A self-dual approach to stochastic partial differential equations

by

Shirin Boroushaki

B.Sc., Ferdowsi University of Mashhad, 2010
M.Sc., Sharif University of Technology, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

January 2018

Abstract

In the first part of this thesis, we use the theory of self-duality to provide a variational approach for the resolution of a number of stochastic partial differential equations. We will be able to address the problem of existence of solutions to a class of semilinear stochastic partial differential equations in the form

$$\begin{cases} du + A(t,u(t))dt = B(t,u(t))dW \\ u(0) = u_0, \end{cases} \tag{1}$$

where for every $t \in [0,T]$, $A(t,\cdot)$ is a maximal monotone operator on a reflexive Banach space V, and B is a linear or non-linear operator with values in a Hilbert space H. We use the fact that any maximal monotone operator A can be expressed as a potential of a self-dual Lagrangian L to associate to the equation (1) a (completely) self-dual functional whose minimizer on a suitable path space yields a solution.

One particular case of (1) which already contains a large number of stochastic PDEs is when A is the subdifferential of a convex function φ. More generally, we can deal with equations of the form

$$\begin{cases} du(t) = -\partial\varphi(t,u(t))dt + B(t,u(t))dW(t) \\ u(0) = u_0. \end{cases}$$

We also prove the existence of solutions to SPDEs in divergence form involving a maximal monotone operator β on $\mathbb{R}^n$, which is not necessarily the gradient of a convex function,

$$\begin{cases} du = \mathrm{div}(\beta(\nabla u(t,x)))dt + B(t)dW(t) & \text{in } [0,T]\times D \\ u(0,x) = u_0 & \text{on } \partial D, \end{cases}$$

where $D \subset \mathbb{R}^n$ is a bounded domain.

In the second part of the thesis, we use methods from optimal transport to address functional inequalities on the n-dimensional sphere $\mathbb{S}^n$. We prove Energy-Entropy duality formulas that yield and improve the celebrated Moser-Onofri inequalities on $\mathbb{S}^2$.

Lay Summary

The evolution of many physical systems can often be described by deterministic differential equations. However, in many models of natural phenomena there are uncertainties that have a considerable effect on the evolution of the systems under study. These naturally appear while studying population and genetic dynamics, neurophysiology, financial markets and turbulence in fluid dynamics. The presence of randomness translates into models described by *stochastic differential equations* expressing the presence of white noise. While variational principles are natural and well developed for the study of deterministic models, their use for stochastic equations have been limited by the complexity of the differential structure of Brownian motion. In this thesis, we show that a probabilistic version of the recently developed *self-dual variational calculus* is as efficient in handling stochastic models as the non-probabilistic counterpart was in dealing with deterministic equations.

Preface

Much of this dissertation is adapted from two of the author's research papers: [1] and [9]. In particular, Chapters 4 and 5, which give a variational resolution of SPDEs, form the main content of [9], *A self-dual variational approach to stochastic partial differential equations*. Chapter 6 is in accordance with [1] (joint work with Dr. Ghoussoub and Dr. Agueh), which was published in Annales de la Faculté des Sciences de Toulouse, Sér. 6, 26 no 2 (2017) p. 217-233. The second manuscript [9] is a joint work with Dr. Ghoussoub. It has been posted on arXiv and is submitted for publication.

Table of Contents

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor, Professor Nassif Ghoussoub whose expertise, understanding and invaluable guidance and support added considerably to my graduate experience. I thank him for being always available to discuss research problems and answer my never-ending questions.

I would also like to express my gratitude to Dr. Martial Agueh, who, although no longer with us, generously assisted me through our collaboration and writing of my first research paper.

I would also like to thank the members of my supervisory committee, Professor Young-Heon Kim and Professor Jun-cheng Wei for their insightful and instructive comments.

Finally, I wish to thank my family for their endless love, support and encouragement throughout my life, and for giving me strength to realize my potential.

Chapter 1

Introduction

This thesis is comprised of two independent parts. Both are based on recent advances in infinite dimensional variational methods: The theory of optimal mass transport, and the self-dual variational calculus.

A standard approach to solve several classes of partial differential systems is to represent them as Euler-Lagrange equations whose solutions are characterized as critical points to the corresponding functionals of the form $I(u) = \int_D F(x, u(x), \nabla u(x))dx$, where D is an open bounded subset of $\mathbb{R}^n$. However, there are a large number of partial differential equations involving non-linear, non-local or non self-adjoint operators which do not fall in the Euler-Lagrange framework. Self-dual variational calculus was developed in the last fifteen years in an effort to construct solutions to such non-variational partial differential equations and evolutions. We refer to the monograph [29] for a comprehensive account of that theory. In the first part, which is indeed the core of this thesis, we show how such a calculus can be applied to solve stochastic partial differential equations, which also do not fit in Euler-Lagrange theory, since their solutions are not known to be critical points of energy functionals. We show here that at least for some of these equations, solutions can be obtained as minima of suitable self-dual functionals on Itô spaces of random paths.

The genesis of self-dual variational calculus can be traced to a 1970 paper of Brezis-Ekeland [12, 13] (see also Nayroles [40, 41]), where they proposed a variational principle for the heat equation and other gradient flows for convex energies. The method consists of minimizing the functional

$$I(u) = \varphi(u) + \varphi^*(-Au) + \langle u, Au \rangle,$$

where $A : V \to V^*$ is a non self-adjoint or non-potential operator on the Banach space V, and φ is a convex functional on V with φ^* its Legendre dual. The basic property of Legendre duality $\varphi(u) + \varphi^*(p) \geq \langle u, p \rangle$ and

$$\varphi(u) + \varphi^*(p) = \langle u, p \rangle \qquad \text{iff} \qquad p \in \partial\varphi(u),$$

yield that $\inf_{u \in V} I(u) \geq 0$. Therefore, the observation is that if the value of the infimum is 0 and is attained at some $\bar{u} \in V$, then the limiting case of

Legendre duality implies that $-A\bar{u} \in \partial\varphi(\bar{u})$, and hence $\bar{u}$ is a solution. However, the main difficulty is to prove that the infimum is actually zero. The conjecture was eventually verified by Ghoussoub-Tzou [34], who identified and exploited the self-dual nature of the Lagrangians involved. Since then, the theory was developed in many directions [26, 27, 31], so as to provide existence results for several stationary and parabolic -but so far deterministic-PDEs, which may or may not be Euler-Lagrange equations.

While in most examples where the approach was used, the self-dual Lagrangians were explicit, an important development in the theory was the realization [30] that in a prior work, Fitzpatrick [23] had associated a (somewhat) self-dual Lagrangian to any given monotone vector field. That meant that the variational theory could apply to any equation involving such operators. We refer to the monograph [29] for a survey and for applications to existence results for solutions of several PDEs and evolution equations. We also note that since the appearance of this monograph, the theory has been successfully applied to the homogenization of periodic non-self adjoint problems (Ghoussoub-Moameni-Zarate [32]).

One of the most important classes of evolutionary equations is the class of (semi-linear) stochastic partial differential equations which are used to model and describe many kinds of dynamics in the natural sciences, as well as in finance. These equations are extensions of Itô stochastic equations introduced in the 1940s by Itô. Basic theoretical questions on *existence* and *uniqueness* of solutions were asked and answered, under various sets of conditions, in the 1970s and 1980s and are still of great interest today. There are basically three approaches to analyzing SPDEs: the "martingale approach" [53], the "semi-group (mild solution) approach" [16] and the "variational approach" [45, 48]. There is an enormously rich literature on all three approaches which cannot be listed here.

In this thesis, we follow the "variational approach" which was initiated in the celebrated thesis of Pardoux [44], and many other subsequent works [45, 46, 48]. However, the application of self-dual variational method to solving SPDEs is long overdue, though V. Barbu [6, 7] did use a Brezis-Ekeland approach to address SPDEs driven by gradients of a convex function and additive noise. We shall deal here with more general situations that cannot be reduced to the deterministic case. We note that the equations studied here have already been solved by other methods, and this work is about presenting a new variational approach, hoping it will lead to progress on other unresolved equations.

In Chapter 2, we present the basics of probability theory in Hilbert and Banach spaces. We list some commonly used concepts from the theory of stochastic processes and stochastic integration with respect to general Hilbert-valued Wiener processes. In Chapter 3, we provide the basic tools of convex analysis required for the foundation of self-dual systems in terms of the class of self-dual Lagrangians and their corresponding vector fields. We then state the variational principles for minimization of self-dual functionals, which also involve the Hamiltonian formulation for the minimization of direct sum of self-dual functionals.

In Chapter 4, we shall tackle basic SPDEs involving additive noise, namely

$$\begin{cases} du(t) = -A(t, u(t))dt + B(t)dW(t) \\ u(0) = u_0, \end{cases} \tag{1.1}$$

where $u_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ for the Hilbert space H, $W(t)$ is a real-valued Wiener process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_t$, and where $B : [0,T] \times \Omega \to H$ is a given Hilbert-space valued progressively measurable process. Here, $A : \Omega \times [0,T] \times V \to 2^{V^*}$ can be a time-dependent adapted –possibly set-valued– maximal monotone map, where V is a Banach space such that $V \subset H \subset V^*$ constitute a Gelfand triple. The simplest example is where the monotone operator A is given by the gradient $\partial\varphi$ of a (possibly random and progressively measurable) function $\varphi : [0,T] \times H \to \mathbb{R} \cup \{+\infty\}$ such that for every $t \in [0,T]$, the function $\varphi(t, \cdot)$ is convex and lower semi-continuous on a Hilbert space H, and the stochastics is driven by a given progressively measurable additive noise coefficient $B : \Omega \times [0,T] \to H$. The equation becomes

$$\begin{cases} du(t) = -\partial\varphi(t, u(t))dt + B(t)dW(t) \\ u(0) = u_0. \end{cases} \tag{1.2}$$

We consider the following *Itô space over* H,

$$\mathcal{A}_H^2 = \left\{ u : \Omega_T \to H;\ u(t) = u(0) + \int_0^t \tilde{u}(s)ds + \int_0^t F_u(s)dW(s) \right\},$$

where $u(0) \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, $\tilde{u} \in L^2(\Omega_T; H)$ and $F_u \in L^2(\Omega_T; H)$, where $\Omega_T = \Omega \times [0,T]$. Here, both the drift $\tilde{u}$ and the diffusive term F_u are progressively measurable.

The key idea is that a solution for (1.2) can be obtained by minimizing the following functional on $\mathcal{A}_H^2$,

$$I(u) = \mathbb{E}\Big\{ \int_0^T \Big(L_\varphi(u(t), -\tilde{u}(t)) + \frac{1}{2} M_B(F_u(t), -F_u(t))\Big)dt + \ell_{u_0}(u(0), u(T))\Big\},$$

where

1. L_φ is the (possibly random) time-dependent Lagrangian on $H \times H$ given by
$$L_\varphi(u,p) = \varphi(w,t,u) + \varphi^*(w,t,p), \tag{1.3}$$
where φ^* is the Legendre transform of φ;

2. ℓ_{u_0} is the time-boundary random Lagrangian on $H \times H$ given by
$$\ell_{u_0}(a,b) := \ell_{u_0(w)}(a,b) = \frac{1}{2}\|a\|_H^2 + \frac{1}{2}\|b\|_H^2 - 2\langle u_0(w), a\rangle_H + \|u_0(w)\|_H^2; \tag{1.4}$$

3. M_B is the random time-dependent diffusive Lagrangian on $H \times H$, given by
$$M_B(G_1,G_2) := \Psi_{B(w,t)}(G_1) + \Psi^*_{B(w,t)}(G_2), \tag{1.5}$$
where $\Psi_{B(w,t)} : H \to \mathbb{R} \cup \{+\infty\}$ is the convex function $\Psi_{B(w,t)}(G) = \frac{1}{2}\|G - 2B(w,t)\|_H^2$.

However, it is not sufficient that I attains its infimum on $\mathcal{A}_H^2$ at some v, but one needs to also show that the infimum is actually equal to zero. By using Itô's formula, we can rewrite $I(v)$ as the sum of 3 non-negative terms
$$\begin{aligned} 0 = I(v) &= \mathbb{E}\int_0^T \Big(\varphi(t,v) + \varphi^*(t,-\tilde{v}(t)) + \langle v(t), \tilde{v}(t)\rangle\Big)\, dt \\ &\quad + 2\,\mathbb{E}\int_0^T \|F_v - B\|_H^2\, dt + \mathbb{E}\,\|v(0) - u_0\|_H^2, \end{aligned}$$
which yields that for almost all $t \in [0,T]$, $\mathbb{P}$-a.s.
$$-\tilde{v}(t) \in \partial\varphi(v(t)), \quad B = F_v, \quad v(0) = u_0.$$
In other words, $v(t) = u_0 - \int_0^t \partial\varphi(s,v(s))ds + \int_0^t B(s)dW(s)$, where the last stochastic integral is in the sense of Itô. Hence, v is a solution to (1.2). The main focus of Chapter 4 is to extend this argument to more general Lagrangians L and consequently resolve Equation (1.1). We then present some applications of the method to classical SPDEs such as stochastic evolution driven by a diffusion and a transport operator, stochastic porous media equation, and quasi-linear equations involving p-Laplacian.

In Chapter 5, we will deal with SPDEs driven by monotone vector fields and involving a non-additive noise. These can take the form
$$\begin{cases} du(t) = -A(t,u(t))dt + B(t,u(t))dW(t) \\ u(0) = u_0, \end{cases} \tag{1.6}$$

where $u \to B(t,u)$ is now a progressively measurable linear or non-linear operator. Analogously, we consider the Itô space this time over the reflexive Banach space V to find a solution that is V-valued. To this end, we shall strengthen the norm on the Itô space over a Gelfand triple, at the cost of losing coercivity, that we shall recover through perturbation methods. By solutions, we mean progressively measurable processes u, valued in suitable Sobolev spaces, that verify the integral equation

$$u(t) = u_0 - \int_0^t A(s,u(s))ds + \int_0^t B(s,u(s))dW(s).$$

To variationally resolve Equation (1.6), we shall use the self-dual approach regarding minimization of direct sum of self-dual functionals (see Theorem 3.5) together with the so-called technique of elliptic regularization to ensure coercivity. We will then let the perturbations go to zero to recover the equation and conclude the existence of a solution. We then apply this result to resolve an equation of the form (1.6) and more generally,

$$\begin{cases} du = \operatorname{div}(\beta(\nabla u(t,x)))dt + B(u(t))dW(t) & \text{in } [0,T]\times D \\ u(0,x) = u_0(x) & \text{on } \partial D, \end{cases}$$

where D is a bounded domain in $\mathbb{R}^n$, the initial position u_0 belongs to $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; L^2(D))$, and the vector field β is a progressively measurable maximal monotone operator on $\mathbb{R}^n$ (see Subsection 5.3.3).

In part (II) of this thesis, we use optimal mass transport to provide a new proof and a dual formula to the Moser-Onofri inequality on $\mathbb{S}^2$,

$$\frac{1}{4}\int_{\mathbb{S}^2} |\nabla u|^2\,\mathrm{d}\omega + \int_{\mathbb{S}^2} u\,\mathrm{d}\omega \geq \log\left(\int_{\mathbb{S}^2} e^u\,\mathrm{d}\omega\right).$$

This is in the same spirit as the approach of Cordero-Erausquin, Nazaret and Villani [15] to the Sobolev and Gagliardo-Nirenberg inequalities and the one of Agueh-Ghoussoub-Kang [2] to more general settings. We use an equivalent formulation of the Onofri inequality which is obtained via stereographic projection of (6.1), namely

$$\frac{1}{16\pi}\int_{\mathbb{R}^2} |\nabla u|^2\,\mathrm{d}x + \int_{\mathbb{R}^2} u\,\mathrm{d}\mu_2 \geq \log\left(\int_{\mathbb{R}^2} e^u\,\mathrm{d}\mu_2\right) \quad \forall u \in H^1(\mathbb{R}^2).$$

However, once a projection on $\mathbb{R}^2$ is performed, several new hurdles appear that are not the case in [15] or [2]: Functions are not necessarily of compact support, hence boundary terms need to be evaluated. Moreover, the corresponding dual free energy of the reference probability density $\mu_2(x) = \frac{1}{\pi(1+|x|^2)^2}$ is not finite on the whole space, which requires the introduction of a renormalized free energy into the dual formula. We also extend this duality to higher dimensions and establish an extension of the Onofri inequality to spheres $\mathbb{S}^n$ with $n \geq 2$. What is remarkable is that the corresponding free energy is again given by $F(\rho) = -n\rho^{1-\frac{1}{n}}$, which means that both the *prescribed scalar curvature problem* and the *prescribed Gaussian curvature problem* lead essentially to the same dual problem whose extremals are stationary solutions of the fast diffusion equations.

Part I

Self-dual variational principles for stochastic partial differential equations

Chapter 2

Basics of stochastic calculus

The prime objective of this chapter is to make this dissertation self-contained in terms of definitions and facts from the measure theoretic foundations of probability theory. A fundamental element in stochastic differential equations is the presence of the diffusion term which in our case corresponds to the stochastic integral with respect to a given Wiener process. Therefore, the goal in this chapter is to provide the reader with sufficient material to develop the probability theory in infinite dimension, namely Hilbert-valued stochastic processes and the construction of stochastic integrals as processes with values in these Hilbert spaces. Our main source for most of the components of this chapter is [16] and [48].

2.1 Random variables and probability space

A measurable space is a pair $(\Omega, \mathcal{F})$ where Ω is a non-empty set and $\mathcal{F}$ is a σ-field, also called a σ-algebra, of subsets of Ω. This means that the family $\mathcal{F}$ contains the set Ω and is closed under the operation of taking complements and countable unions of its elements. If $(\Omega, \mathcal{F})$ and $(V, \mathcal{G})$ are two measurable spaces, then a mapping X from Ω into V such that for arbitrary $A \in \mathcal{G}$, the set $X^{-1}(A) = \{\omega \in \Omega; X(\omega) \in A\} =: \{X \in A\}$ belongs to $\mathcal{F}$ is called is called a measurable mapping or a random variable from $(\Omega, \mathcal{F})$ into $(V, \mathcal{G})$ or a *V-valued random variable.*

Assume that V is a metric space, then the Borel σ-field of V is the smallest σ-field containing all closed (or open) subsets of V and it will be denoted as $\mathcal{B}(V)$.

In fact, if V is a separable Banach space, and V^* is its topological dual, then a mapping $X : \Omega \to V$ is a V-valued random variable if and only if for arbitrary $\zeta \in V^*$, $\zeta(X) : \Omega \to \mathbb{R}$ is an $\mathbb{R}$-valued random variable.

Definition 2.1. *A probability measure on a measurable space $(\Omega, \mathcal{F})$ is a σ-additive function $\mathbb{P}$ from $\mathcal{F}$ into $[0, 1]$ such that $\mathbb{P}(\Omega) = 1$. The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.*

If the triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, we set

$$\bar{\mathcal{F}} = \{A \subset \Omega;\ \exists\ B, C \in \mathcal{F}\ ; B \subset A \subset C,\ \mathbb{P}(B) = \mathbb{P}(C)\}.$$

$\bar{\mathcal{F}}$ is a σ-field, called the *completion* of $\mathcal{F}$. If $\mathcal{F} = \bar{\mathcal{F}}$, then the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be *complete*. Equivalently, $(\Omega, \mathcal{F}, \mathbb{P})$ is complete if any subset N of $B \in \mathcal{F}$ with $\mathbb{P}(B) = 0$ is also in $\mathcal{F}$.

If X is a random variable from $(\Omega, \mathcal{F})$ into $(V, \mathcal{G})$ and $\mathbb{P}$ a probability measure on Ω, then by $\mathbb{L}(X)$ we denote the image of $\mathbb{P}$ by the mapping X;

$$\begin{aligned}\mathbb{L}(X)(A) &= \mathbb{P}(\omega \in \Omega; X(\omega) \in A)\\ &= \mathbb{P}(\{X \in A\}) \qquad\qquad \forall\ A \in \mathcal{G}\end{aligned}$$

The measure $\mu = \mathbb{L}(X)$ is called the *distribution* or the *law* of X.

Similar to the definition of Lebesgue integral of real-valued functions, one can define the integral for a real-valued random variable and analogously the integral of a random variable with values in a separable Banach space V. In fact, the random variable X is said to be *Bochner integrable* or just *integrable* if

$$\int_\Omega \|X(\omega)\| \mathbb{P}(d\omega) < \infty.$$

The integral $\int_\Omega X d\mathbb{P}$ will often be denoted by $\mathbb{E}(X)$, the *expectation* of X and it has many properties of the Lebesgue integral.

Finally, in the last part of this section we introduce some basic function spaces. We denote by $L^1(\Omega, \mathcal{F}, \mathbb{P}; V)$ the set of all equivalence classes of V-valued random variables (with respect to the equivalence relation $X \sim Y \Leftrightarrow X = Y$ a.s.). One can check that $L^1(\Omega, \mathcal{F}, \mathbb{P}; V)$, equipped with the norm

$$\|X\|_1 = \mathbb{E}(\|X\|_V)$$

is a Banach space. In a similar way, one can define $L^p(\Omega, \mathcal{F}, \mathbb{P}; V)$, for arbitrary $p > 1$ with the norm

$$\|X\|_p = (\mathbb{E}\|X\|_V^p)^{1/p}.$$

If $X, Y \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$ and H is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$, we define the *covariance operator* of X by the formula

$$\mathrm{Cov}(X) = \mathbb{E}\Big((X - \mathbb{E}(X)) \otimes (X - \mathbb{E}(X))\Big)$$

$\mathrm{Cov}(X)$ is a symmetric positive and nuclear operator and

$$\mathrm{Tr}[\mathrm{Cov}(X)] = \mathbb{E}(|X - \mathbb{E}(X)|^2)$$

2.2 Gaussian measures in Banach spaces

Definition 2.2. *A Gaussian measure μ on $\mathbb{R}$ is either concentrated at one point $\mu = \delta_m$ or has a density*

$$\frac{1}{\sqrt{2\pi q}} e^{-\frac{1}{2q}(x-m)^2} \qquad \text{for } x \in \mathbb{R},$$

where $q > 0$ and $m \in \mathbb{R}$ are respectively the variance and the mean of the random variable. Such a measure is denoted by $\mathcal{N}(m, q)$.

The Gaussian measure μ on $\mathbb{R}^n$ has the density

$$\frac{1}{\sqrt{(2\pi)^n \det Q}} e^{-\frac{1}{2}\langle Q^{-1}(x-m), x-m\rangle} \qquad \text{for } x \in \mathbb{R}^n,$$

where Q is a positive-definite symmetric $n \times n$ matrix, known as the *covariance matrix*, and $m \in \mathbb{R}^n$ is the mean, and this measure also is denoted by $\mathcal{N}(m, Q)$.

Now let V be a separable Banach space. A probability measure μ on $(V, \mathcal{B}(V))$ is said to be a *Gaussian measure* if and only if the law of an arbitrary linear functional $h \in V^*$, considered as a random variable on $(V, \mathcal{B}(V), \mu)$, is a Gaussian measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. If the law of each $h \in V^*$ is in addition a symmetric (zero mean) Gaussian law on $\mathbb{R}$ then μ is called a *symmetric Gaussian measure.*

2.2.1 Gaussian measures on Hilbert spaces

For Gaussian measures on Hilbert spaces more precise information can be given. Let H be a Hilbert space, according to the general definition on Banach spaces, a probability measure μ on $(H, \mathcal{B}(H))$ is called Gaussian if for arbitrary $h \in H$ there exist $m \in \mathbb{R}$ and $q \geq 0$ such that,

$$\mu\Big(\{x \in H; \langle h, x\rangle \in A\}\Big) = \mathscr{N}(m, q)(A), \qquad \forall A \in \mathcal{B}(\mathbb{R}).$$

It is proved that for a Gaussian measure μ on $(H, \mathcal{B}(H))$ there exist an element $m \in H$ and a linear operator Q, such that

$$\int_H \langle h, x\rangle \mu(dx) = \langle m, h\rangle, \qquad \forall h \in H,$$

$$\int_H \langle h_1, x - m\rangle\langle h_2, x - m\rangle \mu(dx) = \langle Qh_1, h_2\rangle, \qquad \forall h_1, h_2 \in H.$$

The vector m is called the *mean* and Q is called the *covariance operator* of μ. It is clear that the operator Q is symmetric and also non-negative. The Gaussian measure is denoted by $\mathscr{N}(m, Q)$.

The following proposition shows that the covariance operator is nuclear. A proof can be found in [16].

Proposition 2.1. *Let μ be a Gaussian probability measure with mean 0 and covariance Q. Then Q has finite trace.*

2.3 Stochastic processes

Assume that V is a separable Banach space and let $\mathcal{B}(V)$ be the σ-field of its Borel subsets. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let I be an interval of $\mathbb{R}$. An arbitrary family $X = \{X(t)\}_{t\in I}$, of V-valued random variables $X(t)$, $t \in I$ defined on Ω is called a *stochastic process*. We also say that $X(t)$ is a stochastic process on I. We set $X(t, \omega) = X(t)(\omega)$ for all $t \in I$ and $\omega \in \Omega$. Functions $X(\cdot, \omega)$ are called the *trajectories* of $X(t)$.

A stochastic process Y is called a *modification* or a *version* of X if

$$\mathbb{P}(\omega \in \Omega; X(t, \omega) \neq Y(t, \omega)) = 0 \qquad \forall\ t \in I.$$

A process X is called *measurable* if the mapping $X(\cdot, \cdot) : I \times \Omega \to V$ is $\mathcal{B}(I) \otimes \mathcal{F}$-measurable.

X is called *stochastically continuous* at $t_0 \in I$ if for all $\epsilon > 0$ and all $\delta > 0$ there exists $\rho > 0$ such that

$$\mathbb{P}(\|X(t) - X(0)\| \geq \epsilon) \leq \delta \qquad \forall\ t \in [t_0 - \rho, t_0 + \rho] \cap [0, T].$$

2.3.1 Processes with filtration

Definition 2.3. *Assume that* $I = [0, T]$ *or* $[0, +\infty]$. *A filtration on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is an increasing family of* σ*-fields* $\{\mathcal{F}_t\}$, $t \in I$ *such that, for all* $s \le t$ *in* I, *we have* $\mathcal{F}_s \subset \mathcal{F}_t \subset \{\mathcal{F}_t\}_{t \in I}$.

Denote by $\mathcal{F}_{t^+}$ *the intersection of all* $\mathcal{F}_s$ *where* $s > t$. *The filtration is said to be normal if*

1. $\mathcal{F}_0$ *contains all* $A \in \mathcal{F}$ *such that* $\mathbb{P}(A) = 0$.
2. $\mathcal{F}_t = \mathcal{F}_{t^+} = \bigcap_{s>t} \mathcal{F}_s$ *for all* $t \in I$.

If for arbitrary $t \in I$ the random variable $X(t)$ is $\mathcal{F}_t$-measurable then the process X is said to be *adapted* (to the family $\mathcal{F}_t$).

X is *progressively measurable* if for every $t \in [0, T]$ the mapping

$$[0, t] \times \Omega \to V,$$
$$(s, \omega) \to X(s, \omega)$$

is $\mathcal{B}([0, t]) \otimes \mathcal{F}_t$-measurable.

Proposition 2.2. *Let* $X(t)$, $t \in [0, T]$ *be a stochastically continuous and adapted process with values in a separable Banach space* V. *Then* X *has a progressively measurable version.*

Definition 2.4. *Let* $\mathcal{P}_\infty$ *be the* σ*-field generated by the sets of the form* $(s, t] \times F$, *where* $0 \le s < t < \infty$, $F \in \mathcal{F}_s$, *and* $\{0\} \times F$ *for* $F \in \mathcal{F}_0$. *This* σ*-field is called a predictable* σ*-field and its elements are predictable sets. The restriction of the* σ*-field* $\mathcal{P}_\infty$ *to* $[0, T] \times \Omega$ *is denoted by* $\mathcal{P}_T$.

An arbitrary measurable mapping from $([0, \infty] \times \Omega, \mathcal{P}_\infty)$ or $([0, T] \times \Omega, \mathcal{P}_T)$ into $(V, \mathcal{B}(V))$ is called a *predictable process.* A predictable process is necessarily an adapted one. Predictable processes form a large class of processes, as can be seen in the next proposition.

Proposition 2.3. *Assume that* Φ *is an adapted and stochastically continuous process on an interval* $[0, T]$. *Then the process* Φ *has a predictable version on* $[0, T]$.

2.4 Martingales

Let $X(t)$ be a V-valued process. If $\mathbb{E}\|X(t)\|_V < +\infty$ for all $t \in [0,T]$ then the process is called *integrable*. An integrable and adapted V-valued process $X(t)$, $t \in [0,T]$ is said to be a *martingale* if for arbitrary $t, s, \in [0,T],\ t \geq s$

$$\mathbb{E}(X(t)|\mathcal{F}_s) = X(s), \qquad \mathbb{P} - a.s.$$

For fixed $T > 0$, we denote by $\mathscr{M}_T^2(V)$ the space of all V-valued continuous, square integrable martingales M, such that $M(0) = 0$.

Proposition 2.4. *The space $\mathscr{M}_T^2(V)$ equipped with the norm*

$$\|M\|_{\mathscr{M}_T^2(V)} = \left(\mathbb{E} \sup_{t\in[0,T]} \|M(t)\|^2\right)^{\frac{1}{2}}$$

is a Banach space.

If $M \in \mathscr{M}_T^2(\mathbb{R})$ then there exists a unique increasing predictable process $\langle\langle M(\cdot)\rangle\rangle$, starting from 0, such that the process

$$M^2(t) - \langle\langle M(\cdot)\rangle\rangle, \qquad t \in [0,T]$$

is a continuous martingale. The process $\langle\langle M(\cdot)\rangle\rangle$ is called the *quadratic variation* of M.

To define the quadratic variation process for $M \in \mathscr{M}_T^2(H)$ where H is a separable Hilbert space, denote by $\mathbb{L}_1 := \mathbb{L}_1(H)$ the space of all nuclear operators on H equipped with the nuclear norm. An $\mathbb{L}_1$-valued continuous, adapted and increasing process Z such that $Z(0) = 0$ is said to be a *quadratic variation* process of the martingale $M(\cdot)$ if and only if for arbitrary $a, b \in H$ the process

$$\langle M(t), a\rangle\langle M(t), b\rangle - \langle Z(t)a, b\rangle, \qquad t \in [0,T]$$

is an $\{\mathcal{F}_t\}$-martingale. $Z(\cdot)$ is uniquely determined and is also denoted by $\langle\langle M(t)\rangle\rangle$.

Proposition 2.5. *An arbitrary $M \in \mathscr{M}_T^2(H)$ has exactly one quadratic variation process.*

A non-negative random variable τ defined on $(\Omega, \mathcal{F})$ is said to be an $\mathcal{F}_t$-*stopping time* if for arbitrary $t \geq 0$, $\{\omega \in \Omega; \tau(\omega) \leq t\} \in \mathcal{F}_t$.

2.5 Wiener processes

Let U be a Hilbert space with inner product $\langle\ ,\ \rangle_U$, and let Q be a trace class non-negative operator on a Hilbert space U.

Definition 2.5. *A U-valued stochastic process $W(t)$, $t \in [0,T]$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a Q-Wiener process if*

i. $W(0) = 0$

ii. *W has $\mathbb{P}$-a.s. continuous trajectories.*

iii. *W has independent increments.*

iv. $\mathbb{L}(W(t) - W(s)) = \mathscr{N}(0, (t-s)Q), \qquad t \geq s \geq 0.$

Note that there exists a complete orthonormal system $\{e_k\}$ in U and a bounded sequence of non-negative real numbers $\{\lambda_k\}$ such that

$$Qe_k = \lambda_k e_k, \qquad k \in \mathbb{N}.$$

The following proposition shows that the operator Q can completely characterize the distribution of W.

