SINGULAR-DEGENERATE MULTIVALUED STOCHASTIC FAST DIFFUSION EQUATIONS

BENJAMIN GESS
Department of Mathematics
University of Chicago
USA

MICHAEL RÖCKNER
Faculty of Mathematics
University of Bielefeld
Germany

Abstract. We consider singular-degenerate, multivalued stochastic fast diffusion equations with multiplicative Lipschitz continuous noise. In particular, this includes the stochastic sign fast diffusion equation arising from the Bak-Tang-Wiesenfeld model for self-organized criticality. A well-posedness framework based on stochastic variational inequalities (SVI) is developed, characterizing solutions to the stochastic sign fast diffusion equation, previously obtained in a limiting sense only. Aside from generalizing the SVI approach to stochastic fast diffusion equations we develop a new proof of well-posedness, applicable to general diffusion coefficients. In case of linear multiplicative noise, we prove the existence of (generalized) strong solutions, which entails higher regularity properties of solutions than previously known.

1. Introduction

We consider singular-degenerate, multivalued stochastic fast diffusion equations (SFDE) of the type

\[ dX_t \in \Delta(|X_t|^{m-1}X_t)dt + B(t, X_t)dW_t, \]
\[ X_0 = x_0 \]

with \( m \in [0, 1] \), on bounded, smooth domains \( \Omega \subseteq \mathbb{R}^d \) with zero Dirichlet boundary conditions, where \( |r|^{-1}r := \text{Sgn}(r) \) denotes the maximal monotone extension of the sign function. In particular, we include the multivalued case \( m = 0 \) and general diffusion coefficients \( B \). In the following \( W \) is a cylindrical Wiener process on some separable Hilbert space \( U \) and the diffusion coefficients \( B : [0, T] \times H^{-1} \times \]

---

E-mail addresses: gess@uchicago.edu, roeckner@mathematik.uni-bielefeld.de.
Date: December 24, 2014.
2010 Mathematics Subject Classification. Primary: 60H15; Secondary: 76S05.
Key words and phrases. singular degenerate SPDE, multivalued SPDE, self-organized criticality, stochastic fast diffusion, sign fast diffusion, regularity, stochastic variational inequalities.

Acknowledgements: We thank the DFG for support through the research project “Random dynamical systems and regularization by noise for stochastic partial differential equations” and through CRC 701. B.G. thanks Jonas Tölle for stimulating discussion.
\[ \Omega \to L^2(\Omega, H^{-1}) \] take values in the space of Hilbert-Schmidt operators \( L^2(U, H^{-1}) \), where \( H^{-1} \) is the dual of \( H^1_0(\Omega) \).

Our main results are twofold: First, in the case of general diffusion coefficients and initial data we introduce a notion of stochastic variational inequalities (SVI) to (1.1) and establish a new method of proof of well-posedness. In particular, this new method allows treatment of general diffusion coefficients, whereas previously the approach of SVI solutions was restricted to additive or linear multiplicative noise (cf. [1.3], [1.4] below). In this sense, our results generalize those of [3, 7]. The second main result yields regularity properties of solutions in the case of linear multiplicative noise (cf. [1.5] below). In particular, we prove the existence of strong solutions, which extends the results from [8] from the degenerate case \( m > 1 \) to the singular case \( m \in [0, 1] \).

In the case \( m > 0 \) a variational approach to (1.1) has been developed in [14] for \( x_0 \in L^2(\Omega; H^{-1}) \) based on the coercivity property

\[ \nabla \cdot (|v|^{m-1}v), v) \geq c\|v\|^{m+1}_V \quad \forall v \in V = L^{m+1}(\Omega). \]

In the multivalued limiting case \( m = 0 \) two complications appear: First, the reflexivity of the energy space \( L^{m+1}(\Omega) \) is lost, making the variational methods from [14] inapplicable in this case. Second, the operator \( \Delta(|v|^{m-1}v) = \Delta \text{Sgn}(v) \) becomes multivalued. Recently, for regular initial data \( x_0 \in L^2(\Omega; L^2(\Omega)) \) an alternative variational approach to (1.1) has been developed in [11], proving well posedness for \( m \in [0, 1] \). However, for general initial conditions \( x_0 \in L^2(\Omega; H^{-1}) \) solutions could be constructed in a limiting sense only. That is, it has been shown that for each approximating sequence \( x^n_0 \in L^2(\Omega; L^2(\Omega)) \) with \( x^n_0 \to x \) in \( L^2(\Omega; H^{-1}) \) the corresponding variational solutions \( X^n \) converge to a limit \( X \) independent of the chosen approximating sequence \( x^n_0 \). The characterization of \( X \) in terms of a generalized notion of solution to (1.1) remained open. This problem is solved in this paper by introducing a notion of stochastic variational inequalities for (1.1) and proving well-posedness in this framework. The limiting solution \( X \) is thus characterized as an SVI solution to (1.1).

The difficulties for (1.1) as described above are similar to the ones for the stochastic total variation flow

\[ dX_t = \text{div} \left( \frac{\nabla X_t}{|\nabla X_t|} \right) dt + B(t, X_t) dW_t \]

\[ X_0 = x_0. \]

As for (1.1), in case of regular initial data \( x_0 \in L^2(\Omega; H^1_0(\Omega)) \) variational solutions have been constructed in [11]. For general initial data \( x_0 \in L^2(\Omega; L^2(\Omega)) \) solutions to (1.2) could be constructed in a limiting sense only. In the special case of additive noise, i.e.

\[ dX_t \in \text{div} \left( \frac{\nabla X_t}{|\nabla X_t|} \right) dt + dW_t \]

and assuming \( d = 1, 2 \), a notion of SVI solution to (1.3) has been introduced in [3]. Only recently, well-posedness of SVI solutions and thus characterization of limiting solutions in the linear multiplicative case

\[ dX_t \in \text{div} \left( \frac{\nabla X_t}{|\nabla X_t|} \right) dt + \sum_{k=1}^{\infty} f_k X_t d\beta_t^k \]

has been shown in [7]. In this sense, our results on well-posedness of SVI solutions to (1.1) parallel those of [3, 7] in the case of stochastic fast diffusion equations. In both cases (1.3) and (1.4), the SPDE may be transformed into a random PDE, i.e.
a PDE with random coefficients. This technique is a crucial ingredient in the proofs given in [3,7] and requires the restriction to either additive or linear multiplicative noise. In contrast to this, in the first part of this paper (see Section 2) we consider [4,1] for general multiplicative noise and introduce an alternative method to prove well-posedness of SVI solutions that does not rely on a transformation into a random PDE. This allows to treat general noise, while significantly simplifying the proof as compared to [2]. Moreover, in contrast to [3] no restrictions on the dimension d will be required.

In the second part of this paper (see Section 3) we prove regularity properties for solutions to (1.1) in the case of linear multiplicative noise, i.e. for

\begin{equation}
(1.5) \quad dX_t \in \Delta(|X_t|^{m-1}X_t)dt + \sum_{k=1}^{\infty} f_k X_t d\beta_k^k,
\end{equation}

with \( m \in [0, 1], d \in \mathbb{N} \) and \( x_0 \in L^2(\Omega; H^{-1}) \). More precisely, we prove the existence of (generalized) strong solutions (cf. Definition A.1 below), in particular implying that \( X \) takes values in the domain of \( \Delta(|\cdot|^{m-1}) \), \( dt \otimes \mathbb{P} \)-almost everywhere. This extends regularity results obtained in [8] where the degenerate case \( m \geq 1 \) was considered by entirely different methods. The case of singular diffusions \( \{m \in [0, 1]\} \) could not be handled in [8] due to the singularity of the non-linearity \( \phi(r) = |r|^{m-1}r \) at zero. Roughly speaking, this singularity has to be compensated by sufficient decay of the diffusion coefficients at zero; a problem not appearing in degenerate, non-singular cases treated in [10]. This requires a careful choice of approximating problems and leads to entirely different methods than those developed in [8]. In particular, it turns out that different approximations of the nonlinearity \( \phi \) need to be considered in the proof of regularity for (1.5) and in the proof of well-posedness of SVI solutions for (1.1). We underline that also the methods developed in the second part of this paper do not rely on a transformation of (1.5) into a random PDE and therefore depend only loosely on the linear structure of the noise in (1.5).

In fact, as pointed out above, the crucial structural condition in (1.5) is not the linearity of the diffusion coefficients, but their decay behavior at zero.

Stochastic fast diffusion equations of the type (1.1) have been intensively investigated in recent years. For the single-valued case \( m > 0 \) we refer to [12,14] and the references therein. As the multivalued, limiting case \( m = 0 \) is concerned, mostly the case of linear multiplicative noise (1.5) has been considered in the literature. Well-posedness for regular initial data \( x_0 \in L^1(\Omega) \) and \( d = 1, 2, 3 \) was first proven in [4]. Finite time extinction for (1.5) has been investigated in [2,4,6,9,17]. For bounded initial data \( x_0 \in L^\infty(\Omega) \) and finite driving noise, that is \( f_k \equiv 0 \) for all \( k \) large enough, the existence of strong solutions (cf. Definition A.1 below) to (1.5) has been proven in [9] by entirely different methods, relying on a transformation of (1.5) into a random PDE. Well-posedness for (1.1) with \( m = 1 \) and with general multiplicative noise has been obtained in [11] for the first time, proving well-posedness in terms of variational solutions for regular initial data \( x_0 \in L^2(\Omega) \).

For background on the deterministic fast diffusion equation we refer to [18,19] and the references therein.

1.1. Notation. In the following let \( \Omega \subseteq \mathbb{R}^d \) be a bounded set with smooth boundary. \( L^p := L^p(\Omega) \) denotes the usual Lebesgue space with norm \( \| \cdot \|_{L^p} \) and inner product \( \langle \cdot, \cdot \rangle_p \) if \( p = 2 \). Further, \( H^1_0 := H^1_0(\Omega) \) denotes the Sobolev space of order one in \( L^2 \) equipped with the inner product \( \langle v, w \rangle_{H^1_0} = (\nabla v, \nabla w)_2 \) and norm \( \| \cdot \|_{H^1_0} \). Let \( (H^{-1}, \langle \cdot, \cdot \rangle_{H^{-1}}) \) with norm \( \| \cdot \|_{H^{-1}} \) be the dual of \( H^1_0 \). Moreover, we let \( C_0(\Omega) \)
denote the set of all continuous functions on $O$ vanishing at the boundary. In the proofs, as usual, constants may change from line to line.

2. Stochastic variational inequalities

In this section we consider stochastic singular fast diffusion equations of the type
\begin{equation}
\label{eq:2.1}
dX_t \in \Delta \{ |X_t|^{m-1} X_t \} \, dt + B(t, X_t) \, dW_t, \\
X_0 = x_0
\end{equation}
for $m \in [0, 1]$, on a bounded, smooth domain $O \subseteq \mathbb{R}^d$ with Dirichlet boundary conditions and general diffusion coefficients $B$, in particular including additive and linear multiplicative noise. The precise definition of the nonlinear part on the right hand side of \eqref{eq:2.1} including its domain, as well as the definition of a solution to \eqref{eq:2.1} will be given below. We emphasize that the multivalued, limiting case $m = 0$ is included.

Here $W$ is a cylindrical Wiener process in some separable Hilbert space $U$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t \geq 0}$ and the diffusion coefficients $B$ take values in the space of Hilbert-Schmidt operators $L_2(U, H)$. As compared to the regularity results obtained in Section 3 below, for this general choice of diffusion coefficients, we cannot expect (generalized) strong solutions to exist for arbitrary initial conditions $x_0 \in H^{-1}$. Instead we introduce a notion of stochastic variational inequalities for \eqref{eq:2.1} which we prove to uniquely characterize solutions.

We suppose that $B : [0, T] \times H^{-1} \times \Omega$ is progressively measurable and satisfies
\begin{equation}
\label{eq:2.2}
\begin{aligned}
\|B(t, v) - B(t, w)\|_{L_2(U, H^{-1})}^2 &\leq C\|v - w\|_{H^{-1}}^2 \quad \forall v, w \in H^{-1} \\
\|B(t, v)\|_{L_2(U, L_2)}^2 &\leq C(1 + \|v\|_{L_2}^2) \quad \forall v \in L_2,
\end{aligned}
\end{equation}
for some constant $C > 0$ and all $(t, \omega) \in [0, T] \times \Omega$.

Let $\mathcal{M}$ be the space of all signed Radon measures on $O$ with finite total variation. For $\mu \in \mathcal{M}$ we let $|\mu|$ be its variation with total variation $TV(\mu; O) = |\mu|(O)$.

