2}\right)\left(1 \pm O\left(\nu^2\right)\right).$$

where (2.1) ensures $H''(Q_0) < 0 < H''(P_0)$. We thus obtain (2.6), thanks to the definition of τ in (1.10). □

Using the Gibbs function (2.2), we define local equilibrium measures

$$\gamma_j(p) = \mu_j^{-1} \chi_{J_j}(p) \gamma(p), \tag{2.8}$$

where χ_{J_j} denotes the characteristic function of the interval J_j. We also introduce a local relative density w_j^2 by

$$w_j^2(t, x, p) := \mu_j \frac{\varrho(t, x, p)}{\gamma(p)} \qquad \text{for} \quad p \in J_j, \tag{2.9}$$

where the second power on the left-hand side of (2.9) has been introduced for convenience. In terms of w_j, the partial masses from (1.6) can be written as

$$m_j(t, x) = \int_{J_j} w_j^2(t, x, p) \gamma_j(p) \, \mathrm{d}p, \tag{2.10}$$

while the dissipation due to the diffusion in p-space reads

$$\mathcal{D}(t) = 4 \int_{\mathbb{R}^n} D(t, x) \, \mathrm{d}x, \tag{2.11}$$

with

$$D(t, x) := \sum_{j \in \mathbb{Z}} D_j(t, x), \qquad D_j(t, x) := \int_{J_j} \left(\partial_p w_j(t, x, p)\right)^2 \gamma_j(p) \, \mathrm{d}p. \tag{2.12}$$

In particular, m_j and D_j are naturally related to the weighted L^2- and H^1-norms of w_j, where the weight function γ_j is a normalised and localised variant of γ.

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

2.2 Substitute masses and their dynamics

As already outlined in Section 1, our asymptotic analysis is based on suitably defined substitutes to the partial masses m_j from (1.6). The first approximation stems from the evaluation of the relative density; this means we set

$$\overline{m}_j(t,\,x) := w_j^2\big(t,\,x,\,P_j\big). \tag{2.13}$$

with w_j as in (2.9). This definition is motivated by the observation that γ_j from (2.8) is strongly localised near P_j for small ν and that w_j is basically constant for $p \approx P_j$ provided that the partial dissipation D_j from (2.12) is sufficiently small.

The second substitute mass is given by

$$\widetilde{m}_j(t,\,x) := \int_{\mathbb{R}} \big(\psi_{j-1}(p) - \psi_j(p)\big)\varrho(t,\,x,\,p)\,\mathrm{d}p\,, \tag{2.14}$$

where the weight function ψ_j is uniquely determined by

$$\psi_j'(p) := \frac{1}{\eta_j \gamma(p)} \quad \text{for} \quad p \in K_j, \tag{2.15}$$

and

$$\psi_j(p) = 0 \quad \text{for} \quad p < P_{j-1}\,, \qquad \psi_j(p) = 1 \quad \text{for} \quad p > P_j\,. \tag{2.16}$$

These definitions imply

$$\sum_{j\in\mathbb{Z}} \psi_{j-1}(p) - \psi_j(p) = 1,$$

for all $p \in \mathbb{R}$ and hence

$$\sum_{j\in\mathbb{Z}} \widetilde{m}_j(t,\,x) = \sum_{j\in\mathbb{Z}} m_j(t,\,x) = m(t,\,x). \tag{2.17}$$

for all $t \geq 0$ and any $x \in \mathbb{R}^n$.

As illustrated in Figure 5, the weight function $p \mapsto \big(\psi_{j-1}(p) - \psi_j(p)\big)$ approximates, for small $\nu > 0$, the indicator function of the interval J_j but the main point is that the transition layers near Q_{j-1} and Q_j take a particular form which enables us to compute the time derivative of $\widetilde{m}_j$ up to high accuracy.

Proposition 2.2 (balance of substitute masses) *The masses from* (2.13) *and* (2.14) *satisfy*

$$(1+\theta)\Big(\partial_t \widetilde{m}_j(t,\,x) - \Delta_x \widetilde{m}_j(t,\,x)\Big) = \overline{m}_{j-1}(t,\,x) - (1+\kappa)\,\overline{m}_j(t,\,x) + \kappa\,\overline{m}_{j+1}(t,\,x)\,. \tag{2.18}$$

where the constants κ and θ depend on ν as is Lemma 2.1.

Proof By construction – see (2.4), (2.15) and (2.16) – the function ψ_j is continuous and piecewise smooth, and satisfies on $\mathbb{R}$ the singular ODE

$$\nu^2 \psi_j''(p) - \big(H'(p) - \sigma\big)\psi_j'(p) = \alpha_{-,j}\,\delta_{P_j}(p) - \alpha_{+,j}\,\delta_{P_{j+1}}(p),$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

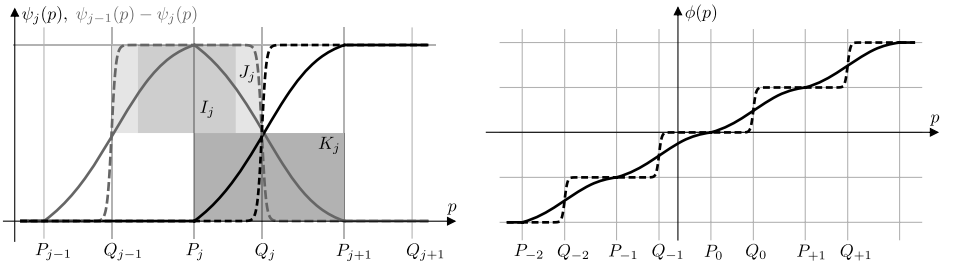

FIGURE 5. *Left panel.* Piecewise smooth moment weights as used in the definition of the substitute mass $\widetilde{m}_j(t)$ in (2.14) for a small and a moderate value of ν (dashed and solid lines, respectively). The grey boxes indicate the intervals I_j, J_j and K_j from (2.3) and Lemma 2.4. *Right panel.* Moment weight ϕ for the definition of $\mathcal{K}$; see (2.21) and (2.20), again for two values of ν. The mean slope of ϕ is $1/L$.

with Dirac weights

$$\alpha_{-,j} := \nu^2 \psi_j'\big(P_j{+}0\big) = \frac{\nu^2}{\eta_j \gamma\big(P_j\big)}\,, \qquad \alpha_{+,j} := \nu^2 \psi_j'\big(P_{j+1}{-}0\big) = \frac{\nu^2}{\eta_j \gamma\big(P_{j+1}\big)}\,.$$

Using the PDE (1.1) and integration by parts with respect to p, we thus verify

$$\begin{aligned} \tau\big(\partial_t - \Delta_x\big) \int_{\mathbb{R}} \psi_j(p)\varrho(t,\,x,\,p)\,\mathrm{d}p &= \alpha_{-,j}\varrho\big(t,\,x,\,P_j\big) - \alpha_{+,j}\varrho\big(t,\,x,\,P_{j+1}\big) \\ &= \frac{\nu^2}{\mu_j\eta_j}\overline{m}_j(t,\,x) - \frac{\nu^2}{\mu_{j+1}\eta_j}\overline{m}_{j+1}(t,\,x) \\ &= \frac{\nu^2}{\mu_0\eta_0}\big(\overline{m}_j(t,\,x) - \kappa\overline{m}_{j+1}(t,\,x)\big). \end{aligned} \tag{2.19}$$

thanks to (2.5), (2.8), (2.9) and (2.13). The claim thus follows thanks to (2.14) and the definition of θ in (2.6). □

Proposition 2.2 is at the very heart of asymptotic analysis as it provides a dynamic relation between the different substitute masses which does not involve the small parameter τ in front of the time derivative. In particular, (2.18) implies the validity of the limit model from Section 1 provided that we can control the approximation errors $m_j - \bar{m}_j$ and $m_j - \widetilde{m}_j$, and this will be done below using the Wasserstein gradient structure.

