Chemical Reactions as Γ-Limit of Diffusion*

Mark A. Peletier[†]
Giuseppe Savaré[‡]
Marco Veneroni[‡]

Abstract. We study the limit of *high activation energy* of a special Fokker–Planck equation known as the Kramers–Smoluchowski equation (KS). This equation governs the time evolution of the probability density of a particle performing a Brownian motion under the influence of a chemical potential H/ε. We choose H having two wells corresponding to two chemical states A and B. We prove that after a suitable rescaling the solution to the KS converges, in the limit of high activation energy ($\varepsilon \to 0$), to the solution of a simpler system modeling the spatial diffusion of A and B combined with the reaction $A \rightleftharpoons B$. With this result we give a rigorous proof of Kramers's formal derivation, and we show how chemical reactions and diffusion processes can be embedded in a common framework. This allows one to derive a chemical reaction as a singular limit of a diffusion process, thus establishing a connection between two worlds often regarded as separate. The proof rests on two main ingredients. One is the formulation of the two disparate equations as evolution equations for measures. The second is a variational formulation of both equations that allows us to use the tools of variational calculus and, specifically, Γ-convergence.

In this paper we prove that in the limit $\varepsilon \to 0$ solutions of the parabolic partial differential equation in spatial variables x and ξ,

$$(0.1) \quad \partial_t \rho - \Delta_x \rho - \tau_\varepsilon \partial_\xi \left(\partial_\xi \rho + \frac{1}{\varepsilon} \rho \, \partial_\xi H \right) = 0 \qquad \begin{array}{l} \text{for } (x,\xi) \in \Omega \times [-1,1] \text{ and } t > 0, \\ \text{with no-flux boundary conditions,} \end{array}$$

converge to solutions of the system of reaction-diffusion equations in spatial variable x,

$$(0.2) \qquad \begin{array}{l} \partial_t \alpha - \Delta_x \alpha = k(\beta - \alpha), \\ \partial_t \beta - \Delta_x \beta = k(\alpha - \beta), \end{array} \qquad \text{for } x \in \Omega \text{ and } t > 0.$$

In what follows we first describe the history of this problem and the modeling context. The definition of the constants τ_ε and k and the function H are given in section 2.

1. Introduction.

1.1. Chemical Reaction as a Diffusion Process.
By the 1930s it was generally accepted that quantum theory describes a wide variety of atomic and subatomic phenomena, including chemical reactions, the topic of this paper. A major step forward was the Born–Oppenheimer approximation [6], which allows one to separate the behavior of the nuclei from that of the electrons: the nuclei can be treated as classical particles that move in a high-dimensional potential landscape, in which the potential is defined through the solution of a quantum mechanical problem for the electrons under the assumption of fixed nuclei.

This approximation brought the calculation of chemical reaction rates within reach, at least theoretically. A chemical reaction event corresponds to the movement of this high-dimensional system of all nuclei from one local minimum of the potential to another, typically driven by fluctuations that arise from collisions with other particles. If the potential landscape can be calculated with sufficient accuracy, then the reaction rate should follow from a classical mechanical problem for the nuclei with stochastic forcing—what we would now call a stochastic differential equation (SDE).

In a seminal paper in 1940, Hendrik Anthony Kramers derived the Fokker–Planck equation for this SDE [17], under the assumption of Gaussian noise, and proceeded to study various limit cases. One of these cases is the "large-friction" or "Smoluchovski" limit, and the corresponding equation is (0.1). Mathematically this corresponds to the motion of a Brownian particle in a potential landscape, as illustrated in Figure 1.
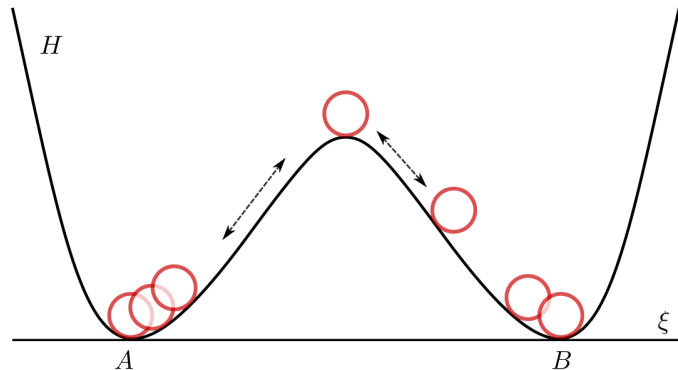
Fig. 1 *Motion of a Brownian particle in a one-dimensional potential landscape.*

Given the dynamics described by an SDE or the corresponding Fokker–Planck equation (0.1), there are many ways of defining transition (reaction) rates (see, e.g., [13] for a nice overview). These are all impossible to calculate explicitly, even for simple choices of the potential function. As a result effort has been directed at determining the rate in the limit of *large activation energy*, when the energy barrier separating the wells is large compared to the noise. There are many successful results in this direction (see [13, 5] for an overview).

1.2. Aim of this Paper.
The aim of this paper is related, but different. It starts from the observation that we have two competing descriptions for particles that are

what we macroscopically call both "diffusing" and "reacting." First, a typical modeling argument leads to equations of the form (0.2) for a finite number of chemical species, with second-order diffusion terms and zero-order reaction terms. Here we limit ourselves to a simple $A \rightleftharpoons B$ system, which arises, for instance, when A and B are two forms of the same molecule, such that the molecule can change from one form into the other. A typical example is a molecule with spatial asymmetry, which might exist in two distinct, mirror-image spatial configurations (but see section 3.3 for more general reactions). The variables α and β are the volume fractions of the two species and the constant k is the reaction rate, which for simplicity we assume is the same for both reactions $A \to B$ and $B \to A$.

On the other hand, the discussion above suggests that the same system should be described by an equation of the form (0.1), in the following way. The variable x models the spatial degrees of freedom of the molecule, i.e., the different positions in space where the molecule can be; the variable ξ models the "chemical" degrees of freedom, corresponding to different arrangements of the atoms inside the molecule. For simplicity (and following Kramers) we assume ξ to be one-dimensional. The energy of a state (x, ξ) is given by a potential function H that we assume to be independent of x and that has a double-well structure as in Figure 1, corresponding to the two stable states A and B. Since the particle undergoes an SDE in (x, ξ), driven by Gaussian noise and the potential H, the probability distribution ρ on (x, ξ)-space satisfies (0.1).

Therefore, the following questions arise: How are (0.1) and (0.2) related? Can they be seen as two signs of the same coin? Is it possible to take a limit (which will be the limit of large activation energy) such that solutions of (0.1) converge to (0.2)?

In this paper we give an answer to these questions by proving two convergence theorems that link the two descriptions. The limit of large activation energy is implemented by replacing H by H/ε, leading to an activation energy of order $O(\varepsilon)$; this explains the factor $1/\varepsilon$ in (0.1). The convergence results are then in the limit $\varepsilon \to 0$.
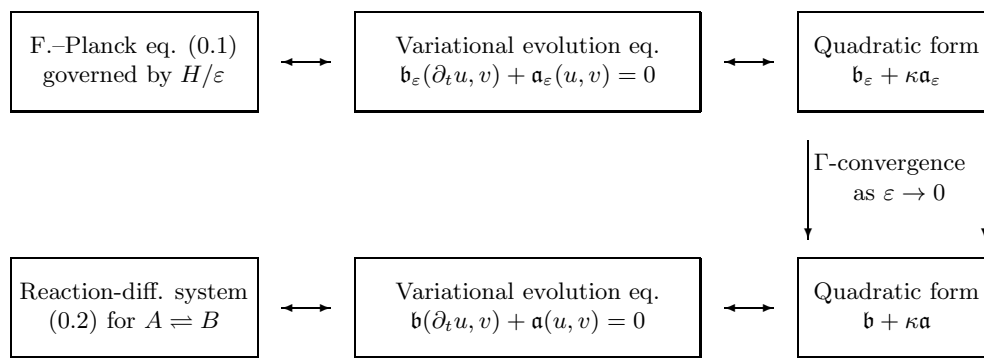
In addition, the unusual aspects of this question prompted us to develop a method that not only allows us to address this question, but may be more generally useful. We describe this in the next section.

1.3. Structure of the Proof. Any attempt to prove a rigorous convergence result faces an important difficulty: the mathematical objects used in the two descriptions are of very different types.

Let us assume for simplicity that ξ parametrizes an imaginary "optimal path" connecting the states A and B such that $\xi = -1$ corresponds to A and $\xi = 1$ to B. At $\varepsilon > 0$, the system is described by a partial differential equation in the space $\Omega \times [-1, 1]$, while in the limit $\varepsilon = 0$ the chemical (ξ-) degrees of freedom are reduced to the two possibilities A and B, which correspond to only $\xi = \pm 1$. Therefore, the question arises in what sense solutions of one could ever converge to the other.

As it turns out, both systems can be described by variational evolution equations in a common space of measures. Measures generalize functions and allow for a rigorous description of concentration phenomena: while functions in Lebesgue or Sobolev spaces cannot represent concentration of finite mass onto a point, in the space of measures this behavior can be described in terms of Dirac distributions $\xi \mapsto \delta(\xi \pm 1)$. Equation (0.1) is already a measure-valued equation; by also reformulating (0.2) as a measure-valued equation we unite the two equations in a common structure. The space of measures carries a natural concept of convergence, the convergence against test functions, and this will be the basis for the first of our convergence results.

A second ingredient in the proof is the use of variational methods to pass to the limit. The ideas for this method go back to [25, 26, 12] (see also [7, 4]), and our use is a generalization of their results. Both systems can be written as gradient flows in weighted L^2-spaces, which are characterized by two quadratic forms. One of the advantages of this variational, gradient-flow formulation is the possibility to shift the study of the convergence of the solutions to the convergence of (linear combinations of) the quadratic forms. This is illustrated in Figure 2. The fundamental tool in this step is the notion of Γ-convergence, introduced by E. De Giorgi (see [11] and [10] for a thorough exposition) in order to study the convergence of minimum problems, and it is nowadays one of the most powerful and flexible instruments for the rigorous study of singular variational problems.

| F.–Planck eq. (0.1) governed by H/ε | ⟷ | Variational evolution eq. $\mathfrak{b}_\varepsilon(\partial_t u, v) + \mathfrak{a}_\varepsilon(u, v) = 0$ | ⟷ | Quadratic form $\mathfrak{b}_\varepsilon + \kappa \mathfrak{a}_\varepsilon$ |

$$\Gamma\text{-convergence as } \varepsilon \to 0$$

| Reaction-diff. system (0.2) for $A \rightleftharpoons B$ | ⟷ | Variational evolution eq. $\mathfrak{b}(\partial_t u, v) + \mathfrak{a}(u, v) = 0$ | ⟷ | Quadratic form $\mathfrak{b} + \kappa \mathfrak{a}$ |

Fig. 2 *Structure of the passage to the limit. The top line is microscopic ($\varepsilon > 0$), the bottom line macroscopic ($\varepsilon = 0$). Each of the horizontal arrows indicates a mathematical equivalence between different formulations. The crucial limiting step $\varepsilon \to 0$ is done by Γ-convergence on the quadratic forms $\mathfrak{b}_\varepsilon + \kappa \mathfrak{a}_\varepsilon$, for arbitrary $\kappa > 0$.*

1.4. Plan of the Paper. In the next section we describe (0.1) and (0.2) and their derivation in more detail. The main results (Theorems 3.1 and 3.2) are given in sections 3.1 and 3.2, and we discuss various aspects in section 3.3.

Section 4 describes the main arguments and tools of the proof: the variational formulation of the equations in weighted L^2-spaces and in measure spaces (sections 4.1 and 4.3), simple regularization estimates (section 4.2), and a few results on Γ-convergence of quadratic forms (section 4.4) related to a weak-strong convergence principle which turns out to be extremely useful in dealing with evolution problems according to the scheme of Figure 2. The applications of this approach to variational evolution problems are briefly discussed in section 4.5 and then further developed in section 6.

Section 5 contains the basic Γ-convergence results (Theorem 5.1). The proof of Theorems 3.1 and 3.2 is then concluded with a general argument showing the link between Γ-convergence of the quadratic forms $\mathfrak{a}_\varepsilon, \mathfrak{b}_\varepsilon$ and the convergence of the solutions to the evolution problems (see the comments in section 4.5): its precise statement is presented and proved in section 6 in an abstract form which is in fact independent of the specific problems under consideration and can be easily applied to other situations.

2. The Model.

2.1. The Setup: Enthalpy. We now describe the systems of this paper in more detail. We consider the unimolecular reaction $A \rightleftharpoons B$. We assume that the observed forms A and B correspond to the wells of a double-well enthalpy function H. (Since it is common in the chemical literature to denote by "enthalpy difference" the release or uptake of heat as a particle A is converted into a particle B, we shall adopt the same language.)