Proposition 2.6. *Assume that $W(t)$ is a Q-Wiener process, then*

1. *$W(t)$ is a Gaussian process on U, and*

$$\mathbb{E}(W(t)) = 0, \quad \operatorname{Cov}(W(t)) = tQ, \qquad t \geq 0$$

2. ***(Representation of the Q-Wiener process)***
For arbitrary $t \geq 0$, W has the expansion

$$W(t) = \sum_{k\in\mathbb{N}} \sqrt{\lambda_k}\, \beta_k(t)\, e_k = \sum_{k\in\mathbb{N}} \beta_k(t) Q^{\frac{1}{2}} e_k, \qquad t \geq 0 \tag{2.1}$$

where

$$\beta_k(t) = \frac{1}{\sqrt{\lambda_k}} \langle W(t), e_k \rangle, \qquad k \in \mathbb{N}$$

are mutually independent real-valued Brownian motions on $(\Omega, \mathcal{F}, \mathbb{P})$. The series converges in $L^2(\Omega, \mathcal{F}, \mathbb{P}; C([0,T]; U))$.

Note that the quadratic variation of a Q-Wiener process in U, with $\operatorname{Tr} Q < +\infty$, is given by $\langle\langle W(t) \rangle\rangle = tQ$, $t \geq 0$.

Definition 2.6. *A Q-Wiener process $W(t)$, $t \in [0, T]$ is called a Wiener process with respect to a filtration $\mathcal{F}_t$ if*

- *$W(t)$ is adapted to $\mathcal{F}_t$, $t \in [0, T]$ and*
- *$W(t) - W(s)$ is independent of $\mathcal{F}_s$ for all $0 \leq s \leq t \leq T$.*

In this thesis, we only consider real-valued Brownian motions, i.e. $U = \mathbb{R}$, and hence $Q = I$.

2.6 Stochasic integral

2.6.1 Operator-valued random variables

One important concept and of great interest is operator-valued random variables which are one of the main elements in construction of stochastic integrals. Let U and H be two separable Hilbert spaces and denote by $\mathbb{L} = \mathbb{L}(U, H)$ the set of all linear bounded operators from U into H. The set $\mathbb{L}$ is a linear space and, equipped with the operator norm, becomes a Banach space. However, if both spaces are infinite dimensional, then $\mathbb{L}$ is not a separable space.

The lack of separability of $\mathbb{L}$ implies also that Bochners definition cannot be applied directly to the $\mathbb{L}$-valued functions. To overcome these difficulties, one option is restrict our investigation to smaller spaces - the space $\mathbb{L}_1(U, H)$ of all nuclear operators from U into H, or the space $\mathbb{L}_2(U, H)$ of all Hilbert-Schmidt operators from U into H, and then the non-measurability problem arisen from non-separability does not appear, that is because these smaller spaces are separable Banach spaces ($\mathbb{L}_2(U, H)$ is a Hilbert space).

A function $\Phi(\cdot)$ from Ω into $\mathbb{L}$ is said to be *strongly measurable* if for arbitrary $u \in U$ the function $\Phi(\cdot)u$ is measurable as a mapping from $(\Omega, \mathcal{F})$ into $(H, \mathcal{B}(H))$. Let $\mathcal{L}$ be the smallest σ-field of subsets of $\mathbb{L}$ containing all sets of the form $\{\Phi \in \mathbb{L}; \text{for } \Phi u \in A\}, \quad u \in U,\ A \in \mathcal{B}(H)$, then $\Phi : \Omega \to \mathbb{L}$ is a strongly measurable mapping from $(\Omega, \mathcal{F})$ into $(\mathbb{L}, \mathcal{L})$. Elements of $\mathcal{L}$ are called *strongly measurable.*

Hilbert-Schmidt operators

Definition 2.7. *Let $T \in \mathbb{L}(U)$ and let $\{e_k\}, k \in \mathbb{N}$ be an ONB of U. We define*

$$\operatorname{Tr} T := \sum_{k=1}^{\infty} \langle Te_k, e_k \rangle_U$$

if the series is convergent.

Note that the trace definition above is independent of the choice of the orthonormal basis.

Definition 2.8. *A bounded linear operator $T : U \to H$ is called Hilbert-Schmidt if*

$$\|T\|_{\mathbb{L}_2(U,H)} := \sum_{k=1}^{\infty} \|Te_k\|_H^2 < \infty$$

where $\{e_k\}$ is an ONB of U.

We denote the space of all Hilbert-Schmidt operators from U to H by $\mathbb{L}_2(U, H)$.

The space of Hilbert-Schmidt operators forms a separable Hilbert space with the inner product defined as the following:

For $S, T \in \mathbb{L}_2(U, H)$ and $\{e_k\}$ an ONB of U, we set

$$\langle T, S \rangle_{\mathbb{L}_2} := \sum_{k=1}^{\infty} \langle Te_k, Se_k \rangle_H.$$

2.6.2 Construction of the stochastic integral

We first state the definition of the stochastic integral for an elementary process, and we then briefly present the steps to extend it to more general processes.

We are given $W(t)$ a Q-Wiener process in $(\Omega, \mathcal{F}, \mathbb{P})$ having values in the Hilbert space U with respect to a normal filtration $\mathcal{F}_t, t \in [0, T]$.

Definition 2.9. *An $\mathbb{L} = \mathbb{L}(U, H)$-valued process $\Phi(t)$, $t \in [0, T]$ taking only a finite number of values is said to be elementary if there exists a sequence $0 = t_0 < t_1 < ... < t_k = T$ and a sequence $\Phi_0, \Phi_1, ..., \Phi_{k-1}$ of $\mathbb{L}$-valued random variables taking on only a finite number of values such that Φ_m are $\mathcal{F}_{t_m}$-measurable and*

$$\Phi(t) = \Phi_m \qquad t \in (t_m, t_{m+1}],\ m = 0, 1, ..., k - 1.$$

For elementary processes Φ, one defines the stochastic integral by the formula

$$(\Phi\cdot W)(t) := \int_0^t \Phi(s)dW(s) = \sum_{m=0}^{k-1} \Phi_m(W_{t_{m+1}\wedge t} - W_{t_m\wedge t}), \quad t\in[0,T] \quad (2.2)$$

In the construction of stochastic integral in a general setting, we need to work with the subspace $U_0 = Q^{1/2}(U)$ of U which is a Hilbert space endowed with the inner product

$$\langle u,v\rangle_0 = \sum_{k=1}^{\infty} \frac{1}{\lambda_k}\langle u,e_k\rangle\langle v,e_k\rangle = \langle Q^{-1/2}u, Q^{-1/2}v\rangle, \qquad u,v\in U_0.$$

Furthermore, another key concept in the construction of the stochastic integral is the space of all Hilbert-Schmidt operators from U_0 to H, $\mathbb{L}_2^0 = \mathbb{L}_2(U_0,H)$, which is a separable Hilbert space equipped with the norm

$$\begin{aligned} ||\Psi||^2_{\mathbb{L}_2^0} = \sum_{i,j=1}^{\infty} |\langle \Psi g_i, f_j\rangle|^2 &= \sum_{i,j=1}^{\infty} \lambda_i |\langle \Psi e_i, f_j\rangle|^2 \\ &= ||\Psi Q^{1/2}||^2_{\mathbb{L}_2(U,H)} \\ &= \text{Tr } [(\Psi Q^{1/2})(\Psi Q^{1/2})^*] \end{aligned}$$

where $\{g_i\}$, with $g_i = \sqrt{\lambda_i}e_i$, $\{e_i\}$ and $\{f_i\}$ are complete ONB in U_0, U, and H respectively.

Let $\Phi(t)$, $t\in[0,T]$ be a measurable $\mathbb{L}_2^0$-valued process; we define the following norm on Φ by

$$\begin{aligned} |||\Phi|||_t &= \left[\mathbb{E}\int_0^t ||\Phi(s)||^2_{\mathbb{L}_2^0} ds\right]^{1/2} \\ &= \left[\mathbb{E}\int_0^t \text{Tr } [(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*]ds\right]^{1/2} \end{aligned}$$

Proposition 2.7. *If a process Φ is elementary and $|||\Phi|||_T < \infty$, then the process $\Phi\cdot W$ is a continuous, square integrable H-valued martingle on $[0,T]$, and*

$$\mathbb{E}|\Phi\cdot W|^2 = |||\Phi|||_t^2, \quad 0\le t\le T \quad (2.3)$$

In order to take into account the adaptability of the considered processes we need to regard the class of integrands that are predictable with values in $\mathbb{L}_2^0$, more precisely, the mapping from $(\Omega \times [0,T], \mathcal{P}_T)$ into $(\mathbb{L}_2^0, \mathcal{B}(\mathbb{L}_2^0))$.

Proposition 2.8. *If Φ is an $\mathbb{L}_0^2$-predictable process such that $|||\Phi|||_T < \infty$, then there exists a sequence $\{\Phi_n\}$ of elementary processes such that $|||\Phi_n - \Phi|||_T \to 0$ as $n \to \infty$.*

Consider the set

$$\mathcal{N}_W^2(0,T;\mathbb{L}_2^0) := \{\Phi : \Omega \times [0,T] \to \mathbb{L}_2^0;\ \Phi \text{ is predictable, and } |||\Phi|||_T < \infty\}$$

$$= L^2(\Omega \times [0,T], \mathcal{P}_T, \mathbb{P} \otimes dt; \mathbb{L}_0^2)$$

or in short $\mathcal{N}_W^2$. By Propositio 2.8, the set of elementary processes is a dense subset of $\mathcal{N}_W^2$, thus with Proposition 2.7 we have that the stochastic integral $\Phi \cdot W$ is an isometric transformation from the dense set of elementary processes into $\mathscr{M}_T^2(H)$, and this is the key fact to extend the integral to all elements of $\mathcal{N}_W^2$. The so-called *localization* procedure provides the final step to extend the definition of the stochastic integral where one can relax the finite norm condition to the following

$$\mathbb{P}\left(\int_0^T \|\Phi(s)\|_{\mathbb{L}_2^0}^2 ds < \infty\right) = 1.$$

All such processes are called *stochastically integrable* on $[0,T]$.

For $\Phi \in \mathcal{N}_W^2$, one defines $\tau_n := \inf\{t \in [0,T];\ \int_0^t \|\Phi(s)\|_{\mathbb{L}_2^0}^2 ds > n\} \wedge T$, which is an increasing sequence of stopping times with respect to $\mathcal{F}_t$, $t \in [0,T]$, such that

$$\mathbb{E}\left(\int_0^T \|1_{(0,\tau_n]}(s)\Phi(s)\|_{\mathbb{L}_2^0}^2 ds\right) \le n < \infty.$$

In addition, the processes $1_{(0,\tau_n]}\Phi$, $n \in \mathbb{N}$ are still $\mathbb{L}_2^0$-predictable since $1_{(0,\tau_n]}$ is left-continuous and $\mathcal{F}_t$-adapted. Thus we get that for $t \in [0,T]$ the stochastic integrals $\int_0^t 1_{(0,\tau_n]}(s)\Phi(s)dW(s)$ are well defined for all $n \in \mathbb{N}$, and hence we set

$$\int_0^t \Phi(s)dW(s) := \int_0^t 1_{(0,\tau_n]}(s)\Phi(s)dW(s),$$

where n is an arbitrary natural number such that $\tau_n \ge t$.

Now finally the following Lemma shows that for arbitrary $m < n$ and $t \in [0,T]$ on $\{\tau_m \geq t\} \subset \{\tau_n \geq t\}$, we have

$$\int_0^t 1_{(0,\tau_n]}(s)\Phi(s)dW(s) = \int_0^t 1_{(0,\tau_m]}(s)\Phi(s)dW(s), \qquad \mathbb{P} - a.s.$$

Lemma 2.1. *Assume that $\Phi \in \mathcal{N}_W^2$ and that τ is an $\mathcal{F}_\tau$-stopping time such that $\mathbb{P}(\tau \leq T) = 1$. Then*

$$\int_0^t 1_{(0,\tau]}(s)\Phi(s)dW(s) = \Phi \cdot W(\tau \wedge t) = \int_0^{\tau \wedge t} \Phi(s)dW(s).$$

2.7 Stochastic Fubini theorem

Let (V, Σ_V) be a measurable space and let $(\omega, t, u) \to \Phi(\omega, t, u)$ be a measurable mapping from $(\Omega_T \times V, \mathbb{P}_T \otimes \mathcal{B}(V))$ into $(\mathbb{L}_2^0, \mathcal{B}(\mathbb{L}_2^0))$. In particular, for arbitrary $u \in V$, $(\cdot, \cdot, u)$ is a predictable $\mathbb{L}_2^0$-valued process. Let in addition μ be a finite positive measure on (V, Σ_V).

The following stochastic version of the Fubini theorem will be used in the sequel.

Theorem 2.1. *Let Φ described as above also satisfies*

$$\int_V |||\Phi(\cdot, \cdot, u)|||\mu(dx) < +\infty,$$

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

$$\int_V \left(\int_0^T \Phi(t,u)dW(t) \right) \mu(dx) = \int_0^T \left(\int_V \Phi(t,u)\mu(dx) \right) dW(t).$$

2.8 Itô's formula

Assume that Φ is a stochastically integrable $\mathbb{L}_2^0$-valued process in $[0,T]$, $\tilde{u}$ an H-valued predictable process $\mathbb{P}$-a.s., Bochner integrable on $[0,T]$, and $u(0)$ a $\mathcal{F}_0$-measurable H-valued random variable. Then the following process is well defined.

$$u(t) = u(0) + \int_0^t \tilde{u}(s)ds + \int_0^t \Phi(s)dW(s), \qquad t \in [0,T]. \tag{2.4}$$

Theorem 2.2. *Assume that a function $F : [0,T] \times H \to \mathbb{R}$ and its partial derivatives F_t, F_x, F_{xx} are uniformly continuous on bounded subsets of $[0,T] \times H$, then for all $t \in [0,T]$, $\mathbb{P}$-a.s. the following Itô formula holds*

$$\begin{aligned} F(t,u(t)) = F(0,u(0)) &+ \int_0^t \langle F_x(s,u(s)), \Phi(s)dW(s)\rangle \\ &+ \int_0^t \Big(F_t(s,u(s)) + \langle F_x(s,u(s)), \tilde{u}(s)\rangle\Big)ds \\ &+ \frac{1}{2}\int_0^t \mathrm{Tr}\Big(F_{xx}(s,u(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*\Big)ds, \end{aligned} \tag{2.5}$$

where $Q \in \mathbb{L}(U)$ is the covariance operator corresponding to the Q-Wiener process W.

In particular if $F(u) = \|u\|_H^2$, then the formula (2.5) for all $t \in [0,T]$, $\mathbb{P}$-a.s. reduces to the following

$$\begin{aligned} \|u(t)\|_H^2 = \|u(0)\|_H^2 &+ \int_0^t \Big(2\langle u(s), \tilde{u}(s)\rangle_H + \|\Phi(s)\|^2_{\mathbb{L}_2(U,H)}\Big)ds \\ &+ 2\int_0^t \langle u(s), \Phi(s)dW(s)\rangle_H, \end{aligned}$$

and consequently

$$\mathbb{E}(\|u(t)\|_H^2) = \mathbb{E}(\|u(0)\|_H^2) + \mathbb{E}\int_0^t \Big(2\langle u(s), \tilde{u}(s)\rangle_H + \|\Phi(s)\|^2_{\mathbb{L}_2(U,H)}\Big)ds. \tag{2.6}$$

Considering the Itô's formula as (2.6), one can obtain an analogous *integration by parts* formula for two processes u and v of the form (2.4); namely if u takes the form (2.4) and also v is given by

$$v(t) = v(0) + \int_0^t \tilde{v}(s)ds + \int_0^t \Psi(s)dW(s), \qquad t \in [0,T],$$

then the following holds

$$\begin{aligned} \mathbb{E}\int_0^T \langle u(t), \tilde{v}(t)\rangle dt = &- \mathbb{E}\int_0^T \langle v(t), \tilde{u}(t)\rangle dt - \mathbb{E}\int_0^T \langle \Phi(t), \Psi(t)\rangle_{\mathbb{L}_2(U,H)} dt \\ &+ \mathbb{E}\langle u(T), v(T)\rangle - \mathbb{E}\langle u(0), v(0)\rangle. \end{aligned} \tag{2.7}$$

Chapter 3

Self-dual Lagrangians and their variational principle

In this chapter, we state the preliminary material required for variational framework of self-dual systems. We start by recalling the basic concepts and relevant tools of convex analysis that will be frequently used in the sequel. The material in the first section of this chapter is quite standard and can be found in most books on convex analysis, such as [49] or [47]. The following sections on self-dual analysis is mostly taken from [29], which is our main reference in this direction.

In this work, the approach toward stochatic PDEs is based upon convex calculus on "phase space", $V \times V^*$, where V is a reflexive Banach space and V^* is its topological dual. We shall therefore consider Lagrangians on $V \times V^*$ that are convex and lower semi-continuous in both variables. All elements of convex analysis will apply, but the calculus on state space becomes much richer for many reasons, such as the possibility of introducing associated Hamiltonians, which are in fact the Legendre dual but in one variable. We also state the connection of class of self-dual Lagrangians on $V \times V^*$ with maximal monotone operators on V in such a way that these operators can be considered as potentials to the Lagrangians for which a generalized notion of subdifferential will be defined.

The main advantage of self-dual Lagrangians is the fact that various non-variational PDEs can be formulated variationally in the setting of *completely self-dual systems*, and a solution can be obtained by minimizing "self-dual functionals" with 0 as their minimal value.

3.1 Basic tools of convex analysis

Definition 3.1. *A function $\varphi : V \to \mathbb{R} \cup \{+\infty\}$ on a Banach space V is said to be:*

1. *(weakly) lower semi-continuous (l.s.c), if its epigraph* $\mathrm{Epi}(\varphi)$*, is closed for the (weak) norm topology, or equivalently for every sequence x_n in*

V that converges (weakly) strongly to x, then $\varphi(x) \leq \liminf_n \varphi(x_n)$.

2. *convex, if $\mathrm{Epi}(\varphi)$ is a convex subset of $V \times \mathbb{R}$, which is equivalent to $\varphi(\lambda u + (1-\lambda)v) \leq \lambda\varphi(u) + (1-\lambda)\varphi(v)$ for $u, v \in V$ and $\lambda \in \mathbb{R}$.*

3. *proper if its effective domain is non-empty i.e. $Dom(\varphi) = \{u \in V;\ \varphi(u) < +\infty\} \neq \emptyset$.*

3.1.1 Subdifferentiability of convex functions

Definition 3.2. *Let $\varphi : V \to \mathbb{R} \cup \{+\infty\}$ be a convex lower semi-continuous function on a Banach space V. Define the subdifferential $\partial\varphi$ of φ to be the following set-valued function:*

$$\partial\varphi(u) = \{p \in V^*; \langle p, v-u\rangle \leq \varphi(v) - \varphi(u) \quad \forall v \in V\},$$

and if $u \notin Dom(\varphi)$, $\partial\varphi(u) = \emptyset$.

The subdifferential $\partial\varphi(u)$ is a closed convex subset of the dual space V^*. It can, however, be empty even though $u \in Dom(\varphi)$, and we shall write $Dom(\partial\varphi) = \{u \in V; \partial\varphi(u) \neq \emptyset\}$.

Definition 3.3. *A subset $G \in V \times V^*$ is said to be*

1. ***monotone**, if for every (u,p) and (v,q) in G,*

$$\langle u-v, p-q\rangle \geq 0.$$

2. ***maximal monotone**, if it is maximal in the family of monotone subsets of $V \times V^*$ ordered by the set inclusion.*

3. ***cyclically monotone**, provided that for any finite number of points $(x_i, p_i)_{i=0}^n$ in G with $x_0 = x_n$, we have*

$$\sum_{i=1}^n \langle p_i, x_i - x_{i-1}\rangle \geq 0.$$

A set-valued map $A : V \to 2^{V^*}$ is then said to be *monotone* (resp., *maximal monotone*) provided its graph $G(A) = \{(u,p) \in V \times V^*,\ p \in A(u)\}$ is *monotone* (resp., *maximal monotone*).

The following theorem was established by Rockafellar [49].

Theorem 3.1. *Let $\varphi : V \to \mathbb{R} \cup \{+\infty\}$ be a proper convex and lower semi-continuous function on a Banach space V. Then its differential map $u \to \partial\varphi(u)$ is a maximal monotone map.*
Conversely, if $A : V \to 2^{V^*}$ is a maximal cyclically monotone map with a non-empty domain, then there exists a proper convex and lower semi-continuous functional φ on V such that $A = \partial\varphi$.

3.1.2 Legendre duality

Definition 3.4. *For a convex function $\varphi : V \to \mathbb{R} \cup \{+\infty\}$, its Fenchel-Legendre dual, $\varphi^* : V^* \to \mathbb{R} \cup \{+\infty\}$ is defined by*

$$\varphi^*(p) = \sup_{u \in V}\{\langle u, p\rangle - \varphi(u)\}$$

Proposition 3.1. *Let $\varphi : V \to \mathbb{R}\cup\{+\infty\}$ be a proper function on a reflexive Banach space. The following properties then hold:*

1. *φ^* is a proper convex lower semi-continuous function from V^* to $\mathbb{R} \cup \{+\infty\}$.*

2. *$\varphi^{**} := (\varphi^*)^* : V \to \mathbb{R} \cup \{+\infty\}$ is the largest convex lower semi-continuous function below φ. Moreover, $\varphi = \varphi^{**}$ if and only if φ is convex and lower semi-continuous on V.*

3. *For every $(u, p) \in V \times V^*$, we have $\varphi(u) + \varphi^*(p) \geq \langle u, p\rangle$, and the following are equivalent:*
    - $\varphi(u) + \varphi^*(p) = \langle u, p\rangle$
    - $p \in \partial\varphi(u)$
    - $u \in \partial\varphi^*(p)$

For a complete argument on Legendre duality, we refer the interested reader to [47, 49]

3.1.3 Legendre transform of integral functionals

Let D be a Borel subset of $\mathbb{R}^n$ with finite Lebesgue measure, and let V be a separable reflexive Banach space. Consider a bounded below function $\varphi : D \times V \to \mathbb{R} \cup \{+\infty\}$ that is measurable with respect to the σ-field generated by the products of Lebesgue sets in D and Borel sets in V. We can associate to φ a functional Φ defined on $L^\alpha(D;V)$ for $\alpha \in [1,\infty)$ via the formula

$$\Phi(u) = \int_D \varphi(x,u(x))dx, \qquad u \in L^\alpha(D;V)$$

For $x \in D, u \in V$, and $p \in V^*$, with the obvious notation

$$\varphi^*(x,p) = \varphi(x,\cdot)^*(p), \qquad \partial\varphi(x,u) = \partial\varphi(x,\cdot)(u)$$

We then have the following proposition which describes the relation between φ and its 'integral' Φ, and that how their Legendre transforms and subdifferentials are related. A proof can be found in [19].

Proposition 3.2. *Assume V is a reflexive and separable Banach space, that $1 \leq \alpha \leq +\infty$ with $\frac{1}{\alpha}+\frac{1}{\beta}=1$, and that $\varphi : D \times V \to \mathbb{R}\cup\{+\infty\}$ is jointly measurable such that $\int_D |\varphi^*(x,\bar{p}(x))|dx < \infty$ for some $\bar{p} \in L^\beta(D;V)$ which holds in particular if φ is bounded below on $D \times V$.*

1. *If the function $\varphi(x,\cdot)$ is lower semi-continuous on V for almost every $x \in D$, then Φ is lower semi-continuous on $L^\alpha(D;V)$.*
2. *If $\varphi(x,\cdot)$ is convex on V for almost every $x \in D$, then Φ is convex on $L^\alpha(D;V)$.*
3. *If $\varphi(x,\cdot)$ is convex and lower semi-continuous on V for almost every $x \in D$, and if $\Phi(\bar{u}) < \infty$ for some $\bar{u} \in L^\infty(D;V)$, then the Legendre transform of Φ on $L^\beta(D;V)$ is given by*

$$\Phi^*(p) = \int_D \varphi^*(x,p(x))dx \quad \text{for all } p \in L^\beta(D;V).$$

4. *If $\int_D |\varphi(x,\bar{u}(x))|dx < \infty$ and $\int_D |\varphi^*(x,\bar{p}(x))|dx < \infty$ for some $\bar{u}$ and $\bar{p}$ in $L^\infty(D;V)$, then for every $u \in L^\alpha(D;V)$ we have*

$$\partial\Phi(u) = \{p \in L^\beta(D;V);\ p(x) \in \partial\varphi(x,u(x)) \ a.e.\}.$$

3.2 The class of self-dual Lagrangians

Let V be a reflexive Banach space. Functions $L : V \times V^* \to \mathbb{R}\cup\{+\infty\}$ on phase space $V \times V^*$ will be called *Lagrangians*. We consider the class of Lagrangians which are proper, convex and lower semi-continuous (l.s.c) in both variables. For any $(q,v) \in V^* \times V$, the Legendre-Fenchel dual (in both variables) of L is defined by

$$L^*(q,v) = \sup\{\langle q,u\rangle + \langle v,p\rangle - L(u,p);\ u \in V, p \in V^*\}$$

The (partial) domains of a Lagrangian L are defined as

$$\mathrm{Dom}_1(L) = \{u \in V;\ L(u,p) < \infty,\ \text{for some } p \in V^*\}$$

and

$$\mathrm{Dom}_2(L) = \{p \in V^*;\ L(u,p) < \infty,\ \text{for some } u \in V\}$$

Remark 3.1. *To any pair of proper convex l.s.c functions φ and ψ on a Banach space V, one can associate a Lagrangian on state space $V \times V^*$ via the formula $L(u,p) = \varphi(u) + \psi^*(p)$. It's Legendre transform is then $L^*(p,u) = \psi(u) + \varphi^*(p)$ with $\mathrm{Dom}_1(L) = \mathrm{Dom}(\varphi)$ and $\mathrm{Dom}_2(L) = \mathrm{Dom}(\psi)$.*

The following definition represents an important class of Lagrangians on phase space called the *self-dual* Lagrangians that are the building blocks of our approach to solve certain PDEs.

Definition 3.5. *A convex lower semi-continuous Lagrangian $L : V \times V^* \to \mathbb{R} \cup \{+\infty\}$ on a reflexive Banach space V is called self-dual on $V \times V^*$ if*

$$L^*(p,u) = L(u,p),$$

and we denote this class by $\mathscr{L}^{sd}(V)$.

Remark 3.2. *By definition of the Legendre transform of a Lagrangian L, it is clear that $L(u,p) + L^*(p,u) \geq 2\langle u,p\rangle$, and for a Lagrangian $L \in \mathscr{L}^{sd}(V)$ we then have*

$$L(u,p) \geq \langle u,p\rangle, \qquad \forall\ (u,p) \in V \times V^*. \tag{3.1}$$

However, the converse is not true, i.e. if L satisfies property (3.1) it does not necessarily imply that L is self-dual.

Definition 3.6. *The Lagrangian L satisfying (3.1) is called Fenchelian. Additionally, if L (non-self-dual) satisfies*

$$L^*(p,u) \geq L(u,p) \geq \langle u,p\rangle, \qquad \forall\ (u,p) \in V \times V^*,$$

it is called a subself-dual Lagrangian.

Example 1. *The basic self-dual Lagrangians:*

- *A simple but important example of a self-dual Lagrangian on $V \times V^*$ is defined via the formula*

$$L(u,p) = \varphi(u) + \varphi^*(p)$$

where φ is any convex l.s.c function on V.

- *Let $\Gamma : V \to V^*$ be a bounded linear operator which is skew-adjoint, and φ a convex l.s.c function, then $L(u,p) = \varphi(u) + \varphi^*(\Gamma u + p)$ is self-dual.*

More elaborate examples will be devised later, though all constructions will be based on these important elements of $\mathscr{L}^{sd}(V)$.

Operations on Lagrangians

Let L and N be two given Lagrangians, we consider the following operations on $V \times V^*$

- **Addition:** $(L \oplus N)(u,p) = \inf_{r \in V^*} \{L(u,r) + N(u,p-r)\}$.
- **Convolution:** $(L \star N)(u,p) = \inf_{z \in V} \{L(z,p) + N(u-z,p)\}$.

Then we have the following lemma.

Lemma 3.1. ([29] Proposition 3.4) *Let V be a reflexive Banach space, and L and N two Lagrangians, then*

1. *If $Dom_1(L) - Dom_1(N)$ (resp,. $Dom_2(L^*) - Dom_2(N^*)$) contains a neighborhood of the origin, then*

$$(L \oplus N)^* = L^* \star N^*, \qquad (\textit{resp.}, \ (L \star N)^* = L^* \oplus N^*)$$

2. *If $L, N \in \mathscr{L}^{sd}(V)$ such that $Dom_1(L) - Dom_1(N)$ contains a neighborhood of the origin, then $L \oplus N$ and $L \star N$ are also in $\mathscr{L}^{sd}(V)$.*

3.3 Self-dual vector fields

In this section, we introduce the concept of *self-dual vector fields*, which are the natural extension of subdifferential of convex lower semi-continuous functions. Moreover, it is proved that self-dual Lagrangians play the role of potentials of maximal monotone vector fields in a way similar to how convex energies are the potentials of their own subdifferentials.

Definition 3.7. *Let $L : V \times V^* \to \mathbb{R} \cup \{+\infty\}$ be a Lagrangian, its symmetrized vector field at $u \in V$ is defined as the -possibly empty- subset of V^* given by*

$$\bar{\partial} L(u) = \{p \in V^*;\ L(u,p) + L^*(p,u) = 2\langle u,p\rangle\}$$

If L is convex and l.s.c. on $V \times V^*$, then

$$\bar{\partial} L(u) = \{p \in V^*;\ (p,u) \in \partial L(u,p)\}$$

If now L is a self-dual Lagrangian, then

$$\bar{\partial} L(u) = \{p \in V^*;\ L(u,p) - \langle u,p\rangle = 0\}$$

A *self-dual vector field* is any map $F : \mathrm{Dom}(F) \subset V \to 2^{V^*}$ if there exists a self-dual Lagrangian L on $V \times V^*$ such that $F(u) = \bar{\partial} L(u)$ for every $u \in \mathrm{Dom}(F)$. A key point is that self-dual Lagrangians on phase space necessarily satisfy $L(u,p) \geq \langle u,p\rangle$ for all $(u,p) \in V \times V^*$, and therefore the zeros of a self-dual vector field (i.e, $0 \in \bar{\partial} L(u)$) can be obtained by simply minimizing the functional $I(u) = L(u,0)$ and proving that the value of the minimum is actually zero.

For a general Lagrangian L we also define the vector field

$$\delta L(u) = \{p \in V^*;\ L(u,p) - \langle u,p\rangle = 0\}.$$

It is easy to show that if L is Fenchelian i.e. it satisfies $L(u,p) \geq \langle u,p\rangle$, for every $(u,p) \in V \times V^*$, then $\delta L(u) \subset \bar{\partial} L(u)$; and clearly, if L is self-dual then $\delta L = \bar{\partial} L$.

Examples

- For a basic self-dual Lagrangian of the form $L(u,p) = \varphi(u) + \varphi^*(p)$, where φ is a convex l.s.c function on V, and φ^* is its Legendre conjugate on V^*, we have
$$\bar{\partial} L(u) = \partial\varphi(u).$$
In order to solve equations of the form $0 \in \partial\varphi(u)$, the corresponding variational problem reduces to minimizing the convex functional $I(u) = L(u,0) = \varphi(u) + \varphi^*(0)$.

- More interesting examples of self-dual Lagrangians are of the form $L(u,p) = \varphi(u) + \varphi^*(-\Gamma u + p)$, where φ is a convex and l.s.c function on V and $\Gamma : V \to V^*$ is a skew-adjoint operator. The corresponding self-dual vector field is then
$$\bar{\partial} L(u) = \Gamma u + \partial\varphi(u).$$

Proposition 3.3. *For any convex lower semi-continuous Lagrangian L on $V \times V^*$, the map $u \mapsto \bar{\partial} L(u)$ is monotone.*

Proposition 3.4. *Let L be a self-dual Lagrangian on $V \times V^*$, then the following are equivalent:*

1. $p \in \bar{\partial} L(u)$
2. *The infimum of the functional $I_p(u) = L(u,p) - \langle u, p\rangle$ over all $u \in V$ is zero and is attained at some $v \in V$.*

The following lemmas are used to prove how to associate a self-dual Lagrangian to a maximal monotone operator in such a way that the operator coincides with the vector field corresponding to the Lagrangian. The Lagrangian L_A defined in (3.2) is called the *Fitzpatrick function*.