A particular challenge in this context is that the energy $\mathcal{E}$ is not bounded below (for $\sigma \neq 0$) but decreases in t since there is an effective mass transport due to the tilting of the potential. In order to estimate the decrease of $\mathcal{E}$, one has to control the growth of $\mathcal{P}$ from (1.17), but the PDE (1.1) does not give rise to uniform bounds for $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}$. To overcome this difficulty, we introduce the moment

$$\mathcal{K}(t) := \int_{\mathbb{R}^n}\int_{\mathbb{R}} \phi(p)\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x\,, \tag{2.20}$$

whose weight function is uniquely defined by

$$\phi'(p) := \sum_{j\in\mathbb{Z}} \psi_j'(p)\,, \qquad \phi(P_0) := 0. \tag{2.21}$$

and illustrated in the right panel of Figure 5.

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

Lemma 2.3 (evolution of $\mathcal{K}$) *We have*

$$(1+\theta)\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}(t) = (1-\kappa)\sum_{j\in\mathbb{Z}}\int_{\mathbb{R}^n}\overline{m}_j(t,\,x)\,\mathrm{d}x,$$

as well as

$$\left|P_0 + L\,\mathcal{K}(t) - \mathcal{P}(t)\right| \le C.$$

for some constant C which does not dependent on t or v.

Proof The definitions (2.15), (2.15) and (2.21) yield

$$\phi(p) = \sum_{j=0}^{+\infty}\psi_j(p) - \sum_{j=-1}^{-\infty}\big(1-\psi_j(p)\big)\,, \tag{2.22}$$

where the right-hand side is actually a finite sum for any given $p\in\mathbb{R}$. In particular, we have

$$\sup_{p\in\mathbb{R}}\left|P_0 + L\phi(p) - p\right| < \infty\,, \tag{2.23}$$

and this implies the second claim. The first one follows from (2.19) and (2.22) after summation over j and integration with respect to x. □

2.3 Asymptotic error estimates

In this section, we establish the key asymptotic estimates concerning the approximation of m_j from (1.6) by the substitute masses $\overline{m}_j$ and $\widetilde{m}_j$ from (2.13) and (2.14), respectively.

Lemma 2.4 (asymptotic auxiliary result) *For any j, there exists an interval $I_j\subset J_j$ such that*

$$\int_{J_j\setminus I_j}\gamma_j(p) \le C\nu\sqrt{\tau}\,, \qquad \sup_{p\in I_j}\left|\psi_j(p)\right| + \left|1-\psi_{j-1}(p)\right| \le C\nu\sqrt{\tau}.$$

holds for some constant C which depends on σ but not on v.

Proof For any j we can – thanks to the monotonicity properties of H' (see Assumption 1.1 and Figure 6) – choose $\underline{R}_j$ and $\overline{R}_j$ such that

$$Q_{j-1} < \underline{R}_j < P_j\,, \qquad H\big(\underline{R}_j\big) - \sigma\underline{R}_j = \tfrac{1}{2}\Big(H\big(Q_{j-1}\big) + H\big(P_j\big) - \sigma\big(Q_{j-1}+P_j\big)\Big).$$

and

$$P_j < \overline{R}_j < Q_j\,, \qquad H\big(\overline{R}_j\big) - \sigma\overline{R}_j = \tfrac{1}{2}\Big(H\big(Q_j\big) + H\big(P_j\big) - \sigma\big(Q_j+P_j\big)\Big)\,.$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

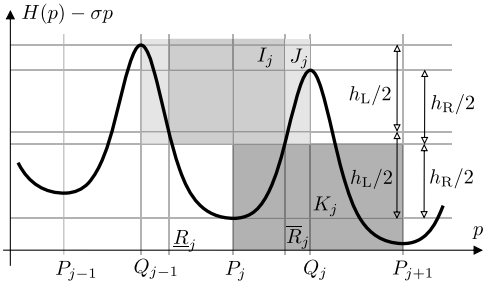

FIGURE 6. Positions $\underline{R}_j$ and $\overline{R}_j$ as used in the proof of Lemma 2.4. The constants h_L and h_R are the Kramers barriers from the scaling law in (1.9) and (1.10), and satisfies $h_R \leq h_L$ thanks to $\sigma \geq 0$.

We define

$$I_j := (\underline{R}_j, \overline{R}_j)\,,$$

and the Laplace method – compare also the asymptotic formula for μ_j in (2.7) – yields

$$\int_{P_j}^{\overline{R}_j} \gamma_j(p)\,\mathrm{d}p = C\nu \exp\left(-\frac{h_L}{2\nu^2}\right)\big(1 \pm O(\nu)\big),$$

as well as

$$\int_{\overline{R}_j}^{P_j} \gamma_j(p)\,\mathrm{d}p = C\nu \exp\left(-\frac{h_R}{2\nu^2}\right)\big(1 \pm O(\nu)\big)\,.$$

These formulas imply the first claim due to the time scaling (1.10) and since $\sigma \geq 0$ guarantees $h_L \geq h_R$. Finally, in view of (2.15)+(2.16), the second claim can be justified along the same lines. □

The main result in this section can be formulated as follows and controls the pointwise approximation error of the substitute masses in terms of the pointwise dissipation D and the total mass m from (1.7) and (2.11), respectively.

Proposition 2.5 (dissipation bounds approximation error) *We have*

$$\sum_{j\in\mathbb{Z}} \big|m_j(t,\,x) - \overline{m}_j(t,\,x)\big| + \big|m_j(t,\,x) - \widetilde{m}_j(t,\,x)\big| \leq C\tau^{-1/2}\nu^2 D(t,\,x) + C\tau^{1/2}\nu^{-2} m(t,\,x).$$

for some constant C independent of ν.

Proof Since all arguments hold pointwise in space and time, we omit both the t- and the x-dependence in all quantities.

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

Local approximation error for $\overline{m}$: By direct computations and Hölders inequality, we find

$$\begin{aligned} e_j &:= \int_{J_j} \left| w_j^2(p) - w_j^2\big(P_j\big) \right| \gamma_j(p)\,\mathrm{d}p \le \int_{Q_{j-1}}^{Q_j} \int_{P_j}^{p} \left| 2w_j(q)\partial_p w_j(q) \right| \mathrm{d}q\, \gamma_j(p)\,\mathrm{d}p \\ &\le 2 \int_{Q_{j-1}}^{Q_j} \left(\int_{P_j}^{p} \frac{w_j^2(q)}{\gamma_j(q)}\,\mathrm{d}q \right)^{1/2} \left(\int_{P_j}^{p} \big(\partial_p w_j(q)\big)^2 \gamma_j(q)\,\mathrm{d}q \right)^{1/2} \gamma_j(p)\,\mathrm{d}p\,. \end{aligned} \tag{2.24}$$

Since $1/\gamma_j(p)$ is strictly increasing on the interval $[P_j,\ Q_j]$, we also have

$$\int_{P_j}^{p} \frac{w_j^2(q)}{\gamma_j(q)}\,\mathrm{d}q \le \frac{1}{\gamma_j^2(p)} \int_{P_j}^{p} w_j^2(q)\gamma_j(q)\,\mathrm{d}q \le \frac{m_j}{\gamma_j(p)^2} \qquad \text{for} \qquad p \in [P_j,\ Q_j],$$

due to (2.10), and combining this with the analogous estimate for $p \in [Q_{j-1},\ P_j]$, we demonstrate that (2.24) can be written as

$$e_j \le C\sqrt{m_j D_j}, \tag{2.25}$$

with $D_j(t)$ as in (2.12). This yields

$$\big|m_j - \overline{m}_j\big| = \left| \int_{J_j} \Big(w_j^2(p) - w_j^2\big(P_j\big) \Big) \gamma_j(p)\,\mathrm{d}p \right| \le e_j \le C\sqrt{m_j D_j}. \tag{2.26}$$

thanks to (2.10), (2.13), and since $\int_{J_j} \gamma_j(p)\,\mathrm{d}p = 1$ holds by (2.4) and (2.8).