While the domain of definition of H should be high-dimensional, corresponding to the many degrees of freedom of the atoms of the molecule, we will here make the standard reduction to a one-dimensional dependence. As we mentioned before, the states A and B correspond to $\xi = -1$ and $\xi = +1$, respectively, and the variable ξ takes its values in $[-1, 1]$. A transition between -1 and $+1$ should pass through the "mountain pass," the point which separates the basins of attraction of A and B, and we arbitrarily choose that mountain pass to be at $\xi = 0$, with $H(0) = 1$. We also assume for simplicity that the wells are at equal depth, which we choose to be zero. A typical example of the function H is shown in Figure 3.
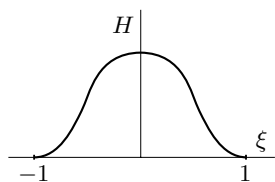
Fig. 3 *A typical function H.*

Specifically, we make the following assumptions about H: $H \in C^\infty([-1, 1])$, and H is even in ξ, maximal at $\xi = 0$ with value 1, and minimal at $\xi = \pm 1$ with value 0; $H(\xi) > 0$ for any $-1 < \xi < 1$; $H'(\pm 1^\mp) = 0$. The assumption of equal depth for the two wells corresponds to an assumption about the rate constants of the two reactions; we comment on this in section 3.3.

2.2. Diffusion in the Chemical Landscape. We now describe the diffusion process, starting with the state space. The "chemical variable" ξ should be interpreted as an internal degree of freedom of the particle, associated with internal changes in configuration. In the case of two alternative states of a molecule, ξ parametrizes all the intermediate states along a connecting path.

In this view the total state of a particle consists of this chemical state ξ together with the spatial position of the particle, represented by a d-dimensional spatial variable x in a Lipschitz, bounded, and open domain $\Omega \subset \mathbb{R}^d$, so that the full state space for the particle is the closure $\overline{D}$ of

$$D := \Omega \times (-1, 1) \qquad \text{with variables } (x, \xi).$$

Taking a probabilistic point of view, and following Kramers, the motion of the particle will be described in terms of its probability density $\rho \in \mathscr{P}(\overline{D})$ in the sense that for Borel sets $X \subset \overline{\Omega}$ and $\Xi \subset [-1, 1]$ the number $\rho(X \times \Xi)$ is the probability of finding the particle at a position $x \in X$ and with a "chemical state" $\xi \in \Xi$.

The particle is assumed to perform a Brownian motion in D, under the influence of the potential landscape described by H. This assumption corresponds to the "large-friction limit" discussed by Kramers. The time evolution of the probability

distribution ρ is then given by the Kramers–Smoluchowski equation

$$(2.1) \qquad \partial_t\rho - \Delta_x\rho - \tau\partial_\xi\big(\partial_\xi\rho + \rho\,\partial_\xi H\big) = 0 \qquad \text{in } \mathscr{D}'(D\times(0,\infty)),$$

with initial condition ρ^0 and Neumann boundary conditions on the lateral boundary ∂D, which imply reflection of the Brownian particle at the boundaries. The coefficient $\tau > 0$ is introduced to parametrize the difference in scales for x and ξ: since x is a rescaled physical distance and ξ is a rescaled "chemical" distance, the units of length in the two variables are different, and the parameter τ can be interpreted as the factor that converts the two scales. Below we shall make an explicit choice for τ.

2.3. The Limit of High Activation Energy. In the setup as described above, there is a continuum of states (i.e., $[-1,1]$) connecting the A state to the B state, and a statement of the type "the particle is in the A state" is therefore not well defined. In order to make a connection with the macroscopic description "$A \rightleftharpoons B$," which requires a clear distinction between the two states, we take the limit of high activation energy, as follows.

We rescale the enthalpy H with a small parameter ε to make it $H(\xi)/\varepsilon$, so that the energy barrier between the wells is exactly $1/\varepsilon$ high. This rescaling has various effects on the behavior of solutions ρ of (2.1). To illustrate one effect, let us consider the invariant measure γ_ε, the unique stationary solution in $\mathscr{P}(\overline{D})$ of (2.1):

$$(2.2) \qquad \gamma_\varepsilon = \lambda_\Omega \otimes \tilde{\gamma}_\varepsilon, \quad \lambda_\Omega := \frac{1}{\mathcal{L}^d(\Omega)}\mathcal{L}^d|_\Omega, \quad \tilde{\gamma}_\varepsilon = Z_\varepsilon^{-1}e^{-H/\varepsilon}\mathcal{L}^1|_{[-1,1]}$$

(where $\mathcal{L}^1, \mathcal{L}^d$ are the one- and d-dimensional Lebesgue measures). The constant Z_ε is fixed by the requirement that $\gamma_\varepsilon(\overline{D}) = \tilde{\gamma}_\varepsilon([-1,1]) = 1$.
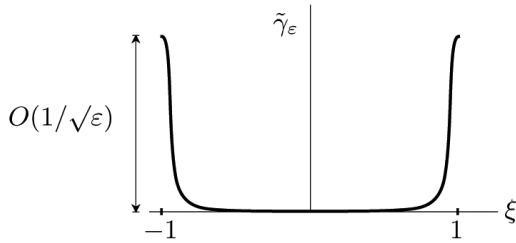
Fig. 4 *The density $\tilde{\gamma}_\varepsilon$.*

Since H is strictly positive at any $-1 < \xi < 1$, the exponential $\exp(-H(\xi)/\varepsilon)$ vanishes, as $\varepsilon \to 0$, at all ξ except for $\xi = \pm 1$ (see Figure 4); therefore, the measure γ_ε concentrates on the lines $\xi = -1$ and $\xi = 1$ and converges weakly-$*$ as $\varepsilon \to 0$ to the limit measure γ given by

$$(2.3) \qquad \gamma = \lambda_\Omega \otimes \tilde{\gamma}, \quad \tilde{\gamma} := \frac{1}{2}\big(\delta_{-1} + \delta_1\big).$$

Here weak-$*$ convergence is to be interpreted in the duality with continuous functions in $\overline{D}$ (thus considering $\mathscr{P}(\overline{D})$ as a weakly-$*$ closed convex subset of the space $\mathscr{M}(\overline{D}) = \big(C^0(\overline{D})\big)'$ of signed Borel measures with finite total variation), i.e.,

$$\lim_{\varepsilon\downarrow 0}\int_{\overline{D}}\phi(x,\xi)\,d\gamma_\varepsilon = \int_{\overline{D}}\phi(x,\xi)\,d\gamma(x,\xi) = \frac{1}{2}\int_\Omega\big(\phi(x,-1)+\phi(x,1)\big)\,d\lambda_\Omega(x)$$

for any $\phi \in C^0(\overline{D})$.

We should interpret the behavior of γ_ε as follows. In the limit $\varepsilon \to 0$, the deep wells at $\xi = \pm 1$ force particles to stay increasingly close to the bottom of the wells. However, at any given $\varepsilon > 0$, there is a positive probability that a particle switches from one well to the other in any given period of time. The rate at which this happens is governed by the local structure of H near $\xi = \pm 1$ and near $\xi = 0$ and becomes very small, of order $\varepsilon^{-1} \exp(-1/\varepsilon)$, as we shall see below.

In the limit $\varepsilon = 0$, the behavior of particles in the ξ-direction is no longer recognizable as diffusional in nature. In the ξ-direction a particle can be in only one of two states, $\xi = \pm 1$, which we have interpreted as the A and B states. Of the diffusional movement in the ξ-direction only a jump process remains, in which a particle at $\xi = -1$ jumps with a certain rate to position $\xi = 1$, or vice versa.

At each time $t \geq 0$ the limit system can thus be described by the nonnegative functions $\alpha(\cdot; t), \beta(\cdot; t) : \Omega \to \mathbb{R}$ representing the densities of particles in the states A and B, respectively; this means that $\int_X \alpha(x; t)\, dx$ is the (normalized) number of particles in the state A contained in the region $X \subset \Omega$, and a similar formula holds for $\beta = 1 - \alpha$.

The time-dependent measures

$$(2.4) \qquad \rho(x, \xi; t) = \alpha(x; t)\mathcal{L}^d(x) \otimes \delta_{-1}(\xi) + \beta(x; t)\mathcal{L}^d(x) \otimes \delta_1(\xi)$$

are thus the limit distributions of particles, and we can expect that $\rho_\varepsilon(\cdot; t)$ converges weakly-$*$ to $\rho(\cdot; t)$ as $\varepsilon \downarrow 0$ in the space $\mathscr{M}(\overline{D})$, i.e.,

$$\lim_{\varepsilon \downarrow 0} \int_{\overline{D}} \phi(x, \xi)\, d\rho_\varepsilon(x, \xi; t) = \int_\Omega \big(\phi(x, -1)\alpha(x; t) + \phi(x, 1)\beta(x; t)\big)\, dx$$

for any $\phi \in C^0(\overline{D})$ and $t > 0$. The remarkable fact is that, under a suitable choice of $\tau = \tau_\varepsilon$ which we will explain in the next section, the limit functions α, β satisfy the reaction-diffusion system (0.2).

2.4. Spatiochemical Rescaling. Since the jumping (chemical reaction) rate at finite $\varepsilon > 0$ is of order $\varepsilon^{-1} \exp(-1/\varepsilon)$, the limiting reaction rate will be zero unless we rescale the system appropriately. This requires us to speed up time by a factor of $\varepsilon \exp(1/\varepsilon)$. At the same time, the diffusion rate in the x-direction remains of order 1 as $\varepsilon \to 0$, and the rescaling should preserve this. In order to obtain a limit in which both diffusion in x and chemical reaction in ξ enter at rates that are of order 1, we use the freedom of choosing the parameter τ that we introduced above.

We therefore choose τ equal to

$$(2.5) \qquad\qquad\qquad \tau_\varepsilon := \varepsilon \exp(1/\varepsilon),$$

and we then find the differential equation

$$(2.6) \qquad \partial_t \rho_\varepsilon - \Delta_x \rho_\varepsilon - \tau_\varepsilon \partial_\xi \big(\partial_\xi \rho_\varepsilon + \tfrac{1}{\varepsilon}\rho_\varepsilon\, \partial_\xi H\big) = 0 \quad \text{in } \mathscr{D}'(D \times (0, \infty)),$$

which clearly highlights the different treatment of x and ξ: the diffusion in x is independent of τ_ε, while the diffusion and convection in the ξ-variable are accelerated by a factor τ_ε.

2.5. Switching to the Density Variable. As is already suggested by the behavior of the invariant measure γ_ε, the solution ρ_ε will become strongly concentrated at the extremities $\{\pm 1\}$ of the ξ-domain $(-1, 1)$. This is the reason why it is useful to

interpret ρ_ε as a family $\rho_\varepsilon(t, \cdot)$ of *time-dependent measures* instead of functions. It turns out that the densities $u_\varepsilon(t, \cdot)$ of $\rho_\varepsilon(t, \cdot)$ with respect to γ_ε,

$$u_\varepsilon(t, \cdot) := \frac{\mathrm{d}\rho_\varepsilon(t, \cdot)}{\mathrm{d}\gamma_\varepsilon},$$

also play a crucial role, and it is often convenient to have both representations at our disposal, freely switching between them. In terms of the variable u_ε, (2.6) becomes

$$(2.7a) \qquad \partial_t u_\varepsilon - \Delta_x u_\varepsilon - \tau_\varepsilon(\partial^2_{\xi\xi} u_\varepsilon - \tfrac{1}{\varepsilon}\partial_\xi H \partial_\xi u_\varepsilon) = 0 \quad \text{in } D \times (0, +\infty),$$

supplemented with the boundary conditions

$$(2.7b) \quad \partial_\xi u_\varepsilon(t, x, \pm 1) = 0 \quad \text{in } \Omega, \quad \nabla_x u_\varepsilon(t, x, \xi) \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega \times [-1, 1], \quad t > 0.$$

We choose an initial condition

$$(2.7c) \qquad u_\varepsilon(0, x, \xi) = u^0_\varepsilon(x, \xi) \qquad \text{for all } (x, \xi) \in D, \quad \text{with} \quad \rho^0_\varepsilon = u^0_\varepsilon \gamma_\varepsilon \in \mathscr{P}(\overline{D}).$$

The work of this paper depends on a special structure of the equation. It is well known (see, e.g., [9]) that the operator $A_\varepsilon := -\Delta_x - \tau_\varepsilon\partial^2_{\xi\xi} + (\tau_\varepsilon/\varepsilon)H'\,\partial_\xi$ with Neumann boundary conditions (2.7b) has a self-adjoint realization in the space $H_\varepsilon := L^2(D; \gamma_\varepsilon)$. Therefore, the weak form of (2.7a) can be written as

$$(2.8) \qquad\qquad b_\varepsilon(\partial_t u(t), v) + a_\varepsilon(u(t), v) = 0 \qquad \text{for all } v \in V_\varepsilon,$$

where the bilinear forms a_ε and b_ε are defined by

$$b_\varepsilon : H_\varepsilon \times H_\varepsilon \to \mathbb{R}, \qquad b_\varepsilon(u, v) := \int_D u\,v\,\mathrm{d}\gamma_\varepsilon,$$

and

$$V_\varepsilon := W^{1,2}(D; \gamma_\varepsilon) := \left\{ u \in L^2(D; \gamma_\varepsilon) \cap W^{1,1}_{\mathrm{loc}}(D) : \int_D |\nabla_{x,\xi} u|^2\,\mathrm{d}\gamma_\varepsilon < +\infty \right\},$$

$$a_\varepsilon : V_\varepsilon \times V_\varepsilon \to \mathbb{R}, \qquad a_\varepsilon(u, v) := \int_D A_\varepsilon u\,v\,\mathrm{d}\gamma_\varepsilon = \int_D \left(\nabla_x u \nabla_x v + \tau_\varepsilon \partial_\xi u\,\partial_\xi v \right) \mathrm{d}\gamma_\varepsilon.$$

Since V_ε is densely and continuously imbedded in H_ε, standard results on variational evolution equations in a Hilbert triplet (see, e.g., [18, 8]) and their regularizing effects show that a unique solution exists in $C([0, \infty); H_\varepsilon) \cap C^\infty((0, \infty); V_\varepsilon)$ for every initial datum $u^0_\varepsilon \in H_\varepsilon$.