We define (cf. Appendix D)
\begin{align*}
L^{m+1} \cap H^{-1} &:= \left\{ v \in L^{m+1} \mid \int_O vh dx \leq C\|h\|_{H_0^1}, \forall h \in C_c^1(O) \text{ for some } C \geq 0 \right\} \\
\mathcal{M} \cap H^{-1} &:= \left\{ \mu \in \mathcal{M} \mid \int_O h(x) d\mu(x) \leq C\|h\|_{H_0^1}, \forall h \in C_c^1(O) \text{ for some } C \geq 0 \right\}.
\end{align*}

Note that $\mathcal{M} \cap H^{-1}$ is known as the space of finite measures of bounded energy (cf. e.g. \cite{16}). Clearly, we have $L^{m+1} \cap H^{-1} \subseteq H^{-1}$ and $\mathcal{M} \cap H^{-1} \subseteq H^{-1}$. For $m > 0$ and $v \in H^{-1}$ we define $\varphi : H^{-1} \to [0, \infty]$ by
\begin{align*}
\varphi(v) := \begin{cases}
\frac{1}{m+1} \|v\|_{L^{m+1}}^{m+1}, & v \in L^{m+1} \cap H^{-1} \\
+\infty, & \text{otherwise}.
\end{cases}
\end{align*}

By Lemma D.1 in Appendix D below, $\varphi$ defines a convex, lower-semicontinuous function on $H^{-1}$. Moreover, by Lemma D.2
\begin{align*}
\varphi(\mu) := \begin{cases}
TV(\mu; O), & \mu \in \mathcal{M} \cap H^{-1} \\
+\infty, & \text{otherwise}.
\end{cases}
\end{align*}
is the lower-semicontinuous hull on $H^{-1}$ of $\| \cdot \|_1$ defined on $L^1 \cap H^{-1}$. For $m \geq 0$ we set
\begin{align*}
\psi(r) = \frac{1}{m+1} |r|^{m+1}, & \quad r \in \mathbb{R},
\end{align*}
and note that
\[ \partial \varphi(r) = |r|m \text{Sgn}(r) =: \phi(r), \quad r \in \mathbb{R} \]
where Sgn is the maximal monotone, multivalued extension of the sign function.

Concerning the subgradient \( \partial \varphi \) of \( \varphi \) we have (see Lemma D.3 below)
\[ \partial \varphi(v) \supset (-\Delta w|w \in H^1_0, w \in \phi(v) \text{ a.e.}), \]
for all \( v \in L^{m+1} \cap H^{-1} \). Hence, we may rewrite (2.1) in the relaxed form
\[ (2.3) \quad dx_t \in -\partial \varphi(x_t) dt + B(t, x_t) dW_t, \]
\[ X_0 = x_0 \]
and define (generalized) strong solutions according to Definition A.1 in Appendix A below.

**Definition 2.1.** Let \( x_0 \in L^2(\Omega; H^{-1}) \). An \( \mathcal{F}_t \)-adapted process \( X \in L^2(\Omega; C([0, T]; H^{-1})) \) is said to be an SVI solution to (2.1) if
i. [Regularity]
\[ \varphi(X) \in L^1([0, T] \times \Omega). \]
ii. [Variational inequality] For each \( \mathcal{F}_t \)-progressively measurable process \( G \in L^2([0, T] \times \Omega; H^{-1}) \) and each \( \mathcal{F}_t \)-adapted process \( Z \in L^2(\Omega; C([0, T]; H^{-1})) \cap L^2([0, T] \times \Omega; L^2) \) solving the equation
\[ Z_t - Z_0 = \int_0^t G_s ds + \int_0^t B(s, Z_s) dW_s, \quad \forall t \in [0, T], \]
we have
\[ \mathbb{E}[\|X_t - Z_t\|_{H^{-1}}^2] + 2\mathbb{E}\int_0^t \varphi(X_r) dr \]
\[ \leq \mathbb{E}[\|x_0 - Z_0\|_{H^{-1}}^2] + 2\mathbb{E}\int_0^t \varphi(Z_r) dr \]
\[ - 2\mathbb{E}\int_0^t (G_r, X_r - Z_r)_{H^{-1}} dr + C\mathbb{E}\int_0^t \|X_r - Z_r\|_{H^{-1}}^2 dr \quad \forall t \geq 0, \]
for some \( C > 0 \).

**Remark 2.2.** If \( (X, \eta) \) is a strong solution to (2.1) satisfying \( \varphi(X) \in L^1([0, T] \times \Omega) \) then \( X \) is an SVI solution to (2.1).

**Proof.** Definition 2.1 (i) is satisfied by assumption. For (ii): Let \( Z \in L^2(\Omega; C([0, T]; H^{-1})) \cap L^2([0, T] \times \Omega; L^2) \) be a solution to
\[ dZ_t = G_t dt + B(t, Z_t) dW_t \]
for some \( \mathcal{F}_t \)-progressively measurable \( G \in L^2([0, T] \times \Omega; H^{-1}) \). Then Itô’s formula implies:
\[ \mathbb{E}[\|X_t - Z_t\|_{H^{-1}}^2] = \mathbb{E}[\|x_0 - Z_0\|_{H^{-1}}^2] + 2\mathbb{E}\int_0^t (\eta_r - G_r, X_r - Z_r)_{H^{-1}} dr \]
\[ + \mathbb{E}\int_0^t \|B(r, X_r) - B(r, Z_r)\|_{L^2(U, H^{-1})}^2 dr \quad \forall t \in [0, T]. \]
Since \( \eta_r \in -\partial \varphi(X_r) \) we have
\[ \langle \eta_r, X_r - Z_r \rangle_{H^{-1}} \leq \varphi(Z_r) - \varphi(X_r), \quad dt \otimes dP \text{ a.e.} \]
which, using (2.2), implies (2.4). \( \square \)

The main result of the current section is the proof of well-posedness of (2.1) in the sense of Definition 2.1.
Theorem 2.3. Let \( x_0 \in L^2(\Omega; H^{-1}) \). Then there is a unique SVI solution \( X \) to (2.1) in the sense of Definition 2.1. For two SVI solutions \( X, Y \) with initial conditions \( x_0, y_0 \in L^2(\Omega; H^{-1}) \) we have

\[
\sup_{t \in [0,T]} E \| X_t - Y_t \|^2_{{H^{-1}}} \leq C E \| x_0 - y_0 \|^2_{{H^{-1}}}.
\]

The unique SVI solution \( X \) coincides with the limiting solution to (2.1) constructed in [11].

Proof. We construct SVI solutions to (2.1) by considering appropriate approximations by strong solutions. The specific form of the construction will also be a crucial ingredient in the proof of uniqueness.

Step 1: Existence

We consider approximating SPDE of the form

\[
dX^\varepsilon_t = \varepsilon \Delta X^\varepsilon_t dt + \Delta \phi^\varepsilon (X^\varepsilon_t) dt + B(t, X^\varepsilon_t) dW_t, \quad X^\varepsilon_0 = x_0,
\]

with \( \varepsilon > 0, \ x_0 \in L^2(\Omega; L^2) \) and \( \psi^\varepsilon, \phi^\varepsilon \) as in Appendix C. By Lemma B.1 there is a unique strong solution \( X^\varepsilon \) to (2.5) satisfying

\[
E \sup_{t \in [0,T]} \| X^\varepsilon_t \|_{H^0}^2 + 2 \varepsilon E \int_0^T \| X^\varepsilon_t \|_{H^0}^2 dt \leq C (E \| x_0 \|_{H^0}^2 + 1),
\]

for some \( C > 0 \) independent of \( \varepsilon > 0 \). For two solutions \( X^{\varepsilon_1}, X^{\varepsilon_2} \) to (2.5) with initial conditions \( x_0^1, x_0^2 \in L^2(\Omega; L^2) \) we have

\[
e^{-Kt} \| X^{\varepsilon_1}_t - X^{\varepsilon_2}_t \|_{H^{-1}}^2 = \| x_0^1 - x_0^2 \|_{H^{-1}}^2 + 2 \int_0^t e^{-Kr} (\varepsilon_1 \Delta X^{\varepsilon_1}_r - \varepsilon_2 \Delta X^{\varepsilon_2}_r, X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} dr + 2 \int_0^t e^{-Kr} (\Delta \phi^{\varepsilon_1}(X^{\varepsilon_1}_r) - \Delta \phi^{\varepsilon_2}(X^{\varepsilon_2}_r), X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} dr + 2 \int_0^t e^{-Kr} (X^{\varepsilon_1}_r - X^{\varepsilon_2}_r, B(r, X^{\varepsilon_1}_r) - B(r, X^{\varepsilon_2}_r))_{H^{-1}} dr + 2 \int_0^t e^{-Kr} \| B(r, X^{\varepsilon_1}_r) - B(r, X^{\varepsilon_2}_r) \|_{L^2}^2 dr - K \int_0^t e^{-Kr} \| X^{\varepsilon_1}_r - X^{\varepsilon_2}_r \|_{H^{-1}}^2 dr.
\]

Using (C.5) we note that

\[
(\Delta \phi^{\varepsilon_1}(X^{\varepsilon_1}_r) - \Delta \phi^{\varepsilon_2}(X^{\varepsilon_2}_r), X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} = - \int (\phi^{\varepsilon_1}(X^{\varepsilon_1}_r) - \phi^{\varepsilon_2}(X^{\varepsilon_2}_r))(X^{\varepsilon_1}_r - X^{\varepsilon_2}_r) dx \\
\leq C(\varepsilon_1 + \varepsilon_2)(1 + \| X^{\varepsilon_1}_r \|_{H^0}^2 + \| X^{\varepsilon_2}_r \|_{H^0}^2)
\]

and

\[
(\varepsilon_1 \Delta X^{\varepsilon_1}_r - \varepsilon_2 \Delta X^{\varepsilon_2}_r, X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} = \int (\varepsilon_1 X^{\varepsilon_1}_r - \varepsilon_2 X^{\varepsilon_2}_r)(X^{\varepsilon_1}_r - X^{\varepsilon_2}_r) dx \\
\leq C(\varepsilon_1 + \varepsilon_2)(\| X^{\varepsilon_1}_r \|_{H^0}^2 + \| X^{\varepsilon_2}_r \|_{H^0}^2),
\]
Using the Burkholder-Davis-Gundy inequality and (2.6) we obtain
\[ e^{-Kt} \|X^x_t - X^z_t\|_{H^{-1}}^2 \leq \|x_0^x - x_0^z\|_{H^{-1}}^2 + C(\varepsilon_1 + \varepsilon_2) \int_0^t (\|X^x_r\|_2^2 + \|X^z_r\|_2^2 + 1)dr \\
+ 2 \int_0^t e^{-Kr} (X^x_r - X^z_r, B(r, X^x_r) - B(r, X^z_r))_{H^{-1}}dW_r \\
+ C \int_0^t e^{-Kr} \|X^x_r - X^z_r\|_{H^{-1}}^2 dr \\
- K \int_0^t e^{-Kr} \|X^x_r - X^z_r\|_{H^{-1}}^2 dr. \]

Using the Burkholder-Davis-Gundy inequality and (2.6) we obtain
\[ (2.7) \quad \mathbb{E} \sup_{t \in [0,T]} e^{-Kt} \|X^x_t - X^z_t\|_{H^{-1}}^2 \leq 2\mathbb{E}\|x_0^x - x_0^z\|_{H^{-1}}^2 + C(\varepsilon_1 + \varepsilon_2)(\mathbb{E}\|x_0^x\|_2^2 + 1), \]
for \( K > 0 \) large enough.

Let now \( X^{x,1}, X^{x,2} \) be strong solutions to (2.5) with the same initial condition \( x_0 \in L^2(\Omega; L^2) \). Then (2.7) implies
\[ \mathbb{E} \sup_{t \in [0,T]} \|X^x_t - X^z_t\|_{H^{-1}}^2 \leq C(\varepsilon_1 + \varepsilon_2)(\mathbb{E}\|x_0^x\|_2^2 + 1) \]
and thus
\[ \mathbb{E} \sup_{t \in [0,T]} \|X^x_t - X^z_t\|_{H^{-1}}^2 \to 0 \quad \text{for} \quad \varepsilon \to 0 \]
for some \( \mathcal{F}_t \)-adapted process \( X \in L^2(\Omega; C([0,T]; H^{-1})) \) with \( X_0 = x_0 \). For \( x_0^1, x_0^2 \in L^2(\Omega; L^2) \) taking the limit \( \varepsilon \to 0 \) in (2.7) yields
\[ (2.8) \quad \mathbb{E} \sup_{t \in [0,T]} e^{-Kt} \|X^1_t - X^2_t\|_{H^{-1}}^2 \leq 2\mathbb{E}\|x_0^1 - x_0^2\|_{H^{-1}}^2. \]

Let now \( X^{x,n} \) be the unique strong solution (cf. Lemma [B.1]) to
\[ (2.9) \quad dX^x_t = \varepsilon \Delta X^{x,n}_t dt + \Delta \phi^y(X^{x,n}_t)dt + B(t, X^{x,n}_t) dW_t, \quad X^{x,n}_0 = x_0^n, \]
for some sequence \( x_0^n \to x_0 \) in \( L^2(\Omega; H^{-1}) \) with \( x_0^n \in L^2(\Omega; L^2) \). Using (2.7) and (2.8) we obtain the existence of a sequence of \( \mathcal{F}_t \)-adapted processes \( X^n \in L^2(\Omega; C([0,T]; H^{-1})) \) with \( X^n_0 = x_0^n \) and an \( \mathcal{F}_t \)-adapted process \( X \in L^2(\Omega; C([0,T]; H^{-1})) \) with \( X_0 = x_0 \) such that
\[ \mathbb{E} \sup_{t \in [0,T]} \|X^{x,n}_t - X^n_t\|_{H^{-1}}^2 \to 0 \quad \text{for} \quad n \to \infty. \]

Let now \( G, Z \) be as in Definition [2.1] Itô’s formula implies
\[ \mathbb{E}\|X^{x,n}_t - Z_t\|_{H^{-1}}^2 \leq \mathbb{E}\|x_0^n - Z_0\|_{H^{-1}}^2 \\
+ 2\mathbb{E} \int_0^t (\varepsilon \Delta X^{x,n}_r + \Delta \phi^y(X^{x,n}_r)) - G_r, X^{x,n}_r - Z_r)_{H^{-1}}dW_r \\
+ \mathbb{E} \int_0^t \|B(r, X^{x,n}_r) - B(r, Z_r)\|_{L^2}^2 dr. \]
For $v \in H^{-1}$ we set

$$
\varphi^\varepsilon(v) = \begin{cases} 
\int_\Omega \psi^\varepsilon(v) \, dx, & v \in L^2 \\
+\infty, & \text{otherwise.}
\end{cases}
$$