Local approximation error for $\widetilde{m}$: With I_j as in Lemma 2.4 and in view of (2.10) and (2.13), we find

$$\begin{aligned} \int_{J_j \setminus I_j} w_j^2(p)\gamma_j(p)\,\mathrm{d}p &= m_j - \int_{I_j} w_j^2(p)\gamma_j(p)\,\mathrm{d}p \\ &= \big(m_j - \overline{m}_j\big) \int_{I_j} \gamma_j(p)\,\mathrm{d}p + m_j \int_{J_j \setminus I_j} \gamma_j(p)\,\mathrm{d}p \\ &\quad + \int_{I_j} \Big(w_j^2\big(P_j\big) - w_j^2(p) \Big) \gamma_j(p)\,\mathrm{d}p \\ &\le 2e_j + m_j \int_{J_j \setminus I_j} \gamma_j(p)\,\mathrm{d}p \\ &\le C\sqrt{m_j D_j} + m_j \int_{J_j \setminus I_j} \gamma_j(p)\,\mathrm{d}p\,, \end{aligned}$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

where we employed (2.25) and (2.26) to derive the estimates. Combining this with Lemma 2.4, we thus obtain

$$\begin{aligned}\int_{J_j} \psi_j(p) w_j^2(p) \gamma_j(p)\,\mathrm{d}p &= \int_{J_j \setminus I_j} \psi_j(p) w_j^2(p) \gamma_j(p)\,\mathrm{d}p + \int_{I_j} \psi_j(p) w_j^2(p) \gamma_j(p)\,\mathrm{d}p \\ &\le \int_{J_j \setminus I_j} w_j^2(p) \gamma_j(p)\,\mathrm{d}p + C\nu\sqrt{\tau} m_j \\ &\le C\sqrt{m_j D_j} + C\nu\sqrt{\tau} m_j,\end{aligned} \tag{2.27}$$

and analogously

$$\int_{J_j} \big(1 - \psi_{j-1}(p)\big) w_j^2(p) \gamma_j(p)\,\mathrm{d}p \le C\sqrt{m_j D_j} + C\nu\sqrt{\tau} m_j\,. \tag{2.28}$$

Moreover, from (1.6), (2.9), (2.14) and the piecewise definition of ψ_j – see (2.16) – we deduce the exact representation formula

$$\begin{aligned} m_j - \widetilde{m}_j = &- \int_{J_{j-1}} \psi_{j-1}(p) w_{j-1}^2(p) \gamma_{j-1}(p)\,\mathrm{d}p + \int_{J_j} \big(1 - \psi_j(p)\big) w_j^2(p) \gamma_j(p)\,\mathrm{d}p \\ &+ \int_{J_j} \psi_{j-1}(p) w_j^2(p) \gamma_j(p)\,\mathrm{d}p - \int_{J_{j+1}} \big(1 - \psi_j(p)\big) w_{j+1}^2(p) \gamma_{j+1}(p)\,\mathrm{d}p\,. \end{aligned} \tag{2.29}$$

Here, the four terms on the right-hand side represent the approximation error from the intervals $[P_{j-1},\ Q_{j-1}]$, $[Q_{j-1},\ P_j]$, $[P_j,\ Q_j]$ and $[Q_j,\ P_{j+1}]$; see Figure 5. From (2.29), we finally obtain the estimate

$$\big|m_j - \widetilde{m}_j\big| \le C \sum_{|i-j|\le 1} \Big(\sqrt{m_i D_i} + \nu\sqrt{\tau} m_i\Big). \tag{2.30}$$

by employing (2.27) on both I_{j-1} and I_j and (2.28) on I_j and I_{j+1}.

Global approximation error: Due to the Cauchy–Schwarz estimate and Young's inequality for products, we have

$$\sum_{j\in\mathbb{Z}} \sqrt{m_j D_j} \le \Big(\sum_{j\in\mathbb{Z}} m_j\Big)^{1/2} \Big(\sum_{j\in\mathbb{Z}} D_j\Big)^{1/2} = \sqrt{mD} \le \tfrac{1}{2}\tau^{1/2}\nu^{-2} m + \tfrac{1}{2}\tau^{-1/2}\nu^2 D. \tag{2.31}$$

so the claim follows from summing up the local estimates (2.26) and (2.30). □

For completeness, we also derive an approximation result for other moments of ϱ.

Corollary 2.6 (approximation of moment integrals) *For any smooth and bounded weight function v, we have*

$$\left| \int_{\mathbb{R}} v(t,\,x,\,p)\varrho(t,\,x,\,p)\,\mathrm{d}p - \sum_{j\in\mathbb{Z}} m_j(t,\,x) v\big(t,\,x,\,P_j\big) \right| \le C\tau^{-1/2}\nu^2 D(t,\,x) + C\nu^2 m(t,\,x).$$

for all $t \ge 0$ and all $x \in \mathbb{R}^n$, where the constant C depends on v but not on ν.

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

Proof To ease the notation, we omit again the t- and the x-dependence. Our definitions in (2.8), (2.9) and (2.10) imply

$$\int_{\mathbb{R}} v(p)\varrho(p)\,\mathrm{d}p = \sum_{j\in\mathbb{Z}} \int_{I_j} v(p)w_j^2(p)\gamma_j(p)\,\mathrm{d}p = \sum_{j\in\mathbb{Z}} \big(v(P_j)m_j + e_{\mathrm{a},j} + e_{\mathrm{b},j}\big)\,,$$

where the error terms are given by

$$e_{\mathrm{a},j} := \int_{I_j} \big(v(p) - v(P_j)\big)\big(w_j^2(p) - w_j^2(P_j)\big)\gamma_j(p)\,\mathrm{d}p,$$

$$e_{\mathrm{b},j} := \int_{I_j} \big(v(p) - v(P_j)\big)w_j^2(P_j)\gamma_j(p)\,\mathrm{d}p\,.$$

Similarly to the proof of Lemma 2.5 – cf. the estimates (2.24) and (2.25) – we show

$$\big|e_{\mathrm{a},j}\big| \le C\sqrt{m_j D_j}\,,$$

where we used that the moment weight v is uniformly bounded on I_j. Moreover, (2.13) and the Laplace method ensure that

$$\big|e_{\mathrm{b},j}\big| \le \overline{m}_j \left| \int_{I_j} \big(v(p) - v(P_j)\big)\gamma_j(p)\,\mathrm{d}p \right| \le C\overline{m}_j v^2.$$

since γ_j is localised near $p = P_j$ and because v is sufficiently smooth. Thanks to (2.31) the desired estimate follows after summation with respect to j from Proposition 2.5 and (1.10). □

2.4 Passage to the limit $v \to 0$

In this section, we pass to the limit v and prove that the partial masses of a solution to the Fokker–Planck equation (1.1) converge to a solution of the limit dynamics as stated in Section 1. To this end, we rely on the following assumption concerning the initial data, where

$$\mathcal{V}(t) := \int_{\mathbb{R}^n} \int_{\mathbb{R}} \big(|x|^2 + p^2\big)\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x. \tag{2.32}$$

refers to the variance of ϱ.