2.6. The Variational Structure of the Limit Problem. The "$\varepsilon = 0$" limit problem (0.2) admits the same variational formulation as the "$\varepsilon > 0$" problem we introduced in section 2.5. To be consistent with the previous setting, we switch to the density functions of the limit distribution ρ in (2.4) with respect to the measure γ defined in (2.3) as the weak limit of γ_ε. We easily get

$$(2.9) \qquad \rho(\cdot; t) = u(\cdot; t)\gamma, \quad \text{where} \quad \begin{cases} u(x, 1; t) = u^+(x, t) = 2\mathcal{L}^d(\Omega)\alpha(x, t), \\ u(x, -1; t) = u^-(x, t) = 2\mathcal{L}^d(\Omega)\beta(x, t), \end{cases}$$

so that the couple (u^-, u^+) coincides with (α, β) up to the normalization factor $2\mathcal{L}^d(\Omega)$ and thus satisfies the same system (0.2). We set $H := L^2(\overline{D}, \gamma)$, and for every

$\rho = u\gamma$ with $u \in H$ we set $u^{\pm}(x) := u(x, \pm 1) \in L^2(\Omega, \lambda_{\Omega})$. Note that for a function $u \in L^2(\overline{D}, \gamma)$ these traces are well defined (in fact, the map $u \mapsto (u^-, u^+)$ is an isomorphism between $L^2(\overline{D}, \gamma)$ and $L^2(\Omega, \frac{1}{2}\lambda_{\Omega}; \mathbb{R}^2)$).

We define

$$(2.10) \qquad b(u, v) := \int_D u(x, \xi) v(x, \xi) \, d\gamma(x, \xi) = \frac{1}{2} \int_{\Omega} \left(u^+ v^+ + u^- v^- \right) d\lambda_{\Omega}.$$

Similarly, we set $V := \left\{ u \in H : u^{\pm} \in W^{1,2}(\Omega) \right\}$, which is continuously and densely imbedded in H, and

$$(2.11) \quad a(u, v) := \frac{1}{2} \int_{\Omega} \left(\nabla_x u^+ \nabla_x v^+ + \nabla_x u^- \nabla_x v^- + k\left(u^+ - u^- \right)(v^+ - v^-) \right) d\lambda_{\Omega}.$$

Then the system (0.2) can be formulated as

$$(2.12) \qquad b(\partial_t u(t), v) + a(u(t), v) = 0 \quad \text{for every } t > 0 \text{ and } v \in V,$$

which has the same structure as (2.8).

3. Main Results and Discussion.

3.1. Main Result I: Weak Convergence of ρ_{ε} and u_{ε}.

The following theorem is the first main result of this paper. It states that for every time $t \geq 0$ the measures $\rho_{\varepsilon}(t)$ that solve (2.6) weakly-$*$ converge as $\varepsilon \downarrow 0$ to a limiting measure $\rho(t)$ in $\mathscr{P}(\overline{D})$, whose density $u(t) = \frac{d\rho(t)}{d\gamma}$ is the solution of the limit system (0.2).

We state our result in a general form, which holds even for signed measures in $\mathscr{M}(\overline{D})$.

THEOREM 3.1. *Let* $\rho_{\varepsilon} = u_{\varepsilon} \gamma_{\varepsilon} \in C^0([0, +\infty); \mathscr{M}(\overline{D}))$ *be the solution of* (2.6)–(2.7c) *with initial datum* ρ_{ε}^0. *If*

$$(3.1) \qquad\qquad \sup_{\varepsilon > 0} \int_D |u_{\varepsilon}^0|^2 \, d\gamma_{\varepsilon} < +\infty$$

and ρ_{ε}^0 *weakly-$*$ converges to*

$$(3.2) \qquad \rho^0 = u^0 \gamma = \frac{1}{2} u^{0-} \lambda_{\Omega} \otimes \delta_{-1} + \frac{1}{2} u^{0+} \lambda_{\Omega} \otimes \delta_{+1} \ \text{ as } \varepsilon \downarrow 0,$$

then $u^0 \in L^2(\overline{D}; \gamma)$, $u^{0, \pm} \in L^2(\Omega)$, *and, for every* $t \geq 0$, *the solution* $\rho_{\varepsilon}(t)$ *weakly-$*$ converges to*

$$(3.3) \qquad \rho(t) = u(t)\gamma = \frac{1}{2} u^-(t) \lambda_{\Omega} \otimes \delta_{-1} + \frac{1}{2} u^+(t) \lambda_{\Omega} \otimes \delta_{+1},$$

whose densities $u^{\pm}$ *belong to* $C^0([0, +\infty); L^2(\Omega)) \cap C^1((0, +\infty); W^{1,2}(\Omega))$ *and solve the system*

$$(3.4a) \qquad\qquad \partial_t u^+ - \Delta_x u^+ = k(u^- - u^+) \qquad \text{in } \Omega \times (0, +\infty),$$

$$(3.4b) \qquad\qquad \partial_t u^- - \Delta_x u^- = k(u^+ - u^-) \qquad \text{in } \Omega \times (0, +\infty),$$

$$(3.4c) \qquad\qquad u^{\pm}(0) = u^{0, \pm} \qquad\qquad\qquad \text{in } \Omega.$$

The positive constant k *in* (3.4a)–(3.4b) *can be characterized as the asymptotic minimal transition cost*

$$(3.5) \qquad k = \frac{1}{\pi} \sqrt{|H''(0)| H''(1)}$$

$$= \lim_{\varepsilon \downarrow 0} \min \left\{ \tau_{\varepsilon} \int_{-1}^1 \left(\varphi'(\xi) \right)^2 d\tilde{\gamma}_{\varepsilon} : \varphi \in W^{1,2}(-1, 1), \ \varphi(\pm 1) = \pm \tfrac{1}{2} \right\}.$$

3.2. Main Result II: A Stronger Convergence of u_ε. Weak-$*$ convergence in the sense of measures is a natural choice in order to describe the limit of ρ_ε, since the densities u_ε and the limit density $u = (u^+, u^-)$ are defined on different domains with respect to different reference measures. Nonetheless, it is possible to consider a stronger convergence which better characterizes the limit and to prove that it is satisfied by the solutions of our problem.

This stronger notion is modeled on Hilbert spaces (or, more generally, on Banach spaces with a locally uniformly convex norm), where strong convergence is equivalent to weak convergence together with the convergence of the norms:

$$(3.6) \qquad x_n \to x \quad \Longleftrightarrow \quad x_n \rightharpoonup x \quad \text{and} \quad \|x_n\| \to \|x\|.$$

In this spirit, the next result states that under the additional request of "strong" convergence of the initial data u_ε^0, we have "strong" convergence of the densities u_ε; we refer the reader to [24, 15] (see also [2, section 5.4]) for further references in a measure-theoretic setting.

THEOREM 3.2. *Let ρ_ε, ρ_ε^0 be as in Theorem 3.1. If, moreover,*

$$(3.7) \qquad \lim_{\varepsilon \downarrow 0} b_\varepsilon(u_\varepsilon^0, u_\varepsilon^0) = b(u^0, u^0),$$

then for every $t > 0$ we have

$$(3.8) \qquad \lim_{\varepsilon \downarrow 0} b_\varepsilon(u_\varepsilon(t), u_\varepsilon(t)) = b(u(t), u(t))$$

and

$$(3.9) \qquad \lim_{\varepsilon \downarrow 0} a_\varepsilon(u_\varepsilon(t), u_\varepsilon(t)) = a(u(t), u(t)).$$

Applying, e.g., [2, Theorem 5.4.4] we can immediately deduce the following result, which clarifies the strengthened form of convergence that we are considering here. This convergence is strong enough to allow us to pass to the limit in nonlinear functions of u_ε.

COROLLARY 3.3. *Under the same assumptions as in Theorem 3.2 we have for every $t > 0$*

$$(3.10) \qquad \lim_{\varepsilon \downarrow 0} \int_D f(x, \xi, u_\varepsilon(x, \xi, t)) \, \mathrm{d}\gamma_\varepsilon(x, \xi) = \int_{\overline{D}} f(x, \xi, u(x, \xi, t)) \, \mathrm{d}\gamma(x, \xi)$$

$$= \frac{1}{2} \int_\Omega \Big(f(x, -1, u^-(x, t)) + f(x, 1, u^+(x, t)) \Big) \, \mathrm{d}\lambda_\Omega(x),$$

where $f : \overline{D} \times \mathbb{R} \to \mathbb{R}$ is an arbitrary continuous function satisfying the quadratic growth condition

$$|f(x, \xi, r)| \leq A + Br^2 \quad \text{for every } (x, \xi) \in \overline{D}, \ r \in \mathbb{R},$$

for suitable nonnegative constants $A, B \in \mathbb{R}$.

3.3. Discussion. The results of Theorems 3.1 and 3.2 are, among other things, rigorous versions of the result of Kramers [17] that was mentioned in the introduction. They show that the simple reaction-diffusion system (3.4) can indeed be viewed as an upscaled version of a diffusion problem in an augmented phase space, or, equivalently,

as an upscaled version of the movement of a Brownian particle in the same augmented phase space.

At the same time, they generalize the work of Kramers by adding the spatial dimension, resulting in a limit system which, for this choice of τ_ε (see below for more on this choice), captures both reaction and diffusion effects.

Measures versus densities. It is interesting to note the roles of the measures ρ_ε, ρ and their densities u_ε, u with respect to $\gamma_\varepsilon, \gamma$. The variational formulations of the equations are given in terms of the densities u_ε, u, but the limit procedure is better understood in terms of the measures ρ_ε, ρ, since a weak-$*$ convergence is involved. This also allows for a unification of two problems with a different structure (a Fokker–Planck equation for u_ε and a reaction-diffusion system for the couple u^-, u^+).

Gradient flows. The weak formulation (2.8) also shows that a solution u_ε can be interpreted as a gradient flow of the quadratic energy $\frac{1}{2}a_\varepsilon(u, u)$ with respect to the $L^2(D; \gamma_\varepsilon)$ distance. Another gradient-flow structure for the solutions of the same problem could be obtained by a different choice of energy functional and distance: for example, as proved in [16], Fokker–Planck equations like (2.6) can be interpreted also as the gradient flow of the relative entropy functional

$$(3.11) \qquad \mathcal{H}(\rho|\gamma_\varepsilon) := \int_D \frac{d\rho}{d\gamma_\varepsilon} \log\left(\frac{d\rho}{d\gamma_\varepsilon}\right) d\gamma_\varepsilon$$

in the space $\mathscr{P}(\overline{D})$ of probability measures endowed with the so-called L^2-Wasserstein distance (see, e.g., [2]). Initiated by the work of Otto [16, 21] and extended in many directions since, this framework provides an appealing variational structure for very general diffusion processes.

Two recent results point toward a connection between the results of this paper and Wasserstein gradient flows. In [1] it was shown how the Wasserstein setting may be the most natural for understanding diffusion as a limit of the motion of Brownian particles. In addition, Mielke uncovered a Wasserstein-type gradient-flow structure for chemical reactions [19].

Fueled by these observations, in the first publication [22] of our results we asked whether a similar convergence result could also be proved within the Wasserstein gradient-flow framework. This question was answered affirmatively in two different ways [14, 3], and we refer the reader to [3] for further discussion of these issues.

The choice of τ_ε. In this paper the time scale τ_ε is chosen to be equal to $\varepsilon \exp(1/\varepsilon)$, and it is natural to ask about the limit behavior for different choices of τ_ε. If the scaling is chosen differently, i.e., if $\tau_\varepsilon \varepsilon^{-1} \exp(-1/\varepsilon)$ converges to 0 or ∞, then completely different limit systems are obtained:

- If $\tau_\varepsilon \ll \varepsilon \exp(1/\varepsilon)$, then the reaction is not accelerated sufficiently as $\varepsilon \to 0$, and the limit system will contain only diffusion (i.e., $k = 0$ in (3.4)).
- If $\tau_\varepsilon \gg \varepsilon \exp(1/\varepsilon)$, on the other hand, then the reaction becomes faster and faster as $\varepsilon \to 0$, resulting in a limit system in which the chemical reaction $A \rightleftharpoons B$ is in continuous equilibrium. Because of this, both A and B have the same concentration u, which u solves the diffusion problem

$$\partial_t u = \Delta u \qquad\qquad\qquad \text{for } x \in \Omega,\ t > 0,$$
$$u(0, x) = \frac{1}{2}\big(u^{0,+}(x) + u^{0,-}(x)\big) \qquad \text{for } x \in \Omega.$$

Note the instantaneous equilibration of the initial data in this system.

While the scaling in terms of ε of τ_ε cannot be chosen differently without obtaining structurally different limit systems, there is still a choice in the prefactor. For $\tau_\varepsilon :=$ $\tilde{\tau}\varepsilon e^{1/\varepsilon}$ with $\tilde{\tau} > 0$ fixed, the prefactor $\tilde{\tau}$ will appear in the definition (3.5) of k.

There is also a modeling aspect to the choice of τ. In this paper we use no knowledge about the value of τ in the diffusion system at finite ε; the choice $\tau = \tau_\varepsilon$ is motivated by the wish to have a limit system that contains both diffusive and reactive terms. If one has additional information about the mobility of the system in the x- and ξ-directions, then the value of τ will follow from this.