Lemma 3.2. *Let $A : \mathrm{Dom}(A) \subset V \to 2^{V^*}$ be a maximal monotone operator. Consider on $V \times V^*$ the Lagrangian L_A defined by*

$$L_A(u,p) = \sup\{\langle p, v\rangle + \langle q, u - v\rangle;\ (v,q) \in G(A)\}. \tag{3.2}$$

Then

1. *if* $\mathrm{Dom}(A) \neq \emptyset$, *$L_A$ is convex and lower semi-continuous on $V \times V^*$.*
2. $L_A = H_A^*$ *where* $H_A(u,p) = \begin{cases} \langle u, p\rangle & \text{if } (u,p) \in G(A) \\ +\infty & \text{elsewhere.} \end{cases}$
3. *for every $u \in \mathrm{Dom}(A)$, we have $A = \bar{\partial} L_A = \delta L_A$. Moreover, L_A is subself-dual i.e.*

$$L_A^*(p,u) \geq L_A(u,p) \geq \langle u, p\rangle, \qquad \text{for every } (u,p) \in V \times V^*.$$

The following lemma is on interpolating convex functions, which is sometimes in the literature called the *proximal average* of two convex functions.

Lemma 3.3. *Let $f_1, f_2 : V \to \mathbb{R} \cup \{+\infty\}$ be two convex l.s.c. functions on a reflexive Banach space V. For $u \in V$, consider the proximal average of f_1 and f_2 as the following*

$$\bar{f}(u) := \inf\Big\{\frac{1}{2} f_1(u_1) + \frac{1}{2} f_2(u_2) + \frac{1}{8}\|u_1 - u_2\|^2;\ u_1, u_2 \in V, u = \frac{1}{2}(u_1 + u_2)\Big\}.$$

Then the following properties hold

1. $(\frac{1}{2}f_1^* + \frac{1}{2}f_2^*)^* \le \bar{f} \le \frac{1}{2}f_1 + \frac{1}{2}f_2$, *which implies that if* $f_1 \le f_2$ *then* $f_1 \le \bar{f} \le f_2$.

2. ***Legendre dual.*** *For* $p \in V^*$, *we have*

$$(\bar{f})^*(p) := \inf\Big\{\frac{1}{2}f_1^*(p_1)+\frac{1}{2}f_2^*(p_2)+\frac{1}{8}\|p_1-p_2\|^2;\, p_1,p_2 \in V^*, p = \frac{1}{2}(p_1+p_2)\Big\}.$$

3. *Denote the proximal average between* f_1 *and* f_2 *by* $\bar{f} =: [f_1, f_2]$, *then*

$$[f_1, f_2]^* = [f_1^*, f_2^*].$$

We are now ready to state the theorem which establishes a one-to-one correspondence between maximal monotone operators and self-dual vector fields.

Theorem 3.2. ***(Self-dual vector fields and maximal monotone operators)*** *If* $A : D(A) \subset V \to 2^{V^*}$ *is a maximal monotone operator with a non-empty domain, then there exists a self-dual Lagrangian* N *on* $V \times V^*$ *such that* $A = \bar{\partial} N$. *Conversely, if* N *is a proper self-dual Lagrangian on a reflexive Banach space* $V \times V^*$, *then the vector field* $u \mapsto \bar{\partial} N(u)$ *is maximal monotone.*

Proof. Since A is maximal monotone, we first via Lemma 3.2 associate the subself-dual Lagrangian L_A so that

$$A = \bar{\partial} L_A \quad \text{and} \quad L_A^*(p,u) \ge L_A(u,p) \ge \langle u,p\rangle. \tag{3.3}$$

Now we claim that the desired self-dual Lagrangian corresponding to the operator A is the proximal average between L_A and L_A^*, namely $N := [L_A, L_A^*]$ given by

$$N(u,p) = \inf\Big\{\frac{1}{2}L_A(u_1,p_1) + \frac{1}{2}L_A^*(p_2,u_2) + \frac{1}{8}\|u_1-u_2\|^2 + \frac{1}{8}\|p_1-p_2\|^2;\, (u,p) = \frac{1}{2}(u_1,p_1) + (u_2,p_2)\Big\}.$$

By Lemma 3.3, since $L_A \le L_A^*$, we have

$$L_A^*(p,u) \ge N(u,p) \ge L_A(u,p) \qquad \text{for every } (u,p) \in V \times V^*,$$

also it is clear that N is self-dual since

$$N^* = [L_A, L_A^*]^* = [L_A^*, L_A^{**}] = [L_A, L_A^*] = N.$$

Now we prove $\bar{\partial}N = \bar{\partial}L_A$. First, take $p \in \bar{\partial}N(u)$, since N is self-dual, this implies that $N(u,p) - \langle u,p\rangle = 0$. Now from $L_A(u,p) \le N(u,p)$ we deduce that $L_A(u,p) \le \langle u,p\rangle$ which by (3.3) implies $L_A(u,p) = \langle u,p\rangle$, thus $p \in \delta L_A(u) \subset \bar{\partial}L_A$.
Now if $p \in \bar{\partial}L_A(u)$, then we have $L_A(u,p) + L_A^*(p,u) = 2\langle u,p\rangle$, also by Lemma 3.2 we have $(u,p) \in G(A)$. Monotonicity of A implies that

$$\langle u,p\rangle \ge \langle v,p\rangle + \langle u-v,q\rangle,$$

which from definition of L_A yields $L_A(u,p) \le \langle u,p\rangle$ and with (3.3) we conclude $L_A(u,p) = \langle u,p\rangle = H_A(u,p)$, thus $L_A^*(p,u) = \langle u,p\rangle$ and therefore

$$N(u,p) = \langle u,p\rangle.$$

Hence $p \in \bar{\partial}N(u)$. By Lemma 3.2 since $A = \bar{\partial}L_A$ then $Au = \bar{\partial}N(u)$ for every $(u,p) \in V \times V^*$ with $u \in \mathrm{Dom}(A)$, and this completes the proof.

For the proof of the converse we refer the reader to [30]. □

3.4 Variational principle for self-dual functionals

Several differential systems often because of lack of self-adjointness or linearity cannot be expressed as Euler-Lagrange equations, but they can be written in the form

$$0 \in \bar{\partial}L(u),$$

where L is a self-dual Lagrangian on phase space $V \times V^*$ and V is a reflexive Banach space. These are the *completely self-dual systems*, and a solution to these systems can be considered as minimizers of a *completely self-dual functional* I for which the minimum value is 0.

The fact that the infimum of the functional I is zero follows from the basic duality theory in convex analysis, which in the self-dual setting leads to a situation where the value of the dual problem is exactly the negative of the value of the primal problem, hence leading to zero as soon as there is no duality gap.

3.4.1 Evolution triples and self-dual Lagrangians

A common framework for PDEs and evolution equations is the so-called *evolution triple* of Gelfand, which is a natural setting to obtain more regular solutions that are valued in suitable Sobolev spaces, as opposed to just

L^2; moreover to relax the restrictive coercivity condition on the underlying Hilbert space. It consists of a Hilbert space H with $\langle\, ,\rangle_H$ as its scalar product, and the space V which is equipped with the norm $\|\cdot\|_V$ that makes it a reflexive Banach space and such that the canonical injection $V \to H$ is continuous. We identify the Hilbert space H with its dual H^* and we inject H in V^* via the following:

Let $\langle\, ,\rangle_*$ denotes the dualization between V^* and V. For each $h \in H$, the map $Th : u \in V \to \langle h, u\rangle_H$ is a continuous linear functional on V in such a way that

$$\langle Th, u\rangle_* = \langle h, u\rangle_H \qquad \forall\, h \in H,\, u \in V.$$

One can easily see that $T : H \to V^*$ is continuous, one-to-one, and that $T(H)$ is dense in V^*. In other words, one can place H in V^* such that $V \subset H \cong H^* \subset V^*$, where the injections are continuous and with dense range. We note that with such an identification the duality $\langle f, u\rangle_*$ coincides with the scalar product $\langle f, u\rangle_H$ as soon as $f \in H$ and $u \in V$.

A typical evolution triple is $V := H_0^1(D) \subset H := L^2(D) \subset V^* := H^{-1}(D)$, where D is a bounded domain in $\mathbb{R}^n$. The following lemma explains the connection between self-duality on H and V.

Lemma 3.4. ([29] Lemma 3.4) *Let $V \subset H \subset V^*$ be an evolution triple, and suppose $L : V \times V^* \to \mathbb{R} \cup \{+\infty\}$ is a self-dual Lagrangian on the Banach space V, that satisfies for some $C_1, C_2 > 0$ and $r_1 \geq r_2 > 1$,*

$$C_2(\|u\|_V^{r_2} - 1) \leq L(u, 0) \leq C_1(1 + \|u\|_V^{r_1}) \qquad \text{for all } u \in V.$$

Then, the Lagrangian defined on $H \times H$ by

$$\bar{L}(u,p) := \begin{cases} L(u,p) & u \in V \\ +\infty & u \in H\backslash V \end{cases}$$

is self-dual on the Hilbert space $H \times H$.

3.4.2 Primal and dual convex optimization problem

Consider the *primal problem* to be the minimization of a convex lower semi-continuous function I that is bounded below on a Banach space V, i.e.

$$\inf_{u \in V} I(u), \tag{3.4}$$

we consider the convex lower semi-continuous Lagrangian $L : V \times Y \to \mathbb{R} \cup \{+\infty\}$ such that I can be written as

$$I(u) = L(u, 0) \qquad \text{for all } u \in V.$$

For any $p \in Y$, we consider the perturbed minimization problem

$$(\mathscr{P}_p) \qquad \inf_{u \in V} L(u, p), \tag{3.5}$$

for which $(\mathscr{P}_0)$ is clearly the initial primal problem (3.4). Now we consider the *dual problem* on $V^* \times Y^*$

$$(\mathscr{P}^*) \qquad \sup_{p^* \in Y^*} -L^*(0, p^*), \tag{3.6}$$

and we define the function $h : Y \to \mathbb{R} \cup \{+\infty\}$ on the space of perturbations Y

$$h(p) = \inf_{u \in V} L(u, p) \qquad \text{for all } p \in Y.$$

The following proposition summarizes the relationship between the primal and dual problems and the behavior of the value function h. A proof can be found in [29].

Proposition 3.5. *Assume L is a proper convex lower semi-continuous Lagrangian that is bounded below on $V \times Y$. Then, the following assertions hold:*

1. $-\infty < \sup_{p^* \in Y^*} - L^*(0, p^*) \leq \inf_{u \in V} L(u, 0) < +\infty.$
2. *h is a convex function on Y such that*

   $$h^*(p^*) = L^*(0, p^*) \qquad \text{for every } p^* \in Y^*,$$

   and $\qquad h^{**}(0) = \sup_{p^* \in Y^*} -L^*(0, p^*).$
3. *h is lower semi-continuous at 0 (i.e. the problem (3.4) is normal) if and only if there is no duality gap, i.e. if*

   $$\sup_{p^* \in Y^*} -L^*(0, p^*) = \inf_{u \in V} L(u, 0).$$
4. *h is subdifferentiable at 0 if and only if $(\mathscr{P}_0)$ is normal and $(\mathscr{P}^*)$ has at least one solution. Moreover, the set of solutions for $(\mathscr{P}^*)$ is equal to $\partial h^{**}(0)$.*

5. *If for some* $u_0 \in V$ *the function* $p \to L(u_0, p)$ *is bounded on a ball centered at* 0 *in* Y*, then* $(\mathscr{P}_0)$ *and* $(\mathscr{P}^*)$ *has at least one solution.*

Before stating the minimization principle of completely self-dual functionals, we should consider the following relaxed version of self-duality which is sufficient for our purpose in this setting.

Definition 3.8. *Let* L *be a Lagrangian on* $V \times V^*$*, we say that* L *is partially self-dual if*

$$L^*(0, u) = L(u, 0).$$

Definition 3.9. *A function* $I : V \to \mathbb{R} \cup \{\infty\}$ *is said to be a completely self-dual functional on the Banach space* V *if there exists a partially self-dual Lagrangian* L *on* $V \times V^*$ *such that*

$$I(u) = L(u, 0), \quad \text{for } u \in V.$$

The following theorem is the key variational principle for minimization of self-dual Lagrangians on $V \times V^*$, and in particular it proves that the value of the minimum is 0.

Theorem 3.3. *Let* L *be partially self-dual (convex lower semi-continuous) on* $V \times V^*$*, and also assume that the mapping* $u \to L(u, 0)$ *is coercive in the sense that*

$$\lim_{\|u\| \to \infty} \frac{L(u, 0)}{\|u\|} = +\infty \tag{3.7}$$

then there exists $\bar{u} \in V$ *such that*

$$I(\bar{u}) = \inf_{u \in V} L(u, 0) = 0$$

Proof. Following the duality theory in convex optimization described in Proposition 3.5, we consider the minimization problem $h(p) = \inf_{u \in V} L(u, p)$ in such a way that $h(0) = \inf_{u \in V} L(u, 0)$ is the initial problem, and then the dual problem is $\sup_{v \in V} - L^*(0, v)$.

By assertion (1) in Proposition 3.5 and noting that L is partially self-dual we have

$$\inf_{u \in V} L(u, 0) \geq \sup_{v \in V} -L^*(0, v) = \sup_{v \in V} -L(v, 0) = -\inf_{v \in V} L(v, 0),$$

thus

$$\inf_{u \in V} L(u, 0) \geq 0.$$

Also by assertion (2) of Proposition 3.5, h is convex on V^* and its Legendre transform on V satisfies $h^*(v) = L^*(0, v) = L(v, 0)$, which by coercivity condition (3.7) implies that h^* is coercive, and this means that h is bounded above on neighborhoods of 0 in V^*. Therefore, h is subdifferentiable at 0, and there exists $\bar{u} \in \partial h(0)$, which is equivalent to $h(0) + h^*(\bar{u}) = 0$.

Now we have

$$\inf_{u \in V} L(u, 0) = h(0) = -h^*(\bar{u}) = -L^*(0, \bar{u}) = -L(\bar{u}, 0) \leq -\inf_{u \in V} L(u, 0),$$

which implies $\inf_{u \in V} L(u, 0) \leq 0$. Hence, we conclude that $\inf_{u \in V} L(u, 0) = L^*(0, \bar{u}) = 0$, and (3.4) is attained at $\bar{u} \in V$, i.e.

$$I(\bar{u}) = L(\bar{u}, 0) = \inf_{u \in V} L(u, 0) = 0.$$

□

3.5 The class of antisymmetric Hamiltonians

There are several examples of PDEs which cannot be expressed in terms of completely self-dual systems but they can be written in the form

$$0 \in \Lambda u + \bar{\partial} L(u),$$

where L is a self-dual Lagrangina on $V \times V^*$, and $\Lambda : D(\Lambda) \subset V \to V^*$ is a, not necessarily linear, operator. They can be solved by minimizing the functional

$$I(u) = L(u, -\Lambda u) + \langle \Lambda u, u \rangle,$$

on V by showing that their infimum is zero and that it is attained. In [29], these are called *self-dual functionals*, which take the form $I(u) = \sup_{v \in V} M(u, v)$ where M is an antisymmetric Hamiltonian on $V \times V$. In the following we first introduce the Hamiltonian corresponding to a self-dual Lagrangian, and in turn the class of antisymmetric Hamiltonians and their variational framework.

For each Lagrangian L on $V \times V^*$, we can define its corresponding Hamiltonian $H_L : V \times V \to \mathbb{R} \cup \{+\infty\}$ by

$$H_L(u, v) = \sup_{p \in V^*} \{\langle v, p \rangle - L(u, p)\},$$

which is the Legendre transform in the second variable and its effective domain is

$$\begin{aligned} Dom_1(H_L) := &\{u \in V; H_L(u,v) > -\infty, \text{ for some } v \in V\} \\ = &\{u \in V; L(u,p) < +\infty, \text{ for some } p \in V^*\} \\ = &\, Dom_1(L). \end{aligned}$$

It is easy to see that if L is a self-dual Lagrangian on $V \times V^*$, then its Hamiltonian on $V \times V$ satisfies the following properties:

- H_L is concave in u and convex lower semi-continuous in v.
- $H_L(v,u) \le -H_L(u,v)$ for all $u, v \in V$.

Definition 3.10. *Let G be a convex subset of a reflexive Banach space V. A functional $M : G \times G \to \mathbb{R} \cup \{+\infty\}$ is said to be an antisymmetric (AS) Hamiltonian on $G \times G$ if it satisfies the following conditions:*

1. *For every $u \in G$, the function $v \to M(u,v)$ is proper concave.*
2. *For every $v \in G$, the function $u \to M(u,v)$ is weakly lower semicontinuous.*
3. *$M(u,u) \le 0$ for every $u \in V$.*

3.5.1 Variational principle for self-dual functionals

Definition 3.11. (Self-dual functional) *Let $I : V \to \mathbb{R} \cup \{+\infty\}$ be a functional on a Banach space V.*

1. *I is self-dual on a convex set $G \subset V$ if it is non-negative and if there exists an AS-Hamiltonian $M : G \times G \to \mathbb{R} \cup \{+\infty\}$ such that*

$$I(u) = \sup_{v \in G} M(u,v), \qquad \text{for every } u \in G.$$

2. *A self-dual functional I is strongly coercive on G if for some $v_0 \in G$ the set $\mathcal{G}_0 = \{u \in G;\ M(u,v_0) \le 0\}$ is bounded in V, where M is the corresponding AS-Hamiltonian in part (1).*

Example 2. *For a self-dual Lagrangian L, the completely self-dual functional $I(u) = L(u, 0)$ as in the preceding section is self-dual. In fact;*

$$I(u) = L(u, 0) = \sup\{H_L(v, u); v \in Dom_1(L)\} \qquad \text{for every } u \in Dom_1(L),$$

and the strong corecivity is equivalent to the condition that, for some $v_0 \in V$ we have

$$\lim_{\|u\| \to +\infty} H_L(v_0, u) = +\infty.$$

In particular, if $L(u, p) = \varphi(u) + \varphi^*(p)$ for φ a proper, convex and lower semicontinuous function, the strong coercivity is simply equivalent to the coercivity of φ.

The following theorem states the main variational principle for self-dual functionals.

Theorem 3.4. *If $I : G \to \mathbb{R} \cup \{+\infty\}$ is a self-dual functional that is strongly coercive on a closed convex subset G of a reflexive Banach space V, then there exists $\bar{u} \in G$ such that*

$$I(\bar{u}) = \inf_{u \in G} I(u) = 0.$$

The proof is based on the important min-max principle that is due to K. Fan [22], and it builds upon the following primary lemma.

Lemma 3.5. *Let G be a closed convex subset of a reflexive Banach space V, and consider $M : G \times G \to \mathbb{R} \cup \{+\infty\}$ to be a functional such that*

- (i) *For each $v \in G$, the map $u \to M(u, v)$ is weakly l.s.c on G.*
- (ii) *For each $u \in G$, the map $v \to M(u, v)$ is concave on G.*
- (iii) *There exists $\gamma \in \mathbb{R}$, such that $M(u, u) \le \gamma$ for every $u \in G$.*
- (iv) *There exists $v_0 \in G$, such that $\mathcal{G}_0 = \{u \in G : M(u, v_0) \le \gamma\}$ is bounded.*

Then there exists $\bar{u} \in G$, such that $M(\bar{u}, v) \le \gamma$ for all $v \in G$.

Proof of Theorem 3.4. I is a self-dual functional which means that it is non-negative and for every $u \in G$, it can be written as

$$I(u) = \sup_{v \in G} M(u, v),$$

where M is an AS-Hamiltonian on $G \times G$. By the definition of an AS-Hamiltonian, M satisfies conditions (i)-(iii) of Lemma 3.5 with $\gamma = 0$. condition (iv) is also satisfied with the strong coercivity assumption. Therefore, Lemma 3.5 yields the existence of $\bar{u} \in G$ such that $M(\bar{u}, v) \leq 0$ for all $v \in G$ and hence

$$0 \geq \sup_{v \in G} M(\bar{u}, v) = I(\bar{u}) \geq 0.$$

□

As established in [28], the Hamiltonian formulation allows for the minimization of direct sums of self-dual functionals. The following variational principle is useful in the case when non-linear and unbounded operators are involved.

Theorem 3.5. (Ghoussoub [29]) *Consider three reflexive Banach spaces Z, X_1, X_2 and operators $A_1 : D(A_1) \subset Z \to X_1$, $\Gamma_1 : D(\Gamma_1) \subset Z \to X_1^*$, $A_2 : D(A_2) \subset Z \to X_2$, and $\Gamma_2 : D(\Gamma_2) \subset Z \to X_2^*$, such that A_1 and A_2 are linear, while Γ_1 and Γ_2 –not necessarily linear– are weak-to-weak continuous. Suppose G is a closed linear subspace of Z such that $G \subset D(A_1) \cap D(A_2) \cap D(\Gamma_1) \cap D(\Gamma_2)$, while the following properties are satisfied:*

1. *The image of $G_0 := \mathrm{Ker}(A_2) \cap G$ by A_1 is dense in X_1.*
2. *The image of G by A_2 is dense in X_2.*
3. *$u \to \langle A_1 u, \Gamma_1 u\rangle + \langle A_2 u, \Gamma_2 u\rangle$ is weakly upper semi-continuous on G.*

Let $L_i, i = 1, 2$ be self-dual Lagrangians on $X_i \times X_i^*$ such that the Hamiltonians H_{L_i} are continuous in the first variable on X_i.

(*i*) *The functional*

$$I(u) = L_1(A_1 u, \Gamma_1 u) - \langle A_1 u, \Gamma_1 u\rangle + L_2(A_2 u, \Gamma_2 u) - \langle A_2 u, \Gamma_2 u\rangle$$

is self-dual on G, and its corresponding AS-Hamiltonian on $G \times G$ is

$$\begin{aligned} M(u, z) = H_{L_1}(A_1 z, A_1 u) + \langle A_1(z - u), \Gamma_1 u\rangle \\ + H_{L_2}(A_2 z, A_2 u) + \langle A_2(z - u), \Gamma_2 u\rangle. \end{aligned}$$

(*ii*) *Consequently under the coercivity condition,*

$$\lim_{\substack{\|u\| \to \infty \\ u \in G}} H_{L_1}(0, A_1 u) - \langle A_1 u, \Gamma_1 u\rangle + H_{L_2}(0, A_2 u) - \langle A_2 u, \Gamma_2 u\rangle = +\infty, \tag{3.8}$$

I attains its minimum at a point $v \in G$ such that $I(v) = 0$, and

$$\begin{aligned}\Gamma_1(v) &\in \bar{\partial} L_1(A_1 v),\\ \Gamma_2(v) &\in \bar{\partial} L_2(A_2 v).\end{aligned} \tag{3.9}$$

Proof. (i) First, we show that $M(u,z)$ is actually an AS-Hamiltonian. To this end, we prove that M satisfies the properties given in Definition 3.10. For each $u \in G$, we have that $z \to M(u,z)$ is concave since A_i's are linear, also $z \to H_{L_i}(A_i z, A_i u)$ is concave. For each $z \in G$, the function $u \to M(u,z)$ is weakly lower semi-continuous due to upper semi-continuity of $u \to \sum_{i=1}^{2} \langle A_i u, \Gamma_i u\rangle$, and the fact that Γ_i's are weak-to-weak continuous and $u \to H_{L_i}(A_i z, A_i u)$ is weakly lower semi-continuous for linear A_i. Finally, since $H_{L_i}(x,y) \le -H_{L_i}(y,x)$ for all $x, y \in X_i$, we deduce that $M(u,u) \le 0$.

Next, we note that for every $u \in D(A_i) \cap D(\Gamma_i) \subset Z$ we have

$$\begin{aligned}L_i(A_i u, \Gamma_i u) &= L_i^*(\Gamma_i u, A_i u)\\ &= \sup_{r \in X_i} \sup_{p \in X_i^*} \{\langle r, \Gamma_i u\rangle + \langle p, A_i u\rangle - L_i(r,p)\}\\ &= \sup_{r \in X_i} \Big\{\langle r, \Gamma_i u\rangle + \sup_{p \in X_i^*}\{\langle p, A_i u\rangle - L_i(r,p)\}\Big\}\\ &= \sup_{r \in X_i} \{\langle r, \Gamma_i u\rangle + H_{L_i}(r, A_i u)\}.\end{aligned} \tag{3.10}$$

Since A_i are liner operators, taking into account condition (1), we write

$$\begin{aligned}\sup_{z \in G} M(u,z) &= \sup_{z \in G} \Big\{\langle A_1 z, \Gamma_1 u\rangle + H_{L_1}(A_1 z, A_1 u) + \langle A_2 z, \Gamma_2 u\rangle\\ &\qquad + H_{L_2}(A_2 z, A_2 u)\Big\} - \langle A_1 u, \Gamma_1 u\rangle - \langle A_2 u, \Gamma_2 u\rangle\\ &= \sup_{z \in G, z_0 \in G_0} \Big\{\langle A_1 z, \Gamma_1 u\rangle + H_{L_1}(A_1 z, A_1 u)\\ &\qquad + \langle A_2(z+z_0), \Gamma_2 u\rangle + H_{L_2}(A_2(z+z_0), A_2 u)\Big\}\\ &\qquad - \langle A_1 u, \Gamma_1 u\rangle - \langle A_2 u, \Gamma_2 u\rangle\\ &= \sup_{w \in G, z_0 \in G_0} \Big\{\langle A_1(w-z_0), \Gamma_1 u\rangle + H_{L_1}(A_1(w-z_0), A_1 u)\\ &\qquad + \langle A_2 w, \Gamma_2 u\rangle + H_{L_2}(A_2 w, A_2 u)\Big\}\\ &\qquad - \langle A_1 u, \Gamma_1 u\rangle - \langle A_2 u, \Gamma_2 u\rangle\end{aligned}$$

$$
\begin{aligned}
&= \sup_{w\in G, r\in X_1} \Big\{ \langle A_1 w + r, \Gamma_1 u\rangle + H_{L_1}(A_1 w + r, A_1 u) \\
&\qquad + \langle A_2 w, \Gamma_2 u\rangle + H_{L_2}(A_2 w, A_2 u) \Big\} \\
&\qquad - \langle A_1 u, \Gamma_1 u\rangle - \langle A_2 u, \Gamma_2 u\rangle \\
&= \sup_{w\in G, x\in X_1} \Big\{ \langle x, \Gamma_1 u\rangle + H_{L_1}(x, A_1 u) + \langle A_2 w, \Gamma_2 u\rangle \\
&\qquad + H_{L_2}(A_2 w, A_2 u) \Big\} - \langle A_1 u, \Gamma_1 u\rangle - \langle A_2 u, \Gamma_2 u\rangle
\end{aligned}
$$

In view of condition (2) and (3.10) we then have

$$
\begin{aligned}
\sup_{z\in G} M(u,z) &= \sup_{x\in X_1} \sup_{y\in X_2} \Big\{ \langle x, \Gamma_1 u\rangle + H_{L_1}(x, A_1 u) + \langle y, \Gamma_2 u\rangle \\
&\qquad + H_{L_2}(y, A_2 u) \Big\} - \langle A_1 u, \Gamma_1 u\rangle - \langle A_2 u, \Gamma_2 u\rangle \\
&= \sup_{x\in X_1} \Big\{ \langle x, \Gamma_1 u\rangle + H_{L_1}(x, A_1 u) \Big\} - \langle A_1 u, \Gamma_1 u\rangle \\
&\qquad + \sup_{y\in X_2} \Big\{ \langle y, \Gamma_2 u\rangle + H_{L_2}(y, A_2 u) \Big\} - \langle A_2 u, \Gamma_2 u\rangle \\
&= L_1(A_1 u, \Gamma_1 u) - \langle A_1 u, \Gamma_1 u\rangle + L_2(A_2 u, \Gamma_2 u) - \langle A_2 u, \Gamma_2 u\rangle \\
&= I(u).
\end{aligned}
$$

For part (ii), it suffices to apply Theorem 3.4 to get that $I(v) = 0$ for some $v \in G$. Now use the fact that for self-dual $L_i, i = 1,2$, we have $L_i(A_i u, \Gamma_i u) - \langle A_i u, \Gamma_i u\rangle \geq 0$ to conclude (3.9). □

Chapter 4

Self-dual variational principle for stochastic partial differential equations with additive noise

4.1 Introduction

In Chapter 3, we provided the reader with the variational framework corresponding to self-dual Lagrangians and we observed that the minimization of (completely) self-dual functionals leads to existence of solutions to certain PDEs. A main application of this approach is to solve equations that are not in the standard Euler-Lagrange form, namely the equations for which the classical variational setting does not apply. One important class of such equations is the family of stochastic partial differential equations on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with normal filtration $(\mathcal{F}_t)_t$. The evolutionary nature of the stochastic equations together with the presence of the stochastic integral with respect to a Wiener process make them of non-variational structure.

To introduce the method, as stated in Chapter 1, we first consider the simple case

$$\begin{cases} du(t) = -\partial\varphi(t, u(t))dt + B(t)dW(t) \\ u(0) = u_0. \end{cases} \tag{4.1}$$

where $\varphi : [0,T] \times H \to \mathbb{R} \cup \{+\infty\}$ is a random and progressively measurable function such that for every $t \in [0,T]$, the function $\varphi(t,\cdot)$ is convex and lower semi-continuous on a Hilbert space H, and the stochastics is driven by a given progressively measurable additive noise coefficient $B : \Omega \times [0,T] \to H$. The main idea is that a solution for (4.1) is in fact the minimum of the following functional on $\mathcal{A}_H^2$,

$$I(u) = \mathbb{E}\Big\{ \int_0^T L_\varphi(u(t), -\tilde{u}(t))\, dt + \frac{1}{2}\int_0^T M_B(F_u(t), -F_u(t))\, dt \\ + \ell_{u_0}(u(0), u(T))\Big\},$$

where

$$\mathcal{A}_H^2 = \left\{ u : \Omega_T \to H;\ u(t) = u_0 + \int_0^t \tilde{u}(s)ds + \int_0^t F_u(s)dW(s) \right\},$$

is the *Itô space over* H, for progressively measurable processes $\tilde{u}$ and F_u, and L_φ, M_B and ℓ_{u_0} are the Lagrangians given by (1.3), (1.4) and (1.5) respectively. We can then apply the self-dual variational principle given in Theorem 3.3 to prove that I attains its infimum on $v \in \mathcal{A}_H^2$, and moreover, the value of the infimum is 0 to obtain

$$0 = I(v) = \mathbb{E}\int_0^T \Big(\varphi(t,v) + \varphi^*(t, -\tilde{v}(t))\Big)\, dt \\ + \mathbb{E}\,\Big(\frac{1}{2}\|v(0)\|_H^2 + \frac{1}{2}\|v(T)\|_H^2 - 2\langle u_0, v(0)\rangle + \|u_0\|_H^2\Big) \\ + \mathbb{E}\int_0^T \Big(\frac{1}{2}\,\|F_v(t) - 2B(t)\|_H^2 + \frac{1}{2}\,\|F_v(t)\|_H^2 - 2\langle F_v(t), B(t)\rangle\Big)\, dt,$$

By using Itô's formula and adding and subtracting the term $\mathbb{E}\int_0^T \langle v(t), \tilde{v}(t)\rangle dt$, we can rewrite $I(v)$ in the form

$$0 = I(v) = \mathbb{E}\int_0^T \Big(\varphi(t, v(t)) + \varphi^*(t, -\tilde{v}(t)) + \langle v(t), \tilde{v}(t)\rangle\Big)\, dt \\ + 2\,\mathbb{E}\int_0^T \|F_v - B\|_H^2\, dt + \mathbb{E}\,\|v(0) - u_0\|_H^2,$$

and the fact that the first integrand is non-negative yields that for almost all $t \in [0,T]$, $\mathbb{P}$-a.s.

$$\varphi(t, v(t)) + \varphi^*(t, -\tilde{v}(t)) + \langle v(t), \tilde{v}(t)\rangle = 0, \text{ hence } -\tilde{v}(t) \in \partial\varphi(v(t)),$$

and the two other identities readily give that $B = F_v$ and $v(0) = u_0$. This implies that $v(t) = u_0 - \int_0^t \partial\varphi(s, v(s))ds + \int_0^t B(s)dW(s)$ is a solution to (4.1).