Assumption 2.7 (initial data) *The initial data are non-negative and satisfy the normalisation condition*

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}} \varrho(0,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x = 1,$$

as well as the estimates

$$\mathcal{V}(0) \le C\,, \qquad \mathcal{E}(0) \le C.$$

for some constant C independent of v, where the moments $\mathcal{V}$ and the energy $\mathcal{E}$ have been defined in (2.32) and (1.14), respectively.

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

The existence, uniqueness and regularity of a smooth solution ϱ are then guaranteed by standard results; see, for instance, [11] for a classical approach. In particular, the solution satisfies (1.2) for all $t \geq 0$ and this implies

$$\sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} m_j(t,\, x)\, \mathrm{d}x = 1\,. \tag{2.33}$$

Our first technical result in this section is to bound the total dissipation in the temporal L^1-norm, which enables us to control the approximation errors from Proposition 2.5 in a time-averaged sense. Notice that such estimates for the dissipation are not granted *a priori* because the energy is not bounded below but approaches the value $-\infty$ as $t \to \infty$. The key ingredients to our proof are the Wasserstein gradient structure as well as the estimates from Lemma 2.3 for the moment $\mathcal{K}$. The latter ensure that the moment $\mathcal{P}$ grows nicely in time although we are not able to bound its time derivative independently of ν.

Lemma 2.8 (L^1-bound for the dissipation) *There exists a constant C independent of ν such that*

$$\int_0^T \mathcal{D}(t)\, \mathrm{d}t \leq \tau \nu^{-4} C(1+T),$$

holds for all $0 < T < \infty$ and all sufficiently small $\nu > 0$.

Proof *Lower bound for the energy:* Using (1.1) as well as integration by parts, we verify

$$\tau \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{V}(t) = 2\big(\tau + \nu^2\big) - 2\int_{\mathbb{R}^n}\int_{\mathbb{R}} p\big(H'(p)-\sigma\big)\varrho(t,\, x,\, p)\, \mathrm{d}p\, \mathrm{d}x \leq C\tau^{-1} + \tau\mathcal{V}(t)\,,$$

where we also used the Young-type estimate

$$2\left|p\big(H'(p)-\sigma\big)\right| \leq \tau p^2 + C\tau^{-1}\left|H'(p)-\sigma\right|^2 \leq \tau p^2 + C\tau^{-1}.$$

as well as (1.10) and the conservation of mass; see (2.33). The comparison principle for scalar ODEs combined with Assumption 2.7 therefore yields

$$\mathcal{V}(T) \leq C\tau^{-2}\exp(T)\,. \tag{2.34}$$

Since the Gaussian minimises the convex Boltzmann entropy – that is, the integral of $\varrho \ln \varrho$ – with prescribed zeroth and second moment, we verify

$$\begin{aligned}\nu^2 \int_{\mathbb{R}^n}\int_{\mathbb{R}} \varrho(T,\, x,\, p) \ln\big(\varrho(T,\, x,\, p)\big)\, \mathrm{d}p\, \mathrm{d}x &\geq C\nu^2\big(-1-\ln\mathcal{V}(T)\big)\\ &\geq C\nu^2\big(-1-T+\ln\tau\big) = C\big(-1-\nu^2 T\big)\,,\end{aligned}$$

where the first estimate stems from direct computations for Gaussian and the second one is provided by (2.34) and the scaling law (1.10). Moreover, since H is bounded by Assumption 1.1 we find

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}} \big(H(p)-\sigma p\big)\varrho(T,\, x,\, p)\, \mathrm{d}p\, \mathrm{d}x \geq -C - \sigma\mathcal{P}(t),$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

with $\mathcal{P}$ as in (1.17), while the properties of ϕ and $\mathcal{K}$ in (2.20) and (2.23) imply

$$\big|\mathcal{P}(T) - L\,\mathcal{K}(T)\big| \le C\,.$$

In summary, we have

$$\mathcal{E}(T) \ge -C\big(1 + \nu^2 T + \big|\mathcal{K}(T)\big|\big)\,. \tag{2.35}$$

Upper bound for the dissipation: The energy balance (1.15) provides

$$0 \le \int_0^T \big(\mathcal{D}(t) + \tau\,\nu^{-2}\,\mathcal{C}(t)\big)\,\mathrm{d}t \le \tau\,\nu^{-4}\big(\mathcal{E}(0) - \mathcal{E}(T)\big), \tag{2.36}$$

and Lemma 2.3 guarantees

$$\big|\mathcal{K}(T) - \mathcal{K}(0)\big| \le \int_0^T \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} \overline{m}_j(t,\,x)\,\mathrm{d}x\,\mathrm{d}t = T + \int_0^T \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} \big|m_j(t,\,x) - \overline{m}_j(t,\,x)\big|\,\mathrm{d}x\,\mathrm{d}t\,,$$

where we used that the total mass is conserved due to (2.33). Exploiting Proposition 2.5 and the conservation of mass, we further get

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} \big|m_j(t,\,x) - \overline{m}_j(t,\,x)\big|\,\mathrm{d}x &\le C \int_{\mathbb{R}^n} \big(\tau^{-1/2}\nu^2 D(t,\,x) + \tau^{1/2}\nu^{-2} m(t,\,x)\big)\,\mathrm{d}x \\ &\le \tau^{-1/2}\nu^2 \mathcal{D}(t) + \tau^{1/2}\nu^{-2}\,.\end{aligned} \tag{2.37}$$

Assumption 2.7 ensures $\mathcal{E}(0) + |\mathcal{K}(0)| \le C$, so combining (2.35), (2.36) and (2.37) we arrive at

$$\int_0^T \mathcal{D}(t)\,\mathrm{d}t \le C\tau\nu^{-4}\left(1 + T + \int_0^T \left(\tau^{-1/2}\nu^2\mathcal{D}(t) + \tau^{1/2}\nu^{-2}\right)\mathrm{d}t\right).$$

The thesis now follows from rearranging terms and since τ is exponentially small in ν according to Kramers' law (1.10). □

We are now able to prove our main result on the dynamics in the vanishing diffusivity limit $\nu \to 0$. To ease the notation, we restrict ourselves to the case $0 < \sigma < \sigma_*$ but emphasise that all arguments can be easily adapted to the cases $\sigma_* < \sigma < 0$ and $\sigma = 0$.

Theorem 2.9 (limit dynamics) *For $0 < \sigma < \sigma_*$ and fixed $0 < T < \infty$, we have*

$$\sum_{j\in\mathbb{Z}} \int_0^T \int_{\mathbb{R}^n} \big|\breve{m}_j(t,\,x) - m_j(t,\,x)\big|\,\mathrm{d}x\,\mathrm{d}t \le C\nu^2\big(1 + T^2\big)\,, \tag{2.38}$$

where $\breve{m}$ denotes the unique solution to the initial value problem

$$\partial_t \breve{m}_j(t,\,x) - \Delta_x \breve{m}_j(t,\,x) = \breve{m}_{j-1}(t,\,x) - \breve{m}_j(t,\,x)\,, \qquad \breve{m}_j(0,\,x) = \widetilde{m}_j(0,\,x). \tag{2.39}$$

and depends on ν via the initial data.