Equal rate constants. The assumption of equal depth of the two minima of H corresponds to the assumption (or, depending on one's point of view, the result) that the rate constant k in (3.4) is the same for the two reactions $A \to B$ and $B \to A$. The general case requires a slightly different choice for H, as follows.

Let the original macroscopic equations for the evolution of the densities α, β in A and B be

$$(3.12a) \qquad\qquad \partial_t\alpha - \Delta\alpha = k^-\beta - k^+\alpha,$$

$$(3.12b) \qquad\qquad \partial_t\beta - \Delta\beta = k^+\alpha - k^-\beta.$$

Choose a fixed function $H_0 \in C^\infty([-1,1])$ such that $H_0'(\pm 1) = 0$ and $H_0(1) - H_0(-1) = \log k^- - \log k^+$. We then construct the enthalpy H_ε by setting

$$H_\varepsilon := H_0 + \frac{1}{\varepsilon}H,$$

where H is the same enthalpy function as above. The same proof as for the equal-well case then gives convergence of the finite-ε problems to (3.12).

Equal diffusion constants. It is possible to change the setup such that the limiting system has different diffusion rates in A and B. We first write (2.6) as

$$\partial_t\rho - \operatorname{div} D_\varepsilon F_\varepsilon = 0,$$

where the mobility matrix $D_\varepsilon \in \mathbb{R}^{(d+1)\times(d+1)}$ and the flux F_ε are given by

$$D_\varepsilon = \begin{pmatrix} I & 0 \\ 0 & \tau_\varepsilon \end{pmatrix} \qquad \text{and} \qquad F_\varepsilon = F_\varepsilon(\rho) = \begin{pmatrix} \nabla u \\ \nabla\rho + \rho\nabla H \end{pmatrix}.$$

By replacing the identity matrix block I in D_ε by a block of the form $a(\xi)\,I$, the x-directional diffusion can be modified as a function of ξ. This translates into two different diffusion coefficients for A and B.

The function H. The limit result of Theorem 3.1 shows that only a small amount of information about the function H propagates into the limit problem, in particular, the local second-order structure of H around the wells and around the mountain-pass point.

One other aspect of the structure of H is hidden: the fact that we rescaled the ξ variable by a factor of $\sqrt{\tau_\varepsilon}$ can also be interpreted as a property of H, since the effective distance between the two wells, as measured against the intrinsic distance associated with the Brownian motion, is equal to $2\sqrt{\tau_\varepsilon}$ after rescaling.

We also assumed in this paper that H has only "half" wells, in the sense that H is defined on $[-1,1]$ instead of $\mathbb{R}$. This was for practical convenience, and one can do essentially the same analysis for a function H that is defined on $\mathbb{R}$. In this case one will regain a slightly different value of k, namely, $k = \sqrt{|H''(0)|H''(1)}/2\pi$. (For this reason this is also the value found by Kramers [17, equation (17)].)

Single particles versus multiple particles, and concentrations versus probabilities. The description in this paper of the system in terms of a probability measure ρ on D is the description of the probability of a *single* particle. This implies that the limit object (u^-, u^+) should be interpreted as the density (with respect to γ) of a limiting probability measure, again describing a single particle.

This is at odds with common continuum modeling philosophy, where the main objects are concentrations (mass or volume) that represent a large number of particles; in this philosophy the solution (u^-, u^+) of (3.4) should be viewed as such a concentration, which is to say as the projection onto x-space of a *joint probability distribution of a large number of particles.*

For the simple reaction $A \rightleftharpoons B$ these two interpretations are actually equivalent. This arises from the fact that $A \to B$ reaction events in each of the particles are independent of each other; therefore, the joint distribution of a large number N of particles factorizes into a product of N copies of the distribution of a single particle. For the case of this paper, therefore, the distinction between these two views is not important.

More general reactions. The remark above implies that the situation will be different for systems where reaction events cause differences in distributions between the particles, such as the reaction $A + B \rightleftharpoons C$. This can be recognized as follows: a particle A that has just separated from a B particle (in a reaction event of the form $C \to A + B$) has a position that is highly correlated with the corresponding B particle, while this is not the case for all the other A particles. Therefore, the A particles will not have the same distribution. The best one can hope for is that in the limit of a large number of particles the distribution becomes the same in some weak way. This is one of the major obstacles in developing a similar connection as the one in this paper for more complex reaction equations.

Regarding possible extensions toward equations involving an arbitrary number of chemical species, as well as different reaction and diffusion rates, we point out that a formal gradient-flow structure has recently been established in [19], independently of this work.

4. Structure of the Proof: Formulation in Terms of Measures, Regularization Estimates, and Γ-Convergence of Quadratic Forms. We now explain the structure of the proof of Theorems 3.1 and 3.2 in more detail. This will also clarify the use of Γ-convergence and highlight the potential of the method for wider application.

We have already seen that the ε-problem (2.7) and the limit problem (0.2) can be formulated in the highly similar variational forms (2.8) and (2.12). The analogy between (2.8) and (2.12) suggests passing to the limit in these weak formulations, or even better, in their equivalent integrated forms

(4.1)
$$b_\varepsilon(u_\varepsilon(t), v_\varepsilon) + \int_0^t a_\varepsilon(u_\varepsilon(t), v_\varepsilon)\,\mathrm{d}t = b(u_\varepsilon^0, v_\varepsilon) \quad \text{for every } v_\varepsilon \in V_\varepsilon,$$
$$b(u(t), v) + \int_0^t a(u(t), v)\,\mathrm{d}t = b(u^0, v) \quad \text{for every } v \in V.$$

Applying standard regularization estimates for the solutions to (2.8) (see the next section) and a weak coercivity property of b_ε, it is not difficult to prove that $u_\varepsilon(t)$ "weakly" converges to $u(t)$ for every $t > 0$, i.e.,

$$\rho_\varepsilon(t) = u_\varepsilon(t)\gamma_\varepsilon \overset{*}{\rightharpoonup} \rho(t) = u(t)\gamma \quad \text{weakly-}\ast \text{ in } \mathscr{M}(\overline{D}).$$

The concept of weak convergence of densities that we are using here is thus the same as in Theorem 3.1, i.e., weak-∗ convergence of the corresponding measures in $\mathscr{M}(\overline{D})$.

In order to pass to the limit in (4.1), the central property is the following *weak-strong* convergence principle:

$$\text{(4.2)} \qquad \begin{aligned} &\text{For every } v \in V \text{ there exists } v_\varepsilon \in V_\varepsilon \text{ with } v_\varepsilon \rightharpoonup v \text{ as } \varepsilon \to 0 \text{ such that} \\ &b_\varepsilon(u_\varepsilon, v_\varepsilon) \to b(u, v) \qquad \text{and} \qquad a_\varepsilon(u_\varepsilon, v_\varepsilon) \to a(u, v) \quad \text{for every } u_\varepsilon \rightharpoonup u. \end{aligned}$$

Note that the previous property implies, in particular, that the recovery family v_ε converges "strongly" to v, according to the notion considered by Theorem 3.2, i.e., $v_\varepsilon \to v$ iff $v_\varepsilon \rightharpoonup v$ with both $b_\varepsilon(v_\varepsilon, v_\varepsilon) \to b(v, v)$ and $a_\varepsilon(v_\varepsilon, v_\varepsilon) \to a(v, v)$. Lemma 4.2 shows that this weak-strong convergence property can be derived from Γ-convergence in the "weak" topology of the family of quadratic forms,

$$\text{(4.3)} \quad q_\varepsilon^\kappa(u) := b_\varepsilon(u, u) + \kappa\, a_\varepsilon(u, u) \quad \text{to} \quad q^\kappa(u) := b(u, u) + \kappa\, a(u, u) \quad \text{for } \kappa > 0.$$

In order to formulate this property in the standard framework of Γ-convergence, we will extend a_ε and b_ε to lower-semicontinuous quadratic functionals (possibly assuming the value $+\infty$) in the space $\mathscr{M}(\overline{D})$, following the approach of [10, Chapters 11–13].

4.1. The Kramers–Smoluchowski Equation. Let us first briefly summarize the functional framework introduced above. We denote by $(\cdot, \cdot)_\varepsilon$ the scalar product in $\mathbb{R}^d \times \mathbb{R}$ defined by

$$\text{(4.4)} \qquad (\boldsymbol{x}, \boldsymbol{y})_\varepsilon := x \cdot y + \tau_\varepsilon\, \xi\, \eta \quad \text{for every } \boldsymbol{x} = (x, \xi),\ \boldsymbol{y} = (y, \eta) \in \mathbb{R}^d \times \mathbb{R},$$

with the corresponding norm $\|\cdot\|_\varepsilon$. We introduced two Hilbert spaces,

$$H_\varepsilon := L^2(D, \gamma_\varepsilon) \qquad \text{and} \qquad V_\varepsilon = W^{1,2}(D, \gamma_\varepsilon),$$

and the bilinear forms

$$\text{(4.5)} \qquad b_\varepsilon(u, v) := \int_D u\, v\, \mathrm{d}\gamma_\varepsilon \qquad\qquad \text{for every } u, v \in H_\varepsilon,$$

$$\text{(4.6)} \qquad a_\varepsilon(u, v) := \int_D (\nabla_{x,\xi} u, \nabla_{x,\xi} v)_\varepsilon\, \mathrm{d}\gamma_\varepsilon \qquad\qquad \text{for every } u, v \in V_\varepsilon,$$

with which (2.7a) has the variational formulation

$$\text{(4.7)} \qquad b_\varepsilon(\partial_t u_\varepsilon, v) + a_\varepsilon(u_\varepsilon, v) = 0 \quad \text{for every } v \in V_\varepsilon,\ t > 0; \quad u_\varepsilon(0, \cdot) = u_\varepsilon^0.$$

The main technical difficulty in studying the limit behavior of (4.7) as $\varepsilon \downarrow 0$ consists of the ε-dependence of the functional spaces $H_\varepsilon, V_\varepsilon$. Since for our approach it is crucial to work in a fixed ambient space, we embed the solutions of (4.7) in the space of finite Borel measures $\mathscr{M}(\overline{D})$ by associating to u_ε the measure $\rho_\varepsilon := u_\varepsilon \gamma_\varepsilon$. We thus introduce the quadratic forms

$$\text{(4.8)} \qquad \mathfrak{b}_\varepsilon(\rho) := b_\varepsilon(u, u) \qquad\qquad \text{if } \rho \ll \gamma_\varepsilon \text{ and } u = \frac{\mathrm{d}\rho}{\mathrm{d}\gamma_\varepsilon} \in H_\varepsilon,$$

$$\text{(4.9)} \qquad \mathfrak{a}_\varepsilon(\rho) := a_\varepsilon(u, u) \qquad\qquad \text{if } \rho \ll \gamma_\varepsilon \text{ and } u = \frac{\mathrm{d}\rho}{\mathrm{d}\gamma_\varepsilon} \in V_\varepsilon,$$

trivially extended to $+\infty$ when ρ is not absolutely continuous with respect to γ_ε or its density u does not belong to H_ε or V_ε, respectively. Denoting by $\mathsf{Dom}(\mathfrak{a}_\varepsilon)$ and

Dom$(\mathfrak{b}_\varepsilon)$ their proper domains, we still denote by $\mathfrak{a}_\varepsilon(\cdot, \cdot)$ and $\mathfrak{b}_\varepsilon(\cdot, \cdot)$ the corresponding bilinear forms defined on Dom$(\mathfrak{a}_\varepsilon)$ and Dom$(\mathfrak{b}_\varepsilon)$, respectively. Setting $\rho_\varepsilon := u_\varepsilon \gamma_\varepsilon$ and $\sigma := v \gamma_\varepsilon$, (4.7) is equivalent to the integrated form

$$(4.10) \qquad \mathfrak{b}_\varepsilon(\rho_\varepsilon(t), \sigma) + \int_0^t \mathfrak{a}_\varepsilon(\rho_\varepsilon(r), \sigma) \, dr = \mathfrak{b}_\varepsilon(\rho_\varepsilon^0, \sigma) \quad \text{for every } \sigma \in \text{Dom}(\mathfrak{a}_\varepsilon).$$

4.2. Regularization Estimates. We recall here the standard regularization estimates satisfied by solutions to (4.10):

$$(4.11) \qquad \frac{1}{2} \mathfrak{b}_\varepsilon(\rho_\varepsilon(t)) + \int_0^t \mathfrak{a}_\varepsilon(\rho_\varepsilon(r)) \, dr = \frac{1}{2} \mathfrak{b}_\varepsilon(\rho_\varepsilon^0) \qquad \text{for every } t \geq 0,$$

$$(4.12) \qquad t \, \mathfrak{a}_\varepsilon(\rho_\varepsilon(t)) + 2 \int_0^t r \mathfrak{b}_\varepsilon(\partial_t \rho_\varepsilon(r)) \, dr = \int_0^t \mathfrak{a}_\varepsilon(\rho_\varepsilon(r)) \, dr \quad \text{for every } t \geq 0,$$

$$(4.13) \qquad \frac{1}{2} \mathfrak{b}_\varepsilon(\rho_\varepsilon(t)) + t \, \mathfrak{a}_\varepsilon(\rho_\varepsilon(t)) + t^2 \mathfrak{b}_\varepsilon(\partial_t \rho_\varepsilon(t)) \leq \frac{1}{2} \mathfrak{b}_\varepsilon(\rho_\varepsilon^0) \qquad \text{for every } t > 0.$$

Although versions of these expressions appear in various places, for the ease of the reader we briefly describe their proof, and we use the more conventional formulation in terms of the bilinear forms a_ε and b_ε, the density u_ε, and spaces H_ε and V_ε; note that b_ε is an inner product for H_ε, and $b_\varepsilon + a_\varepsilon$ is an inner product for V_ε.