The self-dual variational calculus allows to apply the above approach in much more generality. The special Lagrangians L_φ, ℓ_{u_0} and M can be replaced by much more general self-dual Lagrangians. In fact, we show that due to the correspondence between the class of self-dual Lagrangians and maximal monotone operators, one can consider general equations of the form

$$\begin{cases} du(t) = -A(t,u(t))dt + B(t)dW(t) \\ u(0) = u_0, \end{cases} \tag{4.2}$$

where $u_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ for the Hilbert space H, $W(t)$ is a real-valued Wiener process on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_t$, and where $B : [0,T] \times \Omega \to H$ is a given Hilbert-space valued progressively measurable process. $A : \Omega \times [0,T] \times V \to 2^{V^*}$ is a time-dependent progressively measurable –possibly set-valued– maximal monotone map, where V is a Banach space such that $V \subset H \subset V^*$ forms a Gelfand evolution triple.

In Section 4.2, we show how one can lift self-dual Lagrangians from state space to L^p-spaces and then to Itô spaces of stochastic processes. In Section 4.3, we give a variational resolution for Equation (4.2) by using the basic minimization principle (Theorem 3.3) for self-dual Lagrangians. Section 4.4 contains applications to classical SPDEs such as the following stochastic evolution driven by a diffusion and a transport operator,

$$\begin{cases} du = (\Delta u + \mathbf{a}(x) \cdot \nabla u)dt + B(t)dW & \text{on } [0,T] \times D \\ u(0) = u_0 & \text{on } D, \end{cases} \tag{4.3}$$

where $\mathbf{a} : D \to \mathbb{R}^n$ is a smooth vector field with compact support in D, such that $\mathrm{div}(\mathbf{a}) \geq 0$. Other examples include the stochastic porous media equation, but also quasi-linear equations involving the p-Laplacian ($2 \leq p < +\infty$), that is

$$\begin{cases} du = (\Delta_p u - u|u|^{p-2})dt + B(t)dW & \text{on } D \times [0,T] \\ u(0) = u_0 & \text{on } \partial D. \end{cases} \tag{4.4}$$

4.2 Lifting random self-dual Lagrangians to Itô path spaces

Let V be a reflexive Banach space, and $T \in [0, \infty)$ be fixed. Consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a normal filtration $\mathcal{F}_t,\ t \in [0,T]$, and let $L^\alpha(\Omega \times [0,T]; V)$ be the space of Bochner integrable functions from $\Omega_T := \Omega \times [0,T]$ into V with the norm $\|u\|^\alpha_{L^\alpha_V} := \mathbb{E}\int_0^T \|u(t)\|^\alpha_V \, dt$. We may use the shorter notation $L^\alpha_V(\Omega_T) := L^\alpha(\Omega \times [0,T]; V)$ in the sequel.

Definition 4.1. *A self-dual Ω_T-dependent convex Lagrangian on $V \times V^*$ is a function $L : \Omega_T \times V \times V^* \to \mathbb{R} \cup \{+\infty\}$ such that:*

1. *L is progressively measurable with respect to the σ-field generated by the products of $\mathcal{F}_t$ and Borel sets in $[0,t]$ and $V \times V^*$, i.e. for every $t \in [0,T]$, $L(t,\cdot,\cdot)$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t]) \otimes \mathcal{B}(V) \otimes \mathcal{B}(V^*)$-measurable.*
2. *For each $t \in [0,T]$, $\mathbb{P}$-a.s. the function $L(t,\cdot,\cdot)$ is convex and lower semi-continuous on $V \times V^*$.*
3. *For any $t \in [0,T]$, we have $\mathbb{P}$-a.s. $L^*(t,p,u) = L(t,u,p)$ for all $(u,p) \in V \times V^*$, where L^* is the Legendre transform of L in the last two variables.*

To each Ω_T-dependent Lagrangian L on $\Omega_T \times V \times V^*$, one can associate the corresponding Lagrangian $\mathcal{L}$ on the path space $L^\alpha_V(\Omega_T) \times L^\beta_{V^*}(\Omega_T)$, where $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, to be

$$\mathcal{L}(u,p) := \mathbb{E} \int_0^T L(t,u(t),p(t))\ dt,$$

and with the duality between $L^\alpha_V(\Omega_T)$ and $L^\beta_{V^*}(\Omega_T)$ given by

$$\langle u,p \rangle = \mathbb{E} \int_0^T \langle u(t),p(t) \rangle_{V,V^*} dt.$$

The associated Hamiltonian on $L^\alpha_V(\Omega_T) \times L^\alpha_V(\Omega_T)$ will then be

$$H_{\mathcal{L}}(u,v) = \sup \Big\{ \mathbb{E} \int_0^T \{ \langle v(t),p(t) \rangle - L(t,u(t),p(t)) \} dt\ ;\ p \in L^\beta_{V^*}(\Omega_T) \Big\}.$$

The Legendre dual of a "lifted" Lagrangian in both variables naturally lifts to the space of paths $L^\alpha_V(\Omega_T) \times L^\beta_{V^*}(\Omega_T)$ via

$$\mathcal{L}^*(q,v) = \sup_{\substack{u \in L^\alpha_V(\Omega_T) \\ p \in L^\beta_{V^*}(\Omega_T)}} \Big\{ \mathbb{E} \int_0^T \{ \langle q(t),u(t) \rangle + \langle v(t),p(t) \rangle - L(t,u(t),p(t)) \}\ dt \Big\}.$$

The following proposition is standard. See for example [20].

Proposition 4.1. *Suppose that L is an Ω_T-dependent Lagrangian on $V \times V^*$, and $\mathcal{L}$ is the corresponding Lagrangian on the path space $L^\alpha_V(\Omega_T) \times L^\beta_{V^*}(\Omega_T)$. Then,*

1. *$\mathcal{L}^*(p,u) = \mathbb{E} \int_0^T L^*(t,p(t),u(t)) dt$.*
2. *$H_{\mathcal{L}}(u,v) = \mathbb{E} \int_0^T H_L(t,u(t),v(t)) dt$.*

4.2.1 Self-dual Lagrangians associated to progressively measurable monotone fields

Consider now a progressively measurable –possibly set-valued– maximal monotone map that is a map $A:\Omega_T\times V\to 2^{V^*}$ that is measurable for each t, with respect to the product σ-field $\mathcal{F}_t\otimes\mathcal{B}([0,t])\otimes\mathcal{B}(V)$, and such that for each $t\in[0,T]$, $\mathbb{P}$-a.s., the vector field $A_{\omega,t}:=A(t,\omega,\cdot)$ is maximal monotone on V. By Theorem 3.2, one can associate to the maximal monotone maps $A_{\omega,t}$, self-dual Lagrangians $L_{A_{\omega,t}}$ on $V\times V^*$, in such a way that

$$A_{\omega,t}=\bar{\partial}L_{A_{\omega,t}}\quad\text{for every } t\in[0,T]\text{, and } \mathbb{P}\text{-a.s.}$$

This correspondence can be done measurably in such a way that if A is progressively measurable, then the same holds for the corresponding Ω_T-dependent Lagrangian L. We can then lift the random Lagrangian to the space $L^\alpha_V(\Omega_T)\times L^\beta_{V^*}(\Omega_T)$ via

$$\mathcal{L}_A(u,p)=\mathbb{E}\int_0^T L_{A_{\omega,t}}(u(\omega,t),p(\omega,t))dt.$$

Boundedness and coercivity conditions on A translate into corresponding conditions on the representing Lagrangians as follows. For simplicity, we shall assume throughout that the monotone operators are single-valued, though the results apply for general vector fields.

Lemma 4.1. ([32]) Let $A_{\omega,t}$ be the maximal monotone operator as above with the corresponding potential Lagrangian $L_{A_{\omega,t}}$. Assume that for all $u\in V, dt\otimes\mathbb{P}\,a.s.$, $A_{\omega,t}$ satisfies

$$\langle A_{\omega,t}u,u\rangle\geq\max\left\{c_1(\omega,t)\|u\|^\alpha_V-m_1(\omega,t),\;c_2(\omega,t)\|A_{\omega,t}u\|^\beta_{V^*}-m_2(\omega,t)\right\},\tag{4.5}$$

where $c_1,c_2\in L^\infty(\Omega_T,dt\otimes\mathbb{P})$ and $m_1,m_2\in L^1(\Omega_T,dt\otimes\mathbb{P})$. Then the corresponding Lagrangians satisfy the following:

$$C_1(\omega,t)(\|u\|^\alpha_V+\|p\|^\beta_{V^*}-n_1(\omega,t)\leq L_{A_{\omega,t}}(u,p)\leq C_2(\omega,t)(\|u\|^\alpha_V+\|p\|^\beta_{V^*}+n_2(\omega,t)),$$

for some $C_1,C_2\in L^\infty(\Omega_T)$ and $n_1,n_2\in L^1(\Omega_T)$.
The lifted Lagrangian on the L^α-spaces then satisfy for some $C_1,C_2>0$,

$$C_1(\|u\|^\alpha_{L^\alpha_V(\Omega_T)}+\|p\|^\beta_{L^\beta_{V^*}(\Omega_T)}-1)\leq\mathcal{L}_A(u,p)\leq C_2(1+\|u\|^\alpha_{\mathcal{L}^\alpha_V(\Omega_T)}+\|p\|^\beta_{L^\beta_{V^*}(\Omega_T)}).$$

4.2.2 Itô path spaces over a Hilbert space

In view of the definition of a Hilbert-valued Wiener process given in Section 2.5, we suppose that $U = \mathbb{R}$ and hence W is a real-valued Brownian motion.

We now recall Itô's formula.

Proposition 4.2. ([46], [48]) *Let H be a Hilbert space with $\langle\, ,\, \rangle_H$ as its scalar product. Fix $x_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, and let $y \in L^2(\Omega_T; H)$, $Z \in L^2(\Omega_T; H)$ be two progressively measurable processes. Define the H-valued process u as*

$$u(t) := x_0 + \int_0^t y(s)ds + \int_0^t Z(s)dW(s). \tag{4.6}$$

Then, the following holds:

1. *u is a continuous H-valued adapted process such that $\mathbb{E}\left(\sup_{t\in[0,T]} \|u(t)\|_H^2\right) < \infty$.*
2. (Itô's formula) *For all $t \in [0, T]$,*

$$\|u(t)\|_H^2 = \|x_0\|_H^2 + 2\int_0^t \langle y(s), u(s)\rangle_H ds + \int_0^t \|Z(s)\|_H^2 ds + 2\int_0^t \langle u(s), Z(s)\rangle_H dW(s),$$

and consequently

$$\mathbb{E}(\|u(t)\|_H^2) = \mathbb{E}(\|x_0\|_H^2) + \mathbb{E}\int_0^t \Big(2\langle y(s), u(s)\rangle_H + \|Z(s)\|_H^2\Big) ds.$$

More generally, the following *integration by parts* formula holds. For two processes u and v of the form:

$$u(t) = u(0)+\int_0^t \tilde{u}(s)ds+\int_0^t F_u(s)dW(s),\ v(t) = v(0)+\int_0^t \tilde{v}(s)ds+\int_0^t G_v(s)dW(s),$$

we have

$$\mathbb{E}\int_0^T \langle u(t), \tilde{v}(t)\rangle dt = -\,\mathbb{E}\int_0^T \langle v(t), \tilde{u}(t)\rangle dt - \mathbb{E}\int_0^T \langle F_u(t), G_v(t)\rangle dt + \mathbb{E}\langle u(T), v(T)\rangle_H - \mathbb{E}\langle u(0), v(0)\rangle_H. \tag{4.7}$$

Now we define the *Itô space* $\mathcal{A}^2_H$ consisting of all H-valued processes of the following form:

$$\mathcal{A}^2_H = \Big\{u : \Omega_T \to H;\ u(t) = u(0) + \int_0^t \tilde{u}(s)ds + \int_0^t F_u(s)dW(s),$$
$$\text{for } u(0) \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H),\ \tilde{u} \in L^2(\Omega_T; H),\ F_u \in L^2(\Omega_T; H)\Big\}, \tag{4.8}$$

where $\tilde{u}$ and F_u are both progressively measurable. We equip $\mathcal{A}^2_H$ with the norm

$$\|u\|^2_{\mathcal{A}^2_H} = \mathbb{E}\left(\|u(0)\|^2_H + \int_0^T \|\tilde{u}(t)\|^2_H\, dt + \int_0^T \|F_u(t)\|^2_H\, dt\right),$$

so that it becomes a Hilbert space. Indeed, the following correspondence

$$(x_0, y, Z) \in L^2(\Omega; H) \times L^2(\Omega_T; H) \times L^2(\Omega_T; H)$$
$$\mapsto x_0 + \int_0^t y(s)ds + \int_0^t Z(s)dW(s) \in \mathcal{A}^2_H, \tag{4.9}$$
$$u \in \mathcal{A}^2_H \mapsto (u(0), \tilde{u}, F_u) \in L^2(\Omega; H) \times L^2(\Omega_T; H) \times L^2(\Omega_T; H),$$

induces an isometry, since for two processes $u, v \in \mathcal{A}^2_H$, Itô's formula applied to $u - v \in \mathcal{A}^2_H$ yields that

$$\begin{aligned}\|u(t) - v(t)\|^2_H = \|u(0) - v(0)\|^2_H &+ 2\int_0^t \langle \tilde{u}(s) - \tilde{v}(s), u(s) - v(s)\rangle_H ds \\ &+ \int_0^t \|F_u(s) - F_v(s)\|^2_H ds \\ &+ 2\int_0^t \langle u(s) - v(s), F_u(s) - F_v(s)\rangle_H dW_s,\end{aligned}$$

which means that $u = v$ if and only if $u(0) = v(0)$, $F_u = F_v$ and $\tilde{u} = \tilde{v}$. We therefore can and shall identify the Itô space $\mathcal{A}^2_H$ with the product space $L^2(\Omega; H) \times L^2(\Omega_T; H) \times L^2(\Omega_T; H)$.
The dual space $(\mathcal{A}^2_H)^*$ can also be identified with $L^2(\Omega; H) \times L^2(\Omega_T; H) \times L^2(\Omega_T; H)$. In other words, each $p \in (\mathcal{A}^2_H)^*$ can be represented by the triplet

$$p = (p_0, p_1(t), P(t)) \in L^2(\Omega; H) \times L^2(\Omega_T; H) \times L^2(\Omega_T; H),$$

in such a way that the duality can be written as:

$$\langle u, p\rangle_{\mathcal{A}^2_H \times (\mathcal{A}^2_H)^*} = \mathbb{E}\Big\{\langle p_0, u(0)\rangle + \int_0^T \langle p_1(t), \tilde{u}(t)\rangle dt + \frac{1}{2}\int_0^T \langle P(t), F_u(t)\rangle dt\Big\}. \tag{4.10}$$

4.2.3 Self-dual Lagrangians on Itô spaces of random processes

We now prove the following.

Theorem 4.1. *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a complete probability space with normal filtration, and let L and M be two Ω_T-dependent self-dual Lagrangians on $H \times H$, Assume ℓ is an Ω-dependent function on $H \times H$, such that $\mathbb{P}$-a.s.*

$$\ell(\omega, a, b) = \ell^*(\omega, -a, b), \qquad (a, b) \in H \times H. \tag{4.11}$$

The Lagrangian on $\mathcal{A}_H^2 \times (\mathcal{A}_H^2)^*$ defined by

$$\begin{aligned} \mathcal{L}(u,p) = \mathbb{E}\Big\{ \int_0^T L(u(t) - p_1(t), -\tilde{u}(t))\, dt + \ell(u(0) - p_0, u(T)) \\ + \frac{1}{2}\int_0^T M(F_u(t) - P(t), -F_u(t))\, dt \Big\}, \end{aligned} \tag{4.12}$$

is then partially self-dual. Actually, it is self-dual on the subset $\mathcal{A}_H^2 \times \mathcal{D}$ of $\mathcal{A}_H^2 \times (\mathcal{A}_H^2)^*$, where $\mathcal{D} := (\{0\} \times L_H^2(\Omega_T) \times L_H^2(\Omega_T))$.

Proof. Take $(q, v) \in (\mathcal{A}_H^2)^* \times \mathcal{A}_H^2$ with q an element in the dual space identified with the triple $(0, q_1(t), Q(t))$, then

$$\begin{aligned} \mathcal{L}^*(q,v) = & \sup_{\substack{u \in \mathcal{A}_H^2 \\ p \in (\mathcal{A}_H^2)^*}} \{\langle q, u\rangle + \langle v, p\rangle - \mathcal{L}(u,p)\} \\ = & \sup_{u \in \mathcal{A}_H^2} \sup_{\substack{p_0 \in L_H^2(\Omega) \\ p_1 \in L_H^2(\Omega_T), \\ P \in L_H^2(\Omega_T)}} \mathbb{E}\Big\{ \langle p_0, v(0)\rangle + \int_0^T \Big(\langle q_1(t), \tilde{u}(t)\rangle + \langle p_1(t), \tilde{v}(t)\rangle\Big) dt \\ & + \frac{1}{2}\int_0^T \Big(\langle Q(t), F_u(t)\rangle + \langle P(t), G_v(t)\rangle\Big)\, dt \\ & - \int_0^T L(u(t) - p_1(t), -\tilde{u}(t))\, dt - \ell(u(0) - p_0, u(T)) \\ & - \frac{1}{2}\int_0^T M(F_u(t) - P(t), -F_u(t))\, dt \Big\}. \end{aligned}$$

Make the following substitutions:

$$\begin{aligned} u(t) - p_1(t) &= y(t) \in L_H^2(\Omega_T) \\ u(0) - p_0 &= a \in L_H^2(\Omega) \\ F_u(t) - P(t) &= J(t) \in L_H^2(\Omega_T), \end{aligned}$$

to obtain

$$\begin{aligned}\mathcal{L}^*(q,v) = \sup_{u\in\mathcal{A}_H^2}\sup_{a\in L_H^2(\Omega)}\sup_{y\in L_H^2(\Omega_T)}\sup_{J\in L_H^2(\Omega_T)} \mathbb{E}\Big\{&\langle u(0)-a, v(0)\rangle - \ell(a,u(T))\\ &+\int_0^T\Big(\langle q_1(t),\tilde{u}(t)\rangle + \langle u(t)-y(t),\tilde{v}(t)\rangle - L(y(t),-\tilde{u}(t))\Big)\,dt\\ &+\frac{1}{2}\int_0^T\langle Q(t),F_u(t)\rangle + \langle F_u(t)-J(t),G_v(t)\rangle - M(J(t),-F_u(t))\,dt\Big\}.\end{aligned}$$

Use Itô's formula (4.7) for the processes u and v in $\mathcal{A}_H^2$ to get

$$\begin{aligned}\mathcal{L}^*(q,v) = \sup_{u\in\mathcal{A}_H^2}\sup_{\substack{(a,y,J)\in\\ L_H^2(\Omega)\times L_H^2(\Omega_T)\times L_H^2(\Omega_T)}} \mathbb{E}\Big\{&\langle a,-v(0)\rangle + \langle u(T),v(T)\rangle - \ell(a,u(T))\\ &+\int_0^T\langle v(t)-q_1(t),-\tilde{u}(t)\rangle + \langle y(t),-\tilde{v}(t)\rangle - L(y(t),-\tilde{u}(t))\,dt\\ &+\frac{1}{2}\int_0^T\langle G_v(t)-Q(t),-F_u(t)\rangle + \langle J(t),-G_v(t)\rangle - M(J(t),-F_u(t))\,dt\Big\}.\end{aligned}$$

In view of the correspondence

$$\begin{aligned}&(b,r,Z)\in L^2(\Omega;H)\times L^2(\Omega_T;H)\times L^2(\Omega_T;H)\\ &\qquad\mapsto b+\int_0^t r(s)ds+\int_0^t Z(s)dW(s)\in\mathcal{A}_H^2.\end{aligned}$$

$$u\in\mathcal{A}_H^2\mapsto(u(T),-\tilde{u},-F_u)\in L^2(\Omega;H)\times L^2(\Omega_T;H)\times L^2(\Omega_T;H),$$

it follows that

$$\begin{aligned}\mathcal{L}^*(q,v) &= \sup_{(a,b)\in L_H^2(\Omega)\times L_H^2(\Omega)}\mathbb{E}\Big\{\langle a,-v(0)\rangle + \langle b,v(T)\rangle - \ell(a,b)\Big\}\\ &+\sup_{\substack{y\in L_H^2(\Omega_T)\\ r\in L_H^2(\Omega_T)}}\mathbb{E}\Big\{\int_0^T\langle v(t)-q_1(t),r(t)\rangle + \langle y(t),-\tilde{v}(t)\rangle - L(y(t),r(t))\,dt\Big\}\\ &+\frac{1}{2}\sup_{\substack{J\in L_H^2(\Omega_T)\\ Z\in L_H^2(\Omega_T)}}\mathbb{E}\Big\{\int_0^T\langle G_v(t)-Q(t),Z(t)\rangle + \langle J(t),-G_v(t)\rangle - M(J(t),Z(t))dt\Big\},\end{aligned}$$

and therefore taking into account Proposition 4.1 gives

$$\begin{aligned}\mathcal{L}^*(q,v) = \mathbb{E}\,\ell^*(-v(0),v(T)) + \mathbb{E}\int_0^T L^*(-\tilde{v}(t),v(t)-q_1(t))\,dt\\ +\frac{1}{2}\,\mathbb{E}\int_0^T M^*(-G_v(t),G_v(t)-Q(t))\,dt.\end{aligned}$$

Now with the self-duality assumptions on L and M, and the condition on ℓ, we have $\mathcal{L}^*(0,v) = \mathcal{L}(v,0)$, for every $v \in \mathcal{A}_H^2$, which means that $\mathcal{L}$ is partially self-dual on $\mathcal{A}_H^2 \times (\mathcal{A}_H^2)^*$. □

4.3 Variational resolution of stochastic equations driven by additive noise

For simplicity, we shall work in an L^2-setting in w and in time.

4.3.1 A variational principle on Itô space

The following is now a direct consequence of Theorem 4.1 and Theorem 3.3.

Proposition 4.3. *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a complete probability space with normal filtration and let H be a Hilbert space. Suppose L and M are Ω_T-dependent self-dual Lagrangians on $H \times H$, and ℓ is an Ω-dependent time-boundary Lagrangian on $H \times H$. Assume that for some positive C_1, C_2 and C_3, we have*

$$\begin{aligned} \mathbb{E}\int_0^T L(t,v(t),0)\,dt &\le C_1(1+\|v\|^2_{L^2_H(\Omega_T)}) \qquad \text{for } v \in L^2_H(\Omega_T),\\ \mathbb{E}\,\ell(a,0) &\le C_2(1+\|a\|^2_{L^2_H(\Omega)}) \qquad \text{for } a \in L^2_H(\Omega),\\ \mathbb{E}\int_0^T M(\sigma(t),0)\,dt &\le C_3(1+\|\sigma\|^2_{L^2_H(\Omega_T)}) \qquad \text{for } \sigma \in L^2_H(\Omega_T). \end{aligned} \tag{4.13}$$

Consider on $\mathcal{A}_H^2$ the functional

$$I(u) = \mathbb{E}\Big\{\int_0^T \Big(L(u(t), -\tilde{u}(t)) + \frac{1}{2}\,M(F_u(t), -F_u(t))\Big)\,dt + \ell(u(0), u(T))\Big\}.$$

Then, there exists $v \in \mathcal{A}_H^2$ such that $I(v) = \inf_{u \in \mathcal{A}_H^2} I(u) = 0$, and consequently, $\mathbb{P}$-a.s. and for almost all $t \in [0,T]$, we have

$$-\,\tilde{v}(t) \in \bar{\partial} L(t, v(t)) \tag{4.14}$$

$$(-v(0), v(T)) \in \partial\ell(v(0), v(T))$$

$$-F_v(t) \in \bar{\partial} M(F_v(t)).$$

Moreover, if L is strictly convex, then v is unique.

Proof. The functional I can be written as $I(u) = \mathcal{L}(u,0)$, where $\mathcal{L}$ is the partially self-dual Lagrangian defined by (4.12).
In order to apply Theorem 3.3, we need to verify the coercivity condition. To this end, we use conditions (4.13) to show that the map $p \to \mathcal{L}(0,p)$ is bounded on the bounded sets of $(\mathcal{A}^2_H)^*$. Indeed,

$$\begin{aligned}\mathcal{L}(0,p) &= \mathbb{E}\Big\{\int_0^T L(t,-p_1(t),0)\,dt + \ell(-p_0,0) + \frac{1}{2}\int_0^T M(-P(t),0)\,dt\Big\}\\ &\leq C\Big(3 + \|p_1\|^2_{L^2_H(\Omega)} + \|p_0\|^2_{L^2_H(\Omega_T)} + \|P\|^2_{L^2_H(\Omega_T)}\Big),\end{aligned}$$

and by duality, $\lim_{\|u\|\to\infty} \frac{\mathcal{L}(u,0)}{\|u\|} = +\infty$. By Theorem 3.3, there exists $v \in \mathcal{A}^2_H$ such that $I(v) = 0$. We now rewrite I as follows:

$$\begin{aligned}0 = I(v) = \mathbb{E}\,\Big\{\int_0^T L(t,v(t),-\tilde{v}(t)) + \langle v(t),\tilde{v}(t)\rangle\,dt - \int_0^T \langle v(t),\tilde{v}(t)\rangle\,dt\\ + \ell(v(0),v(T)) + \frac{1}{2}\int_0^T M(F_v(t),-F_v(t))\,dt\Big\}.\end{aligned}$$

By Itô's formula

$$\mathbb{E}\int_0^T \langle v(t),\tilde{v}(t)\rangle = \frac{1}{2}\,\mathbb{E}\|v(T)\|^2_H - \frac{1}{2}\,\mathbb{E}\|v(0)\|^2_H - \frac{1}{2}\,\mathbb{E}\int_0^T \|F_v(t)\|^2_H\,dt,$$

which yields

$$\begin{aligned}0 \;=\; I(v) = \mathbb{E}\,\Big\{\;\int_0^T \Big(L(t,v(t),-\tilde{v}(t)) + \langle v(t),\tilde{v}(t)\rangle\Big)\,dt\Big\}\\ + \mathbb{E}\Big\{\ell(v(0),v(T)) - \frac{1}{2}\|v(T)\|^2_H + \frac{1}{2}\|v(0)\|^2_H\Big\}\\ + \frac{1}{2}\,\mathbb{E}\Big\{\int_0^T \Big(\|F_v\|^2_H + M(F_v(t),-F_v(t))\Big)\,dt\Big\}.\end{aligned}$$

The self-duality of the Lagrangians L and M and the hypothesis on the boundary Lagrangian, yield that for a.e. $t \in [0,T]$ and $\mathbb{P}$-a.s. each of the integrands inside the curly-brackets are non-negative, thus

$$\begin{aligned}L(t,v(t),-\tilde{v}(t)) + \langle v(t),\tilde{v}(t)\rangle = 0,\\ \ell(v(0),v(T)) - \frac{1}{2}\|v(T)\|^2_H + \frac{1}{2}\|v(0)\|^2_H = 0,\\ M(F_v(t),-F_v(t)) + \langle F_v,F_v\rangle = 0,\end{aligned}$$

which translate into the three assertions in (4.14).
Finally, if L is strictly convex, then the functional I is strictly convex and the minimum is attained uniquely. □

4.3.2 Regularization via inf-involution

The boundedness condition (4.13) is quite restrictive and not satisfied by most Lagrangians of interest. One way to deal with such a difficulty is to assume similar bounds on L but in stronger Banach norms. Moreover, we need to find more regular solutions that are valued in more suitable Banach spaces than H. To this end, we consider an evolution triple $V \subset H \subset V^*$, where V is a reflexive Banach space and V^* is its dual. We recall the following easy lemma from [29].

Lemma 4.2. *Let L be a self-dual Lagrangian on $V \times V^*$.*

1. *If for some $r > 1$ and $C > 0$, we have $L(u,0) \le C(1 + \|u\|_V^r)$ for all $u \in V$, then there exists $D > 0$ such that $L(u,p) \ge D(\|p\|_{V^*}^s - 1)$ for all $(u,p) \in V \times V^*$, where $\frac{1}{r} + \frac{1}{s} = 1$.*

2. *If for $C_1, C_2 > 0$ and $r_1 \ge r_2 > 1$, we have*

$$C_2(\|u\|_V^{r_2} - 1) \le L(u,0) \le C_1(1 + \|u\|_V^{r_1}) \qquad \text{for all } u \in V,$$

then, there exists $D_1, D_2 > 0$ such that

$$D_2(\|p\|_{V^*}^{s_1} + \|u\|_V^{r_2} - 1) \le L(u,p) \le D_1(1 + \|u\|_V^{r_1} + \|p\|_{V^*}^{s_2}). \quad (4.15)$$

where $\frac{1}{r_i} + \frac{1}{s_i} = 1$ for $i = 1, 2$, and therefore L is continuous in both variables.