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

Proof *Error terms and bounds*: Proposition (2.2) provides

$$\partial_t \widetilde{m}_j(t,\,x) - \Delta_x \widetilde{m}_j(t,\,x) = \widetilde{m}_{j-1}(t,\,x) - \widetilde{m}_j(t,\,x) + \frac{f_j(t,\,x) + g_j(t,\,x) + h_j(t,\,x)}{1+\theta}\,,$$

where the error terms on the right hand are given by

$$f_j := \big(\overline{m}_{j-1} - \widetilde{m}_{j-1}\big) - \big(\overline{m}_j - \widetilde{m}_j\big),$$

as well as

$$g_j := \kappa\big(\overline{m}_{j+1} - \widetilde{m}_{j+1}\big) - \kappa\big(\overline{m}_j - \widetilde{m}_j\big),$$

and

$$h_j := \kappa\big(\widetilde{m}_{j+1} - \widetilde{m}_j\big) + \theta\big(\widetilde{m}_j - \widetilde{m}_{j-1}\big)\,.$$

From Proposition 2.5, Lemmas 2.8 and (2.5), we infer the estimate

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \int_0^T \int_{\mathbb{R}^n} \big|f_j(t,\,x)\big| + \big|g_j(t,\,x)\big|\,\mathrm{d}x\,\mathrm{d}t &\le C \int_0^T \Big(\tau^{-1/2}\nu^2\mathcal{D}(t) + \tau^{1/2}\nu^{-2}\Big)\,\mathrm{d}t\\ &\le C\tau^{1/2}\nu^{-2}(1+T)\,,\end{aligned}$$

while the conservation of mass combined with (2.17) gives

$$\sum_{j\in\mathbb{Z}} \int_0^T \int_{\mathbb{R}^n} \big|h_j(t,\,x)\big|\,\mathrm{d}x\,\mathrm{d}t \le 2(\kappa+\theta) \sum_{j\in\mathbb{Z}} \int_0^T \int_{\mathbb{R}^n} \widetilde{m}_j(t,\,x)\,\mathrm{d}x\,\mathrm{d}t = C(\kappa+\theta)T\,.$$

Properties of the limit dynamics: The linear limit model gives rise to a well-defined semi-group which is non-expansive with respect to the natural L^1-norm (sums over j and integrals with respect to x) as it preserves the positivity and conserves mass; see also the explicit formula for the fundamental solution in (1.13). We can therefore apply Duhamel's principle to the difference $\breve{m} - \widetilde{m}$ and obtain

$$\sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} \big|\breve{m}_j(t,\,x) - \widetilde{m}_j(t,\,x)\big|\,\mathrm{d}x \le \int_0^t e(s)\,\mathrm{d}s\,,$$

where

$$e(t) := (1+\theta)^{-1} \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} \Big(\big|f_j(t,\,x)\big| + \big|g_j(t,\,x)\big| + \big|h_j(t,\,x)\big|\Big)\,\mathrm{d}x\,.$$

Concluding arguments: All partial results derived so far imply

$$\begin{aligned}\sum_{j\in\mathbb{Z}} \int_0^T \int_{\mathbb{R}^n} \big|m_j(t,\,x) - \widetilde{m}_j(t,\,x)\big|\,\mathrm{d}x\,\mathrm{d}t &\le \int_0^T \int_0^t e(s)\,\mathrm{d}s\,\mathrm{d}t \le T\int_0^T e(t)\,\mathrm{d}t\\ &\le C\big(T+T^2\big)\big(\tau^{1/2}\nu^{-2} + \kappa + \theta\big) \le C\big(T+T^2\big)\nu^2\,,\end{aligned}$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

where the last estimate holds thanks to the scaling laws for τ, κ and θ; see (1.10), (2.5) and (2.6). The thesis now follows since

$$\sum_{j\in\mathbb{Z}} \int_0^T \int_{\mathbb{R}^n} \left| m_j(t,\, x) - \widetilde{m}_j(t,\, x) \right| \mathrm{d}x\, \mathrm{d}t \le C\tau^{1/2}\nu^{-2}(1+T).$$

is another consequence of Proposition 2.5 and Lemma 2.8. □

The rather large error in (2.38) stems from the estimate $|\theta| = O\big(\nu^2\big)$. If we replaced the time scaling (1.10) by the refined but less explicit law

$$\tau = \frac{\nu^2}{\mu_0\eta_0} = c_{\mathrm{K}} \exp\left(-\frac{\min\{h_{\mathrm{L}},\, h_{\mathrm{R}}\}}{\nu^2}\right)\Big(1 + O\big(\nu^2\big)\Big).$$

with ν-dependent integral constants μ_0, η_0 as in (2.5), the approximation error would be of order $O\big(\kappa + \tau^{1/2}\nu^{-2}\big)$ and hence exponentially small in ν. Notice also that the initial data for $\breve{m}$ in (2.39) are defined in terms of $\widetilde{m}_j(0,\, x)$ instead of $m_j(0,\, x)$. The difference $\sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^n} \left|\widetilde{m}_j(0,\, x) - m_j(0,\, x)\right| \mathrm{d}x$ is small for sufficiently nice initial data – for instance, if the initial dissipation $\mathcal{D}(0)$ is small – and can only be large if a non-negligible amount of the initial mass is concentrated in the ν-vicinity of the local maxima of the effective potential, that is, near the Q_j's. In the latter case, a fast transient dynamics can/will produce rapid changes in the masses m_j while the substitute masses $\widetilde{m}_j$ still evolve quite regularly according to the limit dynamics.

We finally mention that the combination of Theorem 2.9 and Corollary 2.6 implies the time-dependent probability measure ϱ can in fact be approximated as in (1.8). Moreover, adapting the arguments from in the proof of Corollary 2.6, we also verify

$$\mathcal{P}(t) \approx \sum_{j\in\mathbb{Z}} P_j \int_{\mathbb{R}^n} m_j(t,\, x)\, \mathrm{d}x\,, \qquad \mathcal{V}(t) \approx \sum_{j\in\mathbb{Z}} P_j^{\,2} \int_{\mathbb{R}^n} m_j(t,\, x)\, \mathrm{d}x + \int_{\mathbb{R}^n} x^2 m_j(t,\, x)\, \mathrm{d}x,$$

for the moment integrals from (1.17) and (2.32), where the error terms can be bounded for $0 \le t \le T$ explicitly in terms of ν and $\int_0^T \mathcal{D}(t)\, \mathrm{d}t$. Of particular interest are the first p-moments in the spatially homogeneous case with

$$M_k(t) = \int_{\mathbb{R}} p^k\, \varrho(t,\, p)\, \mathrm{d}p \approx \sum_{j\in\mathbb{Z}} P_j^k\, m_j(t).$$

for $k = 0, 1, 2$. The mass conservation implies $M_0(t) = 1$ and due to $P_{j+1} = P_j + L$, we infer from the limit dynamics the validity of

$$\dot{M}_1(t) \approx \sum_{j\in\mathbb{Z}} P_j\, \dot{m}_j(t) \approx \sum_{j\in\mathbb{Z}} \big(P_{j+1} - P_j\big)\, m_j(t) = L,$$

and

$$\dot{M}_2(t) \approx \sum_{j\in\mathbb{Z}} P_j^2\, \dot{m}_j(t) \approx \sum_{j\in\mathbb{Z}} \big(P_{j+1}^2 - P_j^2\big)\, m_j(t) \approx 2\, L\, M_1(t) - L^2\,,$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

where the approximation holds at least in a weak sense for $0 < \sigma < \sigma_*$ and on the timescale under consideration. We thus conclude

$$\frac{M_1(t)}{t} \approx L\,, \qquad \frac{M_2(t)\cdot M_0(t) - M_1^2(t)}{t} \approx L^2. \tag{2.40}$$

and recover well-known formulas for the limit $\nu \to 0$. In the literature, cf., for instance, [27, 28, 20, 26, 12], the two quantities on the left-hand side are often called *drift coefficient* (or *mean velocity* or *particle current*) and *diffusion coefficient*, respectively, and the characterisation of their long-time behaviour is usually the main motivation for studying Fokker–Planck equations with tilted periodic coefficients. Recall that the ratios from (2.40) concern the effective drift and diffusion in the p-direction and that our approach does not cover the limit $t \to \infty$. Instead, it is based on the energy-dissipation balance (1.15) for finite times intervals and provides a refined asymptotic model for the limit $\nu \to 0$ as it accounts for the mass exchange between the different wells of the effective potential.