When u_0 is sufficiently smooth, standard results (e.g., [8, Chapter VII]) provide the existence of a solution $u_\varepsilon \in C([0, \infty); V_\varepsilon) \cap C^\infty((0, \infty); V_\varepsilon)$, such that the functions $t \mapsto a_\varepsilon(u_\varepsilon(t))$ and $t \mapsto b_\varepsilon(\partial_t u_\varepsilon(t))$ are nonincreasing; in addition, the solution operator (semigroup) S_t is a contraction in H_ε. For this case (4.11) follows from (4.7) by choosing $v := u_\varepsilon$ and integrating in time. Estimate (4.12) can be obtained by choosing $v := \partial_t u_\varepsilon$ and multiplying (4.7) by $2t$: the identity

$$2t \, a_\varepsilon(u_\varepsilon, \partial_t u_\varepsilon) = \partial_t \Big(t \, a_\varepsilon(u_\varepsilon, u_\varepsilon) \Big) - a(u_\varepsilon, u_\varepsilon)$$

and a further integration in time yield (4.12).

Finally, (4.13) follows by writing the sum of (4.11) and (4.12) as

$$(4.14) \qquad \frac{1}{2} b_\varepsilon(u_\varepsilon(t)) + t \, a_\varepsilon(u_\varepsilon(t)) + 2 \int_0^t r \, b_\varepsilon(\partial_t u_\varepsilon(r)) \, dr = \frac{1}{2} b_\varepsilon(u_\varepsilon^0)$$

and recalling that $r \mapsto \mathfrak{b}_\varepsilon(\partial_t u_\varepsilon(r))$ is nonincreasing.

In order to extend them to all $u_\varepsilon^0 \in H_\varepsilon$, we note that for fixed $t > 0$ the two norms on H_ε given by (the square roots of)

$$(4.15) \qquad u_\varepsilon^0 \mapsto \frac{1}{2} b_\varepsilon(u_\varepsilon^0) \qquad \text{and} \qquad u_\varepsilon^0 \mapsto \frac{1}{2} b_\varepsilon(S_t u_\varepsilon^0) + \int_0^t a_\varepsilon(S_r u_\varepsilon^0) \, dr$$

are identical by (4.11) on an H_ε-dense subset. If we approximate a general $u_\varepsilon^0 \in H_\varepsilon$ by smooth $u_{\varepsilon,n}^0$, then the sequence $u_{\varepsilon,n}^0$ is a Cauchy sequence with respect to both norms; by copying the proof of completeness of the space $L^2(0, \infty; V_\varepsilon)$ (see, e.g., [8, Theorem IV.8]) it follows that the integral in (4.15) converges. This allows us to pass to the limit in (4.11). The argument is similar for (4.14), which yields (4.13).

4.3. The Reaction-Diffusion Limit. We now adopt the same point of view to formulate the limit reaction-diffusion system in the setting of measures. Recall that,

owing to the special form (2.3) of γ, $\rho \ll \gamma$ implies $\rho = 1/2(u\lambda_\Omega \otimes \delta_1 + u\lambda_\Omega \otimes \delta_{-1})$, that for $u \in H := L^2(\overline{D}, \gamma)$ we set $u^\pm(x) := u(x, \pm 1)$, and that we have defined the function space

$$V := \left\{ u \in H : u^\pm \in W^{1,2}(\Omega) \right\}$$

and the bilinear forms

$$(4.16) \quad b(u, v) = \frac{1}{2} \int_\Omega \left(u^+ v^+ + u^- v^- \right) \mathrm{d}\lambda_\Omega,$$

$$(4.17) \quad a(u, v) := \frac{1}{2} \int_\Omega \left(\nabla_x u^+ \nabla_x v^+ + \nabla_x u^- \nabla_x v^- + k \left(u^+ - u^- \right)(v^+ - v^-) \right) \mathrm{d}\lambda_\Omega.$$

As before, we now extend these definitions to arbitrary measures by

$$(4.18) \qquad \mathfrak{b}(\rho) := b(u, u) \qquad \text{if } \rho \ll \gamma \text{ and } u = \frac{\mathrm{d}\rho}{\mathrm{d}\gamma} \in H,$$

$$(4.19) \qquad \mathfrak{a}(\rho) := a(u, u) \qquad \text{if } \rho \ll \gamma \text{ and } u = \frac{\mathrm{d}\rho}{\mathrm{d}\gamma} \in V,$$

with corresponding bilinear forms $\mathfrak{b}(\cdot, \cdot)$ and $\mathfrak{a}(\cdot, \cdot)$; problem (3.4) can be reformulated as

$$\mathfrak{b}(\partial_t \rho(t), \sigma) + \mathfrak{a}(\rho(t), \sigma) = 0 \quad \text{for every } t > 0 \text{ and } \sigma \in \mathsf{Dom}(\mathfrak{a}),$$

or in the integral form

$$(4.20) \qquad \mathfrak{b}(\rho(t), \sigma) + \int_0^t \mathfrak{a}(\rho(r), \sigma) \, \mathrm{d}r = \mathfrak{b}(\rho^0, \sigma) \quad \text{for every } \sigma \in \mathsf{Dom}(\mathfrak{a}).$$

Since both problems (4.10) and (4.20) are embedded in the same measure space $\mathscr{M}(\overline{D})$, we can study the convergence of the solution ρ_ε of (4.10) as $\varepsilon \downarrow 0$.

4.4. Γ-Convergence and the Weak-Strong Convergence Principle (4.2) for Quadratic Forms. We have already mentioned that (4.2) is a crucial tool for proving the convergence of the variational evolution problems (4.10). In this section we show that this property is strongly related to the Γ-convergence of the quadratic forms $\mathfrak{a}_\varepsilon$ and $\mathfrak{b}_\varepsilon$.

First, let us briefly recall the definition of Γ-convergence for a family of equicoercive functionals $(\mathfrak{f}_\varepsilon)_{\varepsilon>0}$ in $\mathscr{M}(\overline{D})$. We can refer here to the sequential definition of Γ-convergence [10, Proposition 8.1(e,f)] since $\mathscr{M}(\overline{D})$ is the dual of a separable space (see the argument of [10, Proposition 8.10]).

DEFINITION 4.1. *Let $\mathfrak{f}_\varepsilon, \mathfrak{f} : \mathscr{M}(\overline{D}) \to (-\infty, +\infty]$ be given functionals, satisfying the equicoercivity assumption*

$$(4.21) \qquad \sup_\varepsilon \mathfrak{f}_\varepsilon(\rho_\varepsilon) < +\infty \quad \Rightarrow \quad \rho_\varepsilon \text{ is bounded in } \mathscr{M}(\overline{D}).$$

We say that $\mathfrak{f}_\varepsilon$ $\Gamma\big(\mathscr{M}(\overline{D})\big)$-converges to $\mathfrak{f}$ as $\varepsilon \downarrow 0$ and we write $\Gamma\big(\mathscr{M}(\overline{D})\big)\text{-}\lim_{\varepsilon\downarrow 0} \mathfrak{f}_\varepsilon = \mathfrak{f}$ if the following two conditions are satisfied:
    (Γ-1) If $\rho_\varepsilon \overset{*}{\rightharpoonup} \rho$ in $\mathscr{M}(\overline{D})$, then

$$(4.22) \qquad \liminf_{\varepsilon\downarrow 0} \mathfrak{f}_\varepsilon(\rho_\varepsilon) \geq \mathfrak{f}(\rho).$$

(Γ-2) *For every $\sigma \in \mathscr{M}(\overline{D})$ with $\mathfrak{f}(\sigma) < \infty$ there exists $\sigma_\varepsilon \in \mathscr{M}(\overline{D})$ converging to σ as $\varepsilon \downarrow 0$ such that*

$$(4.23) \qquad \lim_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\sigma_\varepsilon) = \mathfrak{f}(\sigma).$$

Here we are mainly concerned with nonnegative quadratic forms: this means that $\mathsf{Dom}(\mathfrak{f}_\varepsilon)$ is a linear space and $\mathfrak{f}_\varepsilon$ satisfy the parallelogram identity, so that $\mathfrak{f}_\varepsilon(\rho) = \mathfrak{f}_\varepsilon(\rho, \rho)$ for a (symmetric) bilinear form defined on $\mathsf{Dom}(\mathfrak{f}_\varepsilon)$ by

$$(4.24) \qquad 2\,\mathfrak{f}_\varepsilon(\rho, \sigma) := \mathfrak{f}_\varepsilon(\rho + \sigma) - \mathfrak{f}_\varepsilon(\rho) - \mathfrak{f}_\varepsilon(\sigma) \quad \text{for every } \rho, \sigma \in \mathsf{Dom}(\mathfrak{f}_\varepsilon).$$

It is worth noting that the Γ-limit of nonnegative quadratic forms is still a (weakly-$*$ lower-semicontinuous) nonnegative quadratic form [10, Theorem 11.10]. In this case one of the most useful consequences of $\Gamma\big(\mathscr{M}(\overline{D})\big)$-convergence is contained in the next result (see, e.g., [23, Lemma 3.6]), which shows the link between Γ-convergence and (4.2).

LEMMA 4.2 (weak-strong convergence). *Assume that $\mathfrak{f}_\varepsilon$ is a family of equi-coercive nonnegative quadratic forms $\Gamma\big(\mathscr{M}(\overline{D})\big)$-converging to $\mathfrak{f}$ as $\varepsilon \downarrow 0$. Let $\rho_\varepsilon, \sigma_\varepsilon \in \mathscr{M}(\overline{D})$ be two families weakly-$*$ converging to ρ, σ as $\varepsilon \downarrow 0$ and satisfying the uniform bound*

$$(4.25) \qquad \limsup_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\rho_\varepsilon) < +\infty, \qquad \limsup_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\sigma_\varepsilon) < +\infty,$$

so that ρ, σ belong to the domain of the limit quadratic form $\mathfrak{f}$. We have

$$(4.26) \qquad \lim_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\sigma_\varepsilon) = \mathfrak{f}(\sigma) \qquad \implies \qquad \lim_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\rho_\varepsilon, \sigma_\varepsilon) = \mathfrak{f}(\rho, \sigma).$$

Proof. We reproduce here the proof of [23]. For every positive scalar $r > 0$ we have by (4.24) and the fact that $\mathfrak{f}_\varepsilon(\cdot, \cdot)$ is bilinear that

$$2\mathfrak{f}_\varepsilon(\rho_\varepsilon, \sigma_\varepsilon) = 2\mathfrak{f}_\varepsilon(r\,\rho_\varepsilon, r^{-1}\sigma_\varepsilon) = \mathfrak{f}_\varepsilon(r\rho_\varepsilon + r^{-1}\sigma_\varepsilon) - r^2\mathfrak{f}_\varepsilon(\rho_\varepsilon) - r^{-2}\mathfrak{f}_\varepsilon(\sigma_\varepsilon).$$

Taking the inferior limit as $\varepsilon \downarrow 0$ and recalling (4.25), we get for $A := \limsup_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\rho_\varepsilon)$

$$\liminf_{\varepsilon \downarrow 0} 2\mathfrak{f}_\varepsilon(\rho_\varepsilon, \sigma_\varepsilon) \geq \mathfrak{f}(r\rho + r^{-1}\sigma) - r^2 A - r^{-2}\mathfrak{f}(\sigma) = 2\mathfrak{f}(\rho, \sigma) + r^2\big(\mathfrak{f}(\rho) - A\big).$$

Since $r > 0$ is arbitrary and A is finite by (4.25), we obtain $\liminf_{\varepsilon \downarrow 0} \mathfrak{f}_\varepsilon(\rho_\varepsilon, \sigma_\varepsilon) \geq \mathfrak{f}(\rho, \sigma)$; inverting the sign of σ we get (4.26). □

If one wants to apply the previous lemma for passing to the limit in the integral formulation (4.10) of the Kramers-Smoluchowski equations to get (4.20), it is not hard to guess that for every testing measure $\sigma \in \mathsf{Dom}(\mathfrak{a})$ one needs to find a "joint" recovery family σ_ε converging to σ in $\mathscr{M}(\overline{D})$ such that

$$(4.27) \qquad \lim_{\varepsilon \downarrow 0} \mathfrak{a}_\varepsilon(\sigma_\varepsilon) = \mathfrak{a}(\sigma), \quad \lim_{\varepsilon \downarrow 0} \mathfrak{b}_\varepsilon(\sigma_\varepsilon) = \mathfrak{b}(\rho).$$

This property is not guaranteed by the separate Γ-convergence of $\mathfrak{a}_\varepsilon$ and $\mathfrak{b}_\varepsilon$ to $\mathfrak{a}$ and $\mathfrak{b}$, respectively, since the property (Γ-2) would provide two a priori different families, say, $\sigma_\varepsilon^{\mathfrak{a}}, \sigma_\varepsilon^{\mathfrak{b}}$ satisfying (4.27) for $\mathfrak{a}$ and $\mathfrak{b}$.