Using convexity of $\psi^\varepsilon$ we have

$$
(\Delta \varphi^\varepsilon(X^{\varepsilon,n}_r), X^{\varepsilon,n}_r - Z_r)_{H^{-1}} + \varphi^\varepsilon(X^{\varepsilon,n}_r) \leq \varphi^\varepsilon(Z_r)
$$

dt \otimes \mathbb{P}\text{-a.e.}

Moreover,

$$
(\varepsilon \Delta X^{\varepsilon,n}_r, X^{\varepsilon,n}_r - Z_r)_{H^{-1}} \leq \varepsilon \| \Delta X^{\varepsilon,n}_r \|_{H^{-1}} \| X^{\varepsilon,n}_r - Z_r \|_{H^{-1}}
\leq \varepsilon^{\frac{4}{3}} \| \Delta X^{\varepsilon,n}_r \|_{H^{-1}}^2 + \varepsilon^{\frac{2}{3}} \| X^{\varepsilon,n}_r - Z_r \|_{H^{-1}}^2
$$

dt \otimes \mathbb{P}\text{-a.e.}

Hence,

$$
\mathbb{E} \| X^{\varepsilon,n}_t - Z_t \|_{H^{-1}}^2 + 2 \mathbb{E} \int_0^t \varphi^\varepsilon(X^{\varepsilon,n}_r) \, dr
\leq \mathbb{E} \| x_0^n - Z_0 \|_{H^{-1}}^2 + 2 \mathbb{E} \int_0^t \varphi^\varepsilon(Z_r) \, dr
- 2 \mathbb{E} \int_0^t (G_r, X^{\varepsilon,n}_r - Z_r)_{H^{-1}} \, dr + C \mathbb{E} \int_0^t \| X^{\varepsilon,n}_r - Z_r \|_{H^{-1}}^2 \, dr
+ 2 \mathbb{E} \int_0^t \varepsilon^{\frac{4}{3}} \| \Delta X^{\varepsilon,n}_r \|_{H^{-1}}^2 + \varepsilon^{\frac{2}{3}} \| X^{\varepsilon,n}_r - Z_r \|_{H^{-1}}^2 \, dr.
$$

Due to the definition of $\varphi^\varepsilon$ and \[(C.4)\] we have

$$
| \varphi^\varepsilon(v) - \varphi(v) | \leq C \varepsilon (1 + \varphi(v))
\leq C \varepsilon (1 + \| v \|_2^2),
$$

for all $v \in L^2$. Hence,

$$
\mathbb{E} \int_0^t \varphi^\varepsilon(X^{\varepsilon,n}_r) \, dr \geq \mathbb{E} \int_0^t \varphi(X^{\varepsilon,n}_r) \, dr - C \varepsilon \mathbb{E} \int_0^t (1 + \| X^{\varepsilon,n}_r \|_2^2) \, dr
$$

and thus

$$
\liminf_{\varepsilon \to 0} \mathbb{E} \int_0^t \varphi^\varepsilon(X^{\varepsilon,n}_r) \, dr \geq \mathbb{E} \int_0^t \varphi(X^n) \, dr.
$$

Using $\varphi^\varepsilon \leq \varphi$, due to \[(C.3)\] and \[(2.6)\] we may thus let $\varepsilon \to 0$ and then $n \to \infty$ in \[(2.10)\] to obtain

$$
\mathbb{E} \int_0^t \varphi^\varepsilon(X^{\varepsilon,n}_r) \, dr \geq \mathbb{E} \int_0^t \varphi(X^n) \, dr.
$$

Using $\varphi^\varepsilon \leq \varphi$, due to \[(C.3)\] and \[(2.6)\] we may thus let $\varepsilon \to 0$ and then $n \to \infty$ in \[(2.10)\] to obtain

$$
\mathbb{E} \| X_t - Z_t \|_{H^{-1}}^2 + 2 \mathbb{E} \int_0^t \varphi(X_r) \, dr
\leq \mathbb{E} \| x_0 - Z_0 \|_{H^{-1}}^2 + 2 \mathbb{E} \int_0^t \varphi(Z_r) \, dr
- 2 \mathbb{E} \int_0^t (G_r, X_r - Z_r)_{H^{-1}} \, dr + C \mathbb{E} \int_0^t \| X_r - Z_r \|_{H^{-1}}^2 \, dr,
$$

where $\varphi(X) \in L^1([0,T] \times \Omega)$ follows from lower-semicontinuity of $\varphi$ on $H^{-1}$.

**Step 2: Uniqueness**

Let $X$ be an SVI solution to \[(2.1)\] and let $Y^{\varepsilon,n}$ be the (strong) solution to \[(2.9)\] with initial condition $y_0 \in L^2(\Omega; L^2)$ satisfying $y_0 \to y_0$ in $L^2(\Omega; H^{-1})$. Then \[(2.4)\]
with $Z = Y^{ε,n}$ and $G = εΔY^{ε,n} + Δφ^ε(Y^{ε,n})$ yields
\[
E\|X_t - Y^{ε,n}_t\|_{H^{-1}}^2 + 2E\int_0^t φ(X_r)dr \\
\leq E\|x_0 - y_0\|_{H^{-1}}^2 + 2E\int_0^t φ(Y^{ε,n}_r)dr \\
- 2E\int_0^t (εΔY^{ε,n}_r + Δφ^ε(Y^{ε,n}_r), X_r - Y^{ε,n}_r)_{H^{-1}}dr \\
+ CE\int_0^t \|X_r - Y^{ε,n}_r\|_{H^{-1}}^2dr.
\]

For $x \in L^2$ we have
\[-(Δφ^ε(Y^{ε,n}_r), x - Y^{ε,n}_r)_{H^{-1}} + φ^ε(Y^{ε,n}_r) \leq φ^ε(x) \; dt \otimes dP \text{- a.e.} \]

Due to (2.11) we obtain
\[-(Δφ^ε(Y^{ε,n}_r), x - Y^{ε,n}_r)_{H^{-1}} + φ^ε(Y^{ε,n}_r) \leq φ(x) + Cε(1 + φ(Y^{ε,n}_r)) \; dt \otimes dP \text{- a.e.} \]

Since $φ(X) \in L^1([0,T] \times Ω)$ and since $φ$ is the lower-semicontinuous hull of $φ_{L^2}$ on $H^{-1}$, for a.e. $(t,ω) \in [0,T] \times Ω$ we can find a sequence $x^m \in L^2$ such that $x^m \to X_t(ω)$ in $H^{-1}$ and $φ(x^m) \to φ(X_t(ω))$. Hence,
\[-(Δφ^ε(Y^{ε,n}_r), X - Y^{ε,n}_r)_{H^{-1}} + φ(Y^{ε,n}_r) \leq φ(X) + Cε(1 + φ(Y^{ε,n}_r)) \; dt \otimes dP \text{- a.e.} \]

and (2.12) implies
\[E\|X_t - Y^{ε,n}_t\|_{H^{-1}}^2 \leq E\|x_0 - y_0\|_{H^{-1}}^2 \\
+ 2E\int_0^t ε^2 \|ΔY^{ε,n}_r\|_{H^{-1}}^2 + ε^2 \|X_r - Y^{ε,n}_r\|_{H^{-1}}^2dr \\
+ CE\int_0^t \|X_r - Y^{ε,n}_r\|_{H^{-1}}^2dr + Cε\int_0^t (1 + φ(Y^{ε,n}_r))dr.
\]

Taking $ε \to 0$ then $n \to ∞$ yields
\[E\|X_t - Y_t\|_{H^{-1}}^2 \leq E\|x_0 - y_0\|_{H^{-1}}^2 + Cε\int_0^t \|X_r - Y_r\|_{H^{-1}}^2dr,
\]

which by Gronwall’s inequality concludes the proof.

\[
3. \text{Regularity and Strong solutions }
\]

We consider SPDE of the form
\[
dX_t = Δ(|X_t|^{m-1}X_t)dt + \sum_{k=1}^{∞} g^k X_t dβ^k_t, \\
X_0 = x_0
\]

with $m \in [0,1]$ and zero Dirichlet boundary conditions on a smooth, bounded domain $Ω \subseteq \mathbb{R}^d$, in arbitrary dimension $d \in \mathbb{N}$. Here, $β^k$ are independent Brownian motions on a normal, filtered probability space $(Ω,F,(F_t)_{t≥0},P)$ and

(B) $g^k \in C^1(Ω)$ with
\[∑_{k=1}^{∞} \|g^k\|_{C^1(Ω)}^2 < ∞.
\]

For $v \in H^{-1}$ we set
\[B(v)(h) = ∑_{k=1}^{∞} g^k v(e_k, h)_{H^{-1}},
\]
We define $\psi, \phi, \varphi$ in the relaxed form

$$\mu, \nu, \lambda$$

where $e_k \in H^{-1}$ is an orthonormal basis of $H^{-1}$. Then $B : H^{-1} \to L_2(H^{-1}, H^{-1})$ is Lipschitz continuous, i.e.

$$\|B(u) - B(v)\|_{L_2(H^{-1}, H^{-1})}^2 = \sum_{k=1}^{\infty} \|g_k(u - v)\|_{H^{-1}}^2$$

$$\leq \sum_{k=1}^{\infty} \|g_k\|_{C^1(\Omega)}^2 \|u - v\|^2_{H^{-1}}, \text{ for all } u, v \in H^{-1}$$

and

$$\|B(v)\|_{L_2(H^{-1}, L^2)}^2 = \sum_{k=1}^{\infty} \|g_kv\|_{L^2}^2$$

$$\leq \sum_{k=1}^{\infty} \|g_k\|_{C^1(\Omega)}^2 \|v\|^2_{L^2}, \text{ for all } v \in L^2.$$  

We define $\psi, \phi, \varphi$ and $\mathcal{M}, L^{n+1} \cap H^{-1}, \mathcal{M} \cap H^{-1}$ as in Section 2 and rewrite (3.1) in the relaxed form

$$dX_t \in -\partial \varphi(X_t) dt + B(X_t) dW_t, \quad X_0 = x_0.$$  

(Generalized) strong solutions to (3.1) are then defined as in Definition A.1 (with $H = H^{-1}$).

For initial conditions $x_0 \in L^2(\Omega; H^{-1})$ satisfying $\mathbb{E} \varphi(x_0) < \infty$ we prove the existence of strong solutions to (3.1). Moreover, we will prove regularizing properties with respect to the initial condition due to the subgradient structure of the drift. This allows to characterize solutions for initial conditions $x_0 \in L^2(\Omega; H^{-1})$ as generalized strong solutions.

**Theorem 3.1.** Let $x_0 \in L^2(\Omega; H^{-1}).$

i. There is a unique generalized strong solution $(X, \eta)$ to (3.1) and $X$ satisfies

$$\mathbb{E} \varphi(X_t) + \mathbb{E} \int_0^t \|\eta_r\|_{H^{-1}}^2 dr \leq C \mathbb{E} \|x_0\|_{H^{-1}}^2 + 1.$$  

ii. If $\mathbb{E} \varphi(x_0) < \infty$, then there is a unique strong solution $(X, \eta)$ to (3.1) satisfying

$$\mathbb{E} \varphi(X_t) + \mathbb{E} \int_0^t \|\eta_r\|_{H^{-1}}^2 dr \leq \mathbb{E} \varphi(x_0) + C,$$

The (generalized) strong solution $(X, \eta)$ coincides with the limit solution constructed in 2.7.

The proof of Theorem 3.1 proceeds in several steps. In particular, the singularity of $\psi$ causes the need for a non-singular regularization of $\psi$. This approximation has to be carefully chosen in order to obtain uniform bounds. Passing to the limit will in turn rely on Mosco-convergence of the regularized potentials. We will first consider the case of non-singular potentials, keeping careful track of the arising constants in Section 3.0.1 the limit will then be taken in Section 3.0.2.

3.0.1. Non-singular potential $\psi$. In this section we restrict to the approximating case of a smooth, non-singular nonlinearity $\psi$. We assume that $\psi \in C^3(\mathbb{R}; \mathbb{R}_+)$ is convex with Lipschitz continuous derivatives $\phi = \psi$, $\dot{\phi}$ satisfying $\psi(0) = \phi(0) = 0$ and

$$\phi(r)r \geq c_{\psi} \psi(r) - C_{\psi} \quad \text{ and } \quad \phi(r)r^2 \leq C_{\psi} \psi(r), \quad \forall r \in \mathbb{R},$$  

where $c_k \in H^{-1}$ is an orthonormal basis of $H^{-1}$. Then $B : H^{-1} \to L_2(H^{-1}, H^{-1})$ is Lipschitz continuous, i.e.
for some constants \( C_\psi, c_\psi > 0 \).