Appendix A Mass transport in the supercritical regime

In this appendix, we show that appropriately defined moment integrals are also useful in the supercritical (or ballistic) case $\sigma \notin [\sigma_*, \sigma^*]$, in which the effective potential for (1.1) has no local extrema; see the right panel of Figure 2. For simplicity, we restrict our considerations to

$$\sigma > \sigma^*.$$

and show that the large-time evolution of the first p-moment can be deduced from the balance law of a substitute moment. In this way, we recover the well-known linear grow relation for $\mathcal{P}(t)$, see, for instance, [26], which reveals that the natural choice for the ballistic timescale is $\tau = 1$.

Proposition A.1 (centre of mass in the supercritical regime) *There exists a constant C such that*

$$\tau\big|\mathcal{P}(t) - \lambda\, t\big| \le C\big(1 + \tau + \nu^2 t\big) \qquad \textit{with} \qquad \frac{1}{\lambda} := \frac{1}{L}\int_0^L \frac{\mathrm{d}p}{H'(p) - \sigma}.$$

holds for any t and all sufficiently small $\nu > 0$.

Proof We define a moment weight ψ by the ODE initial value problem

$$\nu^2\psi''(p) = \big(H'(p) - \sigma\big)\psi'(p) + 1\,, \qquad \psi'(0) = c\,, \qquad \psi(0) = 0\,, \tag{A1}$$

where c will be chosen below, and using integration by parts we infer from (1.1) the identity

$$\begin{aligned}\tau\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^n}\int_{\mathbb{R}} \psi(p)\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x &= \int_{\mathbb{R}^n}\int_{\mathbb{R}} \Big(\nu^2\psi''(p) - \big(H'(p) - \sigma\big)\psi'(p)\Big)\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}} \varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x = 1\,.\end{aligned} \tag{A2}$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

By variation of constants, we further demonstrate that ψ satisfies

$$\psi'(p) = \frac{1}{\gamma(p)} \left(c\,\gamma(0) + \int_0^p \frac{\gamma(q)}{\nu^2}\,\mathrm{d}q \right),$$

with γ as in (2.2), and conclude that there is precisely one choice of c, namely

$$c = \frac{1}{\nu^2\big(\gamma(L) - \gamma(0)\big)} \int_0^L \gamma(q)\,\mathrm{d}q > 0\,,$$

such that ψ' is L-periodic. This implies

$$\left| \int_{\mathbb{R}^n} \int_{\mathbb{R}} \psi(p)\varrho(t,\,x,\,p)\,\mathrm{d}p\,\mathrm{d}x - \frac{\mathcal{P}(t)}{L} \int_0^L \psi'(q)\,\mathrm{d}q \right| \le C.$$

since $p \mapsto \psi(p) - p\,L^{-1}\int_0^L \psi'(q)\,\mathrm{d}q$ is bounded, and it remains to compute the integral of ψ' over $[0,\,L]$. Inserting the ansatz

$$\psi' =: u = u_0 + \nu^2 u_1 + \nu^4 u_2 + \cdots\,,$$

into the differential equation (A1), we verify

$$u_0(p) = \frac{1}{\sigma - H'(p)}\,, \qquad u_1(p) = \frac{u_0'(p)}{H'(p) - \sigma} = -\frac{H''(p)}{\big(\sigma - H'(p)\big)^3}\,.$$

and the claim follows after integrating (A2) with respect to t. □

Appendix B Mass exchange in a double-well potential

In this appendix, we apply the asymptotic arguments from above to the case of a double-well potential as illustrated and described in Figure B1. For simplicity, we restrict our considerations to the spatially homogeneous situation and study the Fokker–Planck equation

$$\tau\,\partial_t \varrho(t,\,p) = \partial_p\big(\nu^2 \partial_p \varrho(t,\,p) + H'(p)\varrho(t,\,p)\big)\,, \tag{B1}$$

where the scaling law

$$\tau := \omega_0\,\omega_-\,\exp\left(-\frac{h-}{\nu^2}\right).$$

involves the curvature constants from Figure B1 and is provided by Kramers formula. For the latter, we refer to [18, 4], and to [6, 21] for generalisation to higher dimensions.

B.1 Limit dynamics and known proof strategies

As in the tilted case, one expects that almost all mass of the system is – at least for regular initial data and after a small transient time – concentrated near the stable wells, that is, in the vicinity

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

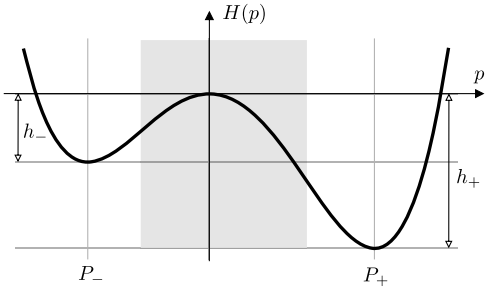

FIGURE B1. In this appendix, H is a smooth double-well potential which grows at least quadratically at infinity, admits the normalised local maximum $H(0) = 0$ and attains two local minima at $P_\pm$ with $P_- < 0 < P_+$. Moreover, we always assume $h_- \le h_+$, where $h_\pm := -H(P_\pm) > 0$, and suppose that H is non-degenerated according to $0 > H''(0) =: -2\pi\omega_0^2$ and $0 < H''(P_\pm) =: 2\pi\omega_\pm^2$.

of the local minima at $p = \pm P$. It is therefore natural to introduce the intervals (or 'phases')

$$J_- := (-\infty,\ 0)\,, \qquad J_+ := (0,\ +\infty),$$

and to define the partial masses by

$$m_\pm(t) := \int_{J_\pm} \varrho(t,\ p)\ \mathrm{d}p \qquad \text{so that} \qquad m_-(t) + m_+(t) = 1\,.$$

Using formal asymptotic analysis, we show – see, for example, [18, 14] for more details – that the two phases exchange mass according to

$$\dot{m}_+(t) = -\dot{m}_-(t) = \begin{cases} m_-(t) & \text{in the generic case with } h_- > h_+\,, \\ m_-(t) - m_+(t) & \text{in the symmetric case of } H(p) = H(-p)\,, \end{cases} \tag{B2}$$

and there exists several ways to derive the limit ODEs rigorously. The first one is to apply large deviation results to the underlying stochastic ODE, see, for instance, [4], but alternatively, PDE-analytic proofs have been given during the last decades by several authors in the framework of gradient flows. However, those proofs have so far been restricted to the symmetric situation with even function H and require non-obvious modification in the general, asymmetric case.