The next result shows that one can overcome the above difficulty by studying the Γ-convergence of the quadratic forms $\mathfrak{q}_\varepsilon^\kappa(\rho) := \mathfrak{b}_\varepsilon(\rho) + \kappa\,\mathfrak{a}_\varepsilon(\rho)$ depending on the parameter $\kappa > 0$.

LEMMA 4.3. *Let $\mathfrak{a}_\varepsilon$, $\mathfrak{b}_\varepsilon$ be two families of nonnegative quadratic forms in $\mathscr{M}(\overline{D})$ such that $\mathfrak{b}_\varepsilon$ is equicoercive. If, for every $\kappa > 0$,*

$$(4.28) \quad \mathfrak{q}_\varepsilon^\kappa(\rho) := \mathfrak{b}_\varepsilon(\rho) + \kappa\,\mathfrak{a}_\varepsilon(\rho) \quad \Gamma\big(\mathscr{M}(\overline{D})\big)\text{-converges to} \quad \mathfrak{q}^\kappa(\rho) := \mathfrak{b}(\rho) + \kappa\,\mathfrak{a}(\rho)$$

as $\varepsilon \downarrow 0$, then the following two properties hold:
   (Γ-1′) If

$$(4.29) \quad \rho_\varepsilon \overset{*}{\rightharpoonup} \rho \; as \; \varepsilon \downarrow 0 \; in \; \mathscr{M}(\overline{D}) \quad and \quad \limsup_{\varepsilon\downarrow 0}\Big(\mathfrak{a}_\varepsilon(\rho_\varepsilon) + \mathfrak{b}_\varepsilon(\rho_\varepsilon)\Big) = C < +\infty,$$

then

$$(4.30) \qquad\qquad \liminf_{\varepsilon\downarrow 0}\mathfrak{a}_\varepsilon(\rho_\varepsilon) \geq \mathfrak{a}(\rho), \quad \liminf_{\varepsilon\downarrow 0}\mathfrak{b}_\varepsilon(\rho_\varepsilon) \geq \mathfrak{b}(\rho).$$

   (Γ-2′) For every $\sigma \in \mathscr{M}(\overline{D})$ such that $\mathfrak{a}(\sigma) + \mathfrak{b}(\sigma) < +\infty$ there exists a family $\sigma_\varepsilon \in \mathscr{M}(\overline{D})$ weakly-$*$ converging to σ such that

$$(4.31) \qquad\qquad \lim_{\varepsilon\downarrow 0}\mathfrak{a}_\varepsilon(\sigma_\varepsilon) = \mathfrak{a}(\sigma), \quad \lim_{\varepsilon\downarrow 0}\mathfrak{b}_\varepsilon(\sigma_\varepsilon) = \mathfrak{b}(\sigma).$$

Conversely, if $\mathfrak{a}_\varepsilon, \mathfrak{b}_\varepsilon$ satisfy (Γ-1′) and (Γ-2′), then (4.28) holds.
   Proof. Let us assume (4.29); the Γ-liminf inequality for $\mathfrak{q}_\varepsilon^\kappa$ yields

$$\liminf_{\varepsilon\downarrow 0}\mathfrak{b}_\varepsilon(\rho_\varepsilon) \geq \liminf_{\varepsilon\downarrow 0}\mathfrak{q}_\varepsilon^\kappa(\rho_\varepsilon) - C\kappa \geq \mathfrak{q}^\kappa(\rho) - C\kappa = \mathfrak{b}(\rho) + \kappa\big(\mathfrak{a}(\rho) - C\big)$$

for every $\kappa > 0$, and therefore the second inequality of (4.30) follows by letting $\kappa \downarrow 0$. A similar argument yields the first inequality of (4.30).
   Concerning (4.31), Γ-convergence of $\mathfrak{q}_\varepsilon^1$ to $\mathfrak{q}^1$ yields a recovery family $\sigma_\varepsilon \overset{*}{\rightharpoonup} \sigma$ such that

$$\lim_{\varepsilon\downarrow 0}\mathfrak{a}_\varepsilon(\sigma_\varepsilon) + \mathfrak{b}_\varepsilon(\sigma_\varepsilon) = \mathfrak{a}(\sigma) + \mathfrak{b}(\sigma) < +\infty.$$

In particular, $\mathfrak{a}_\varepsilon(\sigma_\varepsilon) + \mathfrak{b}_\varepsilon(\sigma_\varepsilon)$ is uniformly bounded, so that (4.30) yields the separate convergence (4.31). $\quad\square$

4.5. Γ-Convergence and Evolution Problems. Since the pioneering papers [25, 26], the link between Γ-convergence and stability of evolution problems of parabolic type has been well known when $b_\varepsilon = b$ is a fixed and coercive bilinear form and can therefore be considered as the scalar product of the Hilbert space $H_\varepsilon \equiv H$: a general result, for Γ-converging families of *convex* functionals, can be found, e.g., in [4, Chapter 3.9.2]. In this case the embedding of the problems in a bigger topological vector space (the role played by $\mathscr{M}(\overline{D})$ in our situation) is no longer needed, and one can deal with the weak and strong topologies of H, obtaining the following equivalent characterizations (see, e.g., [7, Theorem 3.16] and [10, Theorem 13.6]):
   1. Pointwise (strong) convergence in H of the solutions of the evolution problems.
   2. Pointwise convergence in H of the resolvents of the linear operators associated to the bilinear forms a_ε.
   3. Mosco-convergence in H of the quadratic forms associated to a_ε (i.e., Γ-convergence with respect to both the weak and the strong topologies of H; see [20] and [4, section 3.3] for the precise definition).

    4. Γ-convergence in the weak topology of H of the quadratic forms $b + \kappa\, a_\varepsilon$ to $b + \kappa\, a$ for every $\kappa > 0$.

In the present case, where b_ε does depend on ε, Γ-convergence of the extended quadratic forms $\mathfrak{q}_\varepsilon^\kappa = \mathfrak{b}_\varepsilon + \kappa\, \mathfrak{a}_\varepsilon$ with respect to the weak-$*$ topology of $\mathscr{M}(\overline{D})$ is thus a natural extension of the latter condition; we will present in section 6 a simple and general argument showing how to derive the convergence of the evolution problems by the Γ-convergence of $\mathfrak{q}_\varepsilon^\kappa$, thus justifying the scheme of Figure 2.

    Theorem 5.1 in the next section provides the crucial information, i.e., the Γ-convergence of $\mathfrak{q}_\varepsilon^\kappa$ to $\mathfrak{q}^\kappa$.

5. Γ-Convergence Result for the Quadratic Forms $\mathfrak{q}_\varepsilon^\kappa$, $\mathfrak{a}_\varepsilon$, $\mathfrak{b}_\varepsilon$. The aim of this section is to prove the following Γ-convergence result involving the quadratic forms $\mathfrak{a}_\varepsilon, \mathfrak{b}_\varepsilon, \mathfrak{a}, \mathfrak{b}$ defined in $(4.8), (4.9)$ and $(4.18), (4.19)$.

    THEOREM 5.1. *The family of quadratic forms $\mathfrak{b}_\varepsilon$ is equicoercive, according to (4.21), and for every $\kappa > 0$ we have*

$$(5.1) \qquad \mathfrak{q}_\varepsilon^\kappa(\rho) := \mathfrak{b}_\varepsilon(\rho) + \kappa\, \mathfrak{a}_\varepsilon(\rho) \quad \Gamma\big(\mathscr{M}(\overline{D})\big)\text{-converges to} \quad \mathfrak{q}^\kappa(\rho) := \mathfrak{b}(\rho) + \kappa\, \mathfrak{a}(\rho)$$

as $\varepsilon \downarrow 0$, i.e., properties $(\Gamma\text{-}1')$ and $(\Gamma\text{-}2')$ of Lemma 4.3 hold.

    We split the proof of Theorem 5.1 into various steps, focusing directly on the properties $(\Gamma\text{-}1')$ and $(\Gamma\text{-}2')$, which are in fact equivalent to (5.1).

    While the Γ-convergence of $\mathfrak{b}_\varepsilon$ is a direct consequence of the weak convergence of γ_ε to γ, the convergence of $\mathfrak{a}_\varepsilon$ is more subtle, since the convergence of $\mathfrak{a}_\varepsilon$ and the structure of the limit depend critically on the choice of τ_ε (defined in (2.5)): as we will show in section 5.3, the scaling of τ_ε in terms of ε is chosen exactly such that the strength of the "connection" between $\xi = -1$ and $\xi = 1$ is of order $O(1)$ as $\varepsilon \to 0$.

5.1. Equicoercivity. Let us first prove that the quadratic forms $\mathfrak{b}_\varepsilon$ satisfy the equicoercivity condition (4.21).

    LEMMA 5.2 (equicoercivity of $\mathfrak{b}_\varepsilon$). *Every family of measures $\rho_\varepsilon \in \mathscr{M}(\overline{D})$, $\varepsilon > 0$, satisfying*

$$(5.2) \qquad\qquad\qquad \limsup_{\varepsilon > 0} \mathfrak{b}_\varepsilon(\rho_\varepsilon) < +\infty$$

is bounded in $\mathscr{M}(\overline{D})$ and admits a weakly-$*$ converging subsequence.

    Proof. The proof follows immediately by the fact that γ_ε is a probability measure, and therefore

$$|\rho_\varepsilon|(\overline{D}) \leq \Big(\mathfrak{b}_\varepsilon(\rho_\varepsilon)\Big)^{1/2}.$$

Inequality (5.2) thus implies that the total mass of ρ_ε is uniformly bounded, and we can apply the relative weak-$*$ compactness of bounded sets in dual Banach spaces.   ◻

5.2. Estimates near $\Omega \times \{-1, 1\}$.

    LEMMA 5.3. *If $\rho_\varepsilon = u_\varepsilon \gamma_\varepsilon$ satisfies the uniform bound $\mathfrak{a}_\varepsilon(\rho_\varepsilon) \leq C < +\infty$ for every $\varepsilon > 0$, then for every $\delta \in (0,1)$*

$$(5.3) \qquad\qquad\qquad \partial_\xi u_\varepsilon \to 0 \quad \text{in } L^2(\Omega \times \omega_\delta) \quad \text{as } \varepsilon \to 0,$$

where $\omega_\delta := (-1, -\delta) \cup (\delta, 1)$.

Proof. We observe that

$$\tau_\varepsilon \int_D (\partial_\xi u_\varepsilon)^2 \, \mathrm{d}\gamma_\varepsilon \le \mathfrak{a}_\varepsilon(\rho_\varepsilon) \le C < \infty.$$

If $h_\delta = \sup_{\xi \in \omega_\delta} H(\xi) < 1$, then $\inf_{\xi \in \omega_\delta} e^{-H(\xi)/\varepsilon} = e^{-h_\delta/\varepsilon}$, and we find

$$\int_{\Omega \times \omega_\delta} (\partial_\xi u_\varepsilon)^2 \, \mathrm{d}x \, \mathrm{d}\xi \le C \frac{Z_\varepsilon}{\tau_\varepsilon} e^{\frac{h_\delta}{\varepsilon}} = C \frac{Z_\varepsilon}{\varepsilon} e^{\frac{h_\delta - 1}{\varepsilon}}.$$

Taking the limit as $\varepsilon \to 0$, we obtain (5.3). $\quad\Box$

LEMMA 5.4 (convergence of traces). *Let us suppose that $\rho_\varepsilon = u_\varepsilon \gamma_\varepsilon \overset{*}{\rightharpoonup} \rho = u\gamma$ with $\mathfrak{a}_\varepsilon(\rho_\varepsilon) \le C < +\infty$, and let $u_\varepsilon^\pm(x)$ be the traces of u_ε at $\xi = \pm 1$. Then, as $\varepsilon \downarrow 0$,*

(5.4) $$u_\varepsilon^\pm \to u^\pm \quad \text{strongly in } L^2(\Omega),$$

where $u^\pm$ are the functions given by (3.3).

Proof. Let us consider, e.g., the case of u_ε^-. Let us fix $\delta \in (0,1)$; by (5.3) and standard trace results in $W^{1,2}(-1, -1+\delta)$ we know that

(5.5) $$\lim_{\varepsilon \downarrow 0} \int_\Omega \omega_\varepsilon^2(x) \, d\mathcal{L}^d = 0,$$

where

$$\omega_\varepsilon^2(x) := \sup_{-1 \le \xi \le -1+\delta} |u_\varepsilon(x,\xi) - u_\varepsilon^-(x)|^2 \le \delta \int_{-1}^{-1+\delta} |\partial_\xi u_\varepsilon(x,\xi)|^2 \mathrm{d}\xi.$$

Let us fix a function $\phi \in C^0(\Omega)$ and a function $\psi \in C^0[-1,1]$ with $0 \le \psi \le 1$, $\psi(-1) = 1$, $\mathrm{supp}\,\psi \subset [-1, -1+\delta]$; we set

$$J_\varepsilon := \int_{-1}^1 \psi(\xi) \, \mathrm{d}\tilde{\gamma}_\varepsilon(\xi), \quad \tilde{u}_\varepsilon(x) := J_\varepsilon^{-1} \int_{-1}^1 u_\varepsilon(x,\xi)\psi(\xi) \, d\tilde{\gamma}_\varepsilon(\xi),$$

where $\tilde{\gamma}_\varepsilon$ is the measure defined in (2.2). Note that

$$\lim_{\varepsilon \to 0} J_\varepsilon = \langle \psi, \gamma \rangle = \frac{1}{2}\psi(-1) + \frac{1}{2}\psi(1) = \frac{1}{2}.$$

Since ρ_ε weakly-$*$ converge to ρ, we know that

$$\lim_{\varepsilon \downarrow 0} \int_\Omega \phi(x)\tilde{u}_\varepsilon(x) \, \mathrm{d}\lambda_\Omega = \lim_{\varepsilon \downarrow 0} J_\varepsilon^{-1} \int_\Omega \phi(x)\psi(\xi)u_\varepsilon(x,\xi) \, \mathrm{d}\gamma_\varepsilon(x,\xi) = \int_\Omega \phi(x)u^-(x) \, \mathrm{d}\lambda_\Omega,$$

so that $\tilde{u}_\varepsilon$ converges to u^- in the duality with bounded continuous functions. On the other hand,

$$\int_\Omega |\nabla_x \tilde{u}_\varepsilon(x)|^2 \, \mathrm{d}\lambda_\Omega \le J_\varepsilon^{-1} \int_\Omega \int_{-1}^1 |\nabla_x u_\varepsilon(x,\xi)|^2 \psi(\xi) \, \mathrm{d}\tilde{\gamma}(\xi) \, \mathrm{d}\lambda_\Omega(x) \le J_\varepsilon^{-1} \mathfrak{a}_\varepsilon(\rho_\varepsilon) \le 2C,$$

so that $\tilde{u}_\varepsilon \to u^-$ in $L^2(\Omega)$ by the Rellich compactness theorem.