We consider the following non-degenerate, non-singular approximation of (3.1) (cf. Appendix B): 
\[
\begin{align*}
    dX^\varepsilon_t &= \varepsilon \Delta X^\varepsilon_t dt + \Delta \phi(X^\varepsilon_t) dt + B(X^\varepsilon_t) dW_t, \\
    X^\varepsilon_0 &= x_0.
\end{align*}
\]

For \( v \in H^{-1} \) we define
\[
\phi^\varepsilon(v) := \begin{cases}
\frac{\varepsilon}{2} \int_\Omega |v|^2 dx + \int_\Omega \psi(v) dx, & \text{for } v \in L^2 \\
+\infty, & \text{otherwise}.
\end{cases}
\]

Note that \( \phi^\varepsilon \in C^1(L^2) \) with Lipschitz continuous derivative given by
\[
D\phi^\varepsilon(v)(h) = \varepsilon \int_\Omega vhdx + \int_\Omega \phi(v)hdx.
\]

To check the claimed continuity we note that
\[
D\phi^\varepsilon(v)(h) - D\phi^\varepsilon(w)(h) = \varepsilon \int_\Omega (v - w)hdx + \int_\Omega (\phi(v) - \phi(w))hdx
\]
\[
\leq \varepsilon \| v - w \|_2 \| h \|_2 + \| \phi \|_{Lip} \| v - w \|_2 \| h \|_2
\]
\[
\lesssim (\varepsilon + 1) \| v - w \|_2 \| h \|_2.
\]

Moreover, for \( k \in \mathbb{N} \) large enough, we have \( \phi^\varepsilon \in C^2(H^1_0 \cap H^{2k+1}) \) with Lipschitz continuous second derivative given by
\[
D^2\phi^\varepsilon(v)(g, h) = \varepsilon \int_\Omega hgdx + \int_\Omega h\dot{\phi}(v)gdx.
\]

Indeed:
\[
D^2\phi^\varepsilon(v)(g, h) - D^2\phi^\varepsilon(w)(g, h) = \int_\Omega h \left( \dot{\phi}(v) - \dot{\phi}(w) \right) gdx
\]
\[
\leq \| h \|_3 \| \dot{\phi}(v) - \dot{\phi}(w) \|_3 \| g \|_3
\]
\[
\leq \| \dot{\phi} \|_{Lip} \| h \|_{H^{2k+1}} \| v - w \|_{H^{2k+1}} \| g \|_{H^{2k+1}},
\]

where we used the Sobolev embedding \( H^{2k+1} \hookrightarrow L^3 \) for \( k \in \mathbb{N} \) large enough.

Moreover, \( \phi^\varepsilon \) is a convex, lower-semicontinuous function on \( H^{-1} \) with subgradient given by
\[
A^\varepsilon(v) := -\partial \phi^\varepsilon(v) = \varepsilon \Delta v + \Delta \phi(v) \in H^{-1}, \quad \text{for } v \in H^1_0.
\]

Hence, we may write (3.4) as
\[
dX^\varepsilon_t = -\partial \phi^\varepsilon(X^\varepsilon_t) dt + B(X^\varepsilon_t) dW_t.
\]

(Generalized) strong solutions to (3.4) are then defined according to Definition A.1.

By Lemma B.1, for each \( x_0 \in L^2(\Omega; H^{-1}) \) there is a unique variational solution \( X^\varepsilon \) to (3.4) with respect to the Gelfand triple
\[
L^2 \hookrightarrow H^{-1} \hookrightarrow (L^2)^*.
\]

Due to Lemma B.1, \( X^\varepsilon \) is a strong solution if \( x_0 \in L^2(\Omega; L^2) \).

\textbf{Lemma 3.2.} Let \( x_0 \in L^2(\Omega; L^2) \). For each \( \varepsilon > 0 \) we have \( \phi^\varepsilon(X^\varepsilon) \in L^1([0, T] \times \Omega) \) with
\[
\mathbb{E} \int_0^T \phi^\varepsilon(X^\varepsilon_t) dt \leq C(\mathbb{E}\| x_0 \|_{H^{-1}}^2 + 1),
\]
for some constant \( C \) independent of \( \varepsilon > 0 \) and depending on \( \psi \) via the constants \( c_\psi, C_\psi \) only.
Proof. Note that, using (3.3)
\[
\langle A^\varepsilon(v), v \rangle_{H^{-1}} = - \int_{\Omega} (|v|^2 + \phi(v)v) \, dx
\]
\[
\leq - \int_{\Omega} (|v|^2 + c\psi(v) - C) \, dx
\]
\[
\leq -c\varphi^\varepsilon(v) + C,
\]
for all \( v \in H^1_0 \). By Itô’s formula we have
\[
\mathbb{E} e^{-Kt} \|X^\varepsilon_t\|_{H^{-1}}^2
= \mathbb{E} \|x_0\|_{H^{-1}}^2 + 2 \mathbb{E} \int_0^t e^{-Kr} \langle A^\varepsilon(X^\varepsilon_r), X^\varepsilon_r \rangle_{H^{-1}} + e^{-Kr} \|B(X^\varepsilon_r)\|_{L^2(U, H^{-1})}^2 \, dr
- K \int_0^t e^{-Kr} \|X^\varepsilon_r\|_{H^{-1}}^2 \, dr
\leq \mathbb{E} \|x_0\|_{H^{-1}}^2 - 2 \mathbb{E} \int_0^t e^{-Kr} \varphi^\varepsilon(X^\varepsilon_r) + Ce^{-Kr} \|X^\varepsilon_r\|_{H^{-1}}^2 \, dr
- K \int_0^t e^{-Kr} \|X^\varepsilon_r\|_{H^{-1}}^2 \, dr + C.
\]
Choosing \( K \) large enough yields the claim. \( \square \)

Based on the strong solution property of \( X^\varepsilon \) we derive the key estimate in the following

**Lemma 3.3.** Let \( x_0 \in L^2(\Omega; L^2) \). For all \( \varepsilon > 0 \) we have
\[
\mathbb{E}t \varphi^\varepsilon(X^\varepsilon_t) + \mathbb{E} \int_0^t \|\varepsilon X^\varepsilon_r + \phi(X^\varepsilon_r)\|_{H^1_0}^2 \, dr \leq C \left( \mathbb{E} \|x_0\|_{H^{-1}}^2 + 1 \right).
\]

and
\[
\mathbb{E} \varphi^\varepsilon(X^\varepsilon_t) + \mathbb{E} \int_0^t \|\varepsilon X^\varepsilon_r + \phi(X^\varepsilon_r)\|_{H^1_0}^2 \, dr \leq C \mathbb{E} \varphi^\varepsilon(x_0),
\]

for some constant \( C \) independent of \( \varepsilon > 0 \) and depending on \( \psi \) via the constants \( c_\psi, C_\psi \) only.

**Proof.** Let \( J^\lambda \) be the resolvent of \(-\Delta\) on \( H^{-1} \). We let \( G^\lambda = G^{\lambda,k} := J^\lambda \circ \cdots \circ J^\lambda \) the \( k \)-th iteration of \( J^\lambda \). Then \( G^\lambda : H^{-1} \to H^1_0 \cap H^{2k} \) is linear and continuous. Moreover,
\[
\|G^\lambda v\|_2 \leq \|v\|_2, \quad \forall v \in L^2
\]
\[
\|G^\lambda v\|_{H^1_0} \leq \|v\|_{H^1_0}, \quad \forall v \in H^1_0
\]
and
\[
\|G^\lambda v - v\|_2 \leq \|J^\lambda G^{\lambda,k-1} v - J^\lambda v\|_2 + \|J^\lambda v - v\|_2
\]
\[
\leq \|G^{\lambda,k-1} v - v\|_2 + \|J^\lambda v - v\|_2.
\]
Iterating this yields
\[
\|G^\lambda v - v\|_2 \leq k \|J^\lambda v - v\|_2 \to 0 \quad \text{for } \lambda \to 0
\]
for all \( v \in L^2 \). Analogously,
\[
\|G^\lambda v - v\|_{H^1_0} \to 0 \quad \text{for } \lambda \to 0
\]
for all \( v \in H^1_0 \).
We define $\varphi^\varepsilon,\lambda := \varphi^\varepsilon \circ G^\lambda$. Since $G^\lambda : H^{-1} \to H_0^1 \cap H^{2k+1}$ is a linear, continuous operator we have $\varphi^\varepsilon,\lambda \in C^2(H^{-1})$ with Lipschitz continuous derivatives given by
\[
D\varphi^\varepsilon,\lambda(v)(h) = -\varepsilon A G^\lambda v + \Delta\phi(G^\lambda v), G^\lambda h)_{H^{-1}},
\]
\[
D^2\varphi^\varepsilon,\lambda(v)(g, h) = -\varepsilon \int_\Omega (G^\lambda h)(G^\lambda g)dx + \int_\Omega (G^\lambda h)\phi(G^\lambda v)(G^\lambda g)dx.
\]
By Lemma B.1 there is a unique strong solution $X^\varepsilon$ to (3.4) with $X^\varepsilon \in L^2(\Omega; L^\infty([0, T]; L^2)) \cap L^2([0, T] \times \Omega; H_0^1)$. We apply Itô’s formula to $t\varphi^\varepsilon,\lambda(X^\varepsilon_t)$ to get:
\[
\mathbb{E}t\varphi^\varepsilon,\lambda(X^\varepsilon_t)
\]
\[
= -\varepsilon \int_0^t \left(\varepsilon A G^\lambda X^\varepsilon_r + \Delta\phi(G^\lambda X^\varepsilon_r), G^\lambda(\varepsilon A X^\varepsilon_r + \Delta\phi(X^\varepsilon_r))\right)_{H^{-1}}dr
\]
\[
+ \frac{1}{2} \sum_{k=1}^\infty \mathbb{E} \int_0^t r \int_\Omega |G^\lambda(g^k X^\varepsilon_r)|^2 dx dr
\]
\[
+ \frac{1}{2} \sum_{k=1}^\infty \mathbb{E} \int_0^t r \int_\Omega G^\lambda(g^k X^\varepsilon_r)\phi(G^\lambda X^\varepsilon_r)G^\lambda(g^k X^\varepsilon_r)dx dr
\]
\[
+ \mathbb{E} \int_0^t \varphi^\varepsilon,\lambda(X^\varepsilon_r)dr.
\]
We first note that
\[
\int_\Omega |G^\lambda(g^k X^\varepsilon_r)|^2 dx \leq \int_\Omega |g^k X^\varepsilon_r|^2 dx
\]
\[
\leq \|g^k\|_{L^\infty}^2 \|X^\varepsilon_r\|_2^2.
\]
Moreover,
\[
\int_\Omega G^\lambda(g^k X^\varepsilon_r)\phi(G^\lambda X^\varepsilon_r)G^\lambda(g^k X^\varepsilon_r)dx = \int_\Omega (G^\lambda(g^k X^\varepsilon_r))^2 \phi(G^\lambda X^\varepsilon_r)dx
\]
\[
\leq \|\phi\|_{L^\infty} \|g^k\|_{L^\infty}^2 \|X^\varepsilon_r\|_2^2.
\]
We note $G^\lambda v \to v$ in $H_0^1$ for $v \in H_0^1$. Since $\phi$ is Lipschitz we have $\phi(G^\lambda X^\varepsilon) \to \phi(X^\varepsilon)$ in $L^\infty([0, T] \times \Omega; H_0^1)$. Moreover, $G^\lambda g^k X^\varepsilon \to g^k X^\varepsilon$ in $L^2([0, T] \times \Omega; H_0^1)$. Using (3.3) this implies
\[
\lim_{\lambda \to 0} \mathbb{E} \int_0^t r \int_\Omega G^\lambda(g^k X^\varepsilon_r)\phi(G^\lambda X^\varepsilon_r)G^\lambda(g^k X^\varepsilon_r)dx dr
\]
\[
= \mathbb{E} \int_0^t r \int_\Omega (g^k X^\varepsilon_r)^2 \phi(X^\varepsilon_r)dx dr
\]
\[
\leq \|g^k\|_{L^\infty}^2 C_\varepsilon \mathbb{E} \int_0^t \int_\Omega \psi(X^\varepsilon_r)dx dr
\]
\[
= \|g^k\|_{L^\infty}^2 C_\varepsilon \mathbb{E} \int_0^t r \varphi(X^\varepsilon_r)dr.
\]
Hence, by dominated convergence we obtain
\[
\lim_{\lambda \to 0} \frac{1}{2} \sum_{k=1}^\infty \sum_{r=1}^\infty \int_0^t r \int_\Omega |G^\lambda(g^k X^\varepsilon_r)|^2 dx dr
\]
\[
+ \lim_{\lambda \to 0} \frac{1}{2} \sum_{k=1}^\infty \sum_{r=1}^\infty \mathbb{E} \int_0^t r \int_\Omega G^\lambda(g^k X^\varepsilon_r)\phi(G^\lambda(X^\varepsilon_r))G^\lambda(g^k X^\varepsilon_r) dx dr
\]
\[
\leq C \left( 1 + \mathbb{E} \int_0^t r \varphi(X^\varepsilon_r)dr \right).
\]
for some constant $C$ independent of $\epsilon > 0$ and depending on $\psi$ via the constants $c_\psi, C_\psi$ only. We note that
\[
(\varepsilon \Delta G^\lambda X^\varepsilon + \Delta \phi(G^\lambda X^\varepsilon), G^\lambda(\varepsilon \Delta X^\varepsilon + \Delta \phi(X^\varepsilon)))_{H} = (\varepsilon G^\lambda X^\varepsilon + \phi(G^\lambda X^\varepsilon), G^\lambda(\varepsilon X^\varepsilon + \phi(X^\varepsilon)))_{H}.
\]
Since $X^\varepsilon \in L^2([0, T] \times \Omega; H^1_0)$ we have (using dominated convergence)
\[
G^\lambda X^\varepsilon \rightarrow X^\varepsilon \text{ in } L^2([0, T] \times \Omega; H^1_0) \text{ for } \lambda \rightarrow 0.
\]
Since $\phi, \phi$ are Lipschitz continuous this implies $\phi(G^\lambda X^\varepsilon) \rightarrow \phi(X^\varepsilon)$ for $\lambda \rightarrow 0$ in $L^2([0, T] \times \Omega; H^1_0)$. Hence, we obtain
\[
\lim_{\lambda \rightarrow 0} - \mathbb{E} \int_0^t r(\varepsilon G^\lambda X^\varepsilon + \phi(G^\lambda X^\varepsilon), G^\lambda(\varepsilon X^\varepsilon + \phi(X^\varepsilon)))_{H}dr = - \mathbb{E} \int_0^t r\|\varepsilon X^\varepsilon + \phi(X^\varepsilon)\|^2_{H}dr.
\]
Since $\varphi^\varepsilon$ is continuous on $L^2$,
\[
|\varphi^\varepsilon(v)| \leq C(1 + \|v\|^2_2)
\]
and $G^\lambda X^\varepsilon \rightarrow X^\varepsilon$ in $L^2([0, T] \times \Omega; H^1_0)$, $G^\lambda X^\varepsilon \rightarrow X^\varepsilon$ in $L^2(\Omega; L^2)$ for all $t \in [0, T]$, by dominated convergence we get
\[
\lim_{\lambda \rightarrow 0} \mathbb{E} t\varphi^\varepsilon(X^\varepsilon_t) = \mathbb{E} t\varphi(X^\varepsilon_t)
\]
and
\[
\lim_{\lambda \rightarrow 0} \mathbb{E} \int_0^t \varphi^\varepsilon(X^\varepsilon_s)ds = \mathbb{E} \int_0^t \varphi(X^\varepsilon_s)ds.
\]
Putting these estimates together yields
\[
\mathbb{E} t\varphi(X^\varepsilon_t) \leq - \mathbb{E} \int_0^t r\|\varepsilon X^\varepsilon + \phi(X^\varepsilon)\|^2_{H}dr + C \left(1 + \mathbb{E} \int_0^t r\varphi(X^\varepsilon)dr\right) + \mathbb{E} \int_0^t \varphi(X^\varepsilon)ds.
\]
By Lemma 3.2 we conclude
\[
\mathbb{E} t\varphi(X^\varepsilon_t) + \mathbb{E} \int_0^t r\|\varepsilon X^\varepsilon + \phi(X^\varepsilon)\|^2_{H}dr \leq C \left(\mathbb{E}\|x_0\|^2_{H} + 1\right),
\]
for some constant $C$ independent of $\varepsilon > 0$ and depending on $\psi$ via the constants $c_\psi, C_\psi$ only.