A first key observation – both in the symmetric and the asymmetric case – is that the Gibbs function

$$\gamma(p) := \exp\left(-\nu^{-2} H(p)\right),$$

is now integrable, so that (B1) admits the global equilibrium

$$\overline{\gamma}(p) := \frac{\gamma(p)}{\mu_- + \mu_+}\,, \tag{B3}$$

where the constant $\mu_\pm$ will be computed below and ensure that $\int_{\mathbb{R}} \overline{\gamma}(p)\ \mathrm{d}p = 1$. In terms of the relative density

$$u(t,\ p) := \frac{\varrho(t,\ p)}{\overline{\gamma}(p)}\,,$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

the PDE (B1) reads

$$\tau\overline{\gamma}(p)\partial_t u(t,\,p) = \nu^2\partial_p\Big(\overline{\gamma}(p)\partial_p u(t,\,p)\Big). \tag{B4}$$

and can be interpreted as scaled variant of the $\mathsf{H}^0_{\overline{\gamma}}$-gradient flow to the $\mathsf{H}^1_{\overline{\gamma}}$-energy of u, where the lower index indicates that the Sobolev spaces involve the weight function $\overline{\gamma}$. This Hilbert space formulation has – in a slightly different setting – been exploited in [24] for the rigorous derivation of the limit model in the symmetric case with even potential H. In particular, it has been shown that the quadratic metric tensor as well as the quadratic energy for u – which both depend on ν via $\overline{\gamma}$ – Γ-converge to limit objects that provide a linear gradient structure for the partial masses $(m_-,\,m_+)$. Finally, [10] also passes to the limit $\nu\to 0$ in (B4) but exploits more elementary concepts instead of Γ-convergence.

As already mentioned and shown in [16, 17], the Fokker–Planck equation (B1) can also be regarded as the Wasserstein gradient flow to the energy

$$\mathcal{E}(t) := \int_{\mathbb{R}} \big(\nu^2\varrho(t,\,p)\ln\big(\varrho(t,\,p)\big) + H(p)\varrho(t,\,p)\big)\,\mathrm{d}p.$$

in the space of all probability measure on $\mathbb{R}$, and it is reasonable to ask whether one can also pass to the limit $\nu\to 0$ in this non-flat setting with state-dependent metric tensor; cf. [1, 30] for the general theory of such gradient flows. A positive answer – again in a slightly different setting – has been given in [13] and [2] using different concepts of evolutionary Γ-convergence; see especially [2] for a comparative discussion. However, both results are again restricted to the spatial case $h_- = h_+$ because otherwise u cannot be bounded independently of ν.

In what follows we sketch an alternative derivation of the limit models (B2) which combines the dynamics of substitute masses with the *a priori* bounds for the Wasserstein dissipation, does not appeal to any notion of Γ-convergence and covers both symmetric and asymmetric functions H.

B.2 Substitute masses and passage to the limit

In consistency with the case of a tilted periodic potential, we define the scalar quantities

$$\mu_\pm := \int_{J_\pm}\gamma(p)\,\mathrm{d}p\,,\qquad \eta := \int_{P_-}^{P_+}\frac{1}{\gamma(p)}\,\mathrm{d}p\,,\qquad \kappa := \frac{\mu_-}{\mu_+}\,,\qquad \theta := \frac{\tau\mu_-\eta}{\nu^2} - 1,$$

and introduce relative densities $\omega^2_\pm : J_\pm\to\mathbb{R}$ by

$$w^2_\pm(t,\,p) := \frac{\varrho(t,\,p)}{\gamma_\pm(p)}\qquad\text{for}\quad p\in J_\pm\,,$$

where

$$\gamma_\pm(p) := \mu_\pm^{-1}\gamma(p)\chi_{J_\pm}(p).$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

represent the normalised restrictions of γ to $J_\pm$. Moreover, choosing the moment weight ψ according to

$$\psi(P_-) = 0\,, \qquad \psi'(p) = \begin{cases} \dfrac{1}{\eta\gamma(p)} & \text{for } p \in (P_-,\, P_+)\,, \\ 0 & \text{else}\,, \end{cases} \tag{B5}$$

the different substitute masses are given by

$$\widetilde{m}_-(t) := \int_{\mathbb{R}} \big(1 - \psi(p)\big)\varrho(t,\, p)\,\mathrm{d}p\,, \qquad \widetilde{m}_+(t) := \int_{\mathbb{R}} \psi(p)\varrho(t,\, p)\,\mathrm{d}p,$$

and

$$\overline{m}_\pm(t) := w_\pm(t,\, P_\pm)^2\,.$$

We finally observe that

$$\varrho(t,\, p) = w_-(t,\, p)^2\gamma_-(p) + w_+(t,\, p)^2\gamma_-(p),$$

as well as

$$m_\pm(t) = \int_{J_\pm} w_\pm^2(t,\, p)\gamma_\pm(p)\,\mathrm{d}p, \qquad\qquad m_-(t) + m_+(t) = \widetilde{m}_-(t) + \widetilde{m}_+(t) = 1.$$

hold by construction, and that the Wasserstein dissipation can be written as

$$\mathcal{D}(t) = \int_{\mathbb{R}} \frac{\Big(\partial_p\varrho(t,\, p) + \nu^{-2}H'(p)\varrho(t,\, p)\Big)^2}{\varrho(t,\, p)}\,\mathrm{d}p = 4D_-(t) + 4D_+(t),$$

with

$$D_\pm(t) := \int_{J_\pm} \big(\partial_p w_\pm(t,\, p)\big)^2\gamma_\pm(p)\,\mathrm{d}p\,.$$

Thanks to these definitions and employing the techniques from Section 2, we establish the following results on the effective dynamics for $\nu \to 0$.

Proposition B.1 (building blocks for the limit $\nu \to 0$) *The following statements are satisfied for all sufficiently small $\nu > 0$:*

(1) Elementary asymptotics: *The scalar quantities fulfil*

$$\big|\theta\big| + \left|\frac{\eta\,\omega_0}{\nu} - 1\right| + \left|\frac{\mu_\pm\,\omega_\pm}{\nu}\exp\left(-\frac{h_\pm}{\nu^2}\right) - 1\right| + \left|\frac{\kappa\,\omega_-}{\omega_+}\exp\left(\frac{h_+ - h_-}{\nu^2}\right) - 1\right| \le C\nu^2\,.$$

(2) Effective dynamics: *The substitute masses evolve according to*

$$\pm(1+\theta)\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{m}_\pm(t) = \overline{m}_-(t) - \kappa\overline{m}_+(t)\,. \tag{B6}$$

(3) Dissipation bounds error: *We have*

$$\big|\widetilde{m}_\pm(t) - m_\pm(t)\big| + \big|\overline{m}_\pm(t) - m_\pm(t)\big| \le C\left|\tau^{-1/2}\nu^2\mathcal{D}(t) + \tau^{1/2}\nu^{-2}\right|\,.$$

https://doi.org/10.1017/S0956792519000251 Published online by Cambridge University Press

(4) Energy balance: *The total Wasserstein dissipation is bounded by*

$$\int_0^\infty \mathcal{D}(t)\,\mathrm{d}t \le C\tau\nu^{-4}\big(C+\mathcal{E}(0)\big)\,.$$

Here, the constant C depends on H but not on ν.