On the other hand, thanks to (5.5), we have

$$\lim_{\varepsilon \downarrow 0} \int_{\Omega} \left| u_{\varepsilon}^{-}(x) - \tilde{u}_{\varepsilon}(x) \right|^2 \mathrm{d}\lambda_{\Omega}(x)$$

$$= \lim_{\varepsilon \downarrow 0} J_{\varepsilon}^{-2} \int_{\Omega} \left| \int_{-1}^{1} \psi(\xi) \left(u_{\varepsilon}(x,\xi) - u^{-}(x) \right) \mathrm{d}\tilde{\gamma}_{\varepsilon}(\xi) \right|^2 \mathrm{d}\lambda_{\Omega}(x)$$

$$\leq \lim_{\varepsilon \downarrow 0} \int_{D} \psi(\xi) \omega_{\varepsilon}^2(x) \, \mathrm{d}\gamma_{\varepsilon}(x,\xi) = 0,$$

which yields (5.4). $\square$

Remark. A completely analogous argument shows that if ρ_{ε} satisfies a $W^{1,1}(D;\gamma_{\varepsilon})$-uniform bound

$$(5.6) \qquad \int_{D} \|\nabla_{x,\xi} u_{\varepsilon}\|_{\varepsilon} \, \mathrm{d}\gamma_{\varepsilon}(x,\xi) \leq C < +\infty$$

instead of $\mathfrak{a}_{\varepsilon}(\rho_{\varepsilon}) \leq C$, then $u_{\varepsilon}^{\pm} \to u^{\pm}$ in $L^1(\Omega)$.

5.3. Asymptotics for the Minimal Transition Cost. Given $(\varphi^-, \varphi^+) \in \mathbb{R}^2$, let us set

$$(5.7) \quad K_{\varepsilon}(\varphi^-, \varphi^+) := \min\left\{ \tau_{\varepsilon} \int_{-1}^{1} \left(\varphi'(\xi) \right)^2 \mathrm{d}\tilde{\gamma}_{\varepsilon} : \varphi \in W^{1,2}(-1,1), \ \varphi(\pm 1) = \varphi^{\pm} \right\}.$$

It is immediate to check that K_{ε} is a quadratic form depending only on $\varphi^+ - \varphi^-$; i.e.,

$$(5.8) \qquad K_{\varepsilon}(\varphi^-, \varphi^+) = k_{\varepsilon}(\varphi^+ - \varphi^-)^2, \quad k_{\varepsilon} = K_{\varepsilon}(-1/2, 1/2).$$

We call $\mathcal{T}_{\varepsilon}(\varphi^-, \varphi^+)$ the solution of the minimum problem (5.7): it admits the simple representation

$$(5.9) \qquad \mathcal{T}_{\varepsilon}(\varphi^-, \varphi^+) = \frac{1}{2}(\varphi^- + \varphi^+) + (\varphi^+ - \varphi^-)\phi_{\varepsilon},$$

where $\phi_{\varepsilon} = \mathcal{T}_{\varepsilon}(-1/2, 1/2)$. We also set
(5.10)

$$Q_{\varepsilon}(\varphi^-, \varphi^+) := \int_{-1}^{1} \left(\mathcal{T}_{\varepsilon}(\varphi^-, \varphi^+) \right)^2 \mathrm{d}\tilde{\gamma}_{\varepsilon} = \frac{1}{2}\left((\varphi^-)^2 + (\varphi^+)^2 \right) + (q_{\varepsilon} - \tfrac{1}{4})(\varphi^+ - \varphi^-)^2,$$

where

$$(5.11) \qquad q_{\varepsilon} := \int_{-1}^{1} |\phi_{\varepsilon}(\xi)|^2 \, \mathrm{d}\tilde{\gamma}_{\varepsilon}(\xi) = Q_{\varepsilon}(-1/2, 1/2).$$

LEMMA 5.5. *We have*

$$(5.12) \qquad \lim_{\varepsilon \downarrow 0} k_{\varepsilon} = \frac{k}{2} = \frac{\sqrt{-H''(0) \, H''(1)}}{2\pi}$$

and

$$(5.13) \qquad \lim_{\varepsilon \downarrow 0} q_{\varepsilon} = \frac{1}{4}, \quad \text{so that} \quad \lim_{\varepsilon \downarrow 0} Q_{\varepsilon}(\varphi^-, \varphi^+) = \frac{1}{2}(\varphi^-)^2 + \frac{1}{2}(\varphi^+)^2.$$

Proof. ϕ_ε solves the Euler equation

$$(5.14) \qquad \left(e^{-H(\xi)/\varepsilon}\phi_\varepsilon'(\xi)\right)' = 0 \quad \text{on } (-1,1), \quad \phi_\varepsilon(\pm 1) = \pm\tfrac{1}{2}.$$

We can compute an explicit solution of (5.14) by integration:

$$\phi_\varepsilon'(\xi) = Ce^{H(\xi)/\varepsilon}, \qquad \phi_\varepsilon(\xi) = C' + C\int_0^\xi e^{H(\eta)/\varepsilon}\,d\eta.$$

Define $I_\varepsilon := \int_{-1}^1 e^{H(\xi)/\varepsilon}\,d\xi$. The boundary conditions for $\xi = \pm 1$ give

$$C' = 0, \quad C\int_{-1}^1 e^{H(\xi)/\varepsilon}\,d\xi = CI_\varepsilon = 1.$$

It follows that

$$\phi_\varepsilon(\xi) = I_\varepsilon^{-1}\int_0^\xi e^{H(\eta)/\varepsilon}\,d\eta$$

and

$$k_\varepsilon = \tau_\varepsilon I_\varepsilon^{-2}\int_{-1}^1 e^{2H(\xi)/\varepsilon}\,d\tilde\gamma_\varepsilon(\xi) = \tau_\varepsilon Z^{-1}I_\varepsilon^{-1}.$$

We compute, using Laplace's method,

$$I_\varepsilon = \sqrt{\frac{2\pi\varepsilon}{|H''(0)|}}e^{1/\varepsilon}(1 + o(1)) \quad \text{and} \quad Z_\varepsilon = \sqrt{\frac{2\pi\varepsilon}{H''(1)}}(1 + o(1)) \quad \text{as } \varepsilon \to 0,$$

thus obtaining (5.12). Since

$$\phi_\varepsilon' = I_\varepsilon^{-1}e^{H/\varepsilon} \to \delta_0 \quad \text{in } \mathscr{D}'(-1,1)$$

and H is even, we have

$$\phi_\varepsilon(\xi) = I_\varepsilon^{-1}\int_0^\xi e^{H(\eta)/\varepsilon}\,d\eta \to \frac{1}{2}\mathrm{sign}(\xi)$$

uniformly on each compact subset of $[-1,1]$ not containing 0. Since the range of ϕ_ε belongs to $[-1/2, 1/2]$ and $\tilde\gamma_\varepsilon \overset{*}{\rightharpoonup} \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$, we obtain (5.13). $\quad\square$

5.4. End of the Proof of Theorem 5.1. Let us first check the "Γ-liminf" property (Γ-1$'$) of Lemma 4.3. The second limit of (4.30) follows by general lower-semicontinuity results on integral functionals of measures; see, e.g., [2, Lemma 9.4.3]. Concerning the first "lim inf" inequality, we split the quadratic form $\mathfrak{a}_\varepsilon$ into the sum of two parts,

$$(5.15) \qquad \mathfrak{a}_\varepsilon^1(\rho_\varepsilon) := \int_D |\nabla_x u_\varepsilon(x,\xi)|^2\,d\gamma_\varepsilon(x,\xi), \quad \mathfrak{a}_\varepsilon^2(\rho_\varepsilon) := \tau_\varepsilon\int_D (\partial_\xi u_\varepsilon)^2\,d\gamma_\varepsilon(x,\xi).$$

We choose a smooth cutoff function $\eta^- : [-1,1] \to [0,1]$ such that $\eta^-(-1) = 1$ and $\mathrm{supp}(\eta^-) \subset [-1, -1/2]$ and the symmetric function $\eta^+(\xi) := \eta^-(-\xi)$. We also set

$$(5.16) \quad \tilde u_\varepsilon^-(x) := \int_{-1}^1 \eta^-(\xi)u_\varepsilon(x,\xi)\,d\tilde\gamma_\varepsilon(\xi), \quad \tilde u_\varepsilon^+(x) := \int_{-1}^1 \eta^+(\xi)u_\varepsilon(x,\xi)\,d\tilde\gamma_\varepsilon(\xi),$$

and it is easy to check that

$$(5.17) \qquad\qquad \tilde{u}_\varepsilon^\pm \rightharpoonup \frac{1}{2} u^\pm \quad \text{in } \mathscr{D}'(\Omega).$$

We also set $\theta_\varepsilon := \int_{-1}^1 \eta^+(\xi) \, d\tilde{\gamma}_\varepsilon(\xi) \ \big(= \int_{-1}^1 \eta^-(\xi) \, d\tilde{\gamma}_\varepsilon(\xi) \big)$, observing that $\theta_\varepsilon \to 1/2$. We then have by the Jensen inequality and the assumption on the support of $\eta^\pm$

$$\mathfrak{a}_\varepsilon^1(\rho_\varepsilon) \geq \int_\Omega \int_{-1}^1 (\eta^-(\xi) + \eta^+(\xi)) |\nabla_x u_\varepsilon|^2 \, d\tilde{\gamma}_\varepsilon(\xi) \, d\lambda_\Omega \geq \theta_\varepsilon^{-1} \int_\Omega |\nabla \tilde{u}_\varepsilon^-|^2 + |\nabla \tilde{u}_\varepsilon^+|^2 \, d\lambda_\Omega,$$

and, passing to the limit,

$$\liminf_{\varepsilon \downarrow 0} \mathfrak{a}_\varepsilon^1(\rho_\varepsilon) \geq \frac{1}{2} \int_\Omega |\nabla u^-|^2 + |\nabla u^+|^2 \, d\lambda_\Omega.$$

Let us now consider the behavior of $\mathfrak{a}_\varepsilon^2$: applying (5.7) and (5.8) we get

$$\mathfrak{a}_\varepsilon^2(\rho_\varepsilon) = \int_\Omega \left(\tau_\varepsilon \int_{-1}^1 (\partial_\xi u_\varepsilon(x,\xi))^2 \, d\tilde{\gamma}_\varepsilon(\xi) \right) d\lambda_\Omega \geq \int_\Omega k_\varepsilon (\tilde{u}_\varepsilon^-(x) - \tilde{u}_\varepsilon^+(x))^2 \, d\lambda_\Omega,$$

so that by (5.12) and (5.4) we obtain

$$(5.18) \qquad\qquad \liminf_{\varepsilon \downarrow 0} \mathfrak{a}_\varepsilon^2(\rho_\varepsilon) \geq \frac{k}{2} \int_\Omega \left(u^-(x) - u^+(x) \right)^2 d\lambda_\Omega.$$

We want to prove now the "Γ-limsup" property (4.31) of Lemma 4.3. We fix $\sigma = u\gamma$ with u in the domain of the quadratic forms a and b so that $u^\pm = u(\cdot, \pm 1)$ belong to $W^{1,2}(\Omega)$, and we set $\sigma_\varepsilon = u_\varepsilon \gamma_\varepsilon$, where $u_\varepsilon(x,\cdot) = \mathfrak{T}_\varepsilon(u^-(x), u^+(x))$ as in (5.9). We easily have, by (5.13) and the Lebesgue dominated convergence theorem,

$$\lim_{\varepsilon \downarrow 0} \mathfrak{b}_\varepsilon(\sigma_\varepsilon) = \lim_{\varepsilon \downarrow 0} \int_\Omega Q_\varepsilon(u^-(x), u^+(x)) \, d\lambda_\Omega = \int_\Omega \left(\frac{1}{2} |u^-(x)|^2 + \frac{1}{2} |u^+(x)|^2 \right) d\lambda_\Omega = \mathfrak{b}(\sigma).$$

Similarly, since for every $j = 1, \ldots, d$ and almost every $x \in \Omega$

$$\partial_{x_j} u_\varepsilon(x,\xi) = \mathfrak{T}(\partial_{x_j} u^-(x), \partial_{x_j} u^+),$$

we have

$$\lim_{\varepsilon \downarrow 0} \mathfrak{a}_\varepsilon(\sigma_\varepsilon) = \lim_{\varepsilon \downarrow 0} \int_\Omega \left(\sum_{j=1}^d Q_\varepsilon\left(\partial_{x_j} u^-(x), \partial_{x_j} u^+(x) \right) + K_\varepsilon\left(u^-(x), u^+(x) \right) \right) d\lambda_\Omega$$

$$= \int_\Omega \left(\frac{1}{2} |\nabla u^-(x)|^2 + \frac{1}{2} |\nabla u^+(x)|^2 + \frac{k}{2} \left(u^-(x) - u^+(x) \right) \right) d\lambda_\Omega = \mathfrak{a}(\sigma).$$

6. From Γ-Convergence to Convergence of the Evolution Problems: Proof of Theorems 3.1 and 3.2. Having at our disposal the Γ-convergence result of Theorem 5.1 and Lemma 4.2, it is not difficult to pass to the limit in the integrated equation (4.10).