Proposition 4.4. *Consider a Gelfand triple $V \subset H \subset V^*$ and let L be an Ω_T-dependent self-dual Lagrangian on $V \times V^*$. Let M be an Ω_T-dependent self-dual Lagrangian on $H \times H$, and ℓ an Ω-dependent boundary Lagrangian on $H \times H$ satisfying $\ell^*(a,b) = \ell(-a,b)$. Assume the following conditions hold:*

(A_1) *For some $m, n > 1$, $C_1, C_2 > 0$, and for all $v \in L_V^2(\Omega_T)$*

$$C_2(\|v\|_{L_V^2(\Omega_T)}^m - 1) \le \mathbb{E}\int_0^T L(t, v(t), 0)\, dt \le C_1(1 + \|v\|_{L_V^2(\Omega_T)}^n).$$

(A_2) *For some $C_3 > 0$,*

$$\mathbb{E}\, \ell(a,b) \le C_3(1 + \|a\|_{L_H^2(\Omega)}^2 + \|b\|_{L_H^2(\Omega)}^2) \quad \text{for all } a, b \in L^2(\Omega; H).$$

(A_3) *For some* $C_4 > 0$*, and for all* $G_1, G_2 \in L^2_H(\Omega_T)$

$$\mathbb{E}\int_0^T M(G_1(t), G_2(t))dt \le C_4(1 + \|G_1\|^2_{L^2_H(\Omega_T)} + \|G_2\|^2_{L^2_H(\Omega_T)}).$$

Then, there exists $v \in \mathcal{A}^2_H$ *with trajectories in* $L^2(\Omega_T; V)$ *such that* $\tilde{v} \in L^2(\Omega_T; V^*)$*, at which the minimum of the following functional is attained and is equal to* 0*.*

$$I(u) = \mathbb{E}\Big\{\int_0^T \Big(L(u(t), -\tilde{u}(t)) + \frac{1}{2}\, M(F_u(t), -F_u(t))\Big)\, dt + \ell(u(0), u(T))\Big\}.$$

Consequently, $\mathbb{P}$*-a.s. and for almost all* $t \in [0, T]$*, we have*

$$-\tilde{v}(t) \in \bar{\partial} L(t, v(t)) \tag{4.16}$$

$$(-v(0), v(T)) \in \partial\ell(v(0), v(T))$$

$$-F_v(t) \in \bar{\partial} M(F_v(t)).$$

Proof. First, apply Lemma 3.4 to lift L to an Ω_T-dependent self-dual Lagrangian on $H \times H$, then consider for $t \in [0, T]$ and $\mathbb{P}$-a.s., the λ-regularization of L, that is

$$L_\lambda(u, p) = \inf_{z \in H}\Big\{L(z, p) + \frac{\|u - z\|^2_H}{2\lambda} + \frac{\lambda}{2}\|p\|^2_H\Big\}.$$

By Lemma 3.1, L_λ is also an Ω_T-dependent self-dual Lagrangian on $H \times H$ in such a way that conditions (4.13) of Proposition 4.3 hold. Hence, there exists $v_\lambda \in \mathcal{A}^2_H$ such that

$$0 = \mathbb{E}\,\Big\{\int_0^T L_\lambda(v_\lambda(t), -\tilde{v}_\lambda(t))\, dt + \ell(v_\lambda(0), v_\lambda(T)) + \frac{1}{2}\int_0^T M(F_{v_\lambda}(t), -F_{v_\lambda}(t))\, dt\Big\}.$$

Since L is convex and lower semi-continuous, then $dt \otimes \mathbb{P}$ a.s, there exists $J_\lambda(v_\lambda) \in H$ so that

$$L_\lambda(v_\lambda(t), -\tilde{v}_\lambda(t)) = L(J_\lambda(v_\lambda)(t), -\tilde{v}_\lambda(t)) + \frac{\|v_\lambda(t) - J_\lambda(v_\lambda)(t)\|^2_H}{2\lambda} + \frac{\lambda}{2}\|\tilde{v}_\lambda(t)\|^2_H,$$

and hence

$$0 = \mathbb{E}\Big\{\int_0^T \Big(L(J_\lambda(v_\lambda)(t), -\tilde{v}_\lambda(t)) + \frac{\|v_\lambda(t) - J_\lambda(v_\lambda)(t)\|_H^2}{2\lambda} + \frac{\lambda}{2}\|\tilde{v}_\lambda(t)\|_H^2\Big)\,dt$$
$$+ \ell(v_\lambda(0), v_\lambda(T)) + \frac{1}{2}\int_0^T M(F_{v_\lambda}(t), -F_{v_\lambda}(t))\,dt\Big\}. \tag{4.17}$$

From (4.17), condition (A_1) and the assertion of part (2) of Lemma 4.2, we can deduce that $J_\lambda(v_\lambda)$ is bounded in $L^2(\Omega_T; V)$ and $\tilde{v}_\lambda$ is bounded in $L^2(\Omega_T; V^*)$. Also from conditions (A_2) and (A_3), we can deduce the following estimates:

$$\mathbb{E}\int_0^T M(G, H)\,dt \geq C(\|G\|^2_{L^2_H(\Omega_T)} - 1) \quad \text{and} \quad \mathbb{E}\,\ell(a, b) \geq C(\|b\|^2_{L^2_H(\Omega)} - 1).$$

These coercivity properties, together with (4.17), imply that $v_\lambda(0)$ and $v_\lambda(T)$ are bounded in $L^2(\Omega; H)$, and that F_{v_λ} is bounded in $L^2(\Omega_T; H)$. Moreover, since all other terms in (4.17) are bounded below, it follows that

$$\mathbb{E}\int_0^T \|v_\lambda(t) - J_\lambda(v_\lambda)(t)\|^2 dt \leq 2\lambda C \quad \text{for some } C > 0.$$

Hence v_λ is bounded in $\mathcal{A}^2_H$ and there exists a subsequence v_{λ_j} that converges weakly to a path $v \in L^2(\Omega_T; V)$ such that $\tilde{v} \in L^2(\Omega_T; V^*)$, and

$$J_{\lambda_j}(v_{\lambda_j}) \rightharpoonup v \quad \text{in} \quad L^2(\Omega_T; V)$$
$$\tilde{v}_{\lambda_j} \rightharpoonup \tilde{v} \quad \text{in} \quad L^2(\Omega_T; V^*)$$
$$v_{\lambda_j} \rightharpoonup v \quad \text{in} \quad L^2(\Omega_T; H)$$
$$v_{\lambda_j}(0) \rightharpoonup v(0), \quad v_\lambda(T) \rightharpoonup v(T) \quad \text{in} \quad L^2(\Omega; H)$$
$$F_{v_{\lambda_j}} \rightharpoonup F_v \quad \text{in} \quad L^2(\Omega_T; H).$$

Since L, ℓ and M are lower semi-continuous, we have

$$I(v) \leq \liminf_j \mathbb{E}\Big\{\int_0^T \Big(L(J_{\lambda_j}(v_{\lambda_j})(t), -\tilde{v}_{\lambda_j}(t)) + \frac{\|v_{\lambda_j}(t) - J_{\lambda_j}(v_{\lambda_j})(t)\|^2}{2\lambda_j}$$
$$+ \frac{\lambda_j}{2}\|\tilde{v}_{\lambda_j}(t)\|^2\Big)\,dt + \ell(v_{\lambda_j}(0), v_{\lambda_j}(T))$$
$$+ \frac{1}{2}\int_0^T M(F_{v_{\lambda_j}}(t), -F_{v_{\lambda_j}}(t))\,dt\Big\} = 0.$$

For the reverse inequality, we use the self-duality of L and M and the fact that $\ell(-a,b) = \ell^*(a,b)$ to deduce that

$$
\begin{aligned}
I(v) = \mathbb{E}\Big\{ &\int_0^T \Big(L(v(t), -\tilde{v}(t)) + \langle v(t), \tilde{v}(t)\rangle\Big)\,dt\Big\} \\
&+ \mathbb{E}\Big\{\ell(v(0), v(T)) - \frac{1}{2}\|v(T)\|_H^2 + \frac{1}{2}\|v(0)\|_H^2\Big\} \\
&+ \frac{1}{2}\,\mathbb{E}\Big\{\int_0^T \Big(\|F_v\|_H^2 + M(F_v(t), -F_v(t))\Big)\,dt\Big\} \geq 0.
\end{aligned}
$$

Therefore, $I(v) = 0$ and the rest of the proof is similar to the last part of the proof in Proposition 4.3. □

We now deduce the following.

Theorem 4.2. *Consider a Gelfand triple $V \subset H \subset V^*$, and let $A : D(A) \subset V \to V^*$ be an Ω_T-dependent progressively measurable maximal monotone operator satisfying*

$$\langle A_{w,t}u, u\rangle \geq \max\{c_1(\omega,t)\|u\|_V^{\alpha} - m_1(\omega,t),\, c_2(\omega,t)\|Au\|_{V^*}^{\beta} - m_2(\omega,t)\},$$

where $c_1, c_2 \in L^\infty(\Omega_T, dt \otimes \mathbb{P})$ and $m_1, m_2 \in L^1(\Omega_T, dt \otimes \mathbb{P})$. Let B be a given H-valued progressively measurable process in $L^2(\Omega_T; H)$, and u_0 a given random variable in $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Then, the equation

$$
\begin{cases}
du(t) = -A(t, u(t))dt + B(t)dW(t) \\
u(0) = u_0,
\end{cases}
\tag{4.18}
$$

has a solution $u \in \mathcal{A}_H^2$ that is valued in V. It can be obtained by minimizing the functional

$$
\begin{aligned}
I(u) = \; &\mathbb{E}\int_0^T L(u(t), -\tilde{u}(t))\,dt \\
&+ \mathbb{E}\,\Big(\frac{1}{2}\|u(0)\|_H^2 + \frac{1}{2}\|u(T)\|_H^2 - 2\langle u_0, u(0)\rangle_H + \|u_0\|_H^2\Big) \\
&+ \mathbb{E}\int_0^T \Big(\frac{1}{2}\,\|F_u(t) - 2B(t)\|_H^2 \,+\, \frac{1}{2}\,\|F_u(t)\|_H^2 - 2\langle F_u(t), B(t)\rangle_H\Big)\,dt,
\end{aligned}
$$

where L is a self-dual Lagrangian such that $\bar{\partial}L(t,\cdot) = A(t,\cdot)$, $\mathbb{P}$-almost surely.

Proof. It suffices to apply Proposition 4.4 with the self-dual Lagrangian L associated with A, the time boundary Ω-dependent Lagrangian ℓ_{u_0} on $H\times H$ given by

$$\ell_{u_0}(a,b)=\frac{1}{2}\|a\|_H^2+\frac{1}{2}\|b\|_H^2-2\langle u_0(w),a\rangle_H+\|u_0(w)\|_H^2,$$

and the Ω_T-dependent self-dual Lagrangian M on $L^2_H(\Omega_T)$, given by

$$M_B(G_1,G_2)=\Psi_{B(w,t)}(G_1)+\Psi^*_{B(w,t)}(G_2),$$

where $\Psi_{B(w,t)}:H\to\mathbb{R}\cup\{+\infty\}$ is the convex function $\Psi_{B(w,t)}(G)=\frac{1}{2}\|G-2B(w,t)\|_H^2$. □

4.4 Applications to various SPDEs with additive noise

In the following examples, we shall assume D is a smooth bounded domain in $\mathbb{R}^n$, W is a real Brownian motion, and $B:\Omega\times[0,T]\to L^2(D)$ is a fixed progressively measurable stochastic process.

4.4.1 Stochastic evolution driven by diffusion and transport

Consider the following stochastic transport equation:

$$\begin{cases} du=(\Delta u+\mathbf{a}(x)\cdot\nabla u)dt+B(t)dW & \text{on } [0,T]\times D\\ u(0)=u_0 & \text{on } D,\end{cases}\tag{4.19}$$

where $\mathbf{a}:D\to\mathbb{R}^n$ is a smooth vector field with compact support in D, such that $\operatorname{div}(\mathbf{a})\geq 0$. Assume $u_0\in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H^1_0(D))$ such that $\mathbb{P}$-a.s. $\Delta u_0\in L^2(D)$.
Consider the operator $\Gamma u=\mathbf{a}\cdot\nabla u+\frac{1}{2}(\operatorname{div}\mathbf{a})\,u$, which, by Green's formula, is skew-adjoint on $H^1_0(D)$. Also consider the convex function

$$\varphi(u)=\begin{cases}\frac{1}{2}\int_D|\nabla u|^2dx+\frac{1}{4}\int_D(\operatorname{div}\mathbf{a})|u|^2dx & u\in H^1_0(D)\\ +\infty & \text{otherwise,}\end{cases}$$

which is clearly coercive on $H^1_0(D)$. Consider the Gelfand triple $H^1_0(D)\subset L^2(D)\subset H^{-1}(D)$, and the self-dual Lagrangian on $H^1_0(D)\times H^{-1}(D)$, defined by

$$L(u,p)=\varphi(u)+\varphi^*(\Gamma u+p).$$

The corresponding functional on Itô space is then,

$$
\begin{aligned}
I(u) = \ & \mathbb{E}\left\{\int_0^T \Big(\varphi(u(t,\cdot)) + \varphi^*\left(-\tilde{u}(t,\cdot) + \Gamma(u(t,\cdot))\right)\Big)dt\right\} \\
& + \mathbb{E}\left\{\frac{1}{2}\int_0^T \left(\int_D \Big(|F_u(t,x)|^2 + 2|B(t,x)|^2 - 4F_u(t,x)B(t,x)\Big)dx\right)\, dt\right\} \\
& + \mathbb{E}\left\{\int_D \Big(\frac{1}{2}|u(0,x)|^2 + \frac{1}{2}|u(T,x)|^2 - 2u_0(x)u(0,x) + \frac{1}{2}|u_0(x)|^2\Big)\, dx\right\}.
\end{aligned}
$$

Apply Theorem 4.2 to find a path $v \in \mathcal{A}^2_{L^2(D)}$, valued in $H^1_0(D)$, that minimizes I in such a way that $I(v) = 0$, to obtain

$$
-\tilde{v} + \mathbf{a}\cdot\nabla v + \frac{1}{2}(\operatorname{div}\mathbf{a})v \in \partial\varphi(v) = -\Delta v + \frac{1}{2}(\operatorname{div}\mathbf{a})v,
$$

$$
v(0) = u_0, \ F_v = B.
$$

The process $v(t) = u_0 + \int_0^t \Delta v(s)ds + \int_0^t \mathbf{a}\cdot\nabla v(s)ds + \int_0^t B(s)dW(s)$ is therefore a solution to (4.19).

4.4.2 Stochastic porous media

Consider the following SPDE,

$$
\begin{cases} du(t) = \Delta u^p(t)dt + B(t)dW(t) & \text{on } D\times[0,T] \\ u(0) = u_0 & \text{on } D, \end{cases} \tag{4.20}
$$

where $p \geq \frac{n-2}{n+2}$, and $u_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H^{-1}(D))$.
Equip the Hilbert space $H = H^{-1}(D)$ with the inner product

$$
\langle u, v\rangle_{H^{-1}} = \langle u, (-\Delta)^{-1}v\rangle = \int_D u(x)(-\Delta)^{-1}v(x)\, dx.
$$

Since $p \geq \frac{n-2}{n+2}$, $L^{p+1}(D) \subset H^{-1}(D) \subset L^{\frac{p+1}{p}}(D)$ is an evolution triple.
We consider the convex functional

$$
\varphi(u) = \begin{cases} \frac{1}{p+1}\int_D |u(x)|^{p+1}dx & \text{on } L^{p+1}(D) \\ +\infty & \text{elsewhere,} \end{cases}
$$

whose Legendre conjugate is given by

$$
\varphi^*(u^*) = \frac{p}{p+1}\int_D |(-\Delta)^{-1}u^*|^{\frac{p+1}{p}}\, dx.
$$

Now, minimize the following self-dual functional on $\mathcal{A}_H^2$,

$$
\begin{aligned}
I(u) = \ & \mathbb{E}\left\{\frac{1}{p+1}\int_0^T \int_D \Big(|u(x)|^{p+1} + p\, \big|(-\Delta)^{-1}(-\tilde{u}(t))\big|^{\frac{p+1}{p}} \Big) dx\, dt\right\} \\
& + \mathbb{E}\ \left\{\frac{1}{2}\|u(0)\|^2_{H^{-1}} + \frac{1}{2}\|u(T)\|^2_{H^{-1}} + \|u_0\|^2_{H^{-1}} - 2\langle u_0, u(0,\cdot)\rangle_{H^{-1}}\right\} \\
& + \mathbb{E}\left\{\int_0^T \frac{1}{2}\Big(\|F_u(t)\|^2_{H^{-1}} + 2\|B(t)\|^2_{H^{-1}} - 4\langle F_u(t), B(t)\rangle_{H^{-1}}\Big) dt\right\}.
\end{aligned}
$$

Apply Theorem 4.2 to find a process $v \in \mathcal{A}_H^2$ with values in $L^{p+1}(D)$ such that

$$(-\Delta)^{-1}(-\tilde{v}(t)) \in \partial\varphi(v(t)) = v^p, F_v = B, \text{ and } v(0) = u_0.$$

It follows that $v(t) = u_0 + \int_0^t \Delta v^p(s)ds + \int_0^t B(s)dW(s)$, provides a solution for (4.20).

4.4.3 Stochastic PDE involving the p-Laplacian

Consider the equation

$$
\begin{cases}
du = (\Delta_p u - u|u|^{p-2})dt + B(t)dW & \text{on } D \times [0,T] \\
u(0) = u_0 & \text{on } \partial D,
\end{cases}
$$

where $p \in [2, +\infty)$, $\Delta_p u = \operatorname{div}(|\nabla u|^{p-2}\nabla u)$ is the p-Laplacian operator, and u_0 is given such that $u_0 \in W_0^{1,p}(D) \cap \{u; \Delta_p u \in L^p(D)\}$. It is clear that $W_0^{1,p}(D) \subset L^p(D)$ continuously and densely, which ensures that the functional

$$\varphi(u) = \frac{1}{p}\int_D |\nabla u(x)|^p dx + \frac{1}{p}\int_D |u(x)|^p dx,$$

is convex, lower semi-coninuous and coercive on $W_0^{1,p}(D)$ with respect to the evolution triple

$$W_0^{1,p}(D) \subset L^p(D) \subset L^2(D) \subset W_0^{1,p}(D)^* \subset L^q(D),$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Theorem 4.2 applies to the self-dual functional

$$
\begin{aligned}
I(u) = \ & \mathbb{E}\int_0^T \Big(\varphi(t,u) + \varphi^*(t, -\tilde{u})\Big)\, dt \\
& + \mathbb{E}\ \left(\frac{1}{2}\|u(0)\|^2_{L^2(D)} + \frac{1}{2}\|u(T)\|^2_{L^2(D)} - 2\langle u_0, u(0)\rangle + \|u_0\|^2_{L^2(D)}\right) \\
& + \mathbb{E}\int_0^T \left(\frac{1}{2}\,\|F_u(t)\|^2_{L^2(D)} + \|B(t)\|^2_{L^2(D)} - 2\langle F_u(t), B(t)\rangle\right) dt.
\end{aligned}
$$

to yield a $W_0^{1,p}(D)$-valued process $v \in \mathcal{A}^2_{L^2(D)}$, where the null infimum is attained. It follows that

$$-\tilde{v} \in \partial\varphi(v) = -\Delta_p v + v|v|^{p-2},$$

$$v(0) = u_0, \ F_v = B,$$

and hence
$v(t) - u_0 - \int_0^t B(s)dW(s) = \int_0^t \tilde{v}(s)ds = \int_0^t \Delta_p v(s)ds - \int_0^t v(s)|v(s)|^{p-2}ds.$

Remark 4.1. *Note that in the examples given in the this section, the self-dual functional I is strictly convex and hence the solution obtained via minimizing I is unique.*

Chapter 5

Self-dual variational principle for stochastic partial differential equations with non-additive noise

5.1 Introduction

In this chapter, we consider SPDEs driven by monotone vector fields and involving a non-additive noise. These can take the form

$$\begin{cases} du(t) = -A(t, u(t))dt + B(t, u(t))dW(t) \\ u(0) = u_0, \end{cases} \tag{5.1}$$

where $u \to B(t, u)$ is now a progressively measurable linear or non-linear operator. In Chapter 4, we studied the additive case of SPDEs driven by monotone vector field, which can be formulated as completely self-dual functionals and in a Hilbertian setting, a solution can be obtained via a basic self-dual variational principle. In view of the evolution triple $V \subset H \subset V^*$, we then applied an inf-convolution argument to find a solution that is valued in the Sobolev space V. This approach , however, does not work in the non-additive case, since we need to work with stronger topologies on the space of Itô processes that will give the operator B a chance to be completely continuous. We shall therefore strengthen the norm on the Itô space over a Gelfand triple, at the cost of losing coercivity, that we shall recover through perturbation methods.

In order to variationally resolve Equation (5.1), we introduce the Itô space $\mathcal{Y}_V^\alpha$ analogous to $\mathcal{A}_H^2$, but equipped with stronger norms (see Section 5.2), and we show that equation

$$\begin{cases} du = -\bar{\partial}\mathcal{L}(u)(t)\, dt + B(t, u(t))dW \\ u(0) = u_0, \end{cases}$$

has a solution in $\mathcal{Y}_V^\alpha$, where $\mathcal{L}$ is a self-dual Lagrangian on $L^\alpha(\Omega_T; V) \times L^\beta(\Omega_T; V^*)$. Taking into account the representation of maximal monotone operators with self-dual vector fields (Theorem 3.2) and the argument in Section 4.2.1, one can then obtain a variational resolution for Equation (5.1).

The variational principle we use in this chapter is an application of Theorem 3.5 on $L^\alpha(\Omega_T; V) \times L^\beta(\Omega_T; V^*)$. However, we require to perform a stochastic elliptic regularization to perturb the corresponding self-dual Lagrangian so that the coercivity condition (3.8) is satisfied. Finally, we would let the perturbations go to zero to conclude the existence of a solution to the original equation. In Section 5.3, we give some immediate applications of the result to an SPDE driven by the gradient of a convex function, in particular an SPDE of the form

$$\begin{cases} du(t) = \Delta u \, dt + |u|^{q-1} u \, dW \\ u(0) = u_0. \end{cases}$$

where $\frac{1}{2} \leq q < \frac{n}{n-2}$, and we would consider in turn a general SPDE for which the monotone vector field is in divergence form, namely

$$\begin{cases} du = \operatorname{div}(\beta(\nabla u(t,x)))dt + B(u(t))dW(t) & \text{in } [0,T] \times D \\ u(0,x) = u_0 & \text{on } \partial D, \end{cases}$$

where D is a bounded domain in $\mathbb{R}^n$, and $\beta : \mathbb{R}^n \to \mathbb{R}^n$ is a progressively measurable maximal monotone operator.

5.2 Non-additive noise driven by self-dual Lagrangians

In this section, we give a variational resolution for stochastic equations of the form

$$\begin{cases} du = -\bar{\partial}\mathcal{L}(u)(t) \, dt + B(t, u(t))dW \\ u(0) = u_0, \end{cases} \tag{5.2}$$

where $\mathcal{L}$ is a self-dual Lagrangian on $L^\alpha(\Omega_T; V) \times L^\beta(\Omega_T; V^*)$, $1 < \alpha < +\infty$ and β is its conjugate, and where $V \subset H \subset V^*$ is a given Gelfand triple. We shall assume that $\mathcal{L}$ satisfies the following conditions:

$$\begin{aligned} C_2(\|u\|^\alpha_{L^\alpha_V(\Omega_T)} + \|p\|^\beta_{L^\beta_{V^*}(\Omega_T)} - 1) &\leq \mathcal{L}(u,p) \\ &\leq C_1(1 + \|u\|^\alpha_{L^\alpha_V(\Omega_T)} + \|p\|^\beta_{L^\beta_{V^*}(\Omega_T)}), \end{aligned} \tag{5.3}$$

and

$$\|\bar{\partial}\mathcal{L}(u)\|_{L^{\beta}_{V^*}(\Omega_T)} \leq C_3(1 + \|u\|^{\alpha-1}_{L^{\alpha}_V(\Omega_T)}). \tag{5.4}$$

More precisely, we are searching for a solution u of the form

$$u(t) = u(0) + \int_0^t \tilde{u}(s)ds + \int_0^t F_u(s)dW(s), \tag{5.5}$$

where $u \in L^{\alpha}(\Omega_T; V)$, $\tilde{u} \in L^{\beta}(\Omega_T; V^*)$ and $F_u \in L^2(\Omega_T; H)$ are progressively measurable. The space of such processes will be denoted by $\mathcal{Y}^{\alpha}_V$ and will be equipped with the norm,

$$\|u\|_{\mathcal{Y}^{\alpha}_V} = \|u(t)\|_{L^{\alpha}_V(\Omega_T)} + \|\tilde{u}(t)\|_{L^{\beta}_{V^*}(\Omega_T)} + \|F_u(t)\|_{L^2_H(\Omega_T)}.$$

As shown in [48], any such a process $u \in \mathcal{Y}^{\alpha}_V$ has a $dt \otimes \mathbb{P}$-equivalent version $\hat{u}$ that is a V-valued progressively measurable process that satisfies the following Itô's formula:
$\mathbb{P}$-a.s. and for all $t \in [0, T]$,

$$\|u(t)\|^2_H = \|u(0)\|^2_H + 2\int_0^t \langle \tilde{u}(s), \hat{u}(s)\rangle_{V^*,V} ds + \int_0^t \|F_u(s)\|^2_H ds + 2\int_0^t \langle u(s), F_u(s)\rangle_H dW(s). \tag{5.6}$$

In particular, we have for all $t \in [0, T]$,

$$\mathbb{E}(\|u(t)\|^2_H) = \mathbb{E}(\|u(0)\|^2_H) + \mathbb{E}\int_0^t \Big(2\langle \tilde{u}(s), \hat{u}(s)\rangle_{V^*,V} + \|F_u(s)\|^2_H\Big)ds.$$

Furthermore, we have $u \in C([0, T]; H)$. In fact, one can deduce that for any $u \in \mathcal{Y}^{\alpha}_V$, $u \in C([0, T]; V^*)$ and $u \in L^{\infty}(0, T; H)$ $\mathbb{P}$-a.s ([44] and [48]). From now on, a process u in $\mathcal{Y}^{\alpha}_V$ will always be identified with its $dt \otimes \mathbb{P}$-equivalent V-valued version $\hat{u}$.

Theorem 5.1. *Consider a self-dual Lagrangian $\mathcal{L}$ on $L^{\alpha}(\Omega_T; V) \times L^{\beta}(\Omega_T; V^*)$ satisfying (5.3) and (5.4), and let $B : \mathcal{Y}^{\alpha}_V \to L^2(\Omega_T; H)$ be a –not-necessarily linear– weak-to-norm continuous map such that for some $C > 0$ and $0 < \delta < \frac{\alpha+1}{2}$,*

$$\|Bu\|_{L^2_H(\Omega_T)} \leq C\|u\|^{\delta}_{L^{\alpha}_V(\Omega_T)} \quad \text{for any } u \in \mathcal{Y}^{\alpha}_V. \tag{5.7}$$

Let u_0 be a given random variable in $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Equation (5.2) has then a solution u in $\mathcal{Y}^{\alpha}_V$, that is a stochastic process satisfying

$$u(t) = u_0 - \int_0^t \bar{\partial}\mathcal{L}(u)(s)ds + \int_0^t Bu(s)dW(s). \tag{5.8}$$

We would like to apply Theorem 3.5 to $\mathcal{L}$ on $L^{\alpha}(\Omega_T;V) \times L^{\beta}(\Omega_T;V^*)$ and to the following operators acting on $G = \{u \in \mathcal{Y}_V^{\alpha}; u(0) = u_0\}$,

$$A_1 : G \subset \mathcal{Y}_V^{\alpha} \to L^{\alpha}(\Omega_T;V), \qquad \Gamma_1 : G \subset \mathcal{Y}_V^{\alpha} \to L^{\beta}(\Omega_T;V^*)$$
$$A_1(u) = u, \qquad \Gamma_1(u) = -\tilde{u}$$

$$A_2 : G \subset \mathcal{Y}_V^{\alpha} \to L^{2}(\Omega_T;H), \qquad \Gamma_2 : G \subset \mathcal{Y}_V^{\alpha} \to L^{2}(\Omega_T;H)$$
$$A_2(u) = \frac{1}{2}F_u, \qquad \Gamma_2(u) = -F_u + \frac{3}{2}Bu.$$

Unfortunately, the coercivity condition (3.8) required to conclude is not satisfied. We have to therefore perturb the Lagrangian $\mathcal{L}$ (i.e., essentially perform a stochastic elliptic regularization) as well as the operator Γ_1 in order to ensure coercivity. We will then let the perturbations go to zero to conclude.

5.2.1 Stochastic elliptic regularization

To do that, we consider the convex lower semi-continuous function on $L^{\alpha}(\Omega_T, V)$

$$\psi(u) = \begin{cases} \frac{1}{\beta}\mathbb{E}\int_0^T \|\tilde{u}(t)\|_{V^*}^{\beta}dt & \text{if } u \in \mathcal{Y}_V^{\alpha} \\ +\infty & \text{if } u \in L_V^{\alpha}(\Omega_T)\backslash \mathcal{Y}_V^{\alpha}, \end{cases} \tag{5.9}$$

and for any $\mu > 0$, its associated self-dual Lagrangian on $L_V^{\alpha}(\Omega_T) \times L_{V^*}^{\beta}(\Omega_T)$ given by

$$\Psi_{\mu}(u,p) = \mu\psi(u) + \mu\psi^*(\frac{p}{\mu}). \tag{5.10}$$

We also consider a perturbation operator

$$Ku := (\|u\|_{L_V^{\alpha}(\Omega_T)}^{\alpha-1})Du,$$

where D is the duality map between V and V^*. Note that by definition, K is a weak-to-weak continuous operator from $\mathcal{Y}_V^{\alpha}$ to $L_{V^*}^{\beta}(\Omega_T)$.

Lemma 5.1. *Under the above hypothesis on $\mathcal{L}$ and B, there exists a process $u_\mu \in \mathcal{Y}_V^{\alpha}$ such that $u_\mu(0) = u_0$, $\tilde{u}_\mu(T) = \tilde{u}_\mu(0) = 0$, and satisfying*

$$\tilde{u}_\mu + Ku_\mu + \mu\,\partial\psi(u_\mu) \in -\bar{\partial}\mathcal{L}(u_\mu)$$
$$F_{u_\mu} = Bu_\mu.$$

Proof. Apply Theorem 3.5 as follow: Let $Z = \mathcal{Y}_V^\alpha, X_1 = L^\alpha(\Omega_T;V), X_2 = L^2(\Omega_T;H)$ with $G = \{u \in \mathcal{Y}_V^\alpha; u(0) = u_0\}$ which is a closed linear subspace of $\mathcal{Y}_V^\alpha$, and consider the operators

$$A_1 : G \subset \mathcal{Y}_V^\alpha \to L^\alpha(\Omega_T;V), \qquad \Gamma_1 : G \subset \mathcal{Y}_V^\alpha \to L^\beta(\Omega_T;V^*)$$
$$A_1(u) = u, \qquad \Gamma_1(u) = -\tilde{u} - Ku$$

$$A_2 : G \subset \mathcal{Y}_V^\alpha \to L^2(\Omega_T;H), \qquad \Gamma_2 : G \subset \mathcal{Y}_V^\alpha \to L^2(\Omega_T;H)$$
$$A_2(u) = \frac{1}{2}F_u, \qquad \Gamma_2(u) = -F_u + \frac{3}{2}Bu \quad (5.11)$$

where their domain is G, A_1, A_2 are linear, and Γ_1, Γ_2 are weak-weak continuous.
As to the Lagrangians, we take on $L_V^\alpha(\Omega_T) \times L_{V^*}^\beta(\Omega_T)$, the Lagrangian

$$L_1(u,p) = \mathcal{L} \oplus \Psi_\mu(u,p),$$

while on $L_H^2(\Omega_T) \times L_H^2(\Omega_T)$, we take

$$L_2(P,Q) = \mathbb{E}\int_0^T M(P(t,w),Q(t,w))\,dt,$$

where $M(P,Q) = \frac{1}{2}\|P\|_H^2 + \frac{1}{2}\|Q\|_H^2$.
In other words, we are considering the functional

$$\begin{aligned} I_\mu(u) = \mathcal{L} \oplus \Psi_\mu(A_1u,\Gamma_1u) &- \mathbb{E}\int_0^T \langle A_1u,\Gamma_1u\rangle dt \\ &+ \mathbb{E}\int_0^T M(\Gamma_2u, A_2u) - \langle A_2u,\Gamma_2u\rangle\,dt \\ = \mathcal{L} \oplus \Psi_\mu(u, -\tilde{u} - Ku) &- \mathbb{E}\int_0^T \langle u, -\tilde{u} - Ku\rangle\,dt \\ &+ \mathbb{E}\int_0^T M(F_u/2, -F_u + 3Bu/2) - \langle F_u/2, -F_u + 3Bu/2\rangle\,dt. \end{aligned}$$

We now verify the conditions of Theorem 3.5.
We have that $G_0 = \mathrm{Ker}(A_2) \cap G = \{u \in \mathcal{Y}_V^\alpha;\ u(t) = u_0 + \int_0^t \tilde{u}(s)\,ds\}$, where $\tilde{u}$ is some progressively measurable process in $L_{V^*}^\beta(\Omega_T)$. It is clear that $A_1(G_0)$ is dense in $L^\alpha(\Omega_T;V)$. Moreover, $A_2(G)$ is dense in $L^2(\Omega_T;H)$. To check the upper semi-continuity of

$$u \to \mathbb{E}\int_0^T \langle A_1u,\Gamma_1u\rangle + \langle A_2u,\Gamma_2u\rangle\,dt,$$

on $\mathcal{Y}_V^\alpha$ equipped with the weak topology, we apply Itô's formula to obtain that

$$
\begin{aligned}
\mathbb{E}\int_0^T \langle A_1 u, \Gamma_1 u\rangle + \langle A_2 u, \Gamma_2 u\rangle dt &= \mathbb{E}\int_0^T \langle u, -\tilde{u} - Ku\rangle + \langle \frac{F_u}{2}, -F_u + \frac{3Bu}{2}\rangle dt \\
&= \frac{1}{2}\,\mathbb{E}\,\|u_0\|_H^2 - \frac{1}{2}\,\mathbb{E}\,\|u(T)\|_H^2 - \|u\|^{\alpha+1}_{L_V^\alpha(\Omega_T)} \\
&\quad + \frac{3}{4}\,\mathbb{E}\int_0^T \langle F_u(t), Bu(t)\rangle dt.
\end{aligned}
$$

Upper semi-continuity then follows from the compactness of the maps $\mathcal{Y}_V^\alpha \to L^2(\Omega; H)$ given by $u \mapsto (u(0), u(T))$, as well as the weak to norm continuity of B, which makes the functional $u \mapsto \mathbb{E}\int_0^T \langle F_u, Bu\rangle dt$ weakly continuous.