To prove (3.5) we proceed as above but applying Itô’s formula for $\varphi^\varepsilon(X^\varepsilon_t)$ instead of $t\varphi^\varepsilon(X^\varepsilon_t)$.  

3.0.2. Proof of Theorem 3.1

Proof of Theorem 3.1. The proof proceeds via a three-step approximation. First, we approximate the singular potential $\psi(r) = \frac{1}{m+1}(|r|^{m+1}-1)$ by the smooth functions
\[
\psi^\delta(r) = \frac{1}{m+1} (r^2 + \delta) \left(\frac{m+1}{2} - \frac{\delta}{r^2}\right).
\]
Note
\[
\phi^\delta(r) := (\psi^\delta)'(r) = (r^2 + \delta) \left(\frac{m+1}{2} - \frac{\delta}{r^2}\right) r
\]
\[
\phi^\delta(r) := (r^2 + \delta) \left(\frac{m+1}{2} - \frac{\delta}{r^2}\right) r^2
\]
\[
= (r^2 + \delta) \left(\frac{m+1}{2} - \frac{\delta}{m+1}\right) (\delta + mr^2).
\]
Thus, for $\delta \in [0, 1]$  
\[
\phi^\delta(r)r = (r^2 + \delta)^{\frac{m+1}{m+3}} r^2 = (r^2 + \delta)^{\frac{m+1}{m+3} - (r^2 + \delta)^{\frac{m+1}{m+3}} - (m+1)\psi^\delta(r) - 1
\]
and  
\[
\phi^\delta(r)r^2 = (r^2 + \delta)^{\frac{m+3}{m+3} - (m + \delta r^2)r^2}
\]
\[
\leq (m + 1) (r^2 + \delta)^{\frac{m+1}{m+3}}
\]
\[= (m + 1)^2 \psi^\delta(r).
\]
That is, (3.3) is satisfied with constants $c_\phi, C_\psi$ independent of $\delta > 0$. Moreover, we observe  
\[
(3.6) \quad |\psi^\delta(r) - \psi(r)| = \frac{2}{m + 1} \delta^{\frac{m+1}{m+3}}.
\]
Next we consider vanishing viscosity  
\[
(3.7) \quad dX^\varepsilon\delta_t = \varepsilon \Delta X^\varepsilon\delta_t dt + \Delta \phi^\delta(X^\varepsilon\delta_t)dt + B(X^\varepsilon\delta_t)dW_t,
\]
\[X^\varepsilon\delta_0 = x_0.
\]
In a third approximating step we consider smooth approximations of the initial condition, i.e. we first assume $x_0 \in L^2(\Omega; L^2)$.

**Step 1: $\delta \to 0$**

Let $\varepsilon > 0$, $x_0 \in L^2(\Omega; L^2)$ and for $v \in H^{-1}$ set  
\[
\varphi^\varepsilon\delta(v) := \begin{cases} \frac{\varepsilon}{2} \int_{\Omega} |v|^2 dx + \int_{\partial \Omega} \psi^\delta(v) dx, & \text{for } v \in L^2 \\ +\infty, & \text{otherwise}, \end{cases}
\]
Since  
\[
\partial \varphi^\varepsilon\delta(v) = \begin{cases} -\varepsilon \Delta v - \Delta \phi^\delta(v), & \text{for } v \in H^1_0 \\ 0, & \text{otherwise}, \end{cases}
\]
we may rewrite (3.7) as  
\[
dX^\varepsilon\delta_t = -\partial \varphi^\varepsilon\delta(X^\varepsilon\delta_t)dt + B(X^\varepsilon\delta_t)dW_t.
\]
and define strong solutions to (3.7) as in Definition A.2. By Lemma B.1 there is a (unique) strong solution $X^\varepsilon\delta_t$ to (3.7). From Lemma B.1 and Lemma 3.3 we have  
\[
(3.8) \quad \mathbb{E} \sup_{t \in [0, T]} \|X^\varepsilon\delta_t\|_2^2 + 2\varepsilon \mathbb{E} \int_0^T \|X^\varepsilon\delta_t\|_{H^1_0}^2 dt \leq C(\mathbb{E}\|x_0\|_2^2 + 1)
\]
\[
\mathbb{E} \int_0^t \|\varepsilon X^\varepsilon\delta_t + \phi^\delta(X^\varepsilon\delta_t)\|_{H^1_0}^2 dt \leq C(\mathbb{E}\|x_0\|_{H^{-1}}^2 + 1)
\]
and  
\[
(3.9) \quad \mathbb{E} \varphi^\varepsilon\delta(X^\varepsilon\delta_t) + \mathbb{E} \int_0^t \|\varepsilon X^\varepsilon\delta_t + \phi^\delta(X^\varepsilon\delta_t)\|_{H^1_0}^2 dt \leq C\mathbb{E}\varphi^\varepsilon\delta(x_0)
\]
\[\leq C(\mathbb{E}\|x_0\|_2^2 + 1),
\]
with a constant $C > 0$ independent of $\varepsilon, \delta > 0$. Hence, we may extract a sub-sequence $\delta_n \to 0$ such that  
\[
X^\varepsilon\delta_n \to X^\varepsilon, \quad \text{in } L^2([0, T] \times \Omega; H^1_0)
\]
\[
X^\varepsilon\delta_n \to X^\varepsilon, \quad \text{in } L^2(\Omega; L^\infty([0, T]; L^2))
\]
\[
\Delta \phi^\delta_n(X^\varepsilon\delta_n) \to \eta^\varepsilon, \quad \text{in } L^2([0, T] \times \Omega; H^{-1}).
\]
Since $B : H^{-1} \to L_2(H^{-1}; H^{-1})$ is continuous, linear we have $B(\epsilon) \to B(\epsilon)$ in $L^2([0, T] \times \Omega; L_2(H^{-1}; H^{-1}))$ and thus $dt \otimes d\mathbb{P}$-almost everywhere

$$\bar{X}_t^\epsilon = x_0 + \int_0^t (\epsilon \Delta \bar{X}_r^\epsilon + \eta_r^\epsilon) \, dr + \int_0^t B(\bar{X}_r^\epsilon) \, dW_r.$$ Hence, defining

$$X_t^\epsilon := x_0 + \int_0^t (\epsilon \Delta X_r^\epsilon + \eta_r^\epsilon) \, dr + \int_0^t B(X_r^\epsilon) \, dW_r,$$

we have $X^\epsilon = \bar{X}^\epsilon$, $dt \otimes d\mathbb{P}$-almost everywhere and due to [13, Theorem 4.2.5] we have $X^\epsilon \in L^2(\Omega; C([0, T]; H^{-1}))$. We aim to prove that $X^\epsilon$ is a strong solution to

$$dX^\epsilon_t = -\partial \varphi^\epsilon(X^\epsilon_t) \, dt + B(X^\epsilon_t) \, dW_t,$$

where (for $v \in H^{-1}$)

$$\varphi^\epsilon(v) := \begin{cases} \frac{\epsilon}{2} \int_\Omega |v|^2 \, dx + \int_\Omega \psi(v) \, dx, & \text{for } v \in L^2 \\ \infty, & \text{otherwise.} \end{cases}$$

Note that

$$\partial \varphi^\epsilon(v) = \begin{cases} \{\varphi^\epsilon(v) \mid \varphi \in H_0^1 \text{ with } \varphi \in \phi(v) \text{ a.e. }\}, & \text{for } v \in H_0^1 \\ \emptyset, & \text{otherwise.} \end{cases}$$

It thus remains to identify $\epsilon \Delta X^\epsilon + \eta^\epsilon : -\partial \varphi^\epsilon(X^\epsilon)$, $dt \otimes d\mathbb{P}$-almost everywhere. Itô’s formula yields

$$Ee^{-Ct} \|X^\epsilon_t\|^2_{H^{-1}} = E\|x_0\|^2_{H^{-1}} + E \int_0^t e^{-Cr}(\epsilon \Delta X_r^\epsilon + \eta_r^\epsilon, X_r^\epsilon)_{H^{-1}} \, dr$$

$$+ E \int_0^t e^{-Cr} \|B(X_r^\epsilon)\|^2_{L_2} \, dr - CE \int_0^t e^{-Cr} \|X_r^\epsilon\|^2_{H^{-1}} \, dr \quad \forall t \in [0, T].$$

Using Itô’s formula for $X^\epsilon, \delta$ yields (for $C > 0$ large enough)

$$Ee^{-Ct} \|X^\epsilon, \delta_n\|^2_{H^{-1}}$$

$$= E\|x_0\|^2_{H^{-1}} + E \int_0^t e^{-Cr}(\epsilon \Delta X_r^\epsilon, \delta_n + \Delta \phi^\delta_n(X_r^\epsilon, \delta_n), X_r^\epsilon, \delta_n)_{H^{-1}} \, dr$$

$$+ E \int_0^t e^{-Cr} \|B(X_r^\epsilon, \delta_n) - B(X_r^\epsilon)\|^2_{L_2} \, dr - E \int_0^t e^{-Cr} \|B(X_r^\epsilon)\|^2_{L_2} \, dr$$

$$+ 2E \int_0^t e^{-Cr} \|B(X_r^\epsilon, \delta_n), B(X_r^\epsilon)\|_{L_2} \, dr - CE \int_0^t e^{-Cr} \|X_r^\epsilon, \delta_n - X_r^\epsilon\|^2_{H^{-1}} \, dr$$

$$+ CE \int_0^t e^{-Cr} \|X_r^\epsilon\|^2_{H^{-1}} \, dr - 2CE \int_0^t e^{-Cr} \|X_r^\epsilon, \delta_n, X_r^\epsilon\|_{H^{-1}} \, dr$$

$$\leq E\|x_0\|^2_{H^{-1}} + E \int_0^t e^{-Cr}(\epsilon \Delta X_r^\epsilon, \delta_n + \Delta \phi^\delta_n(X_r^\epsilon, \delta_n), X_r^\epsilon, \delta_n)_{H^{-1}} \, dr$$

$$- E \int_0^t e^{-Cr} \|B(X_r^\epsilon)\|^2_{L_2} \, dr + 2E \int_0^t e^{-Cr} \|B(X_r^\epsilon, \delta_n), B(X_r^\epsilon)\|_{L_2} \, dr$$

$$+ CE \int_0^t e^{-Cr} \|X_r^\epsilon\|^2_{H^{-1}} \, dr - 2CE \int_0^t e^{-Cr} \|X_r^\epsilon, \delta_n, X_r^\epsilon\|_{H^{-1}} \, dr.$$
Taking \( \liminf_{n \to 0} \) we obtain (first in distributional sense in \( t \) then a.e. by the Lebesgue Theorem)

\[
Ee^{-Ct}\|X^\varepsilon_t\|_{H^{-1}}^2 \\
\leq \liminf_{\delta_n \to 0} Ee^{-Ct}\|X^{\varepsilon, \delta_n}_t\|_{H^{-1}}^2
\]

\[
\leq E\|x_0\|_{H^{-1}}^2 + \liminf_{\delta_n \to 0} E\int_0^t e^{-Cr}(\varepsilon \Delta X^\varepsilon_{\delta_n} + \Delta \phi_{\delta_n}(X^\varepsilon_{\delta_n}), X^\varepsilon_{\delta_n})_{H^{-1}} \, dr
\]

\[
- \mathbb{E} \int_0^t e^{-Cr}\|B(X^\varepsilon_t)\|_{L^2}^2 \, dr + 2\mathbb{E} \int_0^t e^{-Cr}(B(X^\varepsilon_t), B(X^\varepsilon_t))_{L^2} \, dr
\]

\[
+ \mathbb{C} \mathbb{E} \int_0^t e^{-Cr}\|X^\varepsilon_t\|_{H^{-1}}^2 \, dr - 2\mathbb{C} \mathbb{E} \int_0^t e^{-Cr}(X^\varepsilon_t, X^\varepsilon_t)_{H^{-1}} \, dr
\]

\[
\leq E\|x_0\|_{H^{-1}}^2 + \liminf_{\delta_n \to 0} E\int_0^t e^{-Cr}(\varepsilon \Delta X^\varepsilon_{\delta_n} + \Delta \phi_{\delta_n}(X^\varepsilon_{\delta_n}), X^\varepsilon_{\delta_n})_{H^{-1}} \, dr
\]

\[
+ \mathbb{E} \int_0^t e^{-Cr}\|B(X^\varepsilon_t)\|_{L^2}^2 \, dr - \mathbb{C} \mathbb{E} \int_0^t e^{-Cr}\|X^\varepsilon_t\|_{H^{-1}}^2 \, dr \quad \text{a.e. } t \in [0,T].
\]