Proof The first three assertions can be derived analogously to the proofs of Lemma 2.1 and Propositions 2.2 and 2.5. The justification of the fourth claim is even simpler than in Section 2 because the energy $\mathcal{E}$ is now bounded below due to the existence of the global minimiser $\overline{\gamma}$ from (B3). In particular, we have

$$\mathcal{E}(t) \ge \int_{\mathbb{R}} \overline{\gamma}(p)\Big(\nu^2 \ln\big(\overline{\gamma}(p)\big) + H(p)\Big)\,\mathrm{d}p \ge -\nu^2 \ln\left(\mu_- + \mu_+\right) \ge -C,$$

thanks to $\mu_\pm \sim \exp\left(-h_\pm/\nu^2\right)$. □

Proposition B.1 allows us to pass to the limit $\nu \to 0$ similarly to Theorem 2.9, that is, by means of functions $\breve{m}_\pm$ which solve the limit ODEs and attain the same initial values as $\widetilde{m}_\pm$. The outcome can informally stated as follows.

Corollary B.2 (limit dynamics for $\nu \to 0$) *For sufficiently nice initial data, the asymptotic mass exchange is governed by*

$$\pm\dot{m}_\pm(t) = m_-(t),$$

for $h_- > h_+$ and by

$$\pm\dot{m}_\pm(t) = m_-(t) - \kappa m_+(t).$$

in the non-generic case of $h_- = h_+$, where $\kappa = \omega_+/\omega_-$.

We finally emphasise that the primitive of $1/\gamma$, which defines the moment weight ψ in (B5), features prominently also in [24, 13, 2, 15, 10], but it seems that this function has never been used before to establish a dynamical identity like (B6).

Acknowledgements

The authors gratefully acknowledge the support by the *Deutsche Forschungsgemeinschaft* within the Collaborative Research Center 1060 '*The Mathematics of Emergent Effects*'.

References

[1] Ambrosio, L., Gigli, N. & Savaré, G. (2008) *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich, 2nd ed., Birkhäuser Verlag, Basel.

[2] Arnrich, S., Mielke, A., Peletier, M. A., Savaré, G. & Veneroni, M. (2012) Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction. *Calc. Var. Part. Diff. Eq.* **44**(3–4).

[3] BENDER, C. M. & ORSZAG, S. A. (1999) *Advanced Mathematical Methods for Scientists and Engineers. I*. Springer-Verlag, New York. Asymptotic methods and perturbation theory, Reprint of the 1978 original.

[4] BERGLUND, N. (2013) Kramers' law: validity, derivations and generalisations. *Markov Process. Related Fields* **19**(3), 459–490.

[5] BLANCHET, A., DOLBEAULT, J. & KOWALCZYK, M. (2009) Stochastic Stokes drift, homogenized functional inequalities, and large time behavior of Brownian ratchets. *SIAM J. Math. Anal.* **41**(1), 46–76.

[6] BOVIER, A., ECKHOFF, M., GAYRARD, V. & KLEIN, M. (2004) Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)* **6**(4), 399–424.

[7] CHENG, L. & YIP, N. K. (2015) The long-time behavior of Brownian motion in tilted periodic potentials. *Phys. D* **297**, 1–32.

[8] COLLET, P. & MARTINEZ, S. (2008) Asymptotic velocity of one dimensional diffusions with periodic drift. *J. Math. Biol.* **56**, 765–792.

[9] DENISOVA, S., HÄNGGI, P. & MATEOS, J. L. (2009) AC-driven Brownian motors: A Fokker-Planck treatment. *Am. J. Phys.* **77**(7), 602–606.

[10] EVANS, L. C. & TABRIZIAN, P. R. (2016) Asymptotics for scaled Kramers-Smoluchowski equations. *SIAM J. Math. Anal.* **48**(4), 2944–2961.

[11] FRIEDMAN, A. (1964) *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Inc., Englewood Cliffs, N.J. reprinted 2008 by Dover Publications.

[12] HAIRER, M. & PAVLIOTIS, G. A. (2008) From ballistic to diffusive behavior in periodic potentials. *J. Stat. Phys.* **131**(1), 175–202.

[13] HERRMANN, M. & NIETHAMMER, B. (2011) Kramers' formula for chemical reactions in the context of Wasserstein gradient flows. *Commun. Math. Sci.* **9**(2), 623–635.

[14] HERRMANN, M., NIETHAMMER, B. & VELÁZQUEZ, J. J. L. (2012). Kramers and non-Kramers phase transitions in many-particle systems with dynamical constraint. *Multiscale Model. Simul.* **10**(3), 818–852.

[15] HERRMANN, M., NIETHAMMER, B. & VELÁZQUEZ, J. J. L. (2014) Rate-independent dynamics and Kramers-type phase transitions in nonlocal Fokker-Planck equations with dynamical control. *Arch. Ration. Mech. Anal.* **124**(3), 803–866.

[16] JORDAN, R., KINDERLEHRER, D. & OTTO, F. (1997) Free energy and the Fokker-Planck equation. *Phys. D* **107**(2–4), 265–271. Landscape Paradigms in Physics and Biology (Los Alamos, NM, 1996).

[17] JORDAN, R., KINDERLEHRER, D. & OTTO, F. (1998) The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17.

[18] KRAMERS, H. A. (1940) Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284–304.

[19] LATORRE, J. C., PAVLIOTIS, G. A. & KRAMER, P. R. (2013) Corrections to Einstein's relation for Brownian motion in a tilted periodic potential. *J. Stat. Phys.* **150**(4), 776–803.

[20] LINDNER, B., KOSTUR, M. & SCHIMANSKY-GEIER, L. (2001) Optimal diffusive transport in a tilted periodic potential. *Fluct. Noise Lett.* **1**(1), R25–R39.

[21] MENZ, G. & SCHLICHTING, A. (2014) Poincaré and logarithmic Sobolev inequalities by decomposition of the energy landscape. *Ann. Probab.* **42**(5), 1809–1884.

[22] MICHEL, L. & ZWORSKI, M. (2017) A semiclassical approach to the Kramers-Smoluchowski equation. preprint arXiv:1703.07460.

[23] MIRRAHIMI, S. & SOUGANIDIS, P. E. (2013) A homogenization approach for the motion of motor proteins. *NoDEA Nonlinear Diff. Equations Appl.* **20**(1), 129–147.

[24] PELETIER, M. A., SAVARÉ, G. & VENERONI, M. (2010). From diffusion to reaction via Γ-convergence. *SIAM J. Math. Anal.* **42**(4).

[25] PERTHAME, B. & SOUGANIDIS, P. E. (2009) Asymmetric potentials and motor effect: A homogenization approach. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**(6), 2055–2071.

[26] REIMANN, P., VAN DEN BROECK, C., LINKE, H., HÄNGGI, P., RUBI, J. M. & PÉREZ-MADRID, A. (2002) Diffusion in tilted periodic potentials: Enhancement, universality, and scaling. *Phys. Rev. E* **65**(3), 031104, 1–16.

[27] RISKEN, H. (1989) *The Fokker-Planck Equation*, vol. 18, *Springer Series in Synergetics*, 2nd ed. Springer-Verlag, Berlin. Methods of solution and applications.

[28] RODENHAUSEN, H. (1989). Einstein's relation between diffusion constant and mobility for a diffusion model. *J. Statist. Phys.* **55**(5–6), 1065–1088.

[29] SANCHO, J. M. & LACASTA, A. M. (2010). The rich phenomenology of brownian particles in nonlinear potential landscapes. *Eur. Phys. J. Spec. Top.* **187**(1), 49–62.

[30] VILLANI, C. (2009) *Optimal Transport*, vol. 338. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin. Old and new.