To verify the coercivity, we first note first that condition (5.3) implies that for some (different) $C_1 > 0$,

$$
H_{\mathcal{L}}(0, u) \geq C_1\Big(\|u\|^\alpha_{L_V^\alpha(\Omega_T)} - 1\Big).
$$

By also taking into account condition (5.7) on B, with the fact that $\delta < \frac{\alpha+1}{2}$, we get that

$$
\begin{aligned}
&H_{\mathcal{L}}(0, u) + \mu\,\psi(u) + \mathbb{E}\int_0^T \langle u, \tilde{u} + Ku\rangle\, dt \\
&\qquad + \mathbb{E}\int_0^T H_M(0, \frac{F_u}{2}) - \langle \frac{F_u}{2}, -F_u + \frac{3Bu}{2}\rangle\, dt \\
&= H_{\mathcal{L}}(0, u) + \frac{\mu}{\beta}\|\tilde{u}\|^\beta_{L^\beta_{V^*}(\Omega_T)} - \frac{1}{2}\,\|u_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\,\|u(T)\|^2_{L^2(\Omega;H)} \\
&\quad + \|u\|^{\alpha+1}_{L_V^\alpha(\Omega_T)} + \frac{1}{8}\|F_u(t)\|^2_{L^2(\Omega_T;H)} - \frac{3}{4}\,\mathbb{E}\int_0^T \langle F_u(t), Bu(t)\rangle\, dt \\
&\geq C_1\Big(\|u\|^\alpha_{L_V^\alpha(\Omega_T)} - 1\Big) + \frac{\mu}{\beta}\,\|\tilde{u}\|^\beta_{L^\beta_{V^*}(\Omega_T)} + \|u\|^{\alpha+1}_{L_V^\alpha(\Omega_T)} \\
&\quad + C_2\Big(\|F_u(t)\|^2_{L^2_H(\Omega_T)} - \|F_u\|_{L^2_H(\Omega_T)}\|Bu\|_{L^2_H(\Omega_T)}\Big) + C \\
&\geq C_1\Big(\|u\|^\alpha_{L_V^\alpha(\Omega_T)} - 1\Big) + \frac{\mu}{\beta}\,\|\tilde{u}\|^\beta_{L^\beta_{V^*}(\Omega_T)} + \|u\|^{\alpha+1}_{L_V^\alpha(\Omega_T)} \\
&\quad + C_2\Big(\|F_u(t)\|^2_{L^2_H(\Omega_T)} - \|F_u\|_{L^2_H(\Omega_T)}\|u\|^\delta_{L^2_H(\Omega_T)}\Big) + C \\
&\geq \frac{\mu}{\beta}\,\|\tilde{u}\|^\beta_{L^\beta_{V^*}(\Omega_T)} + \|u\|^{\alpha+1}_{L_V^\alpha(\Omega_T)}\Big(1 + o(\|u\|_{L_V^\alpha(\Omega_T)})\Big) + C_2\|F_u(t)\|^2_{L^2_H}.
\end{aligned}
$$

Therefore, by Theorem 3.5, there exists $u_\mu \in G \subset \mathcal{Y}_V^\alpha$ such that $I_\mu(u_\mu) = 0$, i.e.

$$
\begin{aligned}
0 = \mathcal{L} \oplus \Psi_\mu(u_\mu, -\tilde{u}_\mu - Ku_\mu) - \mathbb{E}\int_0^T \langle u_\mu, -\tilde{u}_\mu - Ku_\mu\rangle \, dt \\
+ \mathbb{E}\int_0^T M(\frac{1}{2}F_{u_\mu}, -F_{u_\mu} + \frac{3}{2}Bu_\mu) - \langle \frac{1}{2}F_{u_\mu}, -F_{u_\mu} + \frac{3}{2}Bu_\mu\rangle \, dt.
\end{aligned}
$$

Since $\mathcal{L} \oplus \Psi_\mu$ is convex and coercive in the second variable, there exists $\bar{r} \in L^\beta_{V^*}(\Omega_T)$ such that

$$
\mathcal{L} \oplus \Psi_\mu(u_\mu, -\tilde{u}_\mu - Ku_\mu) = \mathcal{L}(u_\mu, \bar{r}) + \Psi_\mu(u_\mu, -\tilde{u}_\mu - Ku_\mu - \bar{r}),
$$

hence

$$
\begin{aligned}
0 = \mathcal{L}(u_\mu, \bar{r}) - \langle u_\mu, \bar{r}\rangle + \Psi_\mu(u_\mu, -\tilde{u}_\mu - Ku_\mu - \bar{r}) \\
+ \mathbb{E}\int_0^T \langle u_\mu, \tilde{u}_\mu + Ku_\mu + \bar{r}\rangle \, dt \\
+ \mathbb{E}\int_0^T M(\frac{1}{2}F_{u_\mu}, -F_{u_\mu} + \frac{3}{2}Bu_\mu) - \langle \frac{1}{2}F_{u_\mu}, -F_{u_\mu} + \frac{3}{2}Bu_\mu\rangle \, dt.
\end{aligned}
$$

Due to the self-duality of $\mathcal{L}$, Ψ_μ and M, this becomes the sum of three non-negative terms, and therefore

$$
\begin{aligned}
\mathcal{L}(u_\mu, \bar{r}) - \mathbb{E}\int_0^T \langle u_\mu(t), \bar{r}(t)\rangle dt = 0, \\
\Psi_\mu(u_\mu, -\tilde{u}_\mu - Ku_\mu - \bar{r}) + \mathbb{E}\int_0^T \langle u_\mu(t), \tilde{u}_\mu(t) + Ku_\mu(t) + \bar{r}(t)\rangle \, dt = 0, \\
\mathbb{E}\int_0^T M(\frac{1}{2}F_{u_\mu}(t), -F_{u_\mu}(t) + \frac{3}{2}Bu_\mu(t)) - \langle \frac{1}{2}F_{u_\mu}(t), -F_{u_\mu}(t) + \frac{3}{2}Bu_\mu(t)\rangle \, dt = 0.
\end{aligned}
$$

By the limiting case of Legendre duality, this yields

$$
\begin{aligned}
\tilde{u}_\mu + Ku_\mu + \mu\, \partial\psi(u_\mu) \in -\bar{\partial}\mathcal{L}(u_\mu) \qquad (5.12) \\
-F_{u_\mu}(t) + \frac{3}{2}Bu_\mu(t) \in \bar{\partial}M(t, \frac{1}{2}F_{u_\mu}(t)) = \frac{1}{2}F_{u_\mu}(t).
\end{aligned}
$$

The second line implies that for a.e. $t \in [0,T]$ we have $\mathbb{P}$-a.s. $F_{u_\mu} = Bu_\mu$. Moreover, from (5.12) we have that $\partial\psi(u_\mu) \in L^\beta_{V^*}(\Omega_T)$.
Now for an arbitrary process $v \in \mathcal{Y}_V^\alpha$ of the form $v(t) = v(0) + \int_0^t \tilde{v}(s)ds + \int_0^t F_v(s)dW(s)$, we have $\langle \partial\psi(u_\mu(t)), v\rangle = \langle \|\tilde{u}_\mu\|^{\beta-2}_{V^*} D^{-1}\tilde{u}, \tilde{v}\rangle$. Applying Itô's

formula with the progressively measurable process $X(t) := \|\tilde{u}_\mu\|_{V^*}^{\beta-2} D^{-1}\tilde{u}$, we obtain

$$\begin{aligned}\mathbb{E}\int_0^T \langle \|\tilde{u}_\mu\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu(t), \tilde{v}(t)\rangle = -\mathbb{E}\int_0^T \langle \frac{d}{dt}(\|\tilde{u}_\mu\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu), v(t)\rangle \\ + \mathbb{E}\,\langle \|\tilde{u}_\mu(T)\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu(T), v(T)\rangle \\ - \mathbb{E}\,\langle \|\tilde{u}_\mu(0)\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu(0), v(0)\rangle, \quad (5.13)\end{aligned}$$

which, in view of (5.12), implies that

$$\begin{aligned}0 &= \mathbb{E}\int_0^T \Big[\langle \tilde{u}_\mu(t) + Ku_\mu(t) + \bar{\partial}\mathcal{L}(u_\mu), v\rangle + \mu\langle \|\tilde{u}_\mu\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu, \tilde{v}\rangle\Big]\,dt \\ &= \mathbb{E}\int_0^T \Big\langle \tilde{u}_\mu(t) + Ku_\mu(t) + \bar{\partial}\mathcal{L}(u_\mu) - \mu\frac{d}{dt}(\|\tilde{u}_\mu\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu), v\Big\rangle\,dt \\ &\quad + \mu\,\mathbb{E}\,\langle \|\tilde{u}_\mu(T)\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu(T), v(T)\rangle - \mu\,\mathbb{E}\,\langle \|\tilde{u}_\mu(0)\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu(0), v(0)\rangle,\end{aligned}$$

hence $\tilde{u}_\mu(T) = \tilde{u}_\mu(0) = 0$ and $\tilde{u}_\mu + Ku_\mu - \mu\frac{d}{dt}(\|\tilde{u}_\mu\|_{V^*}^{\beta-2} D^{-1}\tilde{u}_\mu) \in -\bar{\partial}\mathcal{L}(u_\mu)$. □

In the following lemma, we shall remove the regularizing term $\mu\partial\psi$.

Lemma 5.2. *Under the above assumptions on $\mathcal{L}$ and B, there exists $u \in \mathcal{Y}_V^\alpha$ with $u(0) = u_0$, such that*

$$\begin{aligned}\mathcal{L}(u, -\tilde{u} - Ku) + \mathbb{E}\int_0^T \langle u(t), \tilde{u}(t) + Ku(t)\rangle\,dt = 0, \\ F_u = Bu.\end{aligned}$$

Proof. Lemma 5.1 yields that for every $\mu > 0$ there exist $u_\mu \in \mathcal{Y}_V^\alpha$ such that $u_\mu(0) = u_0$, $\tilde{u}_\mu(T) = \tilde{u}_\mu(0) = 0$, and satisfying

$$\begin{aligned}\tilde{u}_\mu + Ku_\mu + \mu\,\partial\psi(u_\mu) \in -\bar{\partial}\mathcal{L}(u_\mu) \\ F_{u_\mu}(t) = Bu_\mu(t).\end{aligned} \tag{5.14}$$

Now we show that u_μ is bounded in $\mathcal{Y}_V^\alpha$ with bounds independent of μ. Indeed, multiplying (5.14) by u_μ and integrating over $\Omega \times [0, T]$, we obtain

$$\mathbb{E}\int_0^T \Big\langle \tilde{u}_\mu(t) + Ku_\mu(t) + \mu\,\partial\psi(u_\mu(t)), u_\mu\Big\rangle = -\mathbb{E}\int_0^T \langle \bar{\partial}\mathcal{L}(u_\mu), u_\mu\rangle dt.$$

Apply Itô's formula and use the fact that $\mathbb{E}\int_0^T \langle \mu\, \partial\psi(u_\mu(t)), u_\mu\rangle\, dt \geq 0$ to get

$$\begin{aligned}-\frac{1}{2}\,\|u_{\mu,0}\|^2_{L^2(\Omega;H)} + \frac{1}{2}\,\|u_\mu(T)\|^2_{L^2(\Omega;H)} - \frac{1}{2}\|F_{u_\mu}\|^2_{L^2_H(\Omega_T)} + \|u_\mu\|^{\alpha+1}_{L^\alpha_V(\Omega_T)} \\ = -\mathbb{E}\int_0^T \langle \mu\,\partial\psi(u_\mu) + \bar{\partial}\mathcal{L}(u_\mu), u_\mu\rangle\, dt \\ \leq -\mathbb{E}\int_0^T \langle \bar{\partial}\mathcal{L}(u_\mu), u_\mu\rangle\, dt.\end{aligned}$$

Since for $u_\mu \in \mathcal{Y}^\alpha_V$ we have $u_\mu \in L^\infty(0,T;H)$, then in view of (5.4), we get

$$\begin{aligned}C_1 + \|u_\mu\|^{\alpha+1}_{L^\alpha_V(\Omega_T)} &\leq \|\bar{\partial}\mathcal{L}(u_\mu)\|_{L^\beta_{V^*}(\Omega_T)}\|u_\mu\|_{L^\alpha_V(\Omega_T)} \\ &\leq C\,\|u_\mu\|^\alpha_{L^\alpha_V(\Omega_T)}.\end{aligned}$$

The above inequality implies that $\|u_\mu\|_{L^\alpha_V(\Omega_T)}$ is bounded.
Next, we multiply (5.14) by $D^{-1}\tilde{u}_\mu$ and integrate over Ω_T to get that

$$0 = \mathbb{E}\int_0^T \Big\langle \tilde{u}_\mu(t) + Ku_\mu(t) + \mu\,\partial\psi(u_\mu(t)) + \bar{\partial}\mathcal{L}(t,u_\mu), D^{-1}\tilde{u}_\mu\Big\rangle\, dt$$

From (5.13), and choosing $v = \|\tilde{u}_\mu\|^{\beta-2}_{V^*}D^{-1}\tilde{u}_\mu$ with $\tilde{v} = \frac{d}{dt}(\|\tilde{u}_\mu\|^{\beta-2}_{V^*}D^{-1}\tilde{u}_\mu)$ and $F_v = 0$, we get that $\mathbb{E}\int_0^T \langle \partial\psi(u_\mu(t)), D^{-1}\tilde{u}_\mu\rangle\, dt = 0$, which together with condition (5.4) imply that

$$\|\tilde{u}_\mu\|^2_{L^\beta_{V^*}(\Omega_T)} \leq \|Ku_\mu\|_{L^\beta_{V^*}(\Omega_T)}\|\tilde{u}_\mu\|_{L^\beta_{V^*}(\Omega_T)} + C\,\|u_\mu\|^{\alpha-1}_{L^\alpha_V(\Omega_T)}\|\tilde{u}_\mu\|_{L^\beta_{V^*}(\Omega_T)},$$

hence

$$\|\tilde{u}_\mu\|_{L^\beta_{V^*}(\Omega_T)} \leq \|Ku_\mu\|_{L^\beta_{V^*}(\Omega_T)} + C\,\|u_\mu\|^{\alpha-1}_{L^\alpha_V(\Omega_T)}$$

which means that $\|\tilde{u}_\mu\|_{L^\beta_{V^*}(\Omega_T)}$ is bounded. From (5.7) and since $F_{u_\mu} = Bu_\mu$ we deduce that $\|F_{u_\mu}\|_{L^2_H(\Omega_T)}$ is also bounded. Now since $(u_\mu)_\mu$ is bounded in $\mathcal{Y}^\alpha_V$, there exists $u \in \mathcal{Y}^\alpha_V$ such that $u_\mu \rightharpoonup u$ weakly in $\mathcal{Y}^\alpha_V$, which means that $u_\mu \rightharpoonup u$ weakly in $L^\alpha_V(\Omega_T)$, $\tilde{u}_\mu \rightharpoonup \tilde{u}$ weakly in $L^\beta_{V^*}(\Omega_T)$, and $F_{u_\mu} \rightharpoonup F_u$ weakly in $L^2_H(\Omega_T)$. From (5.14) and since B is weak-norm continuous we have $F_u = Bu$. Then, by (5.12) we obtain

$$\begin{aligned}0 &= \mathcal{L}(u_\mu, -\tilde{u}_\mu - Ku_\mu - \mu\,\partial\psi(u_\mu)) \\ &\quad + \mathbb{E}\int_0^T \Big\langle u_\mu(t), \tilde{u}_\mu(t) + Ku_\mu(t) + \mu\,\partial\psi(u_\mu(t)\Big\rangle\, dt \\ &\geq \mathcal{L}(u_\mu, -\tilde{u}_\mu - Ku_\mu - \mu\,\partial\psi(u_\mu)) + \mathbb{E}\int_0^T \langle u_\mu(t), \tilde{u}_\mu(t) + Ku_\mu(t)\rangle\, dt.\end{aligned}$$

Since K is weak-to-weak continuous, $\langle \partial\psi(u_\mu)), u_\mu\rangle = \|\tilde{u}_\mu\|^\beta_{L^\beta_{V^*}}$ is uniformly bounded, and $\mathcal{L}$ is weakly lower semi-continuous on $L^\alpha_V \times L^\beta_{V^*}$, we get

$$
\begin{aligned}
0 &\geq \liminf_{\mu\to 0} \mathcal{L}(u_\mu, -\tilde{u}_\mu - Ku_\mu - \mu\,\partial\psi(u_\mu)) + \mathbb{E}\int_0^T \langle u_\mu(t), \tilde{u}_\mu(t) + Ku_\mu(t)\rangle\, dt \\
&\geq \mathcal{L}(u, -\tilde{u} - Ku) + \mathbb{E}\int_0^T \langle u(t), \tilde{u}(t) + Ku(t)\rangle\, dt.
\end{aligned}
$$

Since $\mathcal{L}$ is a self-dual Lagrangian on $L^\alpha_V \times L^\beta_{V^*}$, the reverse inequality is always true, and therefore

$$
\mathcal{L}(u, -\tilde{u} - Ku) + \mathbb{E}\int_0^T \langle u(t), \tilde{u}(t) + Ku(t)\rangle\, dt = 0.
$$

□

5.2.2 A general existence result

We shall work toward eliminating the perturbation K. By Lemma 5.2, for each $\varepsilon > 0$, there exists a $u_\varepsilon \in G$ such that $F_{u_\varepsilon} = Bu_\varepsilon$ and

$$
\mathcal{L}(u_\varepsilon, -\tilde{u}_\varepsilon - \varepsilon Ku_\varepsilon) + \mathbb{E}\int_0^T \langle u_\varepsilon(t), \tilde{u}_\varepsilon(t) + \varepsilon Ku_\varepsilon(t)\rangle\, dt = 0, \tag{5.15}
$$

or equivalently

$$
\tilde{u}_\varepsilon + \varepsilon Ku_\varepsilon \in -\bar{\partial}\mathcal{L}(u_\varepsilon). \tag{5.16}
$$

Similar to the argument in Lemma 5.2 we show that u_ε is bounded in $\mathcal{Y}^\alpha_V$ with bounds independent of ε. First, we multiply (5.16) by u_ε and integrate over Ω_T to obtain

$$
\begin{aligned}
\mathbb{E}\int_0^T \langle \tilde{u}_\varepsilon(t) + \varepsilon Ku_\varepsilon(t), u_\varepsilon(t)\rangle\, dt &= -\mathbb{E}\int_0^T \langle \bar{\partial}\mathcal{L}(u_\varepsilon), u_\varepsilon\rangle\, dt \\
&\leq \|\bar{\partial}\mathcal{L}(u_\varepsilon)\|_{L^\beta_{V^*}(\Omega_T)} \|u_\varepsilon\|_{L^\alpha_V(\Omega_T)} \\
&\leq C\,\|u_\varepsilon\|^\alpha_{L^\alpha_V(\Omega_T)},
\end{aligned}
$$

where we used (5.4). In view of (5.15) and (5.3), this implies that

$$
C(\|u_\varepsilon\|^\alpha_{L^\alpha_V(\Omega_T)} - 1) \leq \mathcal{L}(u_\varepsilon, -\tilde{u}_\varepsilon - \varepsilon Ku_\varepsilon) \leq C\,\|u_\varepsilon\|^\alpha_{L^\alpha_V(\Omega_T)},
$$

from which we deduce that u_ε is bounded in $L^\alpha_V(\Omega_T)$. Next, we multiply (5.16) by $D^{-1}\tilde{u}_\varepsilon$ to obtain

$$\mathbb{E}\int_0^T \langle \tilde{u}_\varepsilon(t) + \varepsilon K u_\varepsilon(t), D^{-1}\tilde{u}_\varepsilon(t)\rangle = -\mathbb{E}\int_0^T \langle \bar{\partial}\mathcal{L}(u_\varepsilon), D^{-1}\tilde{u}_\varepsilon(t)\rangle\, dt,$$

and therefore similar to the reasoning as in Lemma 5.2 we deduce that

$$\|\tilde{u}_\varepsilon\|_{L^\beta_{V^*}(\Omega_T)} \leq \varepsilon \|K u_\mu\|_{L^\beta_{V^*}(\Omega_T)} + C\,\|u_\mu\|^{\alpha-1}_{L^\alpha_V(\Omega_T)}.$$

Hence $\tilde{u}_\varepsilon$ is bounded in $L^\beta_{V^*}(\Omega_T)$, and there exists $u \in \mathcal{Y}^\alpha_V$ such that $u_\varepsilon \rightharpoonup u$ weakly in $L^\alpha_V(\Omega_T)$, and $\tilde{u}_\varepsilon \rightharpoonup \tilde{u}$ weakly in $L^\beta_{V^*}(\Omega_T)$, and $F_{u_\varepsilon} \rightharpoonup F_u$ weakly in $L^2_H(\Omega_T)$. Moreover,

$$\begin{aligned} 0 &= \mathcal{L}(u_\varepsilon, -\tilde{u}_\varepsilon - \varepsilon K u_\varepsilon)) + \mathbb{E}\int_0^T \langle u_\varepsilon(t), \tilde{u}_\varepsilon(t) + \varepsilon K u_\varepsilon(t)\rangle\, dt \\ &\geq \mathcal{L}(u_\varepsilon, -\tilde{u}_\varepsilon - \varepsilon K u_\varepsilon) + \mathbb{E}\int_0^T \langle u_\varepsilon(t), \tilde{u}_\varepsilon(t)\rangle\, dt. \end{aligned}$$

Again, $\mathcal{L}$ is weakly lower semi-continuous on $L^\alpha_V \times L^\beta_{V^*}$, therefore by letting $\varepsilon \to 0$ we get

$$0 \geq \mathcal{L}(u, -\tilde{u}) + \mathbb{E}\int_0^T \langle u(t), \tilde{u}(t)\rangle\, dt.$$

Since the reverse inequality is always true we have

$$\mathcal{L}(u, -\tilde{u}) + \mathbb{E}\int_0^T \langle u(t), \tilde{u}(t)\rangle\, dt = 0,$$

and also $F_u(t) = Bu(t)$. By the limiting case of Legendre duality, we now have for a.e. $t \in [0,T]$, $\mathbb{P}$-a.s. $\tilde{u} \in -\bar{\partial}\mathcal{L}(u)$, integrating over $[0,t]$ with the fact that $\int_0^t \tilde{u}(s)ds = u(t) - u_0 - \int_0^t F_u(s)dW(s)$, and $F_u(t) = Bu(t)$ we obtain

$$u(t) = u_0 - \int_0^t \bar{\partial}\mathcal{L}(u)(s)ds + \int_0^t B(u(s))dW(s).$$

5.3 Non-additive noise driven by monotone vector fields

5.3.1 Non-additive noise driven by gradient of convex energies

The first immediate application is the following case when the equation is driven by the gradient of a convex function.

Theorem 5.2. *Let $V \subset H \subset V^*$ be a Gelfand triple, and let $\phi : V \to \mathbb{R} \cup \{+\infty\}$ be an Ω_T-dependent convex lower semi-continuous function on V such that for $\alpha > 1$ and some constants $C_1, C_2 > 0$, for every $t \in [0,T]$, $\mathbb{P}$-a.s. we have*

$$C_2(\|u\|^{\alpha}_{L^{\alpha}_V(\Omega_T)} - 1) \le \mathbb{E}\int_0^T \phi(t,u(t))\,dt \le C_1(1 + \|u\|^{\alpha}_{L^{\alpha}_V(\Omega_T)}).$$

Consider the equation

$$\begin{cases} du(t) = -\partial\phi(u(t))dt + B(u(t))\,dW(t) \\ u(0) = u_0, \end{cases} \tag{5.17}$$

where $B : \mathcal{Y}^{\alpha}_V \to L^2(\Omega_T; H)$ is a weak-to-norm continuous map satisfying for some $C > 0$ and $0 < \delta < \frac{\alpha+1}{2}$,

$$\|Bu\|_{L^2_H(\Omega_T)} \le C\|u\|^{\delta}_{L^{\alpha}(\Omega_T)} \quad \text{for any } u \in \mathcal{Y}^{\alpha}_V.$$

Let u_0 be a random variable in $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, then Equation (5.17) has a solution u in $\mathcal{Y}^{\alpha}_V$.

Proof. It suffices to apply Theorem 5.1 to the self-dual Lagrangian

$$\mathcal{L}(u,p) = \mathbb{E}\int_0^T \phi(t,u(t,w)) + \phi^*(t,p(t,w))\,dt.$$

□

Example 3. *Let $D \subset \mathbb{R}^n$ be an open bounded domain, then the SPDE*

$$\begin{cases} du(t) = \Delta u\,dt + |u|^{q-1}u\,dW \\ u(0) = u_0. \end{cases} \tag{5.18}$$

has a solution provided $\frac{1}{2} \le q < \frac{n}{n-2}$.

Proof. Applying Theorem 5.1 with $\alpha = 2$, $V = H^1_0(D)$, $H = L^2(D)$, $\varphi(u) = \frac{1}{2}\int_D |\nabla u|^2\,dx$ and $Bu = |u|^{q-1}u$, we see that B is weak-to-norm continuous from $\mathcal{Y}^2_V$ to $L^2(\Omega_T; L^2(D))$, as long as $2q < 2^*$, that is $q < \frac{n}{n-2}$. As to the second condition on B, one notes that

$$\|Bu\|_{L^2_H(\Omega_T)} = \left(\mathbb{E}\int_0^T \|u^q\|^2_{L^2(D)}\,dt\right)^{\frac{1}{2}} \le C\|u\|^{\frac{1}{2}}_{L^2_V}\,\|u\|^{q-\frac{1}{2}}_{L^{4q-2}_V},$$

which means that if $\frac{1}{2} \leq q \leq 1$, then $0 \leq 4q - 2 \leq 2$ and

$$\|Bu\|_{L^2_H(\Omega_T)} \leq C\|u\|^{\frac{1}{2}}_{L^2_V} \|u\|^{q-\frac{1}{2}}_{L^{4q-2}_V} \leq C\|u\|^{\frac{1}{2}}_{L^2_V} \|u\|^{q-\frac{1}{2}}_{L^2_V} \leq C\|u\|^q_{L^2_V},$$

which is the condition required by the above theorem. Note that here, $\delta = q < \frac{3}{2} = \frac{\alpha+1}{2}$.
On the other hand, if $1 < q$, then we apply the theorem with $\alpha = 4q - 2$, then the above computation gives that

$$\|Bu\|_{L^2_H(\Omega_T)} \leq C\|u\|^q_{L^{4q-2}_V},$$

since $2 < 4q - 2$. Note also that $q < 2q - \frac{1}{2} = \frac{\alpha+1}{2}$. However, the Lagrangian (here the convex function φ) is not coercive on the space $\mathcal{Y}^\alpha_V = \mathcal{Y}^{4q-2}_V$. To remedy this, we add a perturbation that makes the Lagrangian coercive on this space by considering the convex function

$$\varphi_\epsilon(u) = \frac{1}{2}\int_D |\nabla u|^2\,dx + \frac{\epsilon}{4q-2}\int_D |\nabla u|^{4q-2}\,dx.$$

By applying Theorem 5.1 with $\alpha = 4q - 2$, $V = W^{1,\alpha}(D), H = L^2(D)$, and φ_ϵ, we get a solution u_ϵ for the equation

$$\begin{cases} du(t) = (\Delta u + \epsilon\Delta_{4q-2}u)\,dt + |u|^{q-1}u\,dW \\ u(0) = u_0. \end{cases} \tag{5.19}$$

An argument like what we have already done (twice) above, then allows us to let ϵ go to zero and get a solution for (5.18). □

5.3.2 Non-additive noise driven by general monotone vector fields

More generally, consider the following type of equations

$$\begin{cases} du(t) = -A(t, u(t))dt + B(t, u(t))dW(t) \\ u(0) = u_0, \end{cases} \tag{5.20}$$

where $V \subset H \subset V^*$ is a Gelfand triple, and $A : \Omega \times [0, T] \times V \to V^*$, and $B : \Omega \times [0, T] \times V \to H$, are progressively measurable.

Theorem 5.3. *Assume $A : D(A) \subset V \to V^*$ is a progressively measurable Ω_T-dependent maximal monotone operator satisfying condition (4.5) with $\alpha > 1$ and its conjugate β, as well as*

$$\|A_{w,t}u\|_{V^*} \leq k(\omega, t)(1 + \|u\|_V) \quad \textit{for all } u \in V,\ dt \otimes \mathbb{P} \textit{ a.s.} \tag{5.21}$$

for some $k \in L^\infty(\Omega_T)$*.*
Let $B : \mathcal{Y}_V^\alpha \to L^2(\Omega_T; H)$ *be a weak-to-norm continuous map such that for some* $C > 0$ *and* $0 < \delta < \frac{\alpha+1}{2}$*,*

$$\|Bu\|_{L^2_H(\Omega_T)} \le C\|u\|^\delta_{L^\alpha(\Omega_T)} \quad \textit{for any } u \in \mathcal{Y}_V^\alpha.$$

Let u_0 *be a given random variable in* $L^2_H(\Omega, \mathcal{F}_0, \mathbb{P}; H)$*, then equation (5.20) has a variational solution in* $\mathcal{Y}_V^\alpha$*.*

Proof. Associate again to $A_{\omega,t}$ an Ω_T-dependent self-dual Lagrangian $L_{A_{\omega,t}}$ on $V \times V^*$ in such a way that for almost every $t \in [0,T]$, $\mathbb{P}$-a.s. we have $A_{\omega,t} = \bar{\partial} L_{A_{\omega,t}}$. Then by Lemma 4.1, the Lagrangian

$$\mathcal{L}_A(u,p) = \mathbb{E}\int_0^T L_{A_{\omega,t}}(u(\omega,t), p(\omega,t))dt,$$

is self-dual on $L^\alpha(\Omega_T; V) \times L^\beta(\Omega_T; V^*)$, and satisfies

$$C_1(\|u\|^\alpha_{L^\alpha_V(\Omega_T)} + \|p\|^\beta_{L^\beta_{V^*}(\Omega_T)} - 1) \le \mathcal{L}(u,p) \le C_2(1 + \|u\|^\alpha_{L^\alpha_V(\Omega_T)} + \|p\|^\beta_{L^\beta_{V^*}(\Omega_T)}).$$

(5.21) also implies that for some $C_3 > 0$,

$$\|\bar{\partial}\mathcal{L}_{\mathcal{A}}(u)\|_{L^\beta_{V^*}(\Omega_T)} \le C_3(1 + \|u\|_{L^\alpha_V(\Omega_T)}).$$

The rest follows from Theorem 5.1. □

5.3.3 Non-additive noise driven by monotone vector fields in divergence form

We now show the existence of a variational solution to the following equation:

$$\begin{cases} du = \operatorname{div}(\beta(\nabla u(t,x)))dt + B(u(t))dW(t) & \text{in } [0,T] \times D \\ u(0,x) = u_0 & \text{on } \partial D, \end{cases} \tag{5.22}$$

where D is a bounded domain in $\mathbb{R}^n$, and where the initial position u_0 belongs to $L^2(\Omega, \mathcal{F}_0, \mathbb{P}; L^2(D))$. We assume that

1. The Ω_T-dependent vector field $\beta : \mathbb{R}^n \to \mathbb{R}^n$ is progressively measurable and maximal monotone such that for functions $c_1, c_2, c_3 \in L^\infty(\Omega_T)$, and $m_1, m_2 \in L^1(\Omega_T)$, it satisfies $dt \otimes \mathbb{P}$-a.s.

$$\langle \beta(x), x\rangle \ge \max\{c_1\|x\|^2_{\mathbb{R}^n} - m_1,\ c_2\|\beta(x)\|^2_{\mathbb{R}^n} - m_2\} \quad \text{for all } x \in \mathbb{R}^n, \tag{5.23}$$

and

$$\|\beta(x)\|_{\mathbb{R}^n} \le c_3(1 + \|x\|_{\mathbb{R}^n}) \quad \text{for all } x \in \mathbb{R}^n, \tag{5.24}$$

2. The operator $B : \mathcal{Y}^2_{H^1_0(D)} \to L^2(\Omega_T; L^2(D))$ is a weak-to-norm continuous map such that for some $C > 0$ and $0 < \delta < \frac{\alpha+1}{2}$,

$$\|Bu\|_{L^2_{L^2}(\Omega_T)} \leq C\|u\|^{\delta}_{L^{\alpha}_{H^1_0}(\Omega_T)} \quad \text{for any } u \in \mathcal{Y}^2_{H^1_0(D)}.$$

Theorem 5.4. *Under the above conditions on β and B, Equation (5.22) has a variational solution.*

We shall need the following lemma, which associates to an Ω_T-dependent self-dual Lagrangian on $\mathbb{R}^n \times \mathbb{R}^n$, a self-dual Lagrangian on $L^2(\Omega_T; H^1_0(D)) \times L^2(\Omega_T; H^{-1}(D))$.

Lemma 5.3. *Let L be an Ω_T-dependent self-dual Lagrangian on $\mathbb{R}^n \times \mathbb{R}^n$, then the Lagrangian defined by*

$$\mathscr{L}(u,p) = \inf\left\{\mathbb{E}\int_0^T \int_D L(\nabla u(t,x), f(t,x))\,dx\,dt;\ f \in L^2_{L^2_{\mathbb{R}^n}(D)}(\Omega_T), -\mathrm{div}(f) = p\right\}$$

is self-dual on $L^2(\Omega_T; H^1_0(D)) \times L^2(\Omega_T; H^{-1}(D))$.