Subtracting (3.12) we obtain

\[
E\int_0^t e^{-Cr}(\varepsilon \Delta X^\varepsilon_t + \eta^\varepsilon_t, X^\varepsilon_t)_{H^{-1}} \, dr
\]

\[
\leq \liminf_{\delta_n \to 0} E\int_0^t e^{-Cr}(\varepsilon \Delta X^\varepsilon_{\delta_n} + \Delta \phi_{\delta_n}(X^\varepsilon_{\delta_n}), X^\varepsilon_{\delta_n})_{H^{-1}} \, dr.
\]

We now consider the convex, lower-semicontinuous functionals \( \tilde{\varphi}^\varepsilon, \tilde{\varphi}^{\varepsilon, \delta} : L^2([0,T] \times \Omega; H^{-1}) \to \mathbb{R} \) defined by

\[
\tilde{\varphi}^\varepsilon(v) := \begin{cases} 
\mathbb{E} \int_0^T e^{-Cr} \int_\Omega (\varepsilon \psi(v) + \tilde{\varphi}(v)) \, dx \, dr, & \text{if } v \in L^2([0,T] \times \Omega; L^2) \\
+\infty, & \text{otherwise}
\end{cases}
\]

\[
= \mathbb{E} \int_0^T e^{-Cr} \varphi^\varepsilon(v_r) \, dr.
\]

and \( \tilde{\varphi}^{\varepsilon, \delta} \) being defined analogously, where we endow \( L^2([0,T] \times \Omega; H^{-1}) \) with the equivalent norm

\[
\|v\|^2_{L^2([0,T] \times \Omega; H^{-1})} := \mathbb{E} \int_0^T e^{-Cr}\|v_r\|_{H^{-1}}^2 \, dr.
\]

Due to the characterization of subgradients of integral functionals proved in [15, Theorem 21] we have

\[
\partial \varphi^\varepsilon(v) = \{\eta \in L^2([0,T] \times \Omega; H^{-1}) | \eta \in \partial \varphi^\varepsilon(v), dt \otimes \mathcal{P}\text{-a.e.} \}
\]

and

\[
\partial \varphi^{\varepsilon, \delta}(v) = \{-\varepsilon \Delta v - \Delta \phi^\delta(v)\} \quad \text{for } v \in L^2([0,T] \times \Omega; H^0_1).
\]

Since \( \varphi^{\varepsilon, \delta} \to \varphi^\varepsilon \) uniformly for \( \delta \to 0 \) (cf. (3.6)), we also have \( \varphi^{\varepsilon, \delta} \to \varphi^\varepsilon \) in Mosco sense. Due to (3.15) we have

\[
(-\varepsilon \Delta X^{\varepsilon, \delta_n} - \Delta \phi_{\delta_n}(X^{\varepsilon, \delta_n}), Y - X^{\varepsilon, \delta_n})_{L^2([0,T] \times \Omega; H^{-1})} + \varphi^{\varepsilon, \delta}(X^{\varepsilon, \delta_n}) \leq \varphi^{\varepsilon, \delta_n}(Y),
\]

for all \( Y \in L^2([0,T] \times \Omega; H^{-1}) \). Using (3.13) and Mosco convergence of \( \varphi^{\varepsilon, \delta} \) to \( \varphi^\varepsilon \)

we may take the \( \liminf_{n \to \infty} \) to get

\[
(-\varepsilon \Delta X^{\varepsilon} - \eta^\varepsilon, Y - X^{\varepsilon})_{L^2([0,T] \times \Omega; H^{-1})} + \varphi^\varepsilon(X^{\varepsilon}) \leq \varphi^\varepsilon(Y).
\]
Hence, \( \varepsilon \Delta X^\varepsilon + \eta^\varepsilon \in -\partial \varphi^\varepsilon (X^\varepsilon) \) and we conclude \( \varepsilon \Delta X^\varepsilon + \eta^\varepsilon \in -\partial \varphi^\varepsilon (X^\varepsilon) \) dt \( \otimes \) d\P - almost everywhere due to (3.14). Then, (3.11) yields
\[
\eta^\varepsilon = \Delta \zeta^\varepsilon
\]
with \( \zeta^\varepsilon \in H^1_0 \) and \( \zeta^\varepsilon \in \phi(X^\varepsilon) \) a.e. In conclusion, \( X^\varepsilon \) is a strong solution to (3.10).

Passing to the limit in (3.8), (3.9) yields
\[
\begin{align*}
(3.17) \quad & \mathbb{E} \sup_{t \in [0,T]} \|X^\varepsilon_t\|^2_2 + 2\varepsilon \mathbb{E} \int_0^T \|X^\varepsilon_r\|^2_{H^2_0} dr \leq C(\mathbb{E}\|x_0\|^2_2 + 1) \\
& \quad \mathbb{E} t \varphi^\varepsilon (X^\varepsilon_t) + \mathbb{E} \int_0^t \|\varepsilon \Delta X^\varepsilon + \eta^\varepsilon\|^2_{H^{-1}} dr \leq C \left( \mathbb{E}\|x_0\|^2_2 + 1 \right)
\end{align*}
\]
and
\[
(3.18) \quad \mathbb{E} \varphi^\varepsilon (X^\varepsilon_t) + \mathbb{E} \int_0^t \varepsilon \Delta X^\varepsilon + \eta^\varepsilon\|^2_{H^{-1}} dr \leq \mathbb{E} \varphi^\varepsilon (x_0) \leq C(\mathbb{E}\|x_0\|^2_2 + 1).
\]

**Step 2:** \( \varepsilon \to 0 \)

For \( \varepsilon_1, \varepsilon_2 > 0 \) let \( (X^{\varepsilon_1}, \eta^{\varepsilon_1}), (X^{\varepsilon_2}, \eta^{\varepsilon_2}) \) be two strong solutions to (3.10) with initial conditions \( x^\varepsilon_0, x^\varepsilon_0 \in L^2(\Omega; L^2) \) respectively. Itô’s formula implies
\[
e^{-Kt}\|X^\varepsilon_{t-} - X^{\varepsilon_2}_{t-}\|^2_{H^{-1}} = \|x^\varepsilon_0 - x^{\varepsilon_2}_{t-}\|^2_{H^{-1}} + \int_0^t 2e^{-Kr}(\varepsilon_1 \Delta X^{\varepsilon_1} + \eta^{\varepsilon_1} - (\varepsilon_2 \Delta X^{\varepsilon_2} + \eta^{\varepsilon_2}), X^{\varepsilon_1}_r - X^{\varepsilon_2}_r)_{H^{-1}} dr
+ \int_0^t e^{-Kr}(X^{\varepsilon_1}_r - X^{\varepsilon_2}_r, B(X^{\varepsilon_1}_r) - B(X^{\varepsilon_2}_r))_{H^{-1}} dW_r
+ \int_0^t e^{-Kr}\|B(X^{\varepsilon_1}_r) - B(X^{\varepsilon_2}_r)\|^2_{L^2} dr - KE \int_0^t e^{-Kr}\|X^{\varepsilon_1}_r - X^{\varepsilon_2}_r\|^2_{H^{-1}} dr.
\]
Due to (3.16) we have
\[
(\varepsilon^{\varepsilon_1} - \varepsilon^{\varepsilon_2}, X^{\varepsilon_1} - X^{\varepsilon_2})_{H^{-1}} \leq 0
\]
and we note that
\[
(\varepsilon_1 \Delta X^{\varepsilon_1} - \varepsilon_2 \Delta X^{\varepsilon_2}, X^{\varepsilon_1} - X^{\varepsilon_2})_{H^{-1}} \leq 2(\varepsilon_1 + \varepsilon_2)(\|X^{\varepsilon_1}\|^2_2 + \|X^{\varepsilon_2}\|^2_2).
\]
Hence, the Burkholder-Davis-Gundy inequality and Lemma [3.1] imply
\[
(3.19) \quad \mathbb{E} \sup_{t \in [0,T]} \|X^{\varepsilon_1}_t - X^{\varepsilon_2}_t\|^2_{H^{-1}} \leq C \mathbb{E}\|x_0\|^2_2 - \mathbb{E}\|x_0\|^2_2 + 1
+ C(\varepsilon_1 + \varepsilon_2) \left( \mathbb{E}\|x_0\|^2_2 + \mathbb{E}\|x_0\|^2_2 + 1 \right).
\]
Let now \( x_0 \in L^2(\Omega; L^2) \) and for each \( \varepsilon > 0 \) let \( (X^\varepsilon, \eta^\varepsilon) \) be a solution to (3.10) with initial condition \( x_0 \). Due to (3.19) there is an \( \mathcal{F}_t \)-adapted process \( X \in L^2([0,T] \times \Omega; H^{-1}) \) with \( X_0 = x_0 \) such that
\[
X^\varepsilon \to X \quad \text{in} \quad L^2(\Omega; C([0,T]; H^{-1})) \quad \text{for} \quad \varepsilon \to 0.
\]
Using (3.18) we can extract a weakly convergent subsequence
\[
\varepsilon_n \Delta X^{\varepsilon_n} + \eta^{\varepsilon_n} \rightharpoonup \eta, \quad \text{in} \quad L^2([0,T] \times \Omega; H^{-1}).
\]
By step one we have \( \varepsilon_n \Delta X^{\varepsilon_n} + \eta^{\varepsilon_n} \in -\partial \varphi^{\varepsilon_n}(X^{\varepsilon_n}) \) a.e., hence
\[
(\varepsilon_n \Delta X^{\varepsilon_n} + \eta^{\varepsilon_n}, X^{\varepsilon_n} - Y)_{L^2([0,T] \times \Omega; H^{-1})} + \varphi^{\varepsilon_n}(X^{\varepsilon_n}) \leq \varphi^{\varepsilon_n}(Y),
\]
for all \( Y \in L^2([0,T] \times \Omega; H^{-1}) \). For \( v \in L^2([0,T] \times \Omega; H^{-1}) \) we define
\[
\varphi(v) = \mathbb{E} \int_0^T \varphi(v_t) dt.
\]
Again, due to [15, Theorem 21] we have
\begin{equation}
\partial \tilde{\varphi}(v) = \{ \eta \in L^2([0, T] \times \Omega; H^{-1}) | \eta \in \partial \varphi(v), \ dt \otimes dP \text{-a.e.} \}.
\end{equation}
For \( v \in L^2 \) we observe that \( \varphi^\varepsilon(v) = \frac{\varepsilon}{2} \| v \|_2^2 + \varphi(v) \) and thus
\begin{equation}
\begin{aligned}
&\langle \varepsilon_n \Delta X^{\varepsilon_n} + \eta^{\varepsilon_n}, X^{\varepsilon_n} - Y \rangle_{L^2([0, T] \times \Omega; H^{-1})} + \varphi(X^{\varepsilon_n}) \\
\leq &\ \tilde{\varphi}(Y) + C \varepsilon_n \| Y \|_{L^2([0, T] \times \Omega; L^2)}^2,
\end{aligned}
\end{equation}
for all \( Y \in L^2([0, T] \times \Omega; L^2) \).