The proof of Theorems 3.1 and 3.2 is a consequence of the following general result.

THEOREM 6.1 (convergence of evolution problems). *Let us consider weakly-$*$ lower-semicontinuous, nonnegative, and extended-valued quadratic forms $\mathfrak{a}_\varepsilon, \mathfrak{b}_\varepsilon, \mathfrak{a}, \mathfrak{b}$ defined on $\mathscr{M}(\overline{D})$, and let us suppose the following.*

(1) Nondegeneracy of the limit forms. $\mathfrak{b}$ *is nondegenerate (i.e., $\mathfrak{b}(\rho) = 0 \Rightarrow \rho = 0$) and* $\mathsf{Dom}(\mathfrak{a})$ *is dense in* $\mathsf{Dom}(\mathfrak{b})$ *with respect to the norm-convergence induced by* $\mathfrak{b}$.

(2) Equicoercivity. $\mathfrak{b}_\varepsilon$ *are equicoercive, i.e., they satisfy the coercivity property stated in Lemma 5.2.*

(3) Joint Γ-convergence. $\mathfrak{q}_\varepsilon^\kappa := \mathfrak{b}_\varepsilon + \kappa \, \mathfrak{a}_\varepsilon$ *satisfy the joint Γ-convergence property (5.1):*

$$(6.1) \qquad \Gamma\left(\mathscr{M}(\overline{D})\right)\text{-}\lim_{\varepsilon \downarrow 0} \mathfrak{q}_\varepsilon^\kappa = \mathfrak{q}^\kappa = \mathfrak{b} + \kappa \, \mathfrak{a} \quad \textit{for every } \kappa > 0.$$

Let $\rho_\varepsilon(t)$, $t \geq 0$, be the solution of the evolution problem (4.10) starting from $\rho_\varepsilon^0 \in \mathsf{Dom}(\mathfrak{b}_\varepsilon)$.
If

$$(6.2) \qquad \rho_\varepsilon^0 \overset{*}{\rightharpoonup} \rho^0 \textit{ weakly-}\ast \textit{ in } \mathscr{M}(\overline{D}) \textit{ as } \varepsilon \downarrow 0 \quad \textit{with} \quad \limsup_{\varepsilon \downarrow 0} \mathfrak{b}_\varepsilon(\rho_\varepsilon^0) < +\infty,$$

then $\rho_\varepsilon(t) \overset{*}{\rightharpoonup} \rho(t)$ weakly-$\ast$ in $\mathscr{M}(\overline{D})$ as $\varepsilon \downarrow 0$ for every $t > 0$, and $\rho(t)$ is the solution of the limit evolution problem (4.20).
If, moreover, $\lim_{\varepsilon \downarrow 0} \mathfrak{b}_\varepsilon(\rho_\varepsilon^0) = \mathfrak{b}(\rho_0)$, then

$$(6.3) \qquad \lim_{\varepsilon \downarrow 0} \mathfrak{b}_\varepsilon(\rho_\varepsilon(t)) = \mathfrak{b}(\rho(t)), \quad \lim_{\varepsilon \downarrow 0} \mathfrak{a}_\varepsilon(\rho_\varepsilon(t)) = \mathfrak{a}(\rho(t)) \quad \textit{for every } t > 0.$$

Proof. Let us first note that by (4.11) and the coercivity property of $\mathfrak{b}_\varepsilon$ the mass of $\rho_\varepsilon(t)$ is bounded uniformly in t. Moreover, (4.13) and the coercivity property show that $\partial_t \rho_\varepsilon$ is a finite measure whose total mass is uniformly bounded in each bounded interval $[t_0, t_1] \subset (0, +\infty)$. By the Arzelà–Ascoli theorem we can extract a subsequence ρ_{ε_n} such that $\rho_{\varepsilon_n}(t) \overset{*}{\rightharpoonup} \rho(t)$ for every $t \geq 0$. The estimates (4.13) and the "Γ-liminf" property (4.30) show that for every $t > 0$, $\rho(t)$ belongs to the domain of the quadratic forms $\mathfrak{a}$ and $\mathfrak{b}$ and satisfies a similar estimate,

$$(6.4) \qquad \frac{1}{2}\mathfrak{b}(\rho(t)) + t \, \mathfrak{a}(\rho(t)) + t^2 \mathfrak{b}(\partial_t \rho(t)) \leq \frac{1}{2} \liminf_{\varepsilon \downarrow 0} \mathfrak{b}(\rho_\varepsilon^0) < +\infty.$$

Let $\sigma \in \mathscr{M}(\overline{D})$ be an arbitrary element of the domains of $\mathfrak{a}$ and $\mathfrak{b}$; by (5.1) we can find a family σ_ε weakly-$\ast$ converging to σ such that (4.31) holds. By (4.10) we have

$$(6.5) \qquad \mathfrak{b}_\varepsilon(\rho_\varepsilon(t), \sigma_\varepsilon) + \int_0^t \mathfrak{a}_\varepsilon(\rho_\varepsilon(r), \sigma_\varepsilon) \, \mathrm{d}r = \mathfrak{b}_\varepsilon(\rho_\varepsilon^0, \sigma_\varepsilon),$$

and (4.13) with the Schwarz inequality yields the uniform bound

$$\left| \mathfrak{a}_\varepsilon(\rho_\varepsilon(t), \sigma_\varepsilon) \right| \leq t^{-1/2} \mathfrak{b}_\varepsilon(\rho_\varepsilon^0)^{1/2} \mathfrak{a}_\varepsilon(\sigma_\varepsilon)^{1/2} \leq C t^{-1/2},$$

where C is independent of ε; we can therefore pass to the limit in (6.5) (actually along the subsequence ε_n) by Lemma 4.2 to find

$$\mathfrak{b}(\rho(t), \sigma) + \int_0^t \mathfrak{a}_\varepsilon(\rho(r), \sigma) \, \mathrm{d}r = \mathfrak{b}(\rho_0^0, \sigma),$$

so that ρ is a solution of the limit equation. Since the limit is uniquely identified by the nondegeneracy and density assumption (1), we conclude that the whole family ρ_ε converges to ρ as $\varepsilon \downarrow 0$. In particular, ρ satisfies the identity

$$(6.6) \qquad \frac{1}{2}\mathfrak{b}(\rho(t)) + \int_0^t \mathfrak{a}(\rho(r))\,\mathrm{d}r = \frac{1}{2}\mathfrak{b}(\rho^0) \quad \text{for every } t \geq 0.$$

This concludes the proof of (6.2) (and of Theorem 3.1).

In order to prove (6.3) (and Theorem 3.2), we note that by (4.11) and (6.6) we easily get

$$\limsup_{\varepsilon \downarrow 0} \frac{1}{2}\mathfrak{b}_\varepsilon(\rho_\varepsilon(t)) + \int_0^t \mathfrak{a}_\varepsilon(\rho_\varepsilon(r))\,\mathrm{d}r \leq \frac{1}{2}\mathfrak{b}(\rho(t)) + \int_0^t \mathfrak{a}(\rho(r))\,\mathrm{d}r.$$

The lower-semicontinuity property (4.30) and Fatou's lemma yield

$$(6.7) \qquad \lim_{\varepsilon \downarrow 0} \mathfrak{b}_\varepsilon(\rho_\varepsilon(t)) = \mathfrak{b}(\rho(t)), \quad \lim_{\varepsilon \downarrow 0} \int_0^t \mathfrak{a}_\varepsilon(\rho_\varepsilon(r))\,\mathrm{d}r = \int_0^t \mathfrak{a}(\rho(r))\,\mathrm{d}r$$

for every $t \geq 0$. Applying the same argument to (4.12) and its "$\varepsilon = 0$" analogue, we conclude that $\mathfrak{a}_\varepsilon(\rho_\varepsilon(t)) \to \mathfrak{a}(\rho(t))$ for every $t > 0$. $\quad\square$

Remark (more general ambient spaces). The particular structure of $\mathscr{M}(\overline{D})$ did not play any role in the previous argument, so the validity of the above result can be easily extended to general topological vector spaces (e.g., the dual of separable Banach spaces with their weak-$*$ topology) once the equicoercivity condition of $\mathfrak{b}_\varepsilon$ (as in Lemma 5.2) is satisfied.

REFERENCES

[1] S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer, *From a large-deviations principle to the Wasserstein gradient flow: A new micro-macro passage*, Comm. Math. Phys., 307 (2011), pp. 791–815.

[2] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Math. ETH Zürich, Birkhäuser, Basel, 2005.

[3] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni, *Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction*, Calc. Var. Partial Differential Equations, to appear.

[4] H. Attouch, *Variational Convergence for Functions and Operators*, Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, 1984.

[5] N. Berglund, *Kramers' Law: Validity, Derivations and Generalisations*, arXiv preprint, arxiv.org/abs/1106.5799/v1 (2011).

[6] M. Born and J. R. Oppenheimer, *Zur Quantentheorie der Molekeln*, Ann. Physik, 84 (1927), pp. 457–484.

[7] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North–Holland, Amsterdam, 1973.

[8] H. Brezis, *Analyse fonctionnelle: Théorie et applications*, Collection Mathématiques Appliquées pour la Maîtrise, Masson, Paris, 1983.

[9] G. Da Prato and A. Lunardi, *On a class of self-adjoint elliptic operators in L^2 spaces with respect to invariant measures*, J. Differential Equations, 234 (2007), pp. 54–79.

[10] G. Dal Maso, *An Introduction to Γ-Convergence*, Progr. Nonlinear Differential Equations Appl. 8, Birkhäuser Boston, Boston, 1993.

[11] E. De Giorgi and T. Franzoni, *Su un tipo di convergenza variazionale*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur., 58 (1975), pp. 842–850.

[12] E. De Giorgi and S. Spagnolo, *Sulla convergenza degli integrali dell'energia per operatori ellittici del 2 ordine*, Boll. Un. Mat. Ital. (4), 8 (1973), pp. 391–411.

[13] P. Hänggi, P. Talkner, and M. Borkovec, *Reaction-rate theory: Fifty years after Kramers*, Rev. Modern Phys., 62 (2) (1990), pp. 251–342.

[14] M. Herrmann and B. Niethammer, *Kramers' formula for chemical reactions in the context of Wasserstein gradient flows*, Commun. Math. Sci., 9 (2011), pp. 623–635.

[15] J. E. Hutchinson, *Second fundamental form for varifolds and the existence of surfaces minimising curvature*, Indiana Univ. Math. J., 35 (1986), pp. 45–71.

[16] R. Jordan, D. Kinderlehrer, and F. Otto, *The variational formulation of the Fokker–Planck equation*, SIAM J. Math. Anal., 29 (1998), pp. 1–17.

[17] H. A. Kramers, *Brownian motion in a field of force and the diffusion model of chemical reactions*, Phys., 7 (1940), pp. 284–304.

[18] J. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications*, Volume I, Springer, New York, Heidelberg, 1972.

[19] A. Mielke, *A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems*, Nonlinearity, 24 (2011), pp. 1329–1346.

[20] U. Mosco, *Approximation of the solutions of some variational inequalities*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3), 21 (1967), pp. 373–394.

[21] F. Otto, *The geometry of dissipative evolution equations: The porous medium equation*, Comm. Partial Differential Equations, 26 (2001), pp. 101–174.

[22] M. A. Peletier, G. Savaré, and M. Veneroni, *From diffusion to reaction via Γ-convergence*, SIAM J. Math. Anal., 42 (2010), pp. 1805–1825.

[23] M. Pennacchio, G. Savaré, and P. Colli Franzone, *Multiscale modeling for the bioelectric activity of the heart*, SIAM J. Math. Anal., 37 (2006), pp. 1333–1370.

[24] J. G. Rešetnjak, *The weak convergence of completely additive vector-valued set functions*, Sibirsk. Mat. Ž., 9 (1968), pp. 1386–1394.

[25] S. Spagnolo, *Sul limite delle soluzioni di problemi di Cauchy relativi all'equazione del calore*, Ann. Scuola Norm. Sup. Pisa (3), 21 (1967), pp. 657–699.

[26] S. Spagnolo, *Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche*, Ann. Scuola Norm. Sup. Pisa (3), 22 (1968), pp. 577–597.