We shall need the following general lemma.

Lemma 5.4. *Let L be a self-dual Lagrangian on a Hilbert space $\mathcal{H} \times \mathcal{H}$, and let $\Pi : \mathcal{V} \to \mathcal{H}$ be a bounded linear operator from a reflexive Banach space $\mathcal{V}$ into $\mathcal{H}$ such that the operator $\Pi^*\Pi$ is an isomorphism from $\mathcal{V}$ into $\mathcal{V}^*$. Then, the Lagrangian*

$$\mathcal{L}(u,p) = \inf\left\{L(\Pi u, f);\ f \in \mathcal{H}, \Pi^*(f) = p\right\},$$

is self-dual on $\mathcal{V} \times \mathcal{V}^*$.

Proof. For a fixed $(q,v) \in \mathcal{V}^* \times \mathcal{V}$, write

$$\begin{aligned}
\mathcal{L}^*(q,v) &= \sup\Big\{\langle q,u\rangle + \langle v,p\rangle - \mathcal{L}(u,p);\ u \in \mathcal{V},\ p \in \mathcal{V}^*\Big\}\\
&= \sup\Big\{\langle q,u\rangle + \langle v,p\rangle - L(\Pi u,f);\ u \in \mathcal{V}, p \in \mathcal{V}^*, f \in \mathcal{H}, \Pi^*(f) = p\Big\}\\
&= \sup\Big\{\langle q,u\rangle + \langle v,\Pi^* f\rangle - L(\Pi u,f);\ u \in \mathcal{V}, f \in \mathcal{H}\Big\}\\
&= \sup\Big\{\langle q,u\rangle + \langle \Pi v, f\rangle - L(\Pi u,f);\ u \in \mathcal{V}, f \in \mathcal{H}\Big\}.
\end{aligned}$$

Since $\Pi^*\Pi$ is an isomorphism, for $q \in \mathcal{V}^*$ there exists a fixed $f_0 \in \mathcal{H}$ such that $\Pi^* f_0 = q$. Moreover, the space

$$\mathcal{E} = \{g \in \mathcal{H};\ g = \Pi u,\ \text{for some } u \in \mathcal{V}\},$$

is closed in $\mathcal{H}$ in such a way that its indicator function $\chi_{\mathcal{E}}$ on $\mathcal{H}$

$$\chi_{\mathcal{E}}(g) = \begin{cases} 0 & g \in \mathcal{E} \\ +\infty & \text{elsewhere,} \end{cases}$$

is convex and lower semi-continuous. Its Legendre transform is then given for each $f \in \mathcal{H}$ by

$$\chi^*_{\mathcal{E}}(f) = \begin{cases} 0 & \Pi^* f = 0 \\ +\infty & \text{elsewhere.} \end{cases}$$

It follows that

$$\begin{aligned} \mathcal{L}^*(q,v) &= \sup\Big\{\langle f_0, \Pi u\rangle + \langle \Pi v, f\rangle - L(\Pi u, f);\ u \in \mathcal{V}, f \in \mathcal{H}\Big\} \\ &= \sup\Big\{\langle f_0, g\rangle + \langle \Pi v, f\rangle - L(g,f) - \chi_{\mathcal{E}}(g);\ g \in \mathcal{H}, f \in \mathcal{H}\Big\} \\ &= (L + \chi_{\mathcal{E}})^*(f_0, \Pi v) \\ &= \inf\Big\{L^*(f_0 - r, \Pi v) + \chi^*_{\mathcal{E}}(r);\, r \in \mathcal{H}\Big\} \end{aligned}$$

where we have used that the Legendre dual of the sum is inf-convolution. Finally taking into account the expression for $\chi^*_{\mathcal{E}}$ we obtain

$$\begin{aligned} \mathcal{L}^*(q,v) &= \inf\Big\{L^*(f_0 - r, \Pi v);\, r \in \mathcal{H}, \Pi^* r = 0\Big\} \\ &= \inf\Big\{L(\Pi v, f_0 - r);\, r \in \mathcal{H}, \Pi^* r = 0\Big\} \\ &= \inf\Big\{L(\Pi v, f);\ f \in \mathcal{H}, \Pi^* f = q\Big\} \\ &= \mathcal{L}(v,q). \end{aligned}$$

□

Proof of Lemma 5.3: This is now a direct application of Lemma 5.4. First, lift the random Lagrangian to define a self-dual Lagrangian on $L^2(\Omega_T; L^2(D;\mathbb{R}^n)) \times L^2(\Omega_T; L^2(D;\mathbb{R}^n))$, via

$$\mathcal{L}(u,p) = \mathbb{E}\int_0^T \int_D L(u(t,x), p(t,x))\, dx\, dt,$$

then use Lemma 5.4 with this Lagrangian and the operators

$$L^2(\Omega_T; H^1_0(D)) \xrightarrow{\Pi = \nabla} L^2(\Omega_T; L^2(D;\mathbb{R}^n)) \xrightarrow{\Pi^* = \nabla^*} L^2(\Omega_T; H^{-1}(D)),$$

to get that $\mathscr{L}$ is a self-dual Lagrangian on $L^2(\Omega_T; H_0^1(D)) \times L^2(\Omega_T; H^{-1}(D))$. Note that $\Pi^*\Pi = \nabla^*\nabla = -\Delta$ induces an isomorphism from $L^2(\Omega_T; H_0^1(D))$ to $L^2(\Omega_T; H^{-1}(D))$.

Proof of Theorem 5.4: Again, by Theorem 3.2 and the discussion in Section 4.2.1, one can associate to the maximal monotone map $\beta_{\omega,t}$, an Ω_T-dependent self-dual Lagrangian $L_{\beta_{\omega,t}}(u,p)$ on $\mathbb{R}^n \times \mathbb{R}^n$ in such a way that

$$\beta_{\omega,t} = \bar{\partial} L_{\beta_{\omega,t}}.$$

If β satisfies (5.23), then the Ω_T-dependent self-dual Lagrangian $L_{\beta_{\omega,t}}$ on $\mathbb{R}^n \times \mathbb{R}^n$ satisfy for almost every $t \in [0,T]$, $\mathbb{P}$-a.s.

$$C_1(\|x\|^2_{\mathbb{R}^n} + \|p\|^2_{\mathbb{R}^n} - n_1) \le L_{\beta_{w,t}}(x,p) \le C_2(\|x\|^2_{\mathbb{R}^n} + \|p\|^2_{\mathbb{R}^n} + n_2), \quad (5.25)$$

where $C_1, C_2 \in L^\infty(\Omega_T)$ and $n_1, n_2 \in L^1(\Omega_T)$.
We can then lift it to the space $L^2(\Omega_T; L^2_{\mathbb{R}^n}(D)) \times L^2(\Omega_T; L^2_{\mathbb{R}^n}(D))$ via

$$\mathcal{L}_\beta(u,p) = \mathbb{E}\int_0^T \int_D L_{\beta_{\omega,t}}(u(t,w,x), p(t,w,x))\, dxdt,$$

in such a way that for positive constants C_1, C_2 and C_3 (different from above)

$$C_2(\|u\|^2_{L^2_H(\Omega_T)} + \|p\|^2_{L^2_H(\Omega_T)} - 1) \le \mathcal{L}_\beta(u,p) \le C_1(1 + \|u\|^2_{L^2_H(\Omega_T)} + \|p\|^2_{L^2_H(\Omega_T)}),$$

where $H := L^2_{\mathbb{R}^n}(D)$. In view of (5.24), we also have

$$\|\bar{\partial}\mathcal{L}_\beta(u)\|_{L^2_H(\Omega_T)} \le C_3(1 + \|u\|_{L^2_H(\Omega_T)}).$$

Use now Lemma 5.3 to lift $\mathcal{L}_\beta$ to a self-dual Lagrangian $\mathscr{L}_\beta$ on $L^2(\Omega_T; H_0^1(D)) \times L^2(\Omega_T; H^{-1}(D))$, via the formula

$$\begin{aligned}\mathscr{L}_\beta(u,p) &= \inf_{\substack{f \in L^2(\Omega_T; L^2_{\mathbb{R}^n}(D)) \\ -\mathrm{div}(f) = p}} \left\{ \mathbb{E}\int_0^T \int_D L_{\beta_{w,t}}(\nabla u(t,x), f(t,x))\, dx\, dt; \right\} \\ &= \inf\left\{ \mathcal{L}_\beta(\nabla u, f);\ f \in L^2(\Omega_T; L^2_{\mathbb{R}^n}(D)), -\mathrm{div}(f) = p \right\}. \quad (5.26)\end{aligned}$$

Apply now Theorem 5.1 to get a process $v \in \mathcal{Y}^2_{H_0^1(D)}$ such that

$$\begin{aligned}\mathscr{L}_\beta(v, -\tilde{v}) + \langle v, \tilde{v}\rangle &= 0 \\ F_v &= B \\ v(0) &= u_0.\end{aligned}$$

Now note that

$$\begin{aligned}
0 &= \mathscr{L}_\beta(v,-\tilde{v}) + \langle v, \tilde{v}\rangle \\
&= \inf_{f\in L^2(\Omega_T;L^2_{\mathbb{R}^n}(D))} \left\{\mathbb{E}\int_0^T\int_D L_{\beta(w,t)}(\nabla v, f)\,dx\,dt; \operatorname{div}(f)=\tilde{v}\right\} \\
&\qquad\qquad + \mathbb{E}\int_0^T \langle v(t),\tilde{v}(t)\rangle_{H^1_0,H^{-1}}\,dt \\
&= \inf_{f\in L^2(\Omega_T;L^2_{\mathbb{R}^n}(D))} \left\{\mathbb{E}\int_0^T\int_D L_{\beta(w,t)}(\nabla v, f) - \langle \nabla v(x,t), f(x,t)\rangle\,dx\,dt\right\} \\
&= \inf_{f\in L^2(\Omega_T;L^2_{\mathbb{R}^n}(D))} J_v(f),
\end{aligned}$$

where

$$J_v(f) := \mathbb{E}\int_0^T\int_D \{L_{\beta(w,t)}(\nabla v, f) - \langle \nabla v(x,t), f(x,t)\rangle\}\,dx\,dt.$$

Note that condition (5.25) implies that $L(y,0)\le C(1+\|y\|^2_{\mathbb{R}^n})$, which means that J_v is coercive on $L^2(\Omega_T;L^2_{\mathbb{R}^n}(D))$, thus there exists $\bar{f}\in L^2(\Omega_T;L^2_{\mathbb{R}^n}(D))$ with $\operatorname{div}(\bar{f})=\tilde{v}$ such that

$$\mathbb{E}\int_0^T\int_D L_{\beta(w,t)}(\nabla v,\bar{f}) - \langle \nabla v(x,t), \bar{f}(x,t)\rangle\,dx\,dt = 0.$$

The self-duality of L then implies that $\bar{f}(x,t)=\bar{\partial}L(\nabla v(x,t))=\beta(\nabla v(x,t))$. Taking divergence leads to $\tilde{v}\in\operatorname{div}(\beta(\nabla v))$. Taking integrals over $[0,t]$ and using the fact that $v\in\mathcal{Y}^2_{H^1_0(D)}$ finally gives

$$\begin{aligned}
\int_0^t \operatorname{div}(\beta(\nabla v(s)))\,ds = \int_0^t \tilde{v}(s)ds &= v(t)-v(0)-\int_0^t F_v(s)dW(s) \\
&= v(t)-u_0-\int_0^t B(v(s))dW,
\end{aligned}$$

which completes the proof.

Part II

Euclidean Moser-Onofri inequality and its extensions to higher dimension

Chapter 6

A dual Moser-Onofri inequality and its extensions to higher dimension

6.1 Introduction

One of the equivalent forms of Moser's inequality [39] on the 2-dimensional sphere $\mathbb{S}^2$ states that the functional

$$I(u) := \frac{1}{4}\int_{\mathbb{S}^2} |\nabla u|^2 \,\mathrm{d}\omega + \int_{\mathbb{S}^2} u \,\mathrm{d}\omega - \log\left(\int_{\mathbb{S}^2} e^u \,\mathrm{d}\omega\right) \tag{6.1}$$

is bounded below on $H^1(\mathbb{S}^2)$, where $\mathrm{d}\omega$ is the Lebesgue measure on $\mathbb{S}^2$, normalized so that $\int_{\mathbb{S}^2} \mathrm{d}\omega = 1$. Later, Onofri [42] showed that the infimum of (6.1) over $H^1(\mathbb{S}^2)$ is actually zero, and that modulo conformal transformations, $u = 0$ is the only optimal function. Note that this inequality is related to the "prescribed Gaussian curvature" problem on $\mathbb{S}^2$,

$$\Delta u + K(x)e^{2u} = 1 \quad \text{on } \mathbb{S}^2, \tag{6.2}$$

where $K(x)$ is the Gaussian curvature associated to the metric $g = e^{2u}g_0$ on $\mathbb{S}^2$, and $\Delta = \Delta_{g_0}$ is the Laplace-Beltrami operator corresponding to the standard metric g_0. Finding g for a given K leads to solving (6.2). Variationally, this reduces to finding the critical points of the functional

$$\mathcal{F}(u) = \int_{\mathbb{S}^2} |\nabla u|^2 \frac{\mathrm{d}V_0}{4\pi} + 2\int_{\mathbb{S}^2} u \frac{\mathrm{d}V_0}{4\pi} - \log\left(\int_{\mathbb{S}^2} K(x)e^{2u}\frac{\mathrm{d}V_0}{4\pi}\right) \quad \text{on } H^1(\mathbb{S}^2), \tag{6.3}$$

where the volume form is such that $\int_{\mathbb{S}^2} \mathrm{d}V_0 = 4\pi$. Onofri's result says that, modulo conformal transformations, $u \equiv 0$ is the only solution of the "prescribed Gaussian curvature" problem (6.2) for $K = 1$, i.e., $\frac{1}{2}\Delta u + e^u = 1$ on $\mathbb{S}^2$, which after rescaling, $u \mapsto 2u$, gives

$$\Delta u + e^{2u} = 1 \quad \text{on } \mathbb{S}^2. \tag{6.4}$$

The proof given by Onofri in [42] makes use of a constrained Moser inequality due to Aubin [3] combined with the invariance of the functional (6.1) under conformal transformations. Other proofs were given by Osgood-Philips-Sarnak [43] and by Hong [36]. See also Ghoussoub-Moradifam [31].

In this part, we use the theory of mass transport to prove that 0 is the infimum of the functional (6.3) at least when $K = 1$. While this approach has by now become standard, there are many reasons why it has not been so far spelled out in the case of the Moser functional. The first is due to the fact that, unlike the case of $\mathbb{R}^n$, optimal mass transport on the sphere is harder to work with. To avoid this difficulty, we use an equivalent formulation of the Onofri inequality (6.1), which is obtained by projecting (6.1) on $\mathbb{R}^2$ via the stereographic projection with respect to the North pole $N = (0, 0, 1)$, i.e., $\Pi : \mathbb{S}^2 \to \mathbb{R}^2$, $\Pi(x) := \left(\frac{x_1}{1-x_3}, \frac{x_2}{1-x_3}\right)$ where $x = (x_1, x_2, x_3)$. The Moser-Onofri inequality becomes the *Euclidean Onofri inequality* on $\mathbb{R}^2$, namely

$$\frac{1}{16\pi}\int_{\mathbb{R}^2} |\nabla u|^2 \,\mathrm{d}x + \int_{\mathbb{R}^2} u \,\mathrm{d}\mu_2 - \log\left(\int_{\mathbb{R}^2} e^u \,\mathrm{d}\mu_2\right) \geq 0 \quad \forall u \in \mathcal{D}(\mathbb{R}^2), \tag{6.5}$$

where μ_2 is the probability density on $\mathbb{R}^2$ defined by $\mu_2(x) = \frac{1}{\pi(1+|x|^2)^2}$, and $\mathrm{d}\mu_2 = \mu_2(x)\,\mathrm{d}x$, and $\mathcal{D}(\mathbb{R}^2) = \{u \in L^1(\mathbb{R}^2, d\mu_2); \nabla u \in L^2(\mathbb{R}^2, \mathrm{d}x)\}$.

One can then try to apply the Cordero-Nazaret-Villani [15] approach as generalized by Agueh-Ghoussoub-Kang [2] and write the *Energy-Entropy production duality* for functions that are of compact support in Ω,

$$\begin{aligned} &\sup\left\{ -\int_\Omega (F(\rho) + \frac{1}{2}|x|^2\rho)\,\mathrm{d}x;\ \rho \in \mathbb{P}(\Omega)\right\} \\ &\qquad = \inf\left\{\int_\Omega \alpha|\nabla u|^2 - G\left(\psi \circ u\right)\,\mathrm{d}x;\ u \in H_0^1(\Omega), \int_\Omega \psi(u)\,\mathrm{d}x = 1\right\}, \end{aligned} \tag{6.6}$$

where $G(x) = (1-n)F(x)+nxF'(x)$ and where ψ and α are also computable from F. Here $\mathbb{P}(\Omega)$ denotes the set of probability densities on Ω.

By choosing $F(x) = -nx^{1-1/n}$ and $\psi(t) = |t|^{2^*}$ where $2^* = \frac{2n}{n-2}$ and $n > 2$, one obtains the following duality formula for the Sobolev inequality

$$\begin{aligned} &\sup\left\{n\int_{\mathbb{R}^n} \rho^{1-1/n}\,\mathrm{d}x - \frac{1}{2}\int_{\mathbb{R}^n} |x|^2\rho\,\mathrm{d}x;\ \rho \in \mathbb{P}(\mathbb{R}^n)\right\} \\ &\qquad = \inf\left\{2\left(\frac{n-1}{n-2}\right)^2 \int_{\mathbb{R}^n} |\nabla u|^2\,\mathrm{d}x; u \in D^{1,2}(\mathbb{R}^n), \int_{\mathbb{R}^n} |u|^{2^*}\,\mathrm{d}x = 1\right\}, \end{aligned} \tag{6.7}$$

where u and ρ have compact support in $\mathbb{R}^n$. The extremal u_∞ and ρ_∞ are then obtained as solutions of

$$\nabla\left(\frac{|x|^2}{2}-\frac{n-1}{\rho_\infty^{1/n}}\right)=0, \quad \rho_\infty=u_\infty^{2^*}\in\mathcal{P}(\mathbb{R}^n). \tag{6.8}$$

The best constants are then obtained by computing ρ_∞ from (6.8) and inserting it into (6.7) in such a way that

$$\inf\left\{2\left(\frac{n-1}{n-2}\right)^2\int_{\mathbb{R}^n}|\nabla u|^2\,\mathrm{d}x;\ u\in D^{1,2}(\mathbb{R}^n), \|u\|_{2^*}=1\right\}$$
$$=n\int_{\mathbb{R}^n}\rho_\infty^{1-1/n}\,\mathrm{d}x-\frac{1}{2}\int_{\mathbb{R}^n}|x|^2\rho_\infty\,\mathrm{d}x.$$

Note that this duality leads to a correspondence between a solution to the Yamabe equation

$$-\Delta u=|u|^{2^*-2}u \quad \text{on } \mathbb{R}^n \tag{6.9}$$

and stationary solution to the rescaled fast diffusion equation

$$\partial_t\rho=\Delta\rho^{1-\frac{1}{n}}+\operatorname{div}(x\rho) \quad \text{on } \mathbb{R}^n. \tag{6.10}$$

The above scheme does not however apply to inequality (6.5). For one, the functions $e^u\mu_2=\frac{e^{u(x)}}{\pi(1+|x|^2)^2}$ do not have compact support, and if one restricts them to bounded domains, we then need to take into consideration various boundary terms. What is remarkable is that a similar program can be carried out provided the dual formula involving the free energy

$$J_\Omega(\rho)=-\int_\Omega(F(\rho)+|x|^2\rho)\,\mathrm{d}x$$

is renormalized by substituting it with $J_\Omega(\rho)-J_\Omega(\mu_2)$.

Another remarkable fact is that the corresponding free energy turned out to be $F(\rho)=-2\rho^{\frac{1}{2}}$, which is the same as the one associated to the critical case of the Sobolev inequality $F(\rho)=-n\rho^{1-\frac{1}{n}}$ when $n\geq 3$. In other words, the Moser-Onofri inequality and the Sobolev inequality "dualize" in the same way, and both the Yamabe problem (6.9) and the prescribed Gaussian curvature problem (6.4) reduce to the study of the fast diffusion equation (6.10), with the caveat that in dimension $n=2$, the above equation needs to be considered only on bounded domains, with Neumann boundary conditions.

More precisely, we shall show that, when restricted to balls B_R of radius R in $\mathbb{R}^2$, there is a duality between the "Onofri functional"

$$I_R(u) = \frac{1}{16\pi}\int_{B_R} |\nabla u|^2\,\mathrm{d}x + \int_{B_R} u\,\mathrm{d}\mu_2$$

$$\text{on } X_R := \{u \in \mathcal{D}(B_R);\ \int_{\mathbb{R}^2} e^u\,\mathrm{d}\mu_2 = 1\},$$

and the free energy

$$J_R(\rho) = \frac{2}{\sqrt{\pi}}\int_{B_R} \sqrt{\rho}\,\mathrm{d}x - \int_{B_R} |x|^2\rho\,\mathrm{d}x$$

$$\text{on } Y_R := \{\rho \in L^1_+(B_R);\ \frac{1}{\mu_2(B_R)}\int_{B_R} \rho\,\mathrm{d}x = 1,$$

where

$$\mu_2(B_R) := \int_{B_R} \mathrm{d}\mu_2 = \frac{R^2}{1+R^2}.$$

Note that if u has its support in B_R, then

$$\int_{\mathbb{R}^2} e^u\,\mathrm{d}\mu_2 = 1 \text{ if and only if } \tfrac{1}{\mu_2(B_R)}\int_{B_R} e^u\,\mathrm{d}\mu_2 = 1.$$

We show that once the free energy is re-normalized by subtracting the free energy of μ_2, we then have

$$\sup\{J_R(\rho) - J_R(\mu_2);\rho \in Y_R\} = 0 = \inf\{I_R(u);\ u \in X_R\}. \tag{6.11}$$

Note that when $R \to +\infty$, the right hand side yields the Onofri inequality

$$\inf\left\{\frac{1}{16\pi}\int_{\mathbb{R}^2} |\nabla u|^2\,\mathrm{d}x + \int_{\mathbb{R}^2} u\,\mathrm{d}\mu_2;\ u \in D^{1,2}(\mathbb{R}^2), \int_{\mathbb{R}^2} e^u\,\mathrm{d}\mu_2 = 1\right\} = 0,$$

while the left-hand side doesn't yield a universal upper bound for $J_R(\rho)$ since

$$J_R(\mu_2) = \log(1+R^2) + \frac{R^2}{1+R^2} \to +\infty \quad \text{as } R \to +\infty.$$

We actually show that our approach extends to higher dimensions. More precisely, if B_R is a ball of radius R in $\mathbb{R}^n$ where $n \geq 2$, and if one considers the probability density μ_n on $\mathbb{R}^n$ defined by

$$\mu_n(x) = \frac{1}{\omega_n(1+|x|^{\frac{n}{n-1}})^n}$$

(ω_n is the volume of the unit sphere in $\mathbb{R}^n$), and the operator $H_n(u, \mu_n)$ on $W^{1,n}(\mathbb{R}^n)$ by

$$H_n(u, \mu_n) := |\nabla u + \nabla(\log \mu_n)|^n - |\nabla(\log \mu_n))|^n - n|\nabla(\log \mu_n))|^{n-2} \nabla(\log \mu_n) \cdot \nabla u,$$

there is then a duality between the functional

$$_R(u) = \frac{1}{\beta(n)} \int_{B_R} H_n(u, \mu_n)\,\mathrm{d}x + \int_{B_R} u\,\mathrm{d}\mu_n$$

$$\text{on } X_R := \{u \in \mathcal{D}(B_R);\ \int_{\mathbb{R}^n} e^u\,\mathrm{d}\mu_n = 1$$

and the free energy – renormalized by again subtracting $J_R(\mu_n)$ –

$$J_R(\rho) = \alpha(n) \int_{B_R} \rho^{\frac{n-1}{n}}\,\mathrm{d}x - \int_{B_R} |x|^{\frac{n}{n-1}} \rho\,\mathrm{d}x$$

$$\text{on } Y_R := \{\rho \in L^1_+(B_R);\ \frac{1}{\mu_n(B_R)} \int_{B_R} \rho\,\mathrm{d}x = 1\}$$

where

$$\alpha(n) = \frac{n}{n-1} {\omega_n}^{-1/n}, \quad \beta(n) = \omega_n \,(\frac{n}{n-1})^{n-1} n^{n+1}$$

and

$$\mu_n(B_R) := \int_{B_R} \mathrm{d}\mu_n = \frac{R^n}{(1 + R^{\frac{n}{n-1}})^{n-1}}.$$

We then deduce the following higher dimensional version of the Onofri inequality: For $n \geq 2$,

$$\frac{1}{\beta(n)} \int_{\mathbb{R}^n} H_n(u, \mu_n)\,\mathrm{d}x + \int_{\mathbb{R}^n} u\,\mathrm{d}\mu_n - \log\left(\int_{\mathbb{R}^n} e^u\,\mathrm{d}\mu_n\right) \geq 0 \tag{6.12}$$

$$\text{for all } u \in D^{1,2}(\mathbb{R}^n).$$

We finish this introduction by mentioning that there was an attempt in [18] to use mass transport to establish the Euclidean Onofri inequality (6.5) in the radial case. In [37], Maggi and Villani also establish Sobolev-type inequalities involving boundary trace terms via mass transport methods. They actually deal with a family of Moser-Trudinger inequalities as a limiting case of Sobolev inequality when the power $p \to n$, in the presence of boundary terms on a Lipschitz domain in $\mathbb{R}^n$. However, to our knowledge, our duality result, the extensions of Onofri's inequality to higher dimensions, as well as the mass transport proof of the general (non-radial) Onofri inequality are new.

In the following, first we recall the mass transport approach to sharp Sobolev inequalities and some consequences. In the next section, we establish the n-dimensional mass transport duality principle, from which we could deduce the two dimensional Euclidean Onofri inequality (6.5).

6.2 Preliminaries

We start by briefly describing the mass transport approach to sharp Sobolev inequalities as proposed by [15]. We will follow here the framework of [2] as it clearly shows the correspondence between the Yamabe equation (6.9) and the rescaled fast diffision equation (6.10).

Let $\rho_0, \rho_1 \in \mathcal{P}(\mathbb{R}^n)$. If T is the optimal map pushing ρ_0 forward to ρ_1 (i.e. $T_\# \rho_0 = \rho_1$) in the mass transport problem for the quadratic cost $c(x-y) = \frac{|x-y|^2}{2}$ (see [52] for details), then $[0,1] \ni t \mapsto \rho_t = (T_t)_\# \rho_0$ is the geodesic joining ρ_0 and ρ_1 in $(\mathcal{P}(\mathbb{R}^n), d_2)$; here $T_t := (1-t)\mathrm{id} + tT$ and d_2 denotes the quadratic Wasserstein distance (see [52]). Moreover, given a function $F : [0,\infty) \to \mathbb{R}$ such that $F(0) = 0$ and $x \mapsto x^n F(x^{-n})$ is convex and non-increasing, the functional $H^F(\rho) := \int_{\mathbb{R}^n} F\left(\rho(x)\right) \mathrm{d}x$ is displacement convex [38], in the sense that $[0,1] \ni t \mapsto H^F(\rho_t) \in \mathbb{R}$ is convex (in the usual sense), for all pairs (ρ_0, ρ_1) in $\mathcal{P}(\mathbb{R}^n)$. A direct consequence is the following convexity inequality, known as "*energy inequality*":

$$H^F(\rho_1) - H^F(\rho_0) \geq \left[\frac{d}{dt} H^F(\rho_t)\right]_{t=0} = \int_{\mathbb{R}^n} \rho_0 \nabla\left(F'(\rho_0)\right) \cdot (T - \mathrm{id})\, \mathrm{d}x,$$

which, after integration by parts of the right hand side term, reads as

$$-H^F(\rho_1) \leq -H^{F+nP_F}(\rho_0) - \int_{\mathbb{R}^n} \rho_0 \nabla\left(F'(\rho_0)\right) \cdot T(x)\, \mathrm{d}x, \tag{6.13}$$

where $P_F(x) = xF'(x) - F(x)$; here id denotes the identity function on $\mathbb{R}^n$. By the Young inequality

$$-\nabla\left(F'(\rho_0)\right) \cdot T(x) \leq \frac{|\nabla F'(\rho_0)|^p}{p} + \frac{|T(x)|^q}{q} \quad \forall p, q > 1 \text{ such that } \quad \frac{1}{p} + \frac{1}{q} = 1, \tag{6.14}$$

(6.13) gives

$$-H^F(\rho_1) \leq -H^{F+nP_F}(\rho_0) + \frac{1}{p}\int_{\mathbb{R}^n} \rho_0 |\nabla F'(\rho_0)|^p \, \mathrm{d}x + \frac{1}{q}\int_{\mathbb{R}^n} \rho_0(x) |T(x)|^q \, \mathrm{d}x,$$

i.e.,

$$-H^F(\rho_1) - \frac{1}{q}\int_{\mathbb{R}^n} |y|^q \rho_1(y)\,\mathrm{d}y \le -H^{F+nP_F}(\rho_0) + \frac{1}{p}\int_{\mathbb{R}^n} \rho_0 |\nabla F'(\rho_0)|^p \,\mathrm{d}x, \tag{6.15}$$

where we use that $T_{\#}\rho_0 = \rho_1$. Furthermore, if $\rho_0 = \rho_1$, then $T = \mathrm{id}$ and equality holds in (6.13). Then equality holds in (6.15) if it holds in the Young inequality (6.14). This occurs when $\rho_0 = \rho_1$ satisfies $\nabla\left(F'\left(\rho_0(x)\right) + \frac{|x|^q}{q}\right) = 0$. Therefore, we have established the following duality:

$$\begin{aligned}\sup\Big\{ -H^F(\rho_1) - \frac{1}{q}\int_{\mathbb{R}^n} |y|^q\rho_1(y)\,\mathrm{d}y;\ \ \rho_1 \in \mathcal{P}(\mathbb{R}^n)\Big\} \\ = \inf\Big\{ -H^{F+nP_F}(\rho_0) + \frac{1}{p}\int_{\mathbb{R}^n} \rho_0|\nabla F'(\rho_0)|^p\,\mathrm{d}x;\ \rho_0 \in \mathcal{P}(\mathbb{R}^n)\Big\},\end{aligned} \tag{6.16}$$

and an optimal function in both problems is $\rho_0 = \rho_1 := \rho_\infty$ solution of

$$\nabla\left(F'\left(\rho_\infty(x)\right) + \frac{|x|^q}{q}\right) = 0. \tag{6.17}$$

In particular, choosing $F(x) = -nx^{1-1/n}$ and $\rho_0 = u^{2^*}$ where $2^* = \frac{2n}{n-2}$ and $n > 2$, then $H^{F+nP_F} = 0$, and (6.16)-(6.17) gives the duality formula for the Sobolev inequality (6.7).

Our goal now is to extend this mass transport proof of the Sobolev inequality to the Euclidean Onofri inequality (6.5). As already mentioned in the introduction, a first attempt on this issue was recently made by [18], but the result produced was only restricted to the radial case. Here we show in full generality (without restricting to radial functions u) that the Euclidean Onofri inequality (6.5) can be proved by mass transport techniques. More precisely, we establish an analogue of the duality (6.7) for Euclidean Onofri inequality in dimension n (see Theorem (6.1)), from which we deduce the n-dimensional Onofri inequality (6.35) (see Theorem 6.2). Furthermore, we obtain –as for the critical Sobolev inequality– a correspondence between the prescribed Gaussian curvature problem (6.4) and the rescaled fast diffusion equation (6.10).

We shall need the following general lemma from the theory of mass transport.