Let \( J^\lambda = (1 - \lambda \Delta)^{-1} \) be the resolvent of \( -\Delta \) on \( H^{-1} \). Then
\[ \| J^\lambda v \|_{H^{-1}} \leq \| v \|_{H^{-1}} \]
and \( J^\lambda v \to v \) in \( H^{-1} \) for \( \lambda \to 0 \). Moreover,
\[ \| J^\lambda v \|_{L^2} \leq \frac{C}{\lambda} \| v \|_{H^{-1}} \]
for all \( v \in H^{-1} \). For \( v \in L^{m+1} \cap H^{-1} \) we have
\[ \varphi(J^\lambda v) = \| J^\lambda v \|_{L^m} \leq \| v \|_{L^m} = \varphi(v). \]
For the case \( m = 0 \), in addition: Let \( v = v_\mu \in \mathcal{M} \cap H^{-1} \). Since \( \varphi \) is the lower-semicontinuous hull of \( \varphi \) restricted to \( L^1 \cap H^{-1} \) (cf. Appendix D), there is a sequence \( v_\mu \in L^1 \cap H^{-1} \) such that \( v_\mu \to v \) in \( H^{-1} \) and \( \varphi(v_\mu) \to \varphi(v) \). By lower-semicontinuity of \( \varphi \) we conclude
\begin{equation}
\begin{aligned}
\varphi(J^\lambda v) &\leq \lim_{n \to \infty} \varphi(J^\lambda v^n) \\
&\leq \lim_{n \to \infty} \varphi(v^n) = \varphi(v), \ \forall v \in \mathcal{M} \cap H^{-1}.
\end{aligned}
\end{equation}
Given \( Y \in L^2([0, T] \times \Omega; H^{-1}) \) we set \( Y^\varepsilon := J^{\varepsilon\frac{4}{3}} Y \in L^2([0, T] \times \Omega; L^2) \). By dominated convergence \( Y^\varepsilon \to Y \) in \( L^2([0, T] \times \Omega; H^{-1}) \). From (3.21) we obtain
\[ \langle \varepsilon_n \Delta X^{\varepsilon_n} + \eta^{\varepsilon_n}, X^{\varepsilon_n} - Y^{\varepsilon_n} \rangle_{L^2([0, T] \times \Omega; H^{-1})} + \tilde{\varphi}(X^{\varepsilon_n}) \]
\[ \leq \tilde{\varphi}(Y^{\varepsilon_n}) + C \varepsilon_n \| Y^{\varepsilon_n} \|_{L^2([0, T] \times \Omega; L^2)}^2. \]
Taking \( n \to \infty \) and using lower semicontinuity of \( \tilde{\varphi} \) we arrive at
\[ \langle \eta, X - Y \rangle_{L^2([0, T] \times \Omega; H^{-1})} + \tilde{\varphi}(X) \leq \tilde{\varphi}(Y) \]
for all \( Y \in L^2([0, T] \times \Omega; H^{-1}) \) and thus \( \eta \in \partial \tilde{\varphi}(X) \), which implies \( \eta \in \partial \varphi(X) \) a.e. by (3.20). In conclusion, \( X \) is a strong solution to
\begin{equation}
dX_t \in -\partial \varphi(X_t)dt + B(X_t)dW_t.
\end{equation}
Taking the limit in (3.17), (3.18) yields
\begin{equation}
\mathbb{E} \sup_{t \in [0, T]} \| X_t \|_2^2 \leq C(\mathbb{E} \| x_0 \|_2^2 + 1)
\end{equation}
\begin{equation}
\mathbb{E} \int_0^t r \| \eta_r \|_{H^{-1}}^2 dr \leq C(\mathbb{E} \| x_0 \|_{H^{-1}}^2 + 1)
\end{equation}
and
\begin{equation}
\mathbb{E} \varphi(X_t) + \mathbb{E} \int_0^t \| \eta_r \|_{H^{-1}}^2 dr \leq C\mathbb{E} \varphi(x_0)
\end{equation}
\[ \leq C(\mathbb{E} \| x_0 \|_2^2 + 1). \]
Moreover, from (3.19) we obtain
(3.26) \[
\mathbb{E} \sup_{t \in [0,T]} \|X_t^1 - X_t^2\|^2_{H^{-1}} \leq C \mathbb{E}\|x_0^1 - x_0^2\|^2_{H^{-1}},
\]
where \(X^1, X^2\) are the corresponding limits for the initial conditions \(x_0^1, x_0^2\) respectively.

**Step 3: Proof of (i)**

Suppose \(x_0 \in L^2(\Omega; H^{-1})\) satisfying \(\mathbb{E}\phi(x_0) < \infty\). We consider the case \(m = 0\), the case \(m > 0\) can be treated analogously. Let \(J^\lambda = (1 - \lambda \Delta)^{-1}\) be the resolvent of \(-\Delta\) on \(H^{-1}\), set \(x_0^n = J^\lambda x_0\) and let \((X^n, \eta^n)\) be the corresponding strong solution to (3.23) constructed in step two. By dominated convergence we have
\[
x_0^n \to x_0, \quad \text{in } L^2(\Omega; H^{-1}).
\]
Due to (3.25), (3.26) and (3.22) we may extract (weakly) convergent subsequences
\[
X^n \to X, \quad \text{in } L^2(\Omega; C([0,T]; H^{-1}))
\]
\[
\eta^n \to \eta, \quad \text{in } L^2([0,T] \times \Omega; H^{-1}).
\]
Since \(\eta^n \in -\partial \phi(X^n)\), strong-weak closedness of the subgradient \(-\partial \phi\) implies \(\eta \in -\partial \phi(X)\) and thus \(\eta \in -\partial \phi(x)\) a.e.. It then easily follows that \((X, \eta)\) is a strong solution to (3.1).

**Step 3: Proof of (ii)**

Let \(x_0 \in L^2(\Omega; H^{-1})\) and \(x_0^n \in L^2(\Omega; L^2)\) with \(x_0^n \to x\) in \(L^2(\Omega; H^{-1})\), \(\mathbb{E}\|x_0^n\|^2_{H^{-1}} \leq \mathbb{E}\|x_0\|^2_{H^{-1}}\) and let \((X^n, \eta^n)\) be the corresponding strong solutions to (3.23) constructed in step two. By (3.26) we have
\[
\mathbb{E} \sup_{t \in [0,T]} \|X^n - X^m\|^2_{H^{-1}} \leq C \mathbb{E}\|x_n - x_m\|^2_{H^{-1}}.
\]
Hence, \(X^n \to X\) in \(L^2(\Omega; C([0,T]; H^{-1}))\). Moreover,
\[
\mathbb{E}\ell \phi(X_t^n) + \mathbb{E} \int_0^t r\|\eta^n\|^2_{H^{-1}} \, dr \leq C \left( \mathbb{E}\|x_0\|^2_{H} + 1 \right).
\]
Hence, there is a map \(\eta\) with \(\eta \in L^2([\tau,T] \times \Omega; H_0^1)\) such that
\[
\eta^n \to \eta, \quad \text{in } L^2([\tau,T] \times \Omega; H^{-1}),
\]
for all \(\tau > 0\). By strong-weak closedness of \(-\partial \phi\) we have \(\eta \in -\partial \phi(X)\) and thus \(\eta \in -\partial \phi(x)\) a.e.. Hence, \(X\) is a generalized strong solution satisfying
\[
\mathbb{E}\ell \phi(X_t) + \mathbb{E} \int_0^t r\|\eta\|^2_{H^{-1}} \, dr \leq C \left( \mathbb{E}\|x_0\|^2_{H^{-1}} + 1 \right).
\]

\[\square\]

**Appendix A. (Generalized) strong solutions to gradient type SPDE**

Let \(\phi : H \to \mathbb{R}\) be a proper, lower-semicontinuous, convex function on a separable real Hilbert space \(H\). We consider SPDE of the type
(A.1) \[
dx_t \in -\partial \phi(X_t) dt + B(t, X_t) dW_t,
\]
\(X_0 = x_0\),

where \(W\) is a cylindrical Wiener process in a separable Hilbert space \(U\) defined on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with normal filtration \((\mathcal{F}_t)_{t \geq 0}\) and \(B : [0,T] \times H \times \Omega \to L_2(U, H)\) is Lipschitz continuous, i.e.
\[
\|B(t, u) - B(t, w)\|^2_{L_2(U, H)} \leq C \|u - w\|^2_H \quad \forall u, w \in H
\]
and all \((t, \omega) \in [0,T] \times \Omega\). Furthermore, we assume that
\[
\|B(\cdot, 0)\|_{L_2(U, H)} \in L^2([0,T] \times \Omega).
\]
We then define

**Definition A.1.** Let $x_0 \in L^2(\Omega; H)$. An $H$-continuous, $\mathcal{F}_t$-adapted process $X \in L^2(\Omega; C([0,T]; H))$ for which there exists a selection $\eta \in -\partial \varphi(X)$, $dt \otimes d\mathbb{P}$-a.e. is said to be a

i. strong solution to (A.1) if

\[
\eta \in L^2([0,T] \times \Omega; H)
\]

and $\mathbb{P}$-a.s.

\[
X_t = x_0 + \int_0^t \eta_r dr + \int_0^t B(r, X_r) dW_r, \quad \forall t \in [0,T].
\]

ii. generalized strong solution to (A.1) if

\[
\eta \in L^2([\tau,T] \times \Omega; H), \quad \forall \tau > 0
\]

and $\mathbb{P}$-a.s.

\[
X_t = X_\tau + \int_\tau^t \eta_r dr + \int_\tau^t B(r, X_r) dW_r, \quad \forall t \in [\tau,T],
\]

for all $\tau > 0$.

**Appendix B.** Non-degenerate, non-singular stochastic fast diffusion equations

In this section we consider non-degenerate, non-singular approximations to (1.1), that is

\[
dX_t = \varepsilon \Delta X_t dt + \Delta \phi(X_t) dt + B(t, X_t) dW_t,
\]

(B.1)

\[
X_0 = x_0,
\]

where $\phi : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous, monotone function satisfying $\phi(0) = 0$. We further assume that $W$ is a cylindrical Wiener process in a separable Hilbert space $U$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t \geq 0}$ and $B : [0,T] \times H^{-1} \times \Omega \to L_2(U, H^{-1})$ is Lipschitz continuous, i.e.

\[
\|B(t, v) - B(t, w)\|_{L_2(U; H^{-1})} \leq C\|v - w\|_{H^{-1}}, \quad \forall v, w \in H^{-1},
\]

for some constant $C > 0$ and all $(t, \omega) \in [0,T] \times \Omega$. We assume

(B.2)

\[
\|B(t, v)\|_{L_2(U; L_2)} \leq C(1 + \|v\|_{L_2}^2), \quad \forall v \in L^2,
\]

and all $(t, \omega) \in [0,T] \times \Omega$. By [13] there is a unique variational solution $X$ to (B.1) with respect to the Gelfand triple

\[
L^2 \hookrightarrow H^{-1} \hookrightarrow (L^2)^*.
\]

Under an additional regularity assumption on the diffusion coefficients $B$ we prove that in fact, these solutions are strong solutions in $H^{-1}$.

**Lemma B.1.** Let $x_0 \in L^2(\Omega; L^2)$. Then

\[
\mathbb{E} \sup_{t \in [0,T]} \|X_t\|_2^2 + \varepsilon \mathbb{E} \int_0^T \|X_r\|_{H_2^0}^2 dr \leq C(\mathbb{E}\|x_0\|_2^2 + 1),
\]

with a constant $C > 0$ independent of $\varepsilon$ and $\phi$.

**Proof.** In the following we let $(e_i)_{i=1}^\infty$ be an orthonormal basis of eigenvectors of $-\Delta$ in $H^{-1}$. We further let $P^n : H^{-1} \to \text{span}\{e_1, \ldots, e_n\}$ be the orthogonal projection onto the span of the first $n$ eigenvectors. We recall that the unique variational
solution $X$ to (B.1) is constructed in [13] as a (weak) limit in $L^2([0,T] \times \Omega; L^2)$ of the solutions to the following Galerkin approximation
\[
\begin{align*}
    dX^n_t &= \varepsilon P^n \Delta X^n_t dt + P^n \Delta \phi(X^n_t) dt + P^n B(t, X^n_t) dW^n_t, \\
    X^n_0 &= P^n x_0.
\end{align*}
\]
We observe that
\[
\|X^n_t\|_2^2 = \|P^n x_0\|_2^2 + 2 \int_0^t \langle X^n_r, \varepsilon P^n \Delta X^n_r + P^n \Delta \phi(X^n_r) \rangle \, dr \\
&\quad + 2 \int_0^t \|X^n_r, P^n B(r, X^n_r) \|_{L_2(U, L^2)}^2 \, dr + 2 \int_0^t \|P^n B(r, X^n_r)\|_{L_2(U, L^2)}^2 \, dr.
\]
Passing to the (weak) limit yields the result. \hfill \Box

**Appendix C. Moreau-Yosida approximation of singular powers**

In this section we collect and prove some facts about the Moreau-Yosida approximation of certain monomials (cf. e.g. [1, Section 2.2] for background on the Moreau-Yosida approximation).