Lemma 6.1. *Let $\rho_0, \rho_1 \in \mathbb{P}(B_R)$, where $\mathbb{P}(B_R)$ denotes the set of probability densities on the ball $B_R \subset \mathbb{R}^n$. Let T be the optimal map pushing ρ_0 forward* *to ρ_1 (i.e. $T_{\#}\rho_0 = \rho_1$) in the mass transport problem corresponding to the quadratic cost. Then*

$$\int_{B_R} \rho_1(y)^{1-\frac{1}{n}}\, dy \leq \frac{1}{n} \int_{B_R} \rho_0(x)^{1-\frac{1}{n}} \mathrm{div}\,(T(x))\, dx. \tag{6.18}$$

Proof. By Brenier's theorem [10], there is a map $T : B_R \to B_R$ such that $T = \nabla\varphi$ where $\varphi : B_R \to \mathbb{R}$ is convex, and $T_{\#}\rho_0 = \rho_1$. We therefore have the following Monge-Ampère equation,

$$\rho_0(x) = \rho_1(T(x)) \det \nabla T(x) \tag{6.19}$$

or equivalently

$$\rho_1(T(x)) = \rho_0(x)[\det \nabla T(x)]^{-1}. \tag{6.20}$$

By the arithmetic-geometric-mean inequality

$$[\det \nabla T(x)]^{\frac{1}{n}} \leq \frac{1}{n} \mathrm{div}(T(x)),$$

(6.20) gives

$$\rho_1(T(x))^{-\frac{1}{n}} \leq \frac{1}{n}\, \rho_0(x)^{-\frac{1}{n}} \mathrm{div}\,(T(x))\,. \tag{6.21}$$

Now using the change of variable $y = T(x)$, we have

$$\int_{B_R} \rho_1(y)^{1-\frac{1}{n}}\, \mathrm{d}y = \int_{B_R} \rho_1\,(T(x))^{1-\frac{1}{n}} \det(\nabla T(x))\, \mathrm{d}x,$$

which implies by (6.19) and (6.21), that

$$\begin{aligned} \int_{B_R} \rho_1(y)^{1-\frac{1}{n}}\, \mathrm{d}y &\leq \frac{1}{n} \int_{B_R} \rho_0(x)^{-\frac{1}{n}} \mathrm{div}(T(x))\rho_0(x)\, \mathrm{d}x \\ &= \frac{1}{n} \int_{B_R} \rho_0(x)^{1-\frac{1}{n}} \mathrm{div}(T(x))\, \mathrm{d}x, \end{aligned}$$

and we are done. □

6.3 Euclidean n-dimensional Onofri inequality: A duality formula

Consider the probability density on $\mathbb{R}^n$, $\mu_n(y) = \frac{1}{\omega_n(1+|y|^{\frac{n}{n-1}})^n}$, where ω_n is the volume of the unit sphere in $\mathbb{R}^n$, and set

$$\theta_R := \int_{B_R} \mu_n(y)\, \mathrm{d}y = \frac{R^n}{(1+R^{\frac{n}{n-1}})^{n-1}}.$$

We shall establish the following duality formula.

Theorem 6.1. (Duality for n-dimensional Euclidean Onofri inequality) *For each ball B_R in $\mathbb{R}^n$ with radius $R > 0$, we consider the following free functional*

$$J_R(\rho) = \alpha(n) \int_{B_R} \rho(y)^{\frac{n-1}{n}} \, dy - \int_{B_R} |y|^{\frac{n}{n-1}} \rho(y) \, dy \qquad \text{for } \rho \in L^1_+(B_R),$$

as well as the "entropy" functional

$$I_R(u) = \frac{1}{\beta(n)} \int_{B_R} H_n(u, \mu_n) \, dx + \int_{B_R} u(x) \, d\mu_n \qquad \text{for } u \in D^{1,2}(B_R),$$

where

$$\alpha(n) = \frac{n}{n-1} \left(\frac{1}{\omega_n} \right)^{1/n}, \qquad \beta(n) = \left(\frac{n}{n-1} \right)^{n-1} n^{(n+1)} \, \omega_n$$

and

$$H_n(u, \mu_n) := |\nabla u + \nabla(\log \mu_n)|^n - |\nabla(\log \mu_n)|^n - n |\nabla(\log \mu_n)|^{n-2} \nabla(\log \mu_n) \cdot \nabla u.$$

The following duality formula then holds:

$$\sup \left\{ J_R(\rho) - J_R(\mu_n); \rho \in L^1_+(\mathbb{R}^n), \int_{B_R} \rho \, dy = \theta_R \right\} \tag{6.22}$$
$$= \inf \left\{ I_R(u); \, u \in \mathcal{D}(B_R), \int_{B_R} e^u \, d\mu_n = \theta_R \right\} = 0.$$

Moreover, the maximum on the l.h.s. is attained only at $\rho_{max} = \mu_n$, and the minimum on the r.h.s. is attained only at $u_{min} = 0$.

Remark 6.1. *Before proving the theorem, we make a few remarks on the operator $H_n(u, \mu_n)$.*

1. *Consider the function $c : \mathbb{R}^n \to \mathbb{R}$, $c(z) = |z|^n$, $n \geq 2$. Clearly c is strictly convex, and $\nabla c(z) = n|z|^{n-2} z$. So we have the convexity inequality*

$$c(z_1) - c(z_0) - \nabla c(z_0) \cdot (z_1 - z_0) \geq 0 \quad \forall z_0, z_1 \in \mathbb{R}^n. \tag{6.23}$$

   Setting $z_0 = \nabla(\log \mu_n)$ and $z_1 = \nabla u + \nabla(\log \mu_n)$, we see that $H_n(u, \mu_n)$ is nothing but the l.h.s of (6.23); we then deduce that

$$H_n(u, \mu_n) \geq 0 \qquad \forall \, u, \mu_n.$$

2. *For all $u \in W_0^{1,n}(B_R)$, the integral of $H_n(u,\mu_n)$ over B_R involves a well-known operator, the n-Laplacian Δ_n, defined by*

$$\Delta_n v := div(|\nabla v|^{n-2}\nabla v). \tag{6.24}$$

Indeed, this can be seen after performing an integration by parts in the last term of $H_n(u,\mu_n)$,

$$\int_{B_R} H_n(u,\mu_n)\,dy = \int_{B_R} |\nabla u + \nabla \log \mu_n|^n\,dy - \int_{B_R} |\nabla(\log \mu_n))|^n\,dy + n\int_{B_R} u\,\Delta_n(\log \mu_n)\,dy.$$

Proof. By applying Lemma 6.1, and using an integration by parts and taking $\frac{1}{m} := 1 - \frac{1}{n}$, we have

$$n\int_{B_R} \rho_1^{1/m}\,\mathrm{d}y \le -\int_{B_R} \nabla(\rho_0^{1/m})\cdot T(x)\,\mathrm{d}x + \int_{\partial B_R} \rho_0^{1/m} T(x)\cdot \nu\,\mathrm{d}S.$$

Use the elementary identity $\nabla(\rho_0^{1/m}) = \frac{1}{m}\rho_0^{1/m}\nabla(\log \rho_0)$ to obtain

$$mn\int_{B_R} \rho_1^{1/m}\,\mathrm{d}y \le -\int_{B_R} \rho_0^{1/m}\nabla(\log \rho_0)\cdot T(x)\,\mathrm{d}x + m\int_{\partial B} \rho_0^{1/m} T(x)\cdot \nu\,\mathrm{d}S. \tag{6.25}$$

Set $\rho_0 = \frac{e^u \mu_n}{\theta_R}$, where $u \in \mathcal{D}(B_R)$ satisfies $\int_{B_R} e^u d\mu_n = \theta_R$, and let ρ_1 be any probability density supported on B_R. By applying (6.25) to ρ_0 and ρ_1, we get

$$mn\int_{B_R} (\theta_R \rho_1)^{1/m}\,\mathrm{d}y \le -\int_{B_R} (e^u\mu_n)^{1/m}\nabla(\log(e^u\mu_n))\cdot T(x)\,\mathrm{d}x + m\int_{\partial B_R} (e^u\mu_n)^{1/m} T(x)\cdot \nu\,\mathrm{d}S.$$

Using Young's inequality, for any $\varepsilon > 0$,

$$-(e^u\mu_n)^{1/m}\,\nabla\left(\log(e^u\mu_n)\right)\cdot T(x) \le \frac{1}{n\varepsilon}|\nabla\left(\log(e^u\mu_n)\right)|^n + \frac{\varepsilon^{m/n}}{m}\,e^u\mu_n|T(x)|^m$$

and the fact that $T_\# \rho_0 = \rho_1$, we get

$$mn^2\varepsilon\int_{B_R} (\theta_R\rho_1)^{1/m}\,\mathrm{d}y - mn\varepsilon\int_{\partial B_R} \mu_n^{1/m} T(x)\cdot\nu\,\mathrm{d}S - \frac{n}{m}\varepsilon^m\int_{B_R} |y|^m \theta_R \rho_1\,\mathrm{d}y \le \int_{B_R} |\nabla u + \nabla \log \mu_n|^n\,\mathrm{d}x. \tag{6.26}$$

We now estimate the boundary term. Since $T : B_R \to B_R$, then $|T(x)| \le R$ for all $x \in B_R$

$$\begin{aligned}\int_{\partial B_R} \mu_n^{1/m} T(x)\cdot\nu \,\mathrm{d}S &= \left(\frac{1}{\omega_n}\right)^{1/m} \frac{1}{(1+R^m)^{n/m}} \int_{\partial B_R} T(x)\cdot \frac{x}{|x|}\,\mathrm{d}S \\ &\le n\omega_n^{1/n} \frac{R^n}{(1+R^m)^{n/m}} = n\omega_n^{1/n}\,\theta_R.\end{aligned} \tag{6.27}$$

where we used the fact that $\nu = \frac{x}{|x|}$.
Inserting (6.27) into (6.26), and setting $\rho := \theta_R \rho_1$, we get for all $\varepsilon > 0$,

$$\begin{aligned}\varepsilon \left[n^2 m \int_{B_R} \rho^{1/m}\,\mathrm{d}y - n^2 m \omega_n^{1/n}\theta_R\right] - \varepsilon^m \frac{n}{m}\int_{B_R} |y|^m \rho\,\mathrm{d}y \\ \le \int_{B_R} |\nabla u + \nabla \log \mu_n|^n\,\mathrm{d}x.\end{aligned} \tag{6.28}$$

Now, we introduce the operator $H_n(u, \mu_n)$ in the r.h.s of (6.28). We have

$$|\nabla u+\nabla \log \mu_n|^n = H_n(u,\mu_n)+|\nabla(\log \mu_n))|^n+n|\nabla(\log \mu_n))|^{n-2}\nabla(\log \mu_n)\cdot\nabla u,$$

which, after an integration by parts, yields

$$\begin{aligned}\int_{B_R} |\nabla u + \nabla \log \mu_n|^n\,\mathrm{d}x = \int_{B_R} H_n(u,\mu_n)\,\mathrm{d}x + \int_{B_R} |\nabla(\log \mu_n))|^n\,\mathrm{d}x \\ - n\int_{B_R} u\,\Delta_n(\log \mu_n)\,\mathrm{d}x,\end{aligned}$$

where Δ_n is the n-Laplacian operator defined by (6.24). By a direct computation, we note that

$$\Delta_n(\log \mu_n) = -n^n m^{n-1}\omega_n \mu_n.$$

It follows that

$$\begin{aligned}\int_{B_R} |\nabla u + \nabla \log \mu_n|^n\,\mathrm{d}x = \int_{B_R} H_n(u,\mu_n)\,\mathrm{d}x + n^{n+1}m^{n-1}\omega_n \int_{B_R} u\,\mathrm{d}\mu_n \\ + \int_{B_R} |\nabla(\log \mu_n))|^n\,\mathrm{d}x,\end{aligned}$$

and so (6.28) becomes for all $\varepsilon > 0$,

$$\begin{aligned}\varepsilon \left[n^2 m \int_{B_R} \rho^{1/m}\,\mathrm{d}y - n^2 m \omega_n^{1/n}\theta_R\right] - \varepsilon^m \frac{n}{m}\int_{B_R} |y|^m \rho\,\mathrm{d}y \le \\ \int_{B_R} H_n(u,\mu_n)\,\mathrm{d}x + n^{n+1}m^{n-1}\omega_n \int_{B_R} u\,\mathrm{d}\mu_n + \int_{B_R} |\nabla(\log \mu_n)|^n\,\mathrm{d}x.\end{aligned} \tag{6.29}$$

Next, we focus on the l.h.s of (6.29). For convenience, we denote

$$A_\rho := n^2 m \int_{B_R} \rho^{1/m}\,\mathrm{d}y - n^2 m \omega_n^{1/n}\theta_R, \quad B_\rho := \frac{n}{m}\int_{B_R} |y|^m \rho\,\mathrm{d}y,$$

and

$$G_\rho(\varepsilon) := \varepsilon A_\rho - \varepsilon^m B_\rho.$$

Then for all $\varepsilon > 0$, (6.29) reads as

$$G_\rho(\varepsilon) \le \int_{B_R} H_n(u,\mu_n)\,\mathrm{d}y + n^{n+1}m^{n-1}\omega_n \int_{B_R} u\,\mathrm{d}\mu_n + \int_{B_R} |\nabla(\log\mu_n))|^n\,\mathrm{d}y. \tag{6.30}$$

Clearly, $G'_\rho(\varepsilon) = A_\rho - m\varepsilon^{m-1}B_\rho$, so $\max_{\varepsilon>0}[G_\varepsilon(\rho)]$ is attained at

$$\varepsilon_{max}(\rho) := \left(\frac{A_\rho}{mB_\rho}\right)^{1/(m-1)}. \tag{6.31}$$

In particular, if $\rho = \mu_n$, we have

$$\varepsilon_{max}(\mu_n) := \left(\frac{A_{\mu_n}}{mB_{\mu_n}}\right)^{1/(m-1)},$$

where

$$A_{\mu_n} = n^2 m\left(\int_{B_R} \mu_n^{1/m}\,\mathrm{d}y - \omega_n^{1/n}\theta_R\right),$$

and

$$B_{\mu_n} = \frac{n}{m\omega_n^{1/n}}\left(\int_{B_R} \mu_n^{1/m}\,\mathrm{d}y - \omega_n^{1/n}\theta_R\right).$$

Note that we have used above the relation

$$\int_{B_R} |y|^m \mu_n\,\mathrm{d}x = \left(\frac{1}{\omega_n}\right)^{1/n}\int_{B_R} \mu_n^{1/m}\,\mathrm{d}x - \theta_R. \tag{6.32}$$

Then

$$\varepsilon_{max}(\mu_n) = (nm\,\omega_n^{1/n})^{1/(m-1)}. \tag{6.33}$$

Choosing $\varepsilon = \varepsilon_{max}(\mu_n)$ in (6.30), we have

$$G_\rho(\varepsilon_{max}(\mu_n)) - \int_{B_R} |\nabla(\log\mu_n))|^n\,\mathrm{d}x \le \int_{B_R} H_n(u,\mu_n)\,\mathrm{d}x + n^{n+1}m^{n-1}\omega_n\int_{B_R} u\,\mathrm{d}\mu_n,$$

that is, after dividing by $\beta(n) = n^{n+1}m^{n-1}\omega_n$,

$$\frac{n^2 m\,\varepsilon_{max}(\mu_n)}{\beta(n)}\int_{B_R}\rho^{1/m}\,\mathrm{d}y - \frac{n(\varepsilon_{max}(\mu_n))^m}{m\beta(n)}\int_{B_R}|y|^m\rho\,\mathrm{d}y$$
$$-\frac{1}{\beta(n)}\left[\int_{B_R}|\nabla(\log\mu_n))|^n\,\mathrm{d}x + n^2 m\omega_n^{1/n}\,\varepsilon_{max}(\mu_n)\,\theta_R\right]$$
$$\leq \frac{1}{\beta(n)}\int_{B_R}H_n(u,\mu_n)\,\mathrm{d}x + \int_{B_R}u\,\mathrm{d}\mu_n = I_R(u). \tag{6.34}$$

We now simplify the l.h.s of (6.34) by using the following basic identities which can be checked by direct computations:

$$\frac{n^2 m\,\varepsilon_{max}(\mu_n)}{\beta(n)} = m(1/\omega_n)^{1/n} = \alpha(n),$$

$$\frac{n(\varepsilon_{max}(\mu_n))^m}{m\beta(n)} = 1,$$

$$n^2 m\omega_n^{1/n}\,\varepsilon_{max}(\mu_n) = m^n n^{n+1}\omega_n,$$

$$\int_{B_R}|\nabla(\log\mu_n)|^n\,\mathrm{d}x = n^n m^n\,\omega_n\int_{B_R}|y|^m\mu_n\,\mathrm{d}y,$$

$$\theta_R = \left(\frac{1}{\omega_n}\right)^{1/n}\int_{B_R}\mu_n^{1/m}\,\mathrm{d}y - \int_{B_R}|y|^m\mu_n\,\mathrm{d}y.$$

Then (6.34) yields

$$J_R(\rho) - J_R(\mu_n) \leq I_R(u)$$

for all functions u and ρ such that $u \in \mathcal{D}(B_R)$, $\int_{B_R} e^u\,\mathrm{d}\mu_n = \theta_R$ and $\int_{B_R}\rho(y)\,\mathrm{d}y = \theta_R$.
We conclude the proof by noting that the left-hand side is equal to 0 for $\rho \equiv \mu_n$, while the right-hand side is equal to 0 for $u \equiv 0$. □

From Theorem 6.1, we obtain the following n-dimensional Onofri inequality.

Theorem 6.2. *(***n-dimensional Euclidean Onofri inequality***) For any $n \geq 2$, the following holds for any $u \in D^{1,2}(\mathbb{R}^n)$,*

$$\frac{1}{\beta(n)}\int_{\mathbb{R}^n}H_n(u,\mu_n)\,dx + \int_{\mathbb{R}^n}u\,d\mu_n - \log\left(\int_{\mathbb{R}^n}e^u\,d\mu_n\right) \geq 0, \tag{6.35}$$

hence the infimum is attained at $u \equiv 0$.

Proof. Take $u \in C_c^1(\mathbb{R}^n)$ such that it has its support in a ball B_R. Let $v = u - C$ on B_R and 0 elsewhere, where C is chosen so that $\int_{B_R} e^v \,\mathrm{d}\mu_n = \mu_n(B_R)$. It follows that $\int_{\mathbb{R}^n} e^v \,\mathrm{d}\mu_n = 1$, hence applying Theorem (6.1) we get that

$$I_R(v) = \frac{1}{\beta(n)} \int_{B_R} H_n(v, \mu_n)\,\mathrm{d}x + \int_{B_R} v\,\mathrm{d}\mu_n(x) - \log \int_{\mathbb{R}^n} e^v \,\mathrm{d}\mu_n \geq 0. \quad (6.36)$$

Since $H_n(v, \mu_n) = H_n(u, \mu_n)$, then (6.36) gives

$$\frac{1}{\beta(n)} \int_{\mathbb{R}^n} H_n(u, \mu_n)\,\mathrm{d}x + \int_{\mathbb{R}^n} u(x)\,\mathrm{d}\mu_n(x) - \log\left(\int_{\mathbb{R}^n} e^u \,\mathrm{d}\mu_n\right) \geq 0.$$

□

From the proof of Theorem 6.1, we can also derive the following inequality.

Corollary 6.1. *Let $n \geq 2$ be an integer. For $v \in C_c^1(\mathbb{R}^n)$ with compact support in $B_R \subset \mathbb{R}^n$ for some $R > 0$, we have*

$$\left(\frac{1}{\omega_n}\right)^{\frac{n-1}{n}} \frac{1}{(1 + R^{\frac{n}{n-1}})^n} \int_{B_R} e^v \,dx + \frac{n-1}{n^2} \int_{B_R} |\nabla v|^n \,dx \geq \int_{B_R} \mu_n^{\frac{n-1}{n}} \,dy. \quad (6.37)$$

In particular, if $n = 2$, then (6.37) gives

$$\int_{B_R} e^v \,dx + \frac{(1 + R^2)^2 \sqrt{\pi}}{4} \int_{B_R} |\nabla v|^2 \,dx \geq \pi(1 + R^2)^2 \log(1 + R^2). \quad (6.38)$$

Proof. Choosing $\rho = \mu_n$ and $\varepsilon = \varepsilon_{max}(\mu_n)$ in (6.28), we have

$$G_{\mu_n}(\varepsilon_{max}(\mu_n)) \leq \int_{B_R} |\nabla(u + \log \mu_n)|^n \,\mathrm{d}x, \quad (6.39)$$

for any u such that $\int_{B_R} e^u \,\mathrm{d}\mu_n = \theta_R$ and $u|_{\partial B_R} = 0$. Using the computations in the proof of Theorem 6.1 and setting $m := \frac{n}{n-1}$, we have

$$G_{\mu_n}(\varepsilon_{max}(\mu_n)) = \frac{A_{\mu_n}^n}{n(mB_{\mu_n})^{n-1}} = mn\left(\int_{B_R} \mu_n^{1/m} \,\mathrm{d}y - \omega_n^{1/n}\,\theta_R\right).$$

This gives

$$mn\left(\int_{B_R} \mu_n^{1/m} \,\mathrm{d}y - \omega_n^{1/n}\,\theta_R\right) \leq \int_{B_R} |\nabla(u + \log \mu_n)|^n \,\mathrm{d}x.$$

Set $v := u + \log \mu_n - \log(\mu_n|_{\partial B_R})$. We have

$$\nabla v = \nabla(u + \log \mu_n), \quad v|_{\partial B_R} = 0, \quad \theta_R = \int_{B_R} e^u \, \mathrm{d}\mu_n = \mu_n|_{\partial B_R} \int_{B_R} e^v \, \mathrm{d}x,$$

where $\mu_n|_{\partial B_R} = \frac{1}{\omega_n(1+R^m)^n}$. Then (6.39) reads as

$$mn \left(\int_{B_R} \mu_n^{1/m} \, \mathrm{d}y - \omega_n^{1/n} \frac{1}{\omega_n(1+R^m)^n} \int_{B_R} e^v \, \mathrm{d}x \right) \leq \int_{B_R} |\nabla v|^n \, \mathrm{d}x. \quad (6.40)$$

This gives (6.37) after simplification. Using $\int_{B_R} \sqrt{\mu_n} = \sqrt{\pi} \log(1 + R^2)$ where $B_R \subset \mathbb{R}^2$, we get (6.38). □

In dimension $n = 2$, the operator H_n becomes $H_2(u, \mu_2) = |\nabla u|^2$, and Theorem (6.1) then yields the 2-dimensional Onofri inequality.

Corollary 6.2. (Duality for the 2-dimensional Euclidean Onofri inequality) *For any ball B_R of radius $R > 0$ in $\mathbb{R}^2$, consider the functionals*

$$I_R(u) = \frac{1}{16\pi} \int_{B_R} |\nabla u(x)|^2 \, dx + \int_{B_R} u(x) \, d\mu_2(x) \quad \text{on } C_c^\infty(B_R),$$

and

$$J_R(\rho) = \frac{2}{\sqrt{\pi}} \int_{B_R} \sqrt{\rho(y)} \, dy - \int_{B_R} |y|^2 \rho(y) \, dy \quad \text{on } L_+^1(\mathbb{R}^2).$$

1. *The following duality formula then holds:*

$$\sup\left\{ J_R(\rho) - J_R(\mu_2);\ \rho \in L_+^1(\mathbb{R}^2),\ \frac{1}{\mu_2(B_R)} \int_{B_R} \rho \, dy = 1 \right\} \quad (6.41)$$
$$= \inf\left\{ I_R(u);\ u \in C_c^\infty(B_R),\ \int_{\mathbb{R}^2} e^u \, d\mu_2 = 1 \right\} = 0,$$

   and the maximum on the l.h.s. (resp. the minimum on the r.h.s.) is only attained at $\rho_{max} = \mu_2$ (resp., at $u_{min} = 0$).

2. *Consequently, the Euclidean Moser-Onofri inequality also holds:*

$$\log\left(\int_{\mathbb{R}^2} e^u \, d\mu_2 \right) - \int_{\mathbb{R}^2} u \, d\mu_2 \leq \frac{1}{16\pi} \int_{\mathbb{R}^2} |\nabla u|^2 dx \quad \forall u \in \mathcal{D}(\mathbb{R}^2). \quad (6.42)$$

Remark 6.2. *It would be interesting to show that $u \equiv 0$ is the only function such that $I_R(u) = 0$, by considering the optimal transport map T that maps the probability measure $\frac{1}{\mu_2(B_R)} e^u \mu_2$ on B_R to $\frac{\mu_2}{\mu_2(B_R)}$, and arguing –by chasing back the inequalities in the proof of the duality– that $I_R(u) = 0$ implies that T is necessarily the identity map. This was the approach used in [15] to find the extremal in the Sobolev inequality.*

Bibliography

[1] M. Agueh, S. Boroushaki, N. Ghoussoub, *A dual MoserOnofri inequality and its extensions to higher dimensional spheres*, Nonlinear Annales de la faculté des sciences de Toulouse Sér. 6, 26 (2), (2017), 217-233.

[2] M. Agueh, N. Ghoussoub, and X. Kang, *Geometric inequalities via a general comparison principle for interacting gases*, Geom. Funct. Anal., 14 (2004), 215-244.

[3] T. Aubin, *Meilleures constantes dans le théorème d'inclusion de Sobolev et un théorème de Fredholm non linéaire pour la transformation conforme de la courbure scalaire*, J. Funct. Anal., 32 (1979), 148-174.

[4] G. Auchmuty, *Duality for non-convex variational principles*, J. Differential Equations, 50 (1), (1983), 80-145.

[5] G. Auchmuty, *Variational principles for operator equations and initial-value problems*, Nonlinear Anal. Theory Methods Appl., 12 (5), (1988), 531-564.

[6] V. Barbu, *A variational approach to stochastic nonlinear parabolic problems*, J. Math. Anal. Appl., 153 (1), (2011), 2-15.

[7] V. Barbu, *Optimal control approach to nonlinear diffusion equations driven by Wiener noise*, J. Optim. Theory Appl., 384 (1), (2012), 1-26.

[8] W. Beckner, *Sharp Sobolev inequalities on the sphere and the Moser-Trudinger inequality*, Anal. Math., 138 (1993), 213-242

[9] S. Boroushaki, N. Ghoussoub, *A Self-dual variational approach to stochastic partial differential equations*, arXiv:1710.01414, (2017).

[10] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. 44 (4) (1991), 375-717.

[11] H. Brézis, *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland, Amsterdam-London, (1973).

[12] H. Brézis, I. Ekeland, *Un principe variationnel associé à certaines equations paraboliques. Le cas independant du temps*, C.R. Acad. Sci. Paris Sér. A 282 (1976), 971–974.

[13] H. Brézis, I. Ekeland, *Un principe variationnel associé à certaines equations paraboliques. Le cas dependant du temps*, C.R. Acad. Sci. Paris Sér. A 282 (1976), 1197–1198.

[14] F. Browder, *Nonlinear maximal monotone operators in Banach space*, Math. Annalen 175 (1968) p. 89-113.

[15] D. Cordero-Erausquin, B. Nazaret, and C. Villani, *A mass transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities*, Adv. Math., 182 (2004), 307-332.

[16] G. Da Prato, J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, (2014).

[17] M. Del-Pino and J. Dolbeault, *Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions*, J. Math. Pures Appl. 90, 81 (2002), 847-875.

[18] J. Dolbeault, M. J. Esteban, and G. Jankowiak, *The Moser-Trudinger-Onofri inequality*, Preprint (2014).

[19] I. Ekeland, *Convexity Methods in Hamiltonian Mechanics*, Springer-Verlag, Berlin, (1990).

[20] I. Ekeland, R. Temam, *Convex Analysis and Variational problems*, Classics in Applied Mathematics, 28 SIAM (1999 Edition).

[21] L.C. Evans, *Partial Differential Equations*,Graduate Studies in Mathematics, vol. 19, Amer. Math. Soc., Providence,(1998).

[22] K. Fan, *Minimax theorems*, Proc. Natl. Acad. Sci. USA., 39 (42), (1953), 42-47.

[23] S.P. Fitzpatrick, *Representing monotone operators by convex functions*, Proc. Centre for Math. Analysis 20 (1989), p. 59-65.

[24] N. Ghoussoub, *A theory of anti-selfdual Lagrangians: Dynamical case*, C.R. Acad. Sci. 340 (4), (2005), 325-330.

[25] N. Ghoussoub, *A variational principle for nonlinear transport equations*, Communications on Pure and Applied Analysis 4 (4), (2005), 735-742.

[26] N. Ghoussoub, *Anti-self-dual Lagrangians: Variational resolutions of non-self-adjoint equations and dissipative evolutions*, AIHP-Analyse non linaire 24 (2), (2007), 171-205.

[27] N. Ghoussoub, *Anti-symmetric Hamiltonians: Variational resolution of Navier-Stokes equations and other nonlinear evolutions*, Comm. Pure & Applied Math., vol. 60, no. 5 (2007) pp. 619-653

[28] N. Ghoussoub: *Superposition of self-dual functionals for non-homogeneous boundary value problems and differential systems*, Journal of Discrete and Continuous Dynamical Systems, Vol. 21, 1 (2008), p. 71-104.

[29] N. Ghoussoub, *Self-dual partial differential systems and their variational principles*, Springer Science & Business Media, (2008).

[30] N. Ghoussoub, *A variational theory for monotone vector fields*, Fixed Point Theory Appl. 4 (1), (2008), 107-135.

[31] N. Ghoussoub, A. Moameni, *Anti-symmetric Hamiltonians (II): Variational resolution of Navier-Stokes equations and other nonlinear evolutions*, AIHP-Analyse non linéaire, 26 (2009), no. 1, 223-255.

[32] N. Ghoussoub, A. Moameni, R. Saiz, *Homogenization of maximal monotone vector fields via selfdual variational calculus*, Advanced Nonlinear Studies 11 (2), (2011), 323-360.

[33] N. Ghoussoub, and A. Moradifam, *Functional inequalities: new perspectives and new applications*, Mathematical Surveys and Monographs, Vol. 187 (2013) (American Mathematical Society, Providence, RI).

[34] N. Ghoussoub, L. Tzou. *A variational principle for gradient flows*, Math. Annalen, Vol 30, 3 (2004) p. 519-549.

[35] N. Ghoussoub, L. Tzou, *Anti-selfdual Lagrangians II: Unbounded non self-adjoint operators and evolution equations*, Annali di Matematica Pura ed Applicata, Vol 187, 2 (2008) p. 323-352

[36] C. W. Hong, *A best constant and the Gaussian curvature*, Proc. Amer. Math. Soc. 97 (4), (1986), 737-747.

[37] F. Maggi, C. Villani, *Balls have the worst best Sobolev inequalities. Part II: Variants and extensions*, Calc. Var. PDE. 31 (1) (2008), 47-74.

[38] R. J. McCann, *A convexity principle for interacting gases*, Adv. Math. 128 (1997) 153-179.

[39] J. Moser, *A sharp form of an inequality by N. Trudinger*, Indianna Univ. Math. J. 20 (1970/71), 1077-1092.

[40] B. Nayroles. Deux théorèmes de minimum pour certains systèmes dissipatifs. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aiv, A1035–A1038, (1976).

[41] B. Nayroles. Un théorème de minimum pour certains systèmes dissipatifs. Variante hilbertienne. *Travaux Sém. Anal. Convexe*, 6(Exp. 2):22, (1976).

[42] E. Onofri, *On the positivity of the effective action in a theory of random surfaces*, Comm. Math. Phys. 86, (1982), 321-326.

[43] B. Osgood, R. Phillips, and P. Sarnak, *Extremals of determinants of Laplacians*, J. Funct. Anal. Phys. 80 (1), (1988), 148-211.

[44] E. Pardoux, *Thesis*, L'universit de Paris Sud, Centre d'Orsay (1975).

[45] E. Pardoux, *Stochastic Partial Differential Equations*, Lectures given in Fudan University, Shanghai. Published by Marseille, France (2007).

[46] E. Pardoux, A. Rascanu, *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*, Stochastic Modelling and Applied Probability, Springer (2014)

[47] R. R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture notes in mathematics, 1364 (1989).

[48] C. Prévôt, M. Röckner, *A concise course on stochastic partial differential equations*, Berlin: Springer 1905, (2007).

[49] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, 28 (1970).

[50] L. C. G. Rogers, D. Williams *Diffusions, Markov processes and martingales: Vol 2, It calculus.*, Cambridge university press, (2000).

[51] N. Trudinger, *On imbeddings into Orlicz spaces and some applications*, Indianna Univ. Math. J. 17, (1968), 473-483.

[52] C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Math, 58, AMS, Providence, RI, (2003).

[53] J. B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour, Lecture Notes in Math. 1180, Springer, Berlin, (1986), 265-439