For $m \in [0, 1]$ let $\psi(r) := \frac{1}{m+1} |r|^{m+1}$, $r \in \mathbb{R}$ and let $\psi^\varepsilon : \mathbb{R} \to \mathbb{R}$ be the Moreau-Yosida approximation of $\psi$, i.e.
\[
\psi^\varepsilon(r) := \inf_{s \in \mathbb{R}} \left\{ \frac{|r-s|^2}{2 \varepsilon} + \psi(s) \right\}.
\]
Then
\[
\partial \psi^\varepsilon = \phi^\varepsilon(r) := \frac{1}{\varepsilon} (r - J^\varepsilon r) \in \phi(J^\varepsilon r) \quad \forall r \in \mathbb{R},
\]
is the Yosida approximation of $\phi = \partial \psi$, where $J^\varepsilon = (I + \varepsilon \phi)^{-1}$ is the resolvent of $\phi$ with $I$ denoting the identity map on $\mathbb{R}$. We note that
\[
|\phi^\varepsilon(r)| \leq |\phi(r)| := \inf \{ |\eta| : \eta \in \phi(r) \} \quad \forall r \in \mathbb{R}.
\]
Moreover,
\[
\psi^\varepsilon(r) = \frac{1}{2 \varepsilon} |r - J^\varepsilon r|^2 + \psi(J^\varepsilon r)
\]
and thus
\[
\psi(J^\varepsilon r) \leq \psi^\varepsilon(r) \leq \psi(r) \quad \forall r \in \mathbb{R}.
\]
By the subgradient inequality we have
\[
\eta(J^\varepsilon r - r) + \psi(r) \leq \psi(J^\varepsilon r)
\]
for all \( \eta \in \phi(r) \). Hence, using the definition of \( \phi^c \)
\[
\psi(r) - \psi(J^c r) \leq -\eta(J^c r - r) \\
\leq |\eta| c(\phi^c(r))|
\]
for every \( \eta \in \phi(r) \). Hence, using (C.3) and (C.4) we obtain
\[
(C.4) \\
|\psi(r) - \psi^c(r)| \leq \varepsilon c(\phi(r))^2 \\
\leq C \varepsilon (1 + \psi(r)) \quad \forall r \in \mathbb{R}.
\]
We note that for all \( a, b \in \mathbb{R} \)
\[
(\phi^c)^{(1)}(a) - \phi^c(z(b)) \cdot (a - b) = (\phi^c(1)^{(1)}(a) - \phi^c(1)(b)) \cdot (J^{1}(a - J^{1}b)) \\
+ (\phi^c(1)(a) - \phi^c(1)(b)) \cdot (a - J^{1}a - (b - J^{1}b)) \\
\geq (\phi^c(1)^{(1)}(a) - \phi^c(1)(b)) \cdot (\varepsilon_1 \phi^c(1)^{(1)}(a) - \varepsilon_2 \phi^c(1)(b)) \\
\geq -\frac{1}{2} (\varepsilon_1 + \varepsilon_2) (|\phi^c(1)^{(1)}(a)|^2 + |\phi^c(1)(b)|^2).
\]
Since
\[
|\phi^c(1)^{(1)}(a)|^2 \leq |\phi(1)|^2 \leq C(1 + |a|^2)
\]
we conclude
\[
(C.5) \\
(\phi^c(1)^{(1)}(a) - \phi^c(1)(b)) \cdot (a - b) \geq -C(\varepsilon_1 + \varepsilon_2)(1 + |a|^2 + |b|^2).
\]

**Appendix D. Relaxation of \( L^{m+1} \) norms on \( H^{-1} \)**

For \( m \geq 0 \) we define
\[
L^{m+1} \cap H^{-1} := \left\{ v \in L^{m+1} \right\} \left( \int vhdx \leq C\|h\|_{H^0}^2, \forall h \in C^1_c(O) \text{ for some } C \geq 0 \right\}.
\]
By continuity, every \( v \in L^{m+1} \cap H^{-1} \) the map \( h \mapsto \int vhdx \) can be uniquely extended from \( C^1_c \) to a bounded linear functional on \( H^0 \). Hence, \( L^{m+1} \cap H^{-1} \subseteq H^{-1} \). We set
\[
\varphi(v) := \left\{ \begin{array}{ll}
\frac{1}{m+1}\|v\|_{m+1}^{m+1}, & v \in L^{m+1} \cap H^{-1} \\
+\infty, & H^{-1} \setminus (L^{m+1} \cap H^{-1}),
\end{array} \right.
\]

**Lemma D.1.** Assume \( m > 0 \). Then \( \varphi \) is lower-semicontinuous on \( H^{-1} \).

**Proof.** Let \( v^n \in L^{m+1} \cap H^{-1} \) with \( \|v^n\|_{m+1} \leq C \) and \( v^n \rightarrow v \) in \( H^{-1} \). Then \( v^n \rightarrow v \) in \( L^{m+1} \) for some subsequence again denoted by \( v^n \). Now \( \| \cdot \|_{m+1} \) is weakly lower-semicontinuous on \( L^{m+1} \) and thus
\[
\|v\|_{m+1} \leq C.
\]

Due to the lack of reflexivity of \( L^1 \) this argument fails in the case \( m = 0 \). We next provide a characterization of the corresponding lower-semicontinuous hull of \( \varphi(\cdot) := \| \cdot \|_1 \) on \( H^{-1} \).

Let \( \mathcal{M} = \mathcal{M}(O) \) be the space of signed Borel measures with finite total variation on \( O \subseteq \mathbb{R}^d \) and
\[
\mathcal{M} \cap H^{-1} := \left\{ \mu \in \mathcal{M} \left| \int O h(x)d\mu(x) \leq C\|h\|_{H^0}^2, \forall h \in C^1_c(O) \text{ for some } C \geq 0 \right\}. \]
By continuity, for every \( v \in \mathcal{M} \cap H^{-1} \) the map \( h \mapsto \int hd\mu \) can be extended from \( C^1_c \) to a (uniquely determined) bounded linear functional on \( H^0 \). The resulting map \( \iota : \mathcal{M} \cap H^{-1} \rightarrow H^{-1} \) thus is injective. Hence, \( \mathcal{M} \cap H^{-1} \subseteq H^{-1} \) and in the following we identify \( \mathcal{M} \cap H^{-1} \) with its embedding into \( H^{-1} \), except for the proof of Lemma
Lemma D.2. Let $m = 0$ and $\varphi$ be the lower-semicontinuous hull of $\varphi_0(\cdot) := \| \cdot \|_1$ on $H^{-1}$. Then

$$
\varphi(v) := \begin{cases} 
TV(\mu), & v = v_\mu \in \mathcal{M} \cap H^{-1} \\
+\infty, & \text{otherwise.}
\end{cases}
$$

Proof. We first prove that $\varphi$ is weakly (hence strongly) lower-semicontinuous on $H^{-1}$. We recall that $\mathcal{M}$ is the dual $C_0(\mathcal{O})^*$ of $C_0(\mathcal{O})$, i.e. the set of all continuous functions vanishing on the boundary $\partial \mathcal{O}$, equipped with the sup-norm. Let $v_n \in \mathcal{M} \cap H^{-1}$ with $\varphi(v_n) \leq C$ and $v_n \to v$ in $H^{-1}$. Since $C_0(\mathcal{O})$ is separable and $\mu^n$ has uniformly bounded total variation we have

$$
\mu^n \to \tilde{\mu}
$$

weakly* in $\mathcal{M}$ for some subsequence of $\mu^n$ and some $\tilde{\mu} \in \mathcal{M}$. Taking the limit $n \to \infty$ in

$$
v_n(h) = \int_{\mathcal{O}} h(x)d\mu^n(x), \quad h \in C_c^1
$$

and using $v_n \to v$ in $H^{-1}$ yields

$$
v(h) = \int_{\mathcal{O}} h(x)d\tilde{\mu}(x), \quad \forall h \in C_c^1
$$

and thus $v = v_\tilde{\mu} \in \mathcal{M} \cap H^{-1}$. Since the total variation norm is lower semicontinuous with respect to weak convergence we obtain

$$
\varphi(v) = TV(\tilde{\mu}) \leq \liminf_{n \to \infty} TV(\mu^n) = \liminf_{n \to \infty} \varphi(v_n) \leq C,
$$

which proves $\varphi$ to be weakly lower-semicontinuous on $H^{-1}$.

It remains to prove that $\varphi$ is the lower-semicontinuous hull on $H^{-1}$ of $\varphi$ restricted to $L^1 \cap H^{-1}$. Let $v \in H^{-1}$ and let $v^\varepsilon := J^\varepsilon v \in L^2$, where $J^\varepsilon := (I - \varepsilon \Delta)^{-1}$ is the resolvent of $-\Delta$ in $H^{-1}$. Note that $v^\varepsilon \to v$ in $H^{-1}$, where we identify $v^\varepsilon \in L^2$ with its embedding into $H^{-1}$ via

$$
v^\varepsilon(h) = \int_{\mathcal{O}} v^\varepsilon hdx, \quad h \in H_0^1.
$$

For $v \in L^2 \subseteq H^{-1}$ we observe

$$
(J^\varepsilon v)(h) = \int_{\mathcal{O}} (J^\varepsilon v)hdx = \int_{\mathcal{O}} vJ^\varepsilon hdx = v(J^\varepsilon h), \quad \forall h \in H_0^1.
$$

By the density of $L^2$ in $H^{-1}$ this yields

$$
J^\varepsilon v = v \circ J^\varepsilon \quad \forall v \in H^{-1}.
$$

Moreover, we recall $J^\varepsilon : C_0(\mathcal{O}) \to C_0(\mathcal{O})$, $\|J^\varepsilon h\|_{C_0} \leq \|h\|_{C_0}$ for all $h \in C_0$ and for $h \in C_c^\infty(\mathcal{O})$ we have

$$
(D.1) \quad J^\varepsilon h \to h \quad \text{in } C_0.
$$

Let $\mu \in \mathcal{M} = (C_0(\mathcal{O}))^*$. We define $\mu^\varepsilon(h) = (J^\varepsilon \mu)(h) := \mu(J^\varepsilon h)$ for $h \in C_0$ and identify $\mu^\varepsilon$ with its representation in $\mathcal{M}$. Hence,

$$
\int_{\mathcal{O}} h d\mu^\varepsilon = \int_{\mathcal{O}} J^\varepsilon h d\mu.
$$
for all \( h \in C_0 \) and thus
\[
TV(\mu^\varepsilon) = \sup_{\|h\|_{C_0} \leq 1} \mu^\varepsilon(h)
\leq TV(\mu) \sup_{\|h\|_{C_0} \leq 1} \|J^\varepsilon h\|_{C_0}
\leq TV(\mu).
\]

Consequently, there is a subsequence \( \mu^{\varepsilon_n} \rightarrow \tilde{\mu} \) weakly* in \( M \) and due to (D.1) this means \( \mu^{\varepsilon_n} \rightarrow \mu \) weakly* in \( M \) and
\[
TV(\mu) \leq \liminf_{n \to \infty} TV(\mu^{\varepsilon_n}) \leq TV(\mu).
\]

A standard contradiction argument then yields
\[
(D.2) \quad \lim_{\varepsilon \to 0} TV(\mu^\varepsilon) = TV(\mu).
\]

Let now \( v_\mu \in M \cap H^{-1} \). Then
\[
(J^\varepsilon v_\mu)(h) = v_\mu(J^\varepsilon h)
= \int_\Omega J^\varepsilon h d\mu
= \int_\Omega h d\mu^\varepsilon,
\]
for all \( h \in C^1_\varepsilon \) and thus \( J^\varepsilon v_\mu \in M \cap H^{-1} \) with \( J^\varepsilon v_\mu = v_{J^\varepsilon \mu} \). Due to (D.2) this implies
\[
\varphi(J^\varepsilon v_\mu) = \varphi(v_{J^\varepsilon \mu})
= TV(J^\varepsilon \mu)
\rightarrow TV(\mu)
= \varphi(v_\mu).
\]

Since \( J^\varepsilon v_\mu \in L^2 \subseteq L^1 \cap H^{-1} \) and \( J^\varepsilon v_\mu \rightarrow v_\mu \) in \( H^{-1} \), \( \varphi \) is the lower-semicontinuous hull on \( H^{-1} \) of \( \varphi \) restricted to \( L^1 \cap H^{-1} \).

**Lemma D.3.** Let \( m \in [0,1] \) and \( v \in L^{m+1} \cap H^{-1} \). Then
\[
\partial \varphi(v) \supseteq \{-\Delta w : w \in H^1_0, w = \phi(v) \ a.e.\}.
\]

**Proof.** Case \( m > 0 \): Let \( J_\varepsilon, \varepsilon > 0 \) be as in the proof of Lemma [D.2]. Assume that \( w = \phi(v) = |v|^{m-1} v \in H^1_0 \). Then \( w \in L^{\frac{m+1}{m}} \) and we have to show that
\[
\varphi(v) \leq (-\Delta w, v - y) + \varphi(y) \quad \forall y \in L^{m+1} \cap H^{-1}.
\]

As in the proof of Lemma [D.2] we have for all \( y \in L^{m+1} \cap H^{-1} \)
\[
\varphi(v) - (-\Delta w, v - y)_{H^{-1}}
= \varphi(v) - \lim_{\varepsilon \to 0} (-\Delta w, J_\varepsilon(v - y))_{H^{-1}}
= \varphi(v) - \lim_{\varepsilon \to 0} (w, J_\varepsilon(v - y))_2
= \varphi(v) - \int_\Omega w(v - y) dx,
\]
where we used that $J\varepsilon v \to v$ in $L^p$ as $\varepsilon \to 0$ for every $v \in L^p$ and all $p \in [1, \infty)$.

The last expression equals
\[
\frac{1}{m+1} \int_{\Omega} |v|^{m+1} dx - \int_{\Omega} w(v - y) dx
\leq \frac{1}{m+1} \int_{\Omega} |v|^{m+1} dx - \int_{\Omega} |v|^{m+1} dx + \int_{\Omega} |v|^{m-1} y dy dx
\leq \varphi(y),
\]
where we used Hölder’s and Young’s inequality in the last step. This finishes the proof.

**Case** $m = 0$: Assume $w \in \phi(v)$ a.e. with $w \in H^1_0$. Arguing as in the case $m > 0$ and using $|w| \leq 1$ a.e. we have
\[
\varphi(v) - (-\Delta w, v - y)_{H^{-1}} = \varphi(v) - \int_{\Omega} w(v - y) dx
\leq \int_{\Omega} |y| dx,
\]
which finishes the proof. □

**